І.Ф.СКІЦЬКО, О.І.СКІЦЬКО

ФІЗИКА

Затверджено Вченою радою КПІ імені Ігоря Сікорського як підручник для студентів, які навчаються за технічними спеціальностями

Київ, КПІ ім.Ігоря Сікорського
2017

К 937
УДК 534.1 – 535.8; 537.1 – 539.18
Ф 48

Гриф надано Вченою радою Національного технічного університету України "КПІ" імені Ігоря Сікорського,
протокол №2 від 13 лютого 2017 року

Автори: Скіцько Іван Федорович, доцент, кандидат.фіз.-мат.наук,
Скіцько Олексій Іванович, с.н.с., к.т.н.

Відповідальний редактор
Авраменко Андрій Олександрович, заступник директора з наукової роботи Інституту технічної теплофізики НАН України, член-кор. НАН України, професор, доктор техн.наук

Рецензенти:
Д.А.Гаврюшенко, професор кафедри молекулярної фізики Київського Національного Університету імені Тараса Шевченка, доктор фіз.-мат.наук, професор
І.Р.Зачек, професор кафедри фізики Національного університету "Львівська політехніка", доктор фіз.-мат.наук, професор

Ф 48 Фізика (Фізика для інженерів): Підручник / І.Ф.Скіцько, О.І Скіцько: – Київ: НТУУ КПІ ім.Ігоря Сікорського, 2017.–513 с.

Підручник написаний у відповідності з програмою по фізиці для студентів вищих технічних навчальних закладів, які вивчають нормативну навчальну дисципліну "Фізика" з циклу математичної та природничо-наукової підготовки за спеціальностями 125 "Кібербезпека " і 172 "Телекомунікації та радіотехніка". Він складається із двадцяти трьох розділів, в яких викладені фізичні основи електромагнетизму, коливань та хвиль, хвильової оптики, квантової механіки, фізики твердого тіла та контактних явищ, оптичного волоконного зв'язку та будови атомного ядра.

Для студентів вищих навчальних закладів технічного профілю.

УДК 534.1–535.8; 537.1–539.18
© Скіцько І.Ф., Скіцько О.І, НТУУ "КПІ" ім.Ігоря Сікорського, 2017

ПЕРЕДМОВА.....**9**

Розділ 1.Електростатичне поле.....**11**

§1.1.Електричний заряд.....11

§1.2.Закон Кулона.....13

§1.3.Напруженість електричного поля.....14

§1.4.Принципи суперпозиції електричних полів.....15

§1.5.Графічне зображення електричних полів.....15

§1.6 Потік вектора напруженості електричного поля.....16

§1.7.Теорема Остроградського-Гаусса.....17

§1.8.Застосування теореми Остроградського-Гаусса до розрахунку електричних полів.....20

§1.9.Робота сил електростатичного поля.....23

§1.10.Різниця потенціалів. Потенціал. Еквіпотенціальні поверхні.....24

§1.11.Циркуляція вектора напруженості поля.....26

§1.12.Зв'язок між напруженістю та потенціалом електростатичного поля.....27

§1.13.Рівняння Пуассона (теорема Остроградського-Гаусса в диференціальній формі).....28

Контрольні питання.....29

Розділ 2. Поляризація діелектриків і поле всередині діелектрика.....**29**

§2.1.Провідники, напівпровідники та діелектрики.....29

§2.2. Електричний диполь в однорідному і неоднорідному електричних полях.....31

§2.3. Поляризація діелектриків та її типи. Вектор поляризації (поляризованість), діелектрична сприйнятливість.....33

§2.4. Електричне поле всередині діелектрика. Діелектрична проникність та її фізичний зміст. Зв'язок між діелектричною сприйнятливістю та діелектричною проникністю.....37

§2.5.Електричне поле на межі двох діелектриків.....39

§2.6.Особливі властивості деяких діелектриків.Сегнетоелектрики. П'єзоелектрики. Електрети.....40

Контрольні питання.....47

Розділ 3.Провідники в електричному полі.....**47**

§3.1.Розподіл зарядів, напруженість та потенціал електричного поля у провіднику.....47

§3.2. Електроємність відокремленого провідника.....51

§3.3. Конденсатори.....53

§3.4 Енергія зарядженого відокремленого провідника та конденсатора. Енергія електростатичного поля.....57

Контрольні питання.....59

Розділ 4. Постійний електричний струм.....**59**

§4.1.Електричний струм та його види. Струм провідності та умови його існування.....59

§4.2.Закон Ома та Джоуля–Ленца в

інтегральній та диференціальній формах.....62
§4.5.Електрорушійна сила.....66
§4.4.Розрахунок розгалужених електричних кіл за допомогою правил Кірхгофа.....71
Контрольні питання.....72

Розділ 5. Магнітне поле.....**72**

§5.1.Взаємодія елементів струму. Закон Ампера. Магнітне поле як форма існування матерії.....72
§5.2.Взаємодія елементів струму.....74
§5.3.Індукція магнітного поля. Закон Біо-Савара-Лапласа.....77
§5.4.Застосування закону Біо-Савара-Лапласа до розрахунку магнітних полів:.....80
1.Магнітне поле прямолінійного провідника зі струмом.....80
2.Поле колового витка зі струмом.....81
3. Магнітне поле соленоїда.....82
4. Магнітне поле електричного заряду, який рухається.....84
§5.5.Циркуляція вектора індукції магнітного поля або закон повного струму.....85
§5.6.Теорема Остроградського-Гаусса для магнітного поля.....87
§5.7.Магнітне поле довгого соленоїда, тороїда та прямого проводу.....88
§5.8.Контур зі струмом у магнітному полі.....90
§5.9.Робота при переміщенні провідника зі струмом в магнітному полі.....93
§5.10.Дія електричного і магнітного полів на рухомий електричний заряд:
-сила Лоренца.....95
-ефект Холла.....97
-визначення питомого заряду електрона методом магнетрона.....98
Контрольні питання.....99

Розділ 6.Постійне магнітне поле в речовині.....**100**

§6.1.Намагнічування середовищ.....100
§6.2.Магнітомеханічні явища.....101
§6.3.Закон повного струму для магнітного поля в речовині. Напруженість магнітного поля. Магнітна сприйнятливість і проникність.....103
§6.4.Магнітне поле на межі двох магнетиків.....105
§6.5.Класифікація магнетиків.....107
§6.5.Діа - і парамагнетизм.....107
§6.7.Феромагнетизм.....111
§6.8.Магнітні матеріали. Ферити.....117
§6.9.Пояснення феромагнетизму.....118
§6.10.Процеси намагнічування феромагнетика.....121
Контрольні питання.....123

Розділ 7.Електромагнітна індукція.....**124**

§7.1. Явище електромагнітної індукції.....124
§7.2. Закон Ленца.....127
§7.3.Основний закон електромагнітної індукції.....128
§7.4.Пояснення закону електромагнітної індукції.....129
§7.5. Самоіндукція. Індуктивність.....133
§7.6. Взаємна індукція.....136
§7.7. Власна енергія струму.....137
§7.8. Енергія магнітного поля.....139
§ 7.9.Взаємна енергія двох струмів.....139
§7.10.Вихрові струми.....140

§7.11.Витіснення змінного струму (скін-ефект).....142
Контрольні питання.....144

Розділ 8. Рівняння Максвелла.....**144**

§8.1. Перше рівняння Максвелла.....144
§8.2. Струм зміщення. Друге рівняння Максвелла.....145
§8.3.Рівняння Максвелла.....148
§8.4.Рівняння Максвелла в диференціальній формі.....149
§8.5.Значення теорії Максвелла.....151
Контрольні питання.....152

Розділ 9. Коливання.....**152**

§9.1.Види коливань. Загальні поняття.....152
§9.2.Власні електричні коливання.....154
§9.3.Рівняння власних коливань. Коливання при відсутності згасання.....155
§9.4.Геометричне зображення гармонічних коливань. Додавання гармонічних коливань однакового напрямку і однакової частоти.....157
§9.5. Додавання гармонічних коливань однакового напрямку з різними частотами. Биття.....158
§9.6. Додавання взаємно перпендикулярних коливань.....159
§9.7.Диференційне рівняння згасаючих коливань і його розв'язання.....161
§9.8.Вимушені електричні коливання. Резонанс.....165
§9.9. Параметричний резонанс.....170
§9.10. Змінний струм.....172
§9.11.Автоколивання.....174
Контрольні питання.....176

Розділ 10.Пружні хвилі.....**176**

§10.1.Утворення хвиль у пружному середовищі. Поздовжні і поперечні хвилі.....177
§10.2.Рівняння плоскої і сферичної хвиль..179
§ 10.3. Хвильове рівняння.....180
§10.4.Енергія пружної хвилі.....181
§10.5.Групова швидкість поширення хвилі183
§10.6.Інтерференція хвиль.....185
§10.7.Звук.....189
§10.8.Ефект Доплера для звукової хвилі.....192
Контрольні питання.....195

Розділ 11.Електромагнітні хвилі.....**195**

§11.1.Загальні відомості про електромагнітні хвилі.....195
§11.2.Доведення необхідності існування електромагнітних хвиль.....196
§11.3. Властивості електромагнітних хвиль.....198
§11.4.Енергія електромагнітних хвиль. Потік енергії. Вектор Пойнтинга.....200
§11.5. Генерування і експериментальне дослідження електромагнітних хвиль (радіохвиль).....201
§11.6.Випромінювання елементарного диполя.....204
Контрольні питання.....208
коливань

Розділ 12. Інтерференція електромагнітних хвиль.....**208**

§ 12.1.Основні відомості про хвилі оптичного діапазону. Інтенсивність електромагнітних хвиль.....208

§12.2. Інтерференція світла.....210
§12.3. Розрахунок інтерференційної картини від двох когерентних джерел.....213
§12.4. Способи спостереження інтерференції світла.....214
§12.5. Інтерференція квазімонохроматичного світла. Часова когерентність.....219
§12.6. Роль кінцевих розмірів джерела світла. Просторова когерентність.....221
§12.7. Практичне застосування інтерференції світла.....224
Контрольні питання.....227

Розділ 13. Дифракція електромагнітних хвиль.....**227**

§13.1. Дифракція хвиль.....227
§13.2.Принцип Гюйгенса–Френеля.....228
§13.3.Метод зон Френеля.Прямолінійне поширення світла.....230
§13.4.Дифракція Френеля на круглому отворі та диску.....234
§13.5.Дифракція Френеля на прямолінійному краю напівплощини.....237
§13.6.Дифракція Фраунгофера.....240
§13.7.Дифракційна решітка.....244
§13.8.Дифракція рентгенівського випромінювання.....248
Контрольні питання.....250

Розділ 14. Поляризація світла.....**250**

§14.1.Природне і поляризоване світло. Закон Малюса.....250
§14.2. Поляризація при відбиванні і заломленні.....254
§14.3. Подвійне променезаломлення.....258
§14.4.Поляризаційні призми і поляроїди....262
§14.5.Штучна оптична анізотропія.....264
§14.6.Обертання площини поляризації.....267
Контрольні питання.....270

Розділ 15. Взаємодія електромагнітних хвиль з речовиною.....**270**

§15.1.Дисперсія світла. Області нормальної і аномальної дисперсії..... 270
§15.2.Групова швидкість світла.....273
§15.3. Поглинання світла.....275
§15.4. Розсіювання світла.....276
§15.5.Ефект Доплера в оптичних явищах...278
Контрольні питання.....280

Розділ 16. Квантова природа випромінювання..... **280**

§16.1.Теплове випромінювання. Його характеристики. Абсолютно чорне тіло. Закон Кірхгофа.....280
§16.2. Закон Стефана-Больцмана. Закон Віна. Формула Релея Джінса.....286
§16.3. Квантова гіпотеза і формула Планка.....289
§16.4. Оптична пірометрія.....292
§16.5. Зовнішній фотоефект і його закони. Рівняння Ейнштейна для зовнішнього фотоефекту.....294
§16.6.Маса та імпульс фотона.....298
§16.7.Тиск світла.....301
§16.8.Ефект Комптона.....303
Контрольні питання.....305

Розділ 17. Будова атома.....**305**

§17.1. Дослід Резерфорда. Планетарна модель атома.....305
§17.2. Атомні спектри.....309
§17.3.Постулати Бора. Теорія атома водню.....312
§17.4. Подальший розвиток теорії Бора. Квантування Бора-Вільсона-Зоммерфельда.....317
§17.5. Просторове квантування.....320
§17.6. Ефект Зеємана.....322
§17.7. Дослід Штерна і Герлаха.....323
§17.8.Власний момент електрона.....324
§17.9.Квантові стани електронів в атомі. Квантові числа.....326
§17.10. Електронна оболонка атома. Принцип Паулі.....327
§17.11. Періодична система.....329
§17.12.Рентгенівські спектри.....330
Контрольні питання.....334

Розділ 18.Елементи квантової механіки.....**334**

§18.1.Протиріччя хвиля – частинка.....334
§18.2.Хвилі де Бройля.....335
§18.3.Експериментальне підтвердження гіпотези де Бройля.....337
§18.4. Співвідношення невизначеностей, як прояв корпускулярно-хвильового дуалізму властивостей матерії.....339
18.5.Статистичне тлумачення хвиль де Бройля.....345
§18.6. Ймовірність місцезнаходження мікрочастинки.....347
§18.7.Рівняння Шредінгера. Принцип причинності в квантовій механіці.....348
§18.8.Рух вільної частинки. Частинка в прямокутній потенціальній ямі.....350
§18.9.Розв'язок рівняння Шредінгера для потенціального бар'єру.....352
§18.10. Холодна емісія електронів із металу.....360
Контрольні питання.....361

Розділ 19.Основи зонної теорії твердих тіл.....**361**

§19.1.Тверде тіло. Аморфні і кристалічні тверді тіла.....361
§19.2. Структура кристалів.....362
§19.3. Основні типи міжатомної взаємодії в твердих тілах.....364
§19.4. Рідкі кристали. Сили Ван дер Ваальса.....367
§19.5. Класифікація твердих тіл за величиною електропровідності та її залежності від температури.....370
§19.6. Енергетичні зони в кристалах.....373
§19.7.Статистика електронів в металах.....378
§19.8.Електропровідність металів та її залежність від температури. Надпровідність 384
§19.9.Статистика електронів і дірок в напівпровідниках.....391
§19.10. Електропровідність напівпровідників та її залежність від температури....393
§19.11. Внутрішній фотоефект.....399
§19.12.Дрейф і дифузія носіїв заряду.....405
§19.13.Ефект Ганна.....411
Контрольні питання.....417

Розділ 20.Термоелектронна емісія.....**418**

§20.1.Історія відкриття термоелектронної емісії.....418
§20.2.Робота виходу електрона із металу...419

§20.3.Закони термоелектронної емісії.....421
§20.4. Контакт двох металів.....424
Контрольні питання.....426

Розділ 21. Контактні явища.....**427**

§21.1 Проникнення зовнішнього електричного поля в метал, діелектрик, напівпровідник.....427
§21.2.Розподіл поля в бар'єрі. Ширина бар'єру.....429
§21.3. *p–n* перехід та його властивості:
-способи отримання *p–n* переходів ...432
-потенціальний бар'єр в області контакту *p*- і *n*- напівпровідників...436
-розподіл електричного поля в бар'єрі. Ширина бар'єру.....439
-зворотний струм.....441
-пряме зміщення.....442
§ 21.4. Ємність електронно–діркового переходу.....444
-варікапи.....446
§ 21.5.Фотодіоди і світлодіоди.....447
§21.6. Інші види контактів.....452
Контрольні питання.....455

Розділ 22. Елементи квантової електроніки.....**456**

§22.1.Фізичні основи квантової електроніки.....456
§22.2. Лазери.....461
§22.3.Властивості лазерного випромінювання.....466
§22.4.Особливості конструкції і роботи деяких типів лазерів.....467
§22.5.Фізичні основи волоконно–оптичних ліній зв'язку.....472
Контрольні питання.....481

Розділ 23. Атомні ядра.....**481**

§23.1. Атомні ядра та їх характеристики....481
§23.2 Ядерні сили.....483
§23.3.Енергія зв'язку ядра 485
§23.4.Моделі атомного ядра.....487
§23.5.Радіоактивність. Основний закон радіоактивного перетворення атомних ядер.....488
§23.6.Ядерні реакції.....492
§ 23.7.Поділ ядер.....493
§23.8. Ланцюгова реакція поділу. Ядерний реактор.....496
§23.9.Реакції синтезу легких ядер.....498
Контрольні питання.....501

Предметний покажчик.....**503**
Список використаної літератури.....**512**

Передмова

Курс фізики у вищому навчальному закладі належить до фундаментальної дисципліни, яка складає основу теоретичної підготовки студентів інженерних спеціальностей, виконує роль тієї бази, без якої не можлива успішна діяльність інженера в будь-якій галузі сучасного виробництва та яка формує його світогляд. Запропонований курс фізики розрахований на двохсеместрове навчання студентів, які спеціалізуються в технічних науках. Ми намагались створити такий курс, в якому особлива увага приділена основам фізики, що використовуються в техніці зв'язку.

Сьогодні швидкими темпами змінюється техніка і технології. Використовуються нові матеріали і технології для створення надсучасних технічних пристроїв. Відбувається перехід від мікроелектроніки до наноелектроніки. Надто цей перехід помітний в області сучасного зв'язку. З'явилися нові спеціальності, як наприклад, "Кібербезпека" та інші. Все це вимагає нових підходів у викладанні й наповнюваності змісту загального курсу фізики.

Знання основ загальної фізики, особливо фізики металів і напівпровідників, фізики контактних явищ металів та напівпровідників, фізичних основ передачі інформації оптичними волокнами дуже важливі для оволодіння сучасними принципами побудови й функціонування засобів захисту інформації та телекомунікаційного обладнання. Тому такі розділи курсу фізики включені до відповідних модулей освітньо−професійних програм підготовки бакалаврів за спеціальностями 125 "Кібербезпека" та 172 "Телекомунікації та радіотехніка". Все це зумовило написання даного підручника: "ФІЗИКА" (Фізика для інженерів).

Для написання підручника використана література, список якої наводиться в наприкінці книги. Головним у відборі матеріалу підручника покладено діючу сьогодні програму з курсу фізики для вище зазначених спеціальностей, а також матеріали курсу лекцій, що на протязі багатьох років читаються слухачам в Державному закладі "Інститут спеціального зв'язку та захисту інформації НТУУ "КПІ" імені Ігоря Сікорського". Підручник складається з двадцяти трьох розділів, в яких викладено фізичні основи електростатики, постійного струму, електромагнетизму, коливань та хвиль, хвильової оптики, квантової природи випромінювання, будови атома та атомного ядра, квантової механіки, фізики твердого тіла та контактних явищ, термоелектронної емісії, квантової електроніки,

розповсюдження світла оптичними волокнами. Матеріал підручника максимально забезпечує теоретичні основи спеціальних курсів для вище зазначених спеціальностей. Особливу увагу приділено визначенню вихідних понять та їх походження.

Автори намагались дотримуватись лаконічного стилю викладу матеріалу, що одночасно поєднував би основи класичної й сучасної фізики, зробити акценти у тексті на основних поняттях і законах, які повинні сприяти кращому засвоєнню матеріалу. Вирішальну роль у роботі над курсом мало постійне активне спілкуання зі студентами на лекціях, практичних і лабораторних заняттях, консультаціях, заліках та іспитах. Неможливо написати підручник і знаходитись у відриві від тих, кому він призначається, спілкуватись з ними тільки в односторонньому порядку на лекціях. У викладанні, як і в будь-якій іншій справі, необхідний експеримент. Із декількох можливих способів викладання того чи іншого питання необхідно надати перевагу тому, який в ході навчання дає кращий результат.

При написанні підручника ми ставили за мету логічно підвести матеріал курсу до такого об'єму інформації, щоб студент зрозумів природу виникнення особливих властивостей контакту двох напівпровідників, контакту метал-напіпровідник, механізм розповсюдження та способи введення й виведення оптичного сигналу в оптоволокнах, принципів роботи оптичних квантових підсилювачів і генераторів, фотоприймачів (пікселів). Розуміння цих питань дуже важливе для становлення інженера-зв'язківця. Автори хотіли представити фізику не як деякий об'єм інформації, який необхідно засвоїти, а як розумну, логічну й красиву науку. Фізику необхідно вивчати під гаслом "Ніякого зазубрювання!" Замість пам'яті формальної повинна використовуватись пам'ять логічна – запам'ятовування повинно досягатись через глибоке розуміння.

Використаний математичний апарат забезпечує достатній рівень строгості викладення матеріалу курсу й відповідає математичній підготовці студентів перших курсів вищих технічних навчальних закладів

Розділ 1.Електростатичне поле

§1.1.Електричний заряд

Сучасній фізиці відомо чотири види фундаментальних взаємодій: гравітаційна, електромагнітна, сильна, слаба. ***Сильна*** взаємодія проявляється між частинками в ядерних масштабах (10^{-15}м). ***Слаба*** – ще більш короткодіюча (порядку $2\cdot10^{-18}$ м), проявляється між елементарними частинками і відіграє важливу роль при їх взаємних перетвореннях. ***Гравітаційна взаємодія*** (тяжіння) універсальна взаємодія між будь-якими тілами. Якщо взаємодія відносно слаба і тіла рухаються повільно (порівняно із швидкістю світла $c = 3\cdot10^8$ м/с), то справедливий закон всесвітнього тяжіння Ньютона (1867р.): *дві будь-які матеріальні частинки з масами* m_1 і m_2 *притягуються у напрямку одна до одної з силою F, яка прямо пропорційна добутку мас і обернено пропорційна квадрату відстані* r *між ними.*

$$F = G\frac{m_1 m_2}{r^2}, \qquad (1.1)$$

(під матеріальними частинками розуміють будь-які тіла за умови, що їхні лінійні розміри набагато менші за відстань між ними). Коефіцієнт пропорційності G називають ***гравітаційною сталою,*** числове значення якої вперше було визначене Г. Кавендішем (1798р.). За сучасними даними $G = 6{,}6745\cdot10^{-11}\,\text{м}^3/(\text{кг}\cdot\text{с}^2)$. Згідно закону (1.1) сила F залежить тільки від положення частинок у даний момент часу. Тобто гравітаційна взаємодія розповсюджується миттєво. *Сила* (1.1) *завжди є силою притягання.*

Електромагнітна взаємодія на багато порядків інтенсивніша гравітаційної й слабкої та є далекодіючою, на противагу сильній взаємодії. Вона найбільш часто зустрічається й багата проявами в мега–, макро– і мікросвіті. Вона обумовлена участю електромагнітного поля. Електромагнітне поле або поглинається, або випромінюється за такої взаємодії, що і обумовлює взаємодію між тілами. *Тому, якщо частинки (тіла) взаємодіють між собою із силами, які повільно зменшуються із збільшенням відстані між ними й набагато перевищують силу всесвітнього тяжіння* (1.1), *то говорять, що ці тіла мають електричний заряд.* Можна сказати, що електричний заряд це фізична величина, яка визначає інтенсивність електромагнітної взаємодії, подібно тому, як маса визначає інтенсивність гравітаційної взаємодії.

Електричний заряд зв'язаний з матеріальним носієм і є джерелом електромагнітного поля. Електричний заряд є внутрішньою характеристикою елементарних частинок і визначає їх електромагнітну взаємодію. *Вся сукупність електричних і магнітних явищ є проявом існування, руху та взаємодії електричних зарядів.*

Електромагнітна взаємодія є далекодіюча й може приводити як до притягання так і до відштовхування між тілами на відміну від гравітаційної взаємодії. Це обумовлює існування двох різнойменних електричних зарядів: позитивного і негативного. Різнойменні заряди притягуються, однойменні відштовхуються. За позитивні приймають ті заряди, які виникають на скляній паличці при натиранні її шовком, а негативні – заряди, які виникають на ебонітовій паличці (або бурштиновій (янтарній)) при натиранні її хутром. В природі частіше всього

зустрічаєтьсявзаємна компенсація позитивних і негативних зарядів (через силу притягання) на макроскопічних тілах, як результат, тіла стають електрично нейтральними й не проявляють електричної взаємодії.

Численними експериментами було встановлено закон збереження електричних зарядів: ***сумарний заряд електрично ізольованої системи залишається незмінним. Заряди можуть лише передаватись від одного тіла системи до іншого або зміщуватись в середині даного тіла.***

Примітка. Система називається електрично ізольованою, якщо через поверхню, яка обмежує систему, не можуть проникати електричні заряди.

Електричні заряди можуть зникати і виникати знову, але завжди зникають або виникають два електричні заряди протилежних знаків. Величина заряду, яка вимірюється в різних інерціальних системах відліку, є однаковою. Це означає, що електричний заряд є релятивістські інваріантним, тобто його величина не залежить від того, рухається він, чи знаходиться в стані спокою.

Думку про те, що електричний заряд дискретний вперше висловив Б. Франклін у 1749р. Перші обгрунтовані припущення про те, що електрика не є неперервною "рідиною", а складається з особливих частинок, були зроблені на основі праць М. Фарадея з дослідження явища електролізу. Цю найменшу порцію електрики, яку часто називають електричним зарядом або "атомом електрики", ірландський фізик Г. Стоней в 1891р. назвав електроном. Пряме експериментальне вимірювання елементарного заряду виконав в 1909 р. американський фізик Р. Міллікен. Він встановив, що заряд q є цілим кратним однієї і тієї ж самої величини $|e|$, тобто $q = n|e|$, де $n = \pm 1, \pm 2, \ldots$. Експериментально було визначено, що $|e| = 1{,}6 \cdot 10^{-19}$ Кл. Ще в 1897 р. Дж. Томсон визначив питомий заряд електрона $e/m_e = 1{,}7588 \cdot 10^{11}$Кл/кг. Звідси маса електрона $m_e = 9{,}0971 \cdot 10^{-31}$ кг. Отже електрон є матеріальним носієм найменшої маси й найменшого електричного заряду в природі. Електричний заряд електрона умовились вважати негативним відповідно до більш ранньої угоди вважати негативним заряд наелектризованого бурштину (янтаря) (грецьке electron – бурштин (янтар)).

Носієм позитивного елементарного електричного заряду є протон. За сучасними експериментальними даними позитивний заряд протона дорівнює за абсолютним значенням негативному елементарному заряду електрона з відносною точністю 10^{-21}. Важко пояснити чому заряди електрона й протона однакові з такою фантастичною точністю. *Мабуть, квантування заряду є тайним і універсальним законом природи.*

Дослід показує, що заряди всіх елементарних частинок є точно однаковими. Факт квантування заряду виходить, звичайно, за рамки класичного електромагнетизму. Як правило, квантування ми не враховуємо й чинимо так, начебто наші точкові заряди q можуть мати будь-яку величину. Це не приводить до труднощів. Однак корисно згадати, що класична теорія не в стані пояснити структури елементарних частинок. (Не можна, між іншим, стверджувати, що сучасна квантова теорія може це зробити!). Що утримує електрон від розпаду – таке ж таємниче, як і те, що визначає точну величину його заряду. Тут повинні існувати якісь сили, відмінні від електричних сил, оскільки електростатичні

сили, ,які діють між різними частинами електрона, приводять до відштовхування.

§1.2.Закон Кулона

Електростатика займається вивченням електричних полів нерухомих зарядів. Спочатку будемо вивчати електростатичні поля у вакуумі. Основний кількісний закон електростатики було відкрито Кулоном в 1785 р. Він формулюється таким чином.

Сила взаємодії F двох точкових зарядів у вакуумі направлена вздовж прямої, яка з'єднує ці заряди, пропорційна добутку їх величин q_1 *і* q_2 *і обернено пропорційна квадрату відстані між ними* r. Вона є силою притягання, якщо знаки зарядів різні і силою відштовхування, якщо знаки зарядів однакові. Математично:

$$F = k\frac{q_1 q_2}{r^2}, \qquad (1.2)$$

де k – коефіцієнт пропорційності.

Точковість зарядів в законі Кулона означає, що лінійні розміри тіл, на яких зосереджені ці заряди, мізерно малі порівняно з відстанню між ними. Силу Кулона називають центральною силою, тому що вона напрямлена вздовж прямої, що з'єднує заряди.

Закон Кулона можна записати у векторній формі:

$$\vec{F} = k\frac{q_1 q_2}{r^2}\vec{e}_r, \qquad (1.3)$$

де $\vec{e}_r$ –одиничний вектор, що має напрямок від заряду q_1 до заряду q_2, якщо визначається сила, яка діє на заряд q_2 зі сторони заряду q_1 (рис.1.1,*а*) і навпаки, $\vec{e}_r$ має напрямок від заряду q_2 до заряду q_1, якщо визначається сила, яка діє на заряд q_1 зі сторони заряду q_2 (рис.1.1,*б*).

Суттєвим фізичним змістом закону (1.2) є твердження про обернену залежність сили взаємодії зарядів від квадрату відстані між ними. В цьому плані закон (1.2) дуже корелює із законом (1.1). Дослідні дані показують, що закон Кулона справедливий для відстаней від 10^{-15}м до, принаймні, декількох кілометрів. Є всі підстави вважати, що для відстаней, які менші 10^{-16}м, закон Кулона перестає бути справедливим.

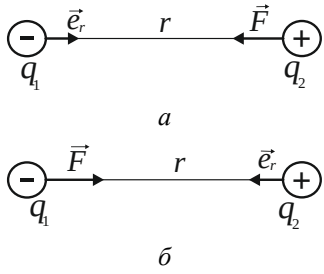


Рис.1.1.

У системі СІ для зарядів у вакуумі коефіцієнт k у формулі (1.2) записують у формі

$$k = 1/(4\pi\varepsilon_o), \qquad (1.4)$$

де $\varepsilon_0 = 8{,}85\cdot 10^{-12}$ (Кл^2/(Н$\cdot$м^2 = Ф/м)– електрична стала і, отже, $k = 9\cdot 10^9$ Н$\cdot$м^2/Кл^2. Множник 4π у виразі (1.4) відображає сферичну симетрію електростатичного поля точкового заряду, оскільки величина 4π числово дорівнює повному тілесному куту в стерадіанах.

Закон Кулона ($F\sim 1/r^2$) піддавався експериментальній перевірці багато разів, аж до наших днів. Це пов'язано з тим, що було побудовано квантову електродинаміку, яка принципово по-іншому, ніж класична теорія, пояснювала взаємодію зарядів. Але це не значить, що закони класичної електродинаміки було повністю відкинуто. Взаємовідносини квантових і класичних законів визначаються

принципом відповідності: класичні закони є граничним випадком квантових, і, значить, між ними повинен існувати зв'язок. Одним із проявів цього зв'язку є твердження: якщо закон Кулона не є точним законом «обернених квадратів» ($F \sim 1/r^2$), то маса спокою фотона m_λ повинна відрізнятися від нуля. Необхідно підкреслити, що питання про рівність нулю m_λ це питання експериментальне.

Оскільки ряд важливих положень сучасної фізики базуються на припущенні про те, що m_λ =0, то проблема експериментального підтвердження цього припущення повинна розглядатись як принципово важлива. Дослідна перевірка закону Кулона – один із шляхів її розв'язку. Про експериментальну перевірку того факту, що маса спокою фотона m_λ= 0, йдеться в §16.6. Досліди, проведені в земних умовах, а також оцінки, зроблені шляхом аналізу астрофізичних даних, не дають підстав вважати масу спокою фотона відмінною від нуля, і тому, закон взаємодії точкових електричних зарядів зберігає статус точного закону «обернених квадратів».

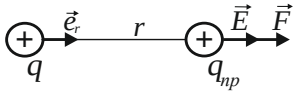
§1.3.Напруженість електричного поля

За сучасними поглядами матеріальним носієм взаємодії нерухомих зарядів є електричне поле. Основною ознакою наявності електричного поля є те, що на будь-який заряд, внесений у це поле, діє сила. Розкриття властивостей поля – одне із найважливіших завдань фізики. Зараз ми розглядатимемо стаціонарні електричні поля, тобто такі поля, що не змінюються з часом і створюються нерухомими електричними зарядами. Такі поля називаються *електростатичними*.

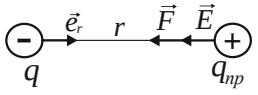
Однак при вивченні змінного електромагнітного поля і особливо електромагнітних хвиль ми переконуємось, що поле може існувати без зарядів (електромагнітні хвилі) і що поняття поля має глибокий фізичний зміст: електромагнітне поле є об'єктивною реальністю. Поле реальне так само, як і речовина і як речовина є одним із видів матерії.

Властивості електричного поля вивчають за допомогою пробного точкового позитивного електричного заряду $q_{пр}$. Пробний заряд мусить бути досить малим, щоб запобігати спотворенню початкового характеру поля. Його присутність не повинна змінити розподіл зарядів, а також не повинна суттєво вплинути на результати вимірювань.

Якщо в поле, що створюється зарядом q, помістити пробний заряд $q_{пр}$, то на нього буде діяти сила (1.3), яка пропорційна до величини $q_{пр}$. Тому ця сила не може бути характеристикою самого поля. Але величина, яка дорівнює відношенню $\vec{F}/q_{пр}$ (питома сила, сила що діє на одиничний позитивний заряд) може бути *силовою характеристикою* поля.



а



б

Рис.1.2.

Тому векторна величина

$$\vec{E} = \vec{F}/q_{пр} \qquad (1.5)$$

називається *напруженістю електричного поля.* Враховуючи закон Кулона (1.3) та відношення (1.5), отримаємо формулу для напруженості поля точкового заряду q на відстані *r*:

$$\vec{E} = \frac{kq}{r^2}\vec{e}_r \cdot \qquad (1.6)$$

Звідси видно, що поле точкового заряду є центрально-симетричним.

За напрямок вектора напруженості $\vec{E}$ беруть напрямок сили, з якою поле діє на пробний позитивний заряд, вміщений у певну точку поля (рис.1.2).

У системі СІ одиниця напруженості електричного поля 1Н/Кл – це напруженість такого поля, яке діє з силою 1Н на точковий заряд 1 Кл. Частіше напруженість поля вимірюють у вольтах на метр (В/м).

§1.4.Принцип суперпозиції електричних полів

Суттєвим фізичним змістом закону Кулона є твердження про адитивну дію електричних зарядів. Згідно цього твердження *сила взаємодії двох зарядів не змінюється при наявності третього заряду.* Незалежно від числа зарядів, які входять в деяку систему, закон Кулона (1.2) можна використовувати для розрахунку взаємодії кожної пари зарядів. Це твердження є основою ***принципу суперпозиції***, який використовується для знаходження величини і напрямку вектора напруженості $\vec{E}$ в кожній точці поля за заданим розподілом у просторі та величиною зарядів. Суть ***принципу суперпозиції*** в тому, що сила, яка діє на пробний заряд $q_{пр}$, що розташований в любій точці поля системи n зарядів, буде представляти собою векторну суму сил, які створюються кожною парою (пробний заряд $q_{пр}$ і заряд системи) і окремо діють на $q_{пр}$ в цій точці.

Математично це можна записати так:

$$\vec{E} = \frac{\vec{F}}{q_{пр}} = \frac{\sum_i^n \vec{F}_i}{q_{пр}} = \frac{\vec{F}_1}{q_{пр}} + \ldots + \frac{\vec{F}_n}{q_{пр}},$$

$$\vec{E} = \vec{E}_1 + \vec{E}_2 + \cdots + \vec{E}_n = \sum_i^n \vec{E}_i, \qquad (1.7)$$

де $\vec{E}_1, \vec{E}_2, \ldots, \vec{E}_n$– напруженості полів, які створюються окремими зарядами в даній точці простору, де знаходиться $q_{пр}$.

Принцип суперпозиції дає можливість обчислювати напруженість поля будь-якої системи зарядів. Подумки поділяючи, наприклад, заряджене тіло скінчених розмірів на точкові заряди, знаходимо складові напруженості в певній точці, створені окремими елементами зарядженого тіла. Потім, згідно з принципом суперпозиції, визначаємо результуючу напруженість.

§1.5.Графічне зображення електричних полів

Електричні поля зображають за допомогою ***ліній напруженості***, які проводять так, щоб дотичні до цих ліній в кожній точці збігалися з напрямками вектора $\vec{E}$ (рис.1.3).

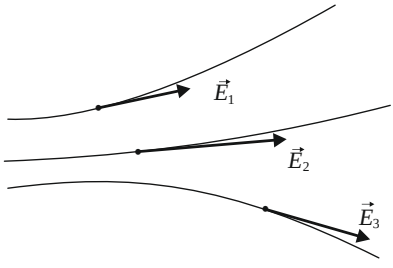


Рис.1.3.

Лінії напруженості мають початок і кінець або йдуть у нескінченність чи з нескінченності. Згідно з формулою (1.6) і рис.1.2 вони починаються на позитивному заряді і закінчуються на негативному заряді. ***Лінії напруженості ніколи не перетинаються.*** В іншому випадку в точці перетину силових ліній були б дві різні напруженості електричного поля. Ці лінії проводять з такою густиною,

щоб кількість ліній, які пронизують одиничну площину, перпендикулярну до вектора $\vec{E}$, чисельно дорівнювала величині вектора $\vec{E}$ електричного поля в місці розміщення площини. На рис.1.4 приведені приклади графічного зображення електричних полів за допомогою ліній напруженості для *а*) і *б*) – позитивного і негативного зарядів, *в*) – системи позитивного і негативного зарядів, *г*) – системи

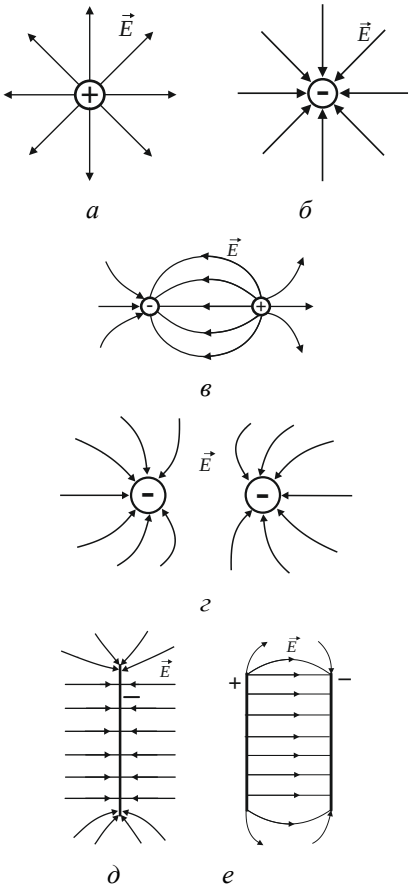


Рис.1.4.

двох негативних зарядів, *д*) – зарядженої негативно площини, *е*) – система двох різнойменно заряджених однаковими за величиною зарядами і однакових за розмірами площин. Поле, у всіх точках якого величина і напрямок вектора напруженості незмінні, називається ***однорідним***. Воно утворюється між зарядженими площинами, якщо вони паралельні і нескінченно великі (рис.14,*е*). Однорідне поле зображується паралельними лініями напруженості з однаковою густотою.

§1.6. Потік вектора напруженості електричного поля

Зв'язок між електричним полем і його джерелами може бути виражений простим способом, який буде мати багато застосувань. Почнемо із визначення величини, яка називається ***потоком*** вектора напруженості електричного поля. Нехай в однорідному електричному полі ($\vec{E} = \text{const}$) знаходиться плоска поверхня *S*, нормаль $\vec{n}$ якої утворює кут α з вектором $\vec{E}$ (рис.1.5). Тоді величину

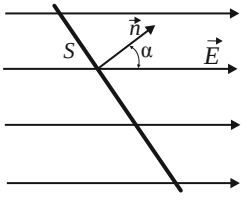


Рис.1.5.

$$\Phi_E = \vec{E}\vec{n}S = |\vec{E}||\vec{n}|S\cos\alpha = ES\cos\alpha = E_n S \qquad (1.8)$$

будемо називати потоком вектора $\vec{E}$ через поверхню *S* (E_n – проекція вектора $\vec{E}$ на нормаль $\vec{n}$). Але в загальному випадку електричні поля неоднорідні $\vec{E} = \vec{E}(x, y, z)$ та й поверхня *S* не є плоскою (рис.1.6). Таку поверхню *S* розбиваємо на елементарні (нескінченно малі) площини dS, які можна вважати плоскими, поле в їх межах –

однорідним, а $\vec{n}$ – нормаль до dS. Вибір додатного напрямку $\vec{n}$ умовний (вважають додатньою зовнішню нормаль до поверхні S).

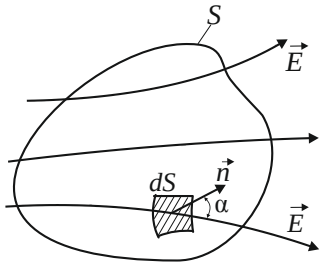


Рис.1.6.

Тоді елементарний потік вектора $\vec{E}$ через dS

$$d\Phi_E=\vec{E}d\vec{S}=\ \vec{E}\vec{n}dS= \\ =|\vec{E}||\vec{n}|dS\cos\alpha=EdS\cos\alpha=E_n dS, \quad (1.9)$$

де формально $d\vec{S}=\vec{n}dS$.

Тепер складемо потоки через всі елементарні площини dS. Ми отримаємо скалярну величину Φ_E, яка є потоком вектора $\vec{E}$ через всю поверхню S:

$$\Phi_E=\int_S \vec{E}\,d\vec{S}=\int_S E_n dS. \quad (1.10)$$

Знак потоку залежить від вибору напрямку нормалі, тобто кута α між $\vec{E}$ і $\vec{n}$.

Для замкнутих поверхонь

$$\Phi_E=\oint_S \vec{E}d\vec{S}, \quad (1.11)$$

де знак $\oint_S$ означає, що сумування елементарних потоків $d\Phi_E$ відбувається по замкнутій поверхні S.

Щоб зрозуміти походження назви потік, виберемо векторну функцію, яка є швидкістю руху рідини, наприклад в річці, де швидкість змінюється від одного місця до іншого, але є постійною з часом в будь-якій точці. Позначимо це векторне поле, яке вимірюється, скажімо, в м/с через $\vec{v}$. Тоді, якщо dS (в м^2) означає елемент площини, зануреної у воду, то добуток $\vec{v}d\vec{S}=|\vec{v}|\cdot|\vec{n}|\cdot\cos\alpha\cdot dS$ є *потоком* води через елемент площини в м^3/с (рис.1.7).

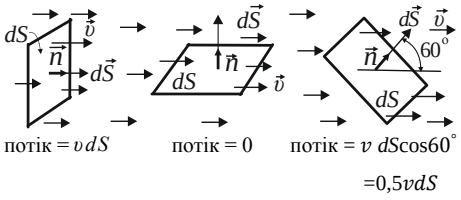


Рис.1.7.

Необхідно підкреслити, що таке визначення потоку можна застосовувати до будь-якої векторної функції, якою б фізичною величиною вона не була.

§1.7.Теорема Остроградського-Гаусса

Розглянемо найбільш простий випадок розрахунку Φ_E. Допустимо, що поле створене ізольованим позитивним електричним зарядом q і що поверхнею S_1 є поверхня сфери, радіус якої r, в центрі якої розташований цей точковий заряд (рис.1.8). Чому дорівнює потік

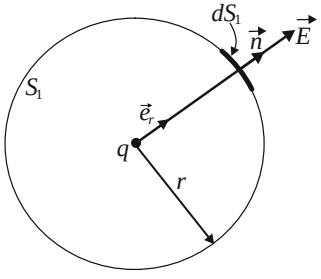


Рис.1.8.

Φ_E через таку поверхню? Відповісти на це питання легко, оскільки величина

напруженості поля $\left|\vec{E}\right|$ в кожній точці сферичної поверхні S_1 дорівнює згідно (1.6) $\frac{q}{4\pi\varepsilon_0 r^2}$ і напрямок $\vec{E}$ співпадає з зовнішньою нормаллю в цій точці. Тоді знаходимо:

$$\begin{cases} \Phi_E = \oint\limits_{S_1} \vec{E}\, d\vec{S}_1 = \oint\limits_{S_1} E dS_1 = \\ = \oint\limits_{S_1} \frac{q}{4\pi\varepsilon_0 r^2} dS_1 = \frac{q}{4\pi\varepsilon_0 r^2} \oint\limits_{S_1} dS_1 = \\ = \frac{q}{4\pi\varepsilon_0 r^2} 4\pi r^2 = q/\varepsilon_0 . \end{cases} \quad (1.12)$$

Це означає, що потік Φ_E не залежить від розмірів сфери і що силові лінії вектора $\vec{E}$ неперервні.

Тепер уявімо собі іншу поверхню S_2, яка охоплює поверхню S_1, але не є сферичної

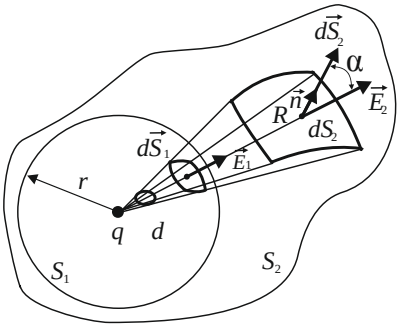


Рис.1.9.

форми (рис.1.9). Доведемо, що потік вектора $\vec{E}$ через поверхню S_2 дорівнює потоку $\vec{E}$ через сферичну поверхню S_1. Для доведення розглянемо конус, який виходить із q і вирізає елементарну площадку dS_1 на поверхні сфери S_1, продовжується до зовнішньої поверхні S_2 і вирізає на ній елемент площі dS_2 на відстані R від заряду q. Знайдемо потоки через dS_1 і dS_2:

$$d\Phi_{E_1} = \vec{E}_1 d\vec{S}_1 = \vec{E}_1 \vec{n}\, dS_1 = \left|\vec{E}_1\right||\vec{n}| \cdot \cos 0° \cdot dS_1 = \frac{q}{4\pi\varepsilon_0 r^2} \cdot dS_1 = \frac{q}{4\pi\varepsilon_0} d\Omega, \quad (1.13)$$

де $d\Omega = dS_1/r^2$–тілесний кут при вершині конуса;

$$d\Phi_{E_2} = \vec{E}_2 d\vec{S}_2 = \vec{E}_2 \vec{n}\, dS_2 = \left|\vec{E}_2\right||\vec{n}| \cdot \cos\alpha \cdot dS_2 = \frac{q}{4\pi\varepsilon_0 R^2} \cdot dS_{2n} = \frac{q}{4\pi\varepsilon_0} d\Omega, \quad (1.14)$$

де $dS_{2n} = dS_2 \cdot \cos\alpha$ – проекція dS_2 на поверхню сфери, радіус якої R. У свою чергу $dS_{2n} / R^2 = d\Omega$ – тілесний кут при вершині конуса, тобто кут, під яким із точки, де знаходиться q, видно елемент поверхні dS_2.

Порівнявши (1.13) і (1.14) бачимо, що $d\Phi_{E_1} = d\Phi_{E_2}$. Кожну частину зовнішньої поверхні S_2 можна сумістити з відповідною частиною внутрішньої сферичної поверхні S_1 таким чином, що повний потік $\vec{E}$ буде однаковий через обидві поверхні S_1 і S_2. Іншими словами, повний потік через поверхню S_2 теж буде дорівнювати q/ε_0. Але поверхня S_2 ***довільної*** форми. *Значить, потік вектора* $\vec{E}$ *через довільну поверхню, яка охоплює заряд* q, *дорівнює* q/ε_0**.** Можна довести, що повний потік вектора $\vec{E}$ через замкнуту поверхню S дорівнює нулю, якщо заряд q розташований зовні поверхні. Це твердження очевидне для точкового заряду і сферичної поверхні (рис.1.10). Потік вектора $\vec{E}_1$ через елементарну площину dS_1 дорівнює

$$d\Phi_1 = \vec{E}_1 \vec{n}\, dS_1 = \left|\vec{E}_1\right||\vec{n}| \cos\alpha_1 \cdot dS_1 = \frac{q}{4\pi\varepsilon_0} \cdot \left(\frac{-dS_{1n}}{r_1^2}\right) = -\frac{q}{4\pi\varepsilon_0} d\Omega,$$

де $dS_1 \cdot \cos\alpha_1 = -dS_{1n}$ – проекція площини dS_1 на поверхню сфери радіусом r_1 (r_1–

відстань від заряду q до площини dS_1), $d\Omega$ – тілесний кут при вершині конуса, яка знаходиться в точці розташування заряду q

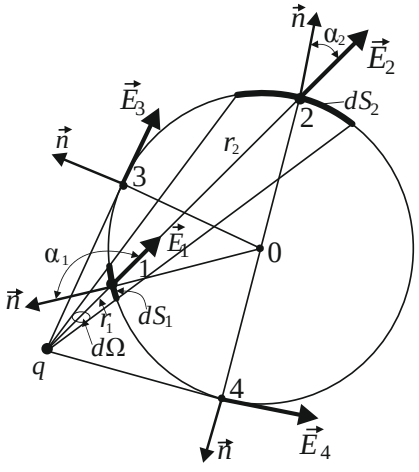


Рис.1.10.

($\cos\alpha_1 < 0$). Аналогічно потік вектора $\vec{E}_2$ через елементарну площину dS_2 буде

$$d\Phi_2 = \vec{E}_2\vec{n}\,dS_2 = |\vec{E}_2||\vec{n}|\cos\alpha_2 \cdot dS_2 = \\ = \frac{q}{4\pi\varepsilon_0}\cdot\left(\frac{dS_{2n}}{r_2^2}\right) = \frac{q}{4\pi\varepsilon_0}d\Omega,$$

де $dS_2 \cdot \cos\alpha_2 = dS_{2n}$– проекція площадки dS_2 на поверхню сфери радіусом r_2 (r_2 – відстань від заряду q до площадки dS_2), $d\Omega$ тілесний кут при вершині конуса (точка розташування заряду q).

Сума $d\Phi_1 + d\Phi_2 = 0.$

Перебираючи попарно відповідні площини на поверхні сфери, доводимо, що сумарний потік вектора $\vec{E}$ в даному випадку дорівнює нулю. Звертаємо увагу на те, що в дотичних точках 3 і 4 сфери скалярний добуток $\vec{E}\cdot\vec{n} = 0,$ так як вектори $\vec{E}$ і $\vec{n}$ взаємно перпендикулярні. Саме в цих точках потік вектора $\vec{E}$ змінює свій знак із негативного на позитивний.

Із вищесказаного можна зробити такий висновок, що потік вектора $\vec{E}$ через замкнуту поверхню не залежить від її розмірів і форми. Це зв'язано із законом «обернених квадратів»у формулі (1.6).

Застосуємо тепер принцип суперпозиції. Будь-яке електричне поле є сума полів окремих джерел (1.7). Ясно, що потік є адитивна величина в тому розумінні, що якщо ми маємо деяке число джерел поля $q_1, q_2, ..., q_n,$ поля кожного із яких мали б напруженість $\vec{E}_1, \vec{E}_2, \dots, \vec{E}_n,$ відповідно, то потік Φ_E через деяку замкнуту поверхню S в реально існуючому полі можна виразити рівністю

$$\Phi_E = \oint_S \vec{E}d\vec{S} = \oint_S (\vec{E}_1 + \vec{E}_2 + \dots + \vec{E}_n)d\vec{S}.$$

Вище ми встановили (формула (1.12), що $\oint_S \vec{E}_i\,d\vec{S} = q_i / \varepsilon_0$, якщо q_i заряд розташований всередині поверхні S, і $\oint_S \vec{E}_i\,d\vec{S}$ дорівнює нулю, якщо q_i, розташований зовні поверхні S. Таким чином, кожний заряд q всередині поверхні S вносить в потік вклад q/ε_0, а всі заряди зовні поверхні не вносять вклад в сумарний потік.

Так ми прийшли до закону (теореми) Остроградського-Гаусса: ***потік вектора напруженості $\vec{E}$ електричного поля у вакуумі через довільну замкнуту поверхню S дорівнює алгебраїчній сумі тих зарядів q, які охоплюються поверхнею S, поділеній на електричну сталу вакууму ε_0:***

$$\oint_S \vec{E}d\vec{S} = \oint_S E_n dS = \frac{1}{\varepsilon_0}\sum_{i=1}^{n} q_i. \quad (1.15)$$

Теорему Остроградського-Гаусса можна називати законом, оскільки вона еквівалентна закону Кулона і може з однаковим успіхом

вважатись ***основним законом електростатичної взаємодії*** після визначення заряду і поля. Закони Остроградського-Гаусса і Кулона не є двома незалежними фізичними законами, а є одним і тим же законом, який виражений в різних формах.

Повертаючись до доведення теореми, ми бачимо, що воно базується на оберненій пропорційності взаємодії зарядів квадрату відстані між ними, і, звичайно, на адитивності взаємодій, тобто на принципі суперпозиції. Теорема Остроградського – Гаусса може застосовуватись до будь-якого фізичного поля, в якому діє закон «обернених квадратів», наприклад до гравітаційного поля, в якому діє сила (1.1).

§1.8.Застосування теореми Остроградського-Гаусса до розрахунку електричних полів

а) *Електричне поле рівномірно зарядженої нескінченної площини.* Площину можна вважати нескінченною, якщо відстанню від точки, в якій визначається напруженість поля, до площини можна знехтувати порівняно з геометричними розмірами площини. Поверхнева густина заряду σ – це кількість заряду, що приходиться на одиничну площу площини ($[\sigma] = \text{Кл}/\text{м}^2$). Електричне поле по обидві сторони площини, яка б не була його величина, повинно бути направлене перпендикулярно площині; в системі не може бути іншого виділеного напрямку (рис.1.11). В точках P і P', які розташовані на однаковій відстані по обидві сторони площини, вектори напруженості $\vec{E}_P$ і $\vec{E}_{P'}$ поля із-за симетрії повинні мати однакову величину, але протилежнті напрямки. Намалюємо циліндр (рис.1.11) з поперечним перерізом S, на одному із кінців якого розташована точка P, а на іншому P'. Потік вектора $\vec{E}$ через бічну поверхню циліндра дорівнює нулю, оскільки лінії напруженості поля паралельні твірним і не перетинають його бічної поверхні. Через основи циліндра потік дорівнює $SE_P + SE_{P'} = 2SE_P$.

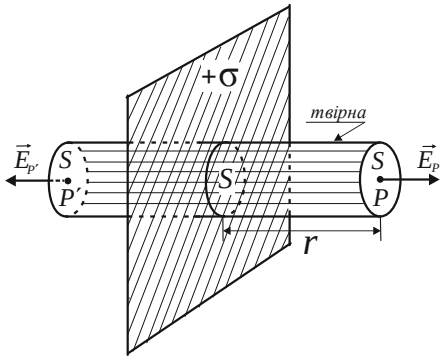


Рис.1.11.

Вибрана циліндрична поверхня охоплює заряд, який розміщено на поверхні S площини, величина якого дорівнює σS. Тоді за теоремою Остроградського–Гаусса (1.15) $2ES = \sigma S/\varepsilon_0$. Звідси

$$E = \sigma/(2\varepsilon_0). \qquad (1.16)$$

Незалежність $\vec{E}$ від відстані r свідчить про те, що поле нескінченої рівномірно зарядженої площини в усьому просторі поза нею однорідне (краєві ефекти не враховуються). Співвідношення (1.16) можна отримати більш складним шляхом, використовуючи принцип суперпозиції (1.7).

б) *Електричне поле між двома різнойменно зарядженими паралельними нескінченними площинами.* Будемо вважати, що поверхневі густини зарядів площин однакові за величиною $|+\sigma| = |-\sigma|$. За принципом суперпозиції можна стверджувати, що в областях за межами площин поле відсутнє (в

кожній точці простору за межами площин значення $\vec{E}$ від обох заряджених площин однакові за величиною і протилежні за знаком, тобто $\vec{E}^+ + \vec{E}^-$=0), а в областях між площинами напрями $\vec{E}$ від обох заряджених площин мають однаковий напрямок (рис.1.12) і отже

$$E = E^+ + E^- = \frac{\sigma}{2\varepsilon_0} + \frac{\sigma}{2\varepsilon_0} = \frac{\sigma}{\varepsilon_0}\cdot \quad (1.17)$$

Електричне поле двох різнойменних заряджених площин локалізоване в об'ємі між цими двома площинами і є однорідним.

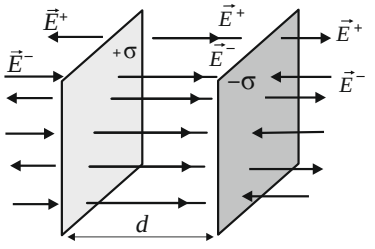


Рис.1.12.

в) *Електричне поле лінійного заряду*. Довгий прямий заряджений провідник, якщо

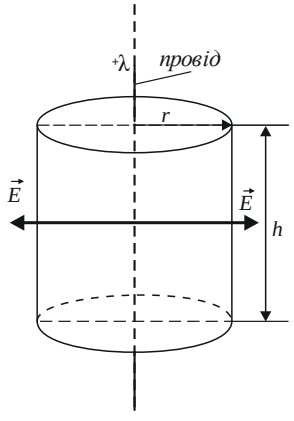


Рис.1.13.

знехтувати його товщиною, можна характеризувати кількістю електрики (заряду) на одиницю довжини. Позначимо цю *лінійну густину заряду,* яка вимірюється в системі СІ в Кл/м, буквою λ. Із міркувань симетрії витікає, що напруженість поля в будь-якій точці повинна бути направлена вздовж радіальної прямої, яка перпендикулярна до проводу, а величина напруженості може залежати від відстані r точки до проводу (рис.1.13). Уявімо собі циліндр, віссю якого є заряджений провід, висота циліндра h, а радіус основи r (рис.1.13). Потік вектора $\vec{E}$ через торці (основи) циліндра дорівнює нулю. Потік через бічну поверхню циліндра дорівнює площі $2\pi rh$ помножений на напруженість поля на поверхні $E(r)$. З іншого боку, заряд, який охоплюється циліндричною поверхнею, дорівнює λh. Таким чином згідно теореми (1.15) можемо записати, що $2\pi rhE(r) = \lambda h/\varepsilon_0$. Звідки

$$E(r) = \frac{\lambda}{2\pi\varepsilon_0 r}. \quad (1.18)$$

Поле лінійного заряду є неоднорідним.

г) *Електричне поле рівномірно зарядженої нескінченної циліндричної поверхні.* Якщо радіус зарядженої циліндричної поверхні *R*, то для відстані *r*>*R* від осі циліндра, отримаємо випадок, що розглядався пункті *в*). Якщо *r*<*R* , то допоміжна циліндрична поверхня не містить всередині зарядів, внаслідок чого $E(r < R) = 0$. Таким чином для зарядженого циліндра для визначення напруженості поля при $r > R$ можна використовувати формулу (1.18). Поле циліндра є неоднорідним.

д) Електричне поле сферичного розподілу заряду. Розглянемо поле сферично-симетричного розподілу заряду, тобто розподілу, в якому об'ємна густина заряду ρ ($\rho = \Delta q/\Delta V$ – заряд одиниці об'єму, який вимірюється в Кл/м^3) залежить тільки від відстані до центральної точки. На (рис.1.14) зображений поперечний переріз такого розподілу. Густина заряду в цьому розподілі

нерівномірна: в центрі вона більша, ніж на деякій відстані від нього, потім зменшується, потім знову збільшується і на відстані, більшій R, дорівнює нулю. Необхідно визначити напруженість електричного поля в деякій точці P_1, яка розташована за межами розподілу заряду і в точці P_2 всередині нього (рис.1.15).

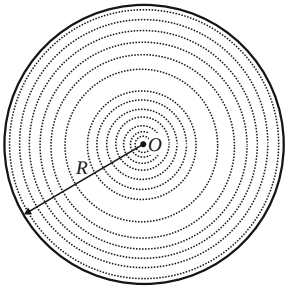


Рис.1.14.

Через сферичну симетрію електричне поле в будь-якій точці повинно бути направлене по радіусу – другого напрямку бути не може. Величина поля E повинна бути однаковою у всіх точках сферичної поверхні S_1, радіус якої r_1, так як всі такі точки є еквівалентними.

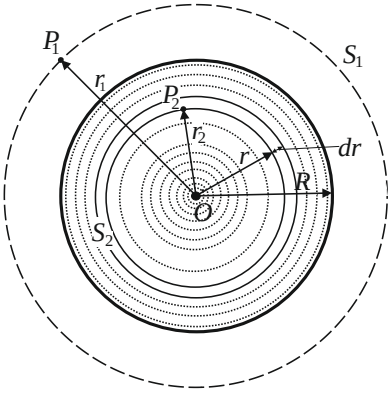


Рис.1.15.

Позначимо величину цього поля E_1. Тоді потік цього поля через поверхню S_1 дорівнює $4\pi r_1^2 E_1$ і, за теоремою Остроградського-Гауса, повинен дорівнювати заряду, цю охоплюється цією поверхнею, поділеному на ε_0. Таким чином

$$E_1 = \frac{\text{заряд всередині } S_1}{4\pi\varepsilon_0 r_1^2}. \quad (1.19)$$

Порівнюючи цю формулу з полем точкового заряду (1.6), ми бачимо, що поле у всіх точках поверхні S_1 є таким же, як якби весь заряд всередині S_1 був зосереджений в центрі O. Аналогічне твердження можна застосувати до будь-якої сфери, проведеної всередині розподілу заряду, центр якої знаходиться в точці O. Поле в будь-якій точці поверхні S_2 є таким, як якби весь заряд всередині S_2 був зосереджений в центрі O, а зовні S_2 зарядів не було би.

Розрахувати заряд всередині поверхонь S_1 або S_2 можна так. Загальна формула для розрахунку заряду в деякому об'ємі V наступна:

$$q = \int_V \rho dV,$$

де dV – елементарний об'єм, в межах якого $\rho = \text{const}$. На рис.1.15 виберемо сфери, радіуси яких r і $r + dr$, де dr – нескінченно малий приріст радіуса r сфери. Тоді за об'єм dV візьмемо простір між цими двома сферами. Цей об'єм $dV=4\pi r^2 dr$. Тому формула для розрахунку заряду всередині сфери, радіус якої r, для сферично-симетричного розподілу заряду така:

$$q = \int_0^r 4\pi\rho(r) r^2\, dr. \quad (1.20)$$

Таким чином заряд всередині поверхні S_1 дорівнює $q = \int_0^{r_1} 4\pi\rho(r) r^2\, dr$, а заряд всередині поверхні S_2 –

$$q = \int_0^{r_2} 4\pi\rho(r) r^2\, dr.$$

Розглянемо окремі частинні випадки.

е) Електричне поле рівномірно зарядженої сфери, радіус якої R (рис.1.16). Так як всередині поверхні S_2 зарядів не має, то напруженість поля всередині сфери дорівнює нулю. Для поверхні S_1 заряд всередині неї дорівнює заряду сфери, то напруженість поля для відстаней $r > R$ можна розрахувати за формулою (1.19).

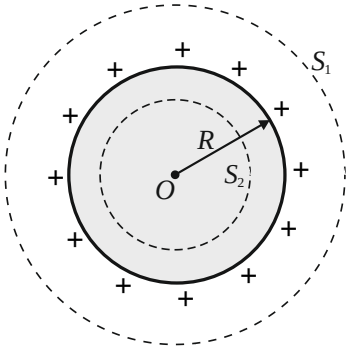


Рис.1.16.

ж) Електричне поле однорідно зарядженої кулі, радіус якої R.

Для цього випадку $\rho = \text{const}$ (рис.1.17). Тоді заряд всередині поверхні S_1 буде:

$$q = \int_0^R 4\pi\rho r^2 dr = \frac{4\pi\rho R^3}{3}$$

і напруженість поля в точках поверхні S_1 визначиться за формулою (1.19).

$$E_1 = \frac{\rho R^3}{3\varepsilon_0 r_1^2} = \frac{q}{4\pi\varepsilon_0 r_1^2}.$$

Заряд всередині поверхні S_2 за формулою (1.20) дорівнює:

$$q_2 = 4\pi\rho\int_0^{r_2} r^2 dr = \frac{4\pi\rho r_2^3}{3}.$$

Тоді за формулою (1.19) напруженість поля в точках на поверхні сфері S_2 буде:

$$E_2 = \frac{q_2}{4\pi\varepsilon_0 r_2^2} = \frac{\rho}{3\varepsilon_0} r_2. \qquad (1.21)$$

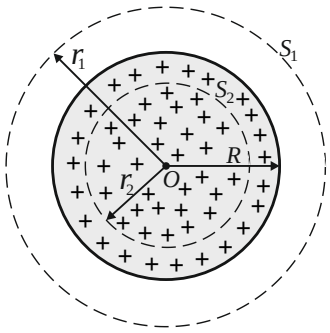


Рис.1.17.

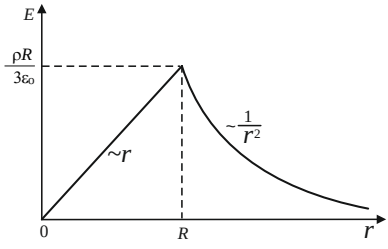


Рис.1.18.

Це означає, що напруженість поля всередині кулі лінійно зростає з відстанню від центра кулі O від нуля до $\rho R/(3\varepsilon_0)$. На рис.1.18 приведено графік залежності E від r для рівномірно зарядженої кулі.

§1.9.Робота сил електростатичного поля

Обчислимо роботу сил електростатичного поля заряду q при переміщенні пробного заряду $q_{пр}$ в цьому полі з точки 1, що перебуває на відстані r_1 від заряду q, в точку 2 на відстані r_2 від нього (рис.1.19). В любій точці цього поля на заряд $q_{пр}$ діє сила Кулона (1.3)

$$\vec{F} = \frac{1}{4\pi\varepsilon_0}\frac{qq_{пр}}{r^2}\vec{e}_r.$$

Робота сили $\vec{F}$ на елементарному переміщенні $d\vec{l}$ дорівнює:

$$dA=\vec{F}d\vec{l}=Fdl\cos\alpha=$$
$$=\frac{1}{4\pi\varepsilon_0}\frac{qq_{np}}{r^2}dl\cos\alpha=\frac{qq_{np}dr}{4\pi\varepsilon_0 r^2},$$

де dr – проекція вектора $d\vec{l}$ на напрямок дії сили $\vec{F}$ (рис.1.19). Робота при переміщенні заряду q_{np} із точки 1 в точку 2 дорівнює:

$$A_{12}=\int_{r_1}^{r_2}dA=\frac{qq_{np}}{4\pi\varepsilon_0}\int_{r_1}^{r_2}\frac{dr}{r^2}=\frac{qq_{np}}{4\pi\varepsilon_0}\left(-\frac{1}{r}\Big|_{r_1}^{r_2}\right).$$

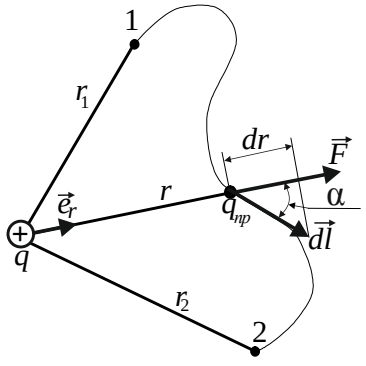


Рис.1.19.

Звідки

$$A_{12}=\frac{qq_{np}}{4\pi\varepsilon_0 r_1}-\frac{qq_{np}}{4\pi\varepsilon_0 r_2}. \quad (1.22)$$

Із (1.22) випливає, що робота сил електростатичного поля не залежить від форми шляху переміщення заряду q_{np} між точками 1–2, а визначається лише розміщенням початкової 1 (r_1) і кінцевої 2 (r_2) точок. Силові поля, які задовольняють такій умові, називаються потенціальними, або консервативними. Отже, електростатичне поле точкового заряду є потенціальним. Цей висновок можна поширити на випадок електричних полів будь-якої системи нерухомих точкових зарядів.

§1.10.Різниця потенціалів. Потенціал. Еквіпотенціальні поверхні

Робота сил консервативного поля (1.22) виконується за рахунок зменшення потенціальної енергії, тобто:

$$A_{12}=W_{p_1}-W_{p_2}=-(W_{p_2}-W_{p_1}) \quad (1.23)$$

де

$$W_{p_1}=\frac{qq_{np}}{4\pi\varepsilon_0 r_1}+\text{const},$$

а

$$W_{p_2}=\frac{qq_{np}}{4\pi\varepsilon_0 r_2}+\text{const}$$

потенціальна енергія заряду q_{np} в полі заряду q на відстанях r_1 і r_2 від нього, відповідно. В загальному випадку

$$W_p=\frac{qq_{np}}{4\pi\varepsilon_0 r}+\text{const}. \quad (1.24)$$

Значення константи у виразі (1.24) вибирають таким чином, щоб при видаленні заряду q_{np} від заряду q на нескінченість (тобто $r\to\infty$) потенціальна енергія дорівнювала нулю. При такій умові (1.24) набуває вигляду:

$$W_p=\frac{qq_{np}}{4\pi\varepsilon_0 r}. \quad (1.25)$$

Якщо заряди q і q_{np} однойменні, то потенціальна енергія їхньої взаємодії (відштовхування) додатня і зростає при зближенні цих зарядів (рис.1.20).

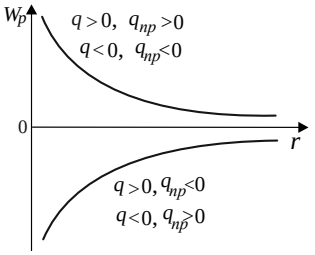


Рис.1.20.

У випадку взаємного притягання різнойменних зарядів потенціальна енергія їхньої взаємодії від'ємна і зменшується при

зближенні зарядів. Для різних пробних зарядів q'_{np}, q''_{np} і т. д. будуть виконуватись різні роботи (1.22). Однак, відношення A_{12}/q_{np} буде однаковим для всіх пробних зарядів. Величина

$$\varphi_1 - \varphi_2 = A_{12}/q_{np} \qquad (1.26)$$

називається різницею потенціалів електростатичного поля між точками 1 і 2. $\varphi_1 - \varphi_2$ однозначно визначається через роботу (1.22):

$$\varphi_1 - \varphi_2 = \frac{q}{4\pi\varepsilon_0 r_1} - \frac{q}{4\pi\varepsilon_0 r_2}, \quad (1.27)$$

де $\varphi_1 = \frac{q}{4\pi\varepsilon_0 r_1} + \text{const}$, а

$$\varphi_2 = \frac{q}{4\pi\varepsilon_0 r_2} + \text{const}$$

називаються потенціалами поля в точках 1 і 2. Потенціал, як і потенціальна енергія (1.24), визначається з точністю до константи, тобто

$$\varphi = \frac{q}{4\pi\varepsilon_0 r} + \text{const}. \qquad (1.28)$$

Якщо прийняти, що на нескінченно великій відстані r ($r \to \infty$) від заряду q потенціал дорівнює нулю, то константа у (1.28) приймається за нуль і

$$\varphi = \frac{q}{4\pi\varepsilon_0 r}. \qquad (1.29)$$

Якщо порівняти формули (1.25) і (1.29), то

$$\varphi = W_p / q_{np}. \qquad (1.30)$$

Із (1.30) витікає, що потенціалом φ будь-якої точки електростатичного поля називають фізичну величину, яка числово дорівнює потенціальній енергії одиничного позитивного заряду, поміщеного в цю точку.

Із формули (1.26) ще можемо записати, що

$$\varphi_1 = A_{1\infty}/q_{np},$$

де враховано, що q_{np} ми із точки 1 переміщуємо на нескінченність ($r_2 \to \infty$).

Тому, ***потенціал*** даної точки електростатичного поля – це така фізична величина, яка числово дорівнює роботі, яку виконують зовнішні сили (проти сил електростатичного поля) при переміщенні одиничного позитивного заряду з нескінченності в дану точку поля.

Одиниця потенціалу – вольт. 1В – це потенціал такої точки поля, в якій заряд величиною 1 Кл має потенціальну енергію 1 Дж.

Робота є адитивна величина. Нехай поле створюється системою n точкових зарядів $q_1, q_2, \ldots, q_n$. Робота, яка виконується силами цього поля над зарядом q_{np}, буде дорівнювати алгебраїчній сумі робіт сил, обумовлених кожним зарядом окремо:

$$A_{12} = \sum_{i=1}^{n} A_i.$$

Згідно (1.22) кожна із робіт A_i дорівнює

$$A_i = \frac{1}{4\pi\varepsilon_0}\left(\frac{q_i q_{np}}{r_{i1}} - \frac{q_i q_{np}}{r_{i2}}\right),$$

де r_{i1} – відстань від заряду q_i до початкового положення заряду q_{np}, r_{i2} – відстань від q_i до кінцевого положення заряду q_{np}. Значить

$$A_{12} = \frac{1}{4\pi\varepsilon_0}\sum_{i=1}^{n}\frac{q_i q_{np}}{r_{i1}} - \frac{1}{4\pi\varepsilon_0}\sum_{i=1}^{n}\frac{q_i q_{np}}{r_{i2}}.$$

Співставивши цей вираз із співвідношенням (1.23), отримаємо для потенціального енергії заряду q_{np} в полі системи зарядів вираз

$$W_p = \frac{1}{4\pi\varepsilon_0}\sum_{i=1}^{n}\frac{q_i q_{np}}{r_i},$$

із якого витікає, що

$$\varphi = \frac{W_p}{q_{np}} = \frac{1}{4\pi\varepsilon_0}\sum_{i=1}^{n}\frac{q_i}{r_i}. \qquad (1.31)$$

Співставляючи формули (1.31) і (1.29), приходимо до висновку, що потенціал поля, яке створене системою зарядів, дорівнює алгебраїчній сумі потенціалів, які створюються

кожним із зарядів окремо. В той час як напруженості полів складаються при накладанні полів векторно, потенціали складаються алгебраїчно. Через цю причину розрахунок потенціалів є більш простим, ніж розрахунок напруженостей електричного поля.

Геометричне місце точок з однаковим потенціалом називається еквіпотенціальною поверхнею. Її рівняння має вигляд:

$$\varphi(x, y, z) = \text{const}\ .$$

При переміщення $q_{пр}$ по еквіпотенціальній поверхні на елементарний відрізок $d\vec{l}$ потенціал не змінюється, а отже, і робота

$$A_{12}=q_{пр}(\varphi_1 - \varphi_2)= \int_1^2 \vec{F}d\vec{l} =\int_1^2 q_{пр}\,\vec{E}d\vec{l} = 0.$$

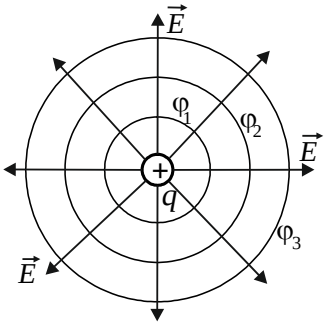


Рис.1.21.

Звідси, скалярний добуток $\vec{E}d\vec{l} = 0$. Так як $|\vec{E}| \neq 0$, $|d\vec{l}| \neq 0$, то $\vec{E}d\vec{l} = 0$ за умови, що кут між $\vec{E}$ і $d\vec{l}$ дорівнює $\pi/2$. Це означає, що вектор $\vec{E}$ напруженості електричного поля в кожній точці напрямлений перпендикулярно до еквіпотенціальної поверхні.

Еквіпотенціальні поверхні точкового заряду – це сферичні оболонки навколо нього (рис.1.21) ($\varphi_1 > \varphi_2 > \varphi_3$).

§1.11.Циркуляція вектора напруженості електричного поля

З формули (1.22) також випливає, що при переміщені точкового заряду в електростатичному полі по довільному замкнутому контуру Г $(r_1 = r_2)$ робота дорівнює нулю. Математично цю умову записують так:

$$A = \oint_\Gamma \vec{F}\,d\vec{l}=\oint_\Gamma q_{пр}\,\vec{E}d\vec{l} =q_{пр}\oint_\Gamma \vec{E}d\vec{l}=0.$$

Так як $q_{пр} \neq 0$, то

$$\oint_\Gamma \vec{E}d\vec{l} = 0. \qquad (1.32)$$

Лінійний інтеграл $\oint_\Gamma \vec{E}d\vec{l}$, обчислений по довільному замкнутому контуру Г, називають *циркуляцією вектора* $\vec{E}$.

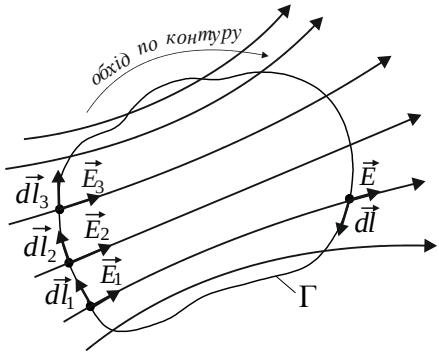


Рис.1.22.

Циркуляція вектора $\vec{E}$ по контуру Г – це сума скалярних добутків $\vec{E}_1 d\vec{l}_1 + \vec{E}_2 d\vec{l}_2 + \vec{E}_3 d\vec{l}_3 + \cdots$, обчислених в кожній точці контуру Г (рис.1.22).

Співвідношення (1.32) є одним із фундаментальних рівнянь електростатики, яке відображає той факт, що силові лінії електростатичного поля незамкнені: вони починаються на позитивних зарядах і закінчуються на негативних або йдуть у нескінченність.

При переміщенні пробного точкового заряду у такому полі по замкнутому контуру на одних ділянках шляху робота буде додатньою, на інших від'ємною, а повна робота завжди дорівнює нулю. Рівняння (1.32) є інтегральною формою запису потенціальності електростатичного поля.

§1.12.Зв'язок між напруженістю та потенціалом електростатичного поля

Оскільки напруженість $\vec{E}$ і потенціал φ є різними за фізичним змістом характеристиками тих самих точок електростатичного поля, то між ними повинен існувати взаємозв'язок.

Нехай в електричному полі знаходиться заряд $q_{пр}$. Робота при переміщенні цього заряду вздовж осі OX між двома нескінченно близькими точками буде:

$$dA = F_x dx = q_{пр} E_x dx .$$

З іншого боку, елементарна робота при переміщенні заряду $q_{пр}$ в електростатичному полі виражається через різницю потенціалів між цими точками:

$$dA = q_{пр}(\varphi_1 - \varphi_2) = -q_{пр}(\varphi_2 - \varphi_1) = \\ = -q_{пр}(\varphi + d\varphi - \varphi) = -q_{пр} d\varphi,$$

де $\varphi_1 = \varphi$, а $\varphi_2 = \varphi + d\varphi$.

Тоді, прирівнявши елементарні роботи, отримуємо:

$$E_x dx = -d\varphi, \text{ або } E_x = -\frac{\partial \varphi}{\partial x} .$$

Знак "–" означає, що під дією сил електричного поля додатній заряд переміщується в бік зменшення потенціалу. Значок похідної $\frac{\partial \varphi}{\partial x}$ означає, що $\varphi = \varphi(x, y, z)$, але в даному випадку похідна береться тільки по змінній x.

Аналогічні міркування можна поширити і на напрямки переміщень вздовж осей OY і OZ:

$$E_y = -\frac{\partial \varphi}{\partial y}; \quad E_z = -\frac{\partial \varphi}{\partial z} .$$

Таким чином ми знаходимо E_x, E_y, E_z – компоненти вектора напруженості $\vec{E}$. Вираз для $\vec{E}$ можемо записати так:

$$\vec{E} = E_x \vec{i} + E_y \vec{j} + E_z \vec{k} = -\left(\frac{\partial \varphi}{\partial x}\vec{i} + \frac{\partial \varphi}{\partial y}\vec{j} + \frac{\partial \varphi}{\partial z}\vec{k}\right).$$

У векторному аналізі градієнтом скалярної величини φ називається така векторна величина, для якої справедливий запис:

$$\text{grad}\varphi = \frac{\partial \varphi}{\partial x}\vec{i} + \frac{\partial \varphi}{\partial y}\vec{j} + \frac{\partial \varphi}{\partial z}\vec{k}.$$

Цей вектор направлений у бік найшвидшого зростання потенціалу φ.

Отже,

$$\vec{E} = -\text{grad}\varphi. \qquad (1.33)$$

Знак "–" вказує на те, що вектор $\vec{E}$ напруженості поля направлений вбік найшвидшого зменшення потенціалу. Напруженість в якій-небудь точці електростатичного поля дорівнює градієнту потенціалу в цій точці поля, взятому з оберненим знаком.

Знаючи потенціал φ в кожній точці поля, за формулою (1.33) можемо обчислити напруженість в кожній точці поля.

Можна розв'язати і обернену задачу, тобто знаючи напруженість поля в кожній точці поля, можна знайти різницю потенціалів між довільними точками. Робота при переміщені заряду $q_{пр}$ з точки 1 в точку 2 дорівнює:

$$A_{12} = \int_1^2 \vec{F} d\vec{l} = \int_1^2 q_{пр} \vec{E} d\vec{l} = q_{пр}(\varphi_1 - \varphi_2).$$

Звідси отримаємо, що

$$\varphi_1 - \varphi_2 = \int_1^2 \vec{E} d\vec{l} . \qquad (1.34)$$

Інтеграл можна брати вздовж довільної лінії, яка з'єднує точки 1 і 2, оскільки електричне поле є потенціальним.

§1.13.Рівняння Пуассона (теорема Остроградського-Гаусса в диференціальній формі)

Співвідношення (1.15) виражає теорему Остроградського-Гаусса в інтегральній формі. Величини, що входять у вираз теореми, характеризують різні точки електричного поля. Так, якщо вектор $\vec{E}$ характеризує точки деякої замкнутої поверхні, то електричний заряд стосується точок об'єму, обмеженого цією поверхнею.

Отримаємо аналогічний вираз цієї теореми в диференційній формі, тобто знайдено співвідношення між фізичними величинами, які характеризують якусь точку поля. Для цього знайдемо потік вектора напруженості поля $\vec{E}$ через поверхню нескінченно малого прямокутного паралелепіпеда з ребрами $dx,\ dy,\ dz,$ який охоплює заряд, розподілений з об'ємною густиною ρ (рис.1.23). Будемо вважати, що $\rho = \rho(x, y, z)$ є неперервна функція координат.

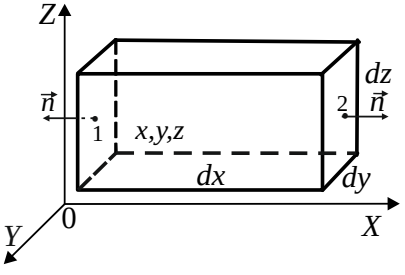


Рис.1.23.

Представлення про неперервний розподіл електрики в просторі є такою ж ідеалізацію, як і представлення про неперервний розподіл речовини. На грані 1 зовнішня нормаль направлена проти позитивного напрямку осі *OX*. Тому потік вектора $\vec{E}$ через цю грань буде $-E_x(x)dydz$. На протилежній грані 2, навпаки, напрям зовнішньої нормалі $\vec{n}$ співпадає з позитивним напрямком осі *OX* і для потока вектора $\vec{E}$ через цю грань необхідно писати $+E_x(x+dx)dydz$. Сума обох потоків буде:

$$(E_x(x+dx) - E_x(x))dydz = \\ = \frac{\partial E_x}{\partial x}dxdydz = \frac{\partial E_x}{\partial x}dV,$$

де $dV = dxdydz$ – об'єм елементарного паралепіпеда. Враховано, що приріст функції

$$E_x(x+dx) - E_x(x) = \frac{\partial E_x}{\partial x}dx,$$

де $\frac{\partial E_x}{\partial x}$ – похідна від $E_x(x, y, z)$ тільки по змінній *x*.

Аналогічно знайдуться потоки через дві пари останніх граней:

$$(E_y(y+dy) - E_y(y))dxdz = \frac{\partial E_y}{\partial y}dV,$$

$$(E_z(z+dz) - E_z(z))dxdy = \frac{\partial E_z}{\partial z}dV.$$

Повний потік через всю поверхню паралелепіпеда

$$d\Phi_E = (\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z})dV = \mathrm{div}\vec{E}\ dV, \quad (1.35)$$

де введено позначення

$$\mathrm{div}\vec{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}\cdot \quad (1.36)$$

За теоремою Остроградського-Гаусса

$$d\Phi_E = dq/\varepsilon_0 = \rho dV/\varepsilon_0, \quad (1.37)$$

де dq – заряд, що знаходяться всередині паралепіпеда. Прирівнюючи (1.35) і (1.37) отримаємо:

$$\mathrm{div}\vec{E} = \rho/\varepsilon_0. \quad (1.38)$$

Ця формула і виражає, теорему Остроградського-Гаусса в диференціальній формі.

Величина, яка визначається виразом (1.36), зберігає зміст незалежно від конкретної

фізичної або геометричної природи вектора $\vec{E}$. Вона називається ***дивергенцією*** вектора $\vec{E}$.

Інтегральна і диференціальна форми теореми Остроградсгькоо-Гаусса повністю еквівалентні. ***Ці теореми справедливі не тільки в електростатиці***, але і в електродинаміці, де розглядаються змінні з часом електромагнітні поля. Вся сукупність дослідних даних підтверджує це твердження. Тим самим рівняння (1.15) і математично еквівалентне рівняння (1.38) перестають бути скромними наслідками закону Кулона, а підносяться в ранг ***основних постулатів*** теорії електрики.

Контрольні питання

1.Закон Кулона та його фундаментальність для електростатичного поля.

2.Вектор напруженості електричного поля: визначення та фізичний зміст.

3.Силова лінія електричного поля: визначення, властивості, фізичний зміст.

4.Потік вектора напруженості електричного поля: визначення, фізичний зміст.

5.Теорема Остроградського-Гаусса для електростатичного поля та її зв'язок із законом Кулона.

6.Застосування теореми Остроградського-Гаусса до розрахунку електричного поля: точкового заряду, безмежної однорідно зарядженої площини, безмежно довгої однорідно зарядженої прямої нитки, однорідно зарядженої кулі.

7.Робота по переміщенню заряду в електричному полі: її властивості та методи обчислення.

8.Різниця потенціалів: визначення, фізичний зміст. Потенціал: зміст поняття «безмежності» у визначенні потенціалу.

9.Принцип суперпозиції електричних полів: визначення, відмінність цього принципу для напруженості і для потенціалу.

10. Зв'язок між вектором напруженості $\vec{E}$ і потенціалом поля φ.

11.Взаємне розташування силових ліній вектора $\vec{E}$ і еквіпотенціальних ліній (поверхонь) потенціалу φ електричного поля. Довести цю властивість їхнього взаємного розташування.

12.Рівняння Пуассона: визначення, фізичний зміст, зв'язок із теоремою Остроградського-Гаусса.

13.Циркуляція вектора напруженості поля $\vec{E}$ по замкнутому контуру: визначення, фізичний зміст.

Розділ 2. Поляризація діелектриків і поле всередині діелектрика

§2.1.Провідники, напівпровідники та діелектрики

За здатністю проводити електричний струм всі тверді тіла поділяються на три великих групи (Табл.2.1): метали, напівпровідники і діелектрики (або ізолятори). Сюди можна добавити і надпровідники – ряд металів та їх сполук, які при дуже низьких температурах проводять струм без втрат енергії – їх електричний опір дорівнює нулю. Найкращими провідниками є метали. Питомий опір ρ металів – порядку $10^{-8} \div 10^{-5}$ Ом·м. У діелектриків питомий опір більший 10^{8} Ом·м.

Діелектрики (англ. “dielectric”, від грецького diá – через, наскрізь і англ. electric – електричний) – це речовини, які погано проводять електричний струм. Термін «діелектрик», введений Фарадеєм для позначення речовин, в які проникає електричне поле. Діелектриками є гази (не іонізовані), деякі рідини і тверді тіла. Електропровідність діелектриків у порівнянні з металами дуже мала. Напівпровідники займають проміжне місце між провідниками та діелектриками.

Матеріал	Питомий опір ρ, Ом·м	Провідні властивості
Провідники	$10^{-8}\div10^{-5}$	Є “вільні” носії заряду (електрони), кількість яких практично не залежить від температури
Напівпро-відники	$10^{-5}\div10^{8}$	Є “вільні” носії заряду, кількість яких залежить (експоненціально) від температури
Діелектрики	$10^{8}\div10^{16}$	При нормальних умовах відсутні “вільні” носії заряду

На рис.2.1 приведені значення питомого опору деяких металів, напівпровідників і діелектриків. Видно, що питомий опір у срібла менший, ніж у тефлона в 10^{24} (!) раз. Ні одна фізична характеристика не змінюється так сильно при переході від речовини до речовини, як питомий опір.

Останнім часом в техніці почали використовувати і надпровідники. Основна перевага надпровідників в тому, що вони проводять електричний струм без втрат енергії.

Недоліком існуючих надпровідників є необхідність охолоджувати їх до дуже низьких температур (один із "високотемпературних" сплавів Nb_3Ge переходить в надпровідний стан тільки при 21,3 К (−251,7°С)).

І діелектрики, і напівпровідники, і метали широко використовуються в техніці. Діелектрики використовуються, наприклад, як ізолятори. Назвемо деякі із них. До твердих

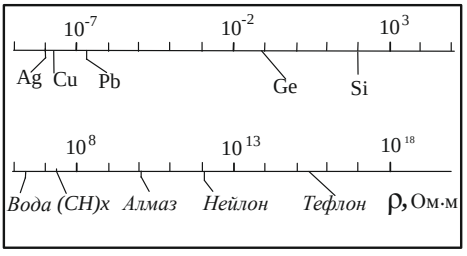


Рис.2.1.

діелектриків відносяться скло, гума, фарфор, слюда, пластмаси і полімерні матеріали, папір, дерево, алмаз, кварц та інші. Рідкими діелектриками є: хімічно чиста (дистильована) вода, ацетон, гліцерин, спирт метиловий, хлорбензол, мінеральні нафтові мастила. Діелектриками є неіонізовані гази і пари речовини, що знаходяться в не дуже сильних електричних полях, повітря, азот та інші.

Напівпровідниками є, як правило, багато твердих речовин (наприклад, германій, кремній) так і хімічних сполук (наприклад, GaAs, CdTe, PbS). Напівпровідникові прилади можна зустріти в багатьох пристроях, починаючи із побутової радіоапаратури і закінчуючи сучасними комп’ютерами, телефонами і т. п., створення яких неможливе без напівпровідникових елементів.

Із металів виготовляють з’єднувальні проводи, а також різні електросилові установки (генератори струму, електромотори, трансформатори і т.п.). В таблиці 2.1 приведені значення питомого опору при Т=293 К і постійній напрузі для електротехнічних

матеріалів, а також їх основна відмінність з точки зору провідних властивостей. Кількісну різницю в електропровідності діелектриків і металів класична фізика пояснює тим, що в металах є «вільні» електрони, а в діелектриках всі електрони міцно зв'язані з атомами.

Варто підкреслити, що різкої границі між групами матеріалів на має. Особливо умовна границя між діелектриками і напівпровідниками. При підвищенні температури опір кристалічного діелектрика різко зменшується і останній стає напівпровідником, навпаки, при зниженні температури напівпровідник стає діелектриком. При певних умовах напівпровідники можуть поводити себе подібно металу.

§2.2. Електричний диполь в однорідному і неоднорідному електричних полях

При внесенні діелектрика в електричне поле відбувається його поляризація, яка призводить до появи на його поверхнях, які перетинаються силовими лініями поля, не-скомпенсованих зв'язаних електричних зарядів протилежних знаків. Тобто дія електричного поля на діелектрик зводиться до зміни розподілу електричних зарядів і в цьому суть явища поляризації діелектриків.

Для розуміння явища поляризації діелектриків необхідно розглянути поведінку електричного диполя в електричному полі. ***Електричний диполь*** – система двох однакових за величиною і протилежних за знаком точкових зарядів $+q$ і $-q$, які знаходяться один від одного на деякій відстані l. Дипольний момент

$$\vec{p} = q\vec{l}\,, \qquad (2.1)$$

де вектор $\vec{l}$ має напрямок від заряду $-q$ до заряду $+q$ (напрямок осі диполя рис.2.2). Сумарний заряд диполя дорівнює нулю, проте він створює електричне поле, яке має осьову симетрію. Тому картина поля в будь-якій площині, яка проходить через вісь диполя, буде однією і тією ж, причому вектор $\vec{E}$ лежить в цій площині. Положення точки відносно диполя будемо характеризувати за допомогою радіуса-вектора $\vec{r} = r\vec{e}_r$ або за допомогою полярних координат r і θ (рис.2.2).

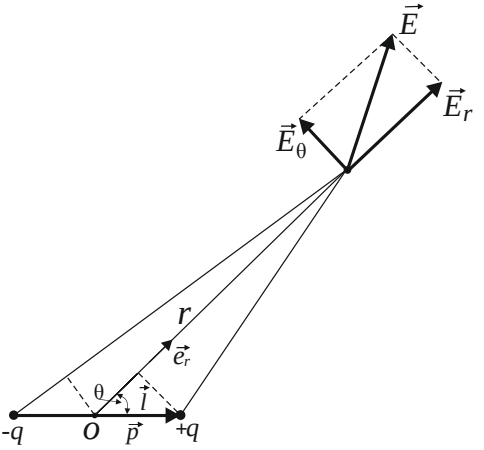


Рис.2.2.

Потенціал і напруженість поля диполя визначаються за формулами:

$$\varphi(r,\theta) = \frac{1}{4\pi\varepsilon_0}\cdot\frac{p\cos\theta}{r^2}, \qquad (2.2)$$

$$E(r,\theta) = \frac{1}{4\pi\varepsilon_0}\cdot\frac{p}{r^3}\sqrt{1+3\cos^2\theta}. \qquad (2.3)$$

Порівняння формули (2.2) з (1.29), а – (2.3) з (1.6) свідчить про те, що потенціал і напруженість поля диполя зменшуються з відстанню швидше, ніж потенціал і напруженість поля точкового заряду ($\varphi \sim 1/r$, $E \sim 1/r^2$). На рис.1.4,*в* показані лінії вектора $\vec{E}$ для диполя.

Диполь в однорідному електричному полі. Результуюча сила, $\vec{F} = \vec{F}_1 + \vec{F}_2$, яка діє

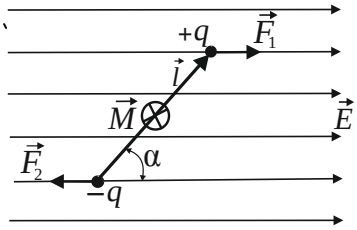


Рис.2.3. Електричний диполь в однорідному зовнішньому електричному полі.

на диполь, дорівнює нулю, так як сили $\vec{F}_1$ і $\vec{F}_2$, які діють на позитивний і негативний заряди диполя, однакові за величиною і протилежні за напрямком (рис.2.3). Ці сили утворюють пару сил. Момент пари сил M дорівнює

$$M = qEl\sin\alpha, \qquad (2.4)$$

де α – кут між векторами $\vec{l}$ і напруженістю поля $\vec{E}$.

Користуючись поняттям момента диполя (2.1), вираз для момента пари сил (2.4) можна записати у вигляді

$$M = pE\sin(\vec{p},\vec{E}). \qquad (2.5)$$

Напрям момента цієї пари сил збігається з напрямком осі обертання диполя, тобто перпендикулярний до $\vec{p}$ і $\vec{E}$. Тому вектор момента пари сил $\vec{M}$, який діє на диполь можна виразити формулою

$$\vec{M} = \vec{p} \times \vec{E}. \qquad (2.6)$$

Бачимо, що в однорідному полі на диполь діє тільки пара сил, яка прагне повернути диполь так, щоб $\vec{p}$ і $\vec{E}$ були паралельні (α = 0).

Щоб повернути диполь в електричному полі на деякий кут, необхідно виконати певну роботу. Через те, що ця робота дорівнює збільшенню потенціальної енергії диполя, то звідси можна знайти вираз для енергії диполя в електричному полі. Приймемо за нуль відліку енергію положення диполя, при якому його вісь перпендикулярна до напрямку поля (α = $= \pi/2$). Тоді енергія диполя, момент $\vec{p}$ якого утворює кут α з напрямком поля, дорівнює:

$$W_P = \int_{\frac{\pi}{2}}^{\alpha} pE\sin\alpha d\alpha = \\ = -pE\cos\alpha = -\vec{p}\,\vec{E}. \qquad (2.7)$$

Зауважимо, що вираз (2.7) не враховує енергію взаємодії зарядів, які утворюють диполь. Розміри диполя невеликі, тому ця формула справедлива і для неоднорідного поля.

Диполь в неоднорідному полі. Для спрощення будемо вважати, що момент $\vec{p}$ диполя паралельний напрямку поля (α = 0) (рис.2.4). У цьому випадку сили, які діють на

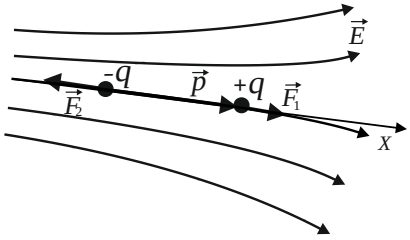


Рис.2.4.Електричний диполь в неоднорідному електричному полі $\vec{E}$ у випадку, коли $\vec{p}$ направлений по полю $\vec{E}$. Сила $\vec{F}_2$ більша сили $\vec{F}_1$ і результуюча сила $\vec{F}_2 - \vec{F}_1$ старається перемістити диполь в область більшої напруженості зовнішнього електричного поля.

кінці диполя, вже неоднакові і тому їх результуюча сила не дорівнює нулю. На диполь в неоднорідному полі діє сила, яка прагне пересунути диполь в область поля з більшою напруженістю. Знайдемо величину цієї сили. Направимо координатну вісь X вздовж момента диполя і вважатимемо, що довжина диполя дуже мала (порядку розмірів молекули). Сила, яка діє на негативний кінець диполя, є $-qE$, де E – напруженість поля в точці знаходження заряду $-q$. Сила, яка діє на позитивний кінець диполя, дорівнює $+q\left(E - \frac{\partial E}{\partial x}l\right)$, де l –

довжина диполя, $\frac{\partial E}{\partial x}-$ градієнт поля. Тому повна сила F_x дорівнює

$$F_x = F_2 - F_1 = qE - q\left(E - \frac{\partial E}{\partial x} l\right) = \\ = ql\frac{\partial E}{\partial x} = p\frac{\partial E}{\partial x}.$$

Для випадку, коли дипольний момент $\vec{p}$ утворює кут α з напрямком поля $\vec{E}$, сила

$$F_x = p\frac{\partial E}{\partial x}\cos\alpha. \qquad (2.8)$$

В однорідному полі градієнт поля $\frac{\partial E}{\partial x} = 0$ і результуюча сила F_x дорівнює нулю.

Якщо диполь перебуває в неоднорідному полі і непаралельний полю, то на нього діє і пара сил, яка прагне повернути диполь паралельно полю, і сила, яка втягує диполь в область сильнішого поля. В кінці кінців диполь опиняється в області сильнішого поля так, щоб вектор $\vec{p}$ був зорієнтований по полю (по напрямку вектора $\vec{E}$).

§2.3.Поляризація діелектриків та її типи. Вектор поляризації (поляризованість), діелектрична сприйнятливість

Явище обмеженого зміщення зарядів в атомах і молекулах або напрямленої орієнтації дипольних моментів жорстких молекул в зовнішньому електричному полі, внаслідок якого на поверхні діелектрика виникають зв'язані електричні заряди, називається поляризацією діелектриків.

Механізми поляризації діелектриків різні і залежать від характеру хімічного зв'язку. Наприклад, в іонних кристалах (NaCl та подібних) поляризація є результатом зсуву одних іонів відносно інших (іонна поляризація: рис.2.5,*а*) і деформації електронних оболонок окремих іонів (електронна поляризація). В кристалах з ковалентним зв'язком (наприклад, алмаз) поляризація обумовлена головним

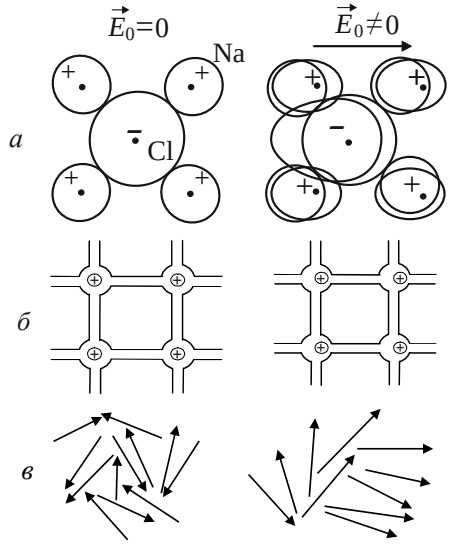


Рис. 2.5.Поляризація діелектриків: *а* – іонна, «+» – іони Na^+, «–» – іони Cl^-; *б* – електронна , + – іони С (вуглецю) кристалічної гратки, суцільні хвилясті лінії – ковалентні (електронні) зв'язки; *в* – орієнтаційна, стрілки дипольні моменти молекул H_2S, $\vec{E}_0$– вектор напруженості зовнішнього електричного поля.

чином зміщенням електронів, які здійснюють хімічний зв'язок (рис.2.5,*б*). В так званих полярних діелектриках (наприклад, твердий H_2S) молекули мають електричні дипольні моменти, які при відсутності зовнішнього електричного поля $\vec{E}_0$ орієнтовані хаотично, а в полі набувають переважну орієнтацію по полю $\vec{E}_0$ (рис.2.5,*в*). Така орієнтаційна поляризація типова для багатьох рідин і газів. Існує ще один важливий з практичної точки зору вид поляризації, який отримав назву спонтанної.

Ця поляризація має місця в кристалічних речовинах – сегнетоелектриках. В деякому інтервалі температур в сегнетоелектричних кристалах існують окремі,

сильно поляризовані області – **домени** (***«Домен»*** *від латинського слова dominicus (господарський, панський). В середньовіччі «доменом» називали особисте володіння великого феодала. В фізиці слово «домен» використовується для позначення області, властивості якої різко відрізняються від властивостей оточуючого її матеріалу*).

Тому діелектрики поділяються на такі типи.

1). ***Неполярні діелектрики*** – це діелектрики, які складаються з молекул, дипольний момент яких дорівнює нулю. Це можливо тоді, коли «центри мас» позитивних і негативних зарядів молекул співпадають (l=0) за відсутності електричного поля.

Прикладом неполярних діелектриків є гази: N_2, H_2, O_2, CO_2, CH_4. В електричному полі сили, які діють на позитивні і негативні заряди молекули, направлені в різні сторони і тому розсувають їх (рис.2.5,*б*). Центр позитивних і центр негативних зарядів молекули зміщується на відстань l тим більшу, чим більша напруженість поля $\vec{E}_0$. Виникає дипольний момент молекули $p = ql$ і таким чином молекула перетворюється в полярну. Так як $l \sim E$, а $p = ql$, то $p \sim E$, або $p = \beta\varepsilon_0 E$, де E – напруженість поля всередині діелектрика, β – поляризованість окремої молекули діелектрика. Величина β, має неоднакові значення для атомів і молекул різних речовин. Поляризованість β характеризує здатність електронів в атомі або в молекулі зміщуватись під дією сил електричного поля. Дипольні моменти $\vec{p}$ молекул неполярних діелектриків називають індукованими або квазіпружними.

При внесенні неполярного діелектрика в електричне поле всі індуковані дипольні моменти розміщуються ланцюжками вздовж лінії напруженості $\vec{E}_0$ (рис.2.6). Внаслідок цього грані діелектрика набувають різнойменних зарядів – діелектрик поляризується. Такого роду поляризація називається ***електронною***.

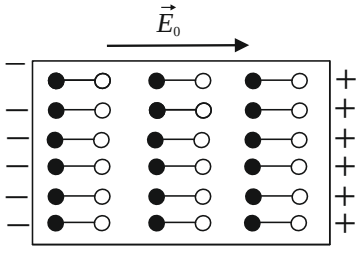


Рис.2.6.
На рисунку позначено :
○— позитивні заряди;
●— негативні заряди.

2). ***Полярні діелектрики – це діелектрики, в яких «центри мас» позитивних і негативних зарядів*** не співпадають. До полярних діелектриків належать гази SO_2, H_2S, NH_3 та ін., рідини – вода H_2O, соляна кислота HCl, бензол C_6H_6 тощо.

Молекули таких діелектриків називаються ***полярними***. Ці молекули за відсутності зовнішнього поля мають дипольні моменти $\vec{p}_0 = q\vec{l}$. Їх називають ***жорсткими диполями***. Величина p_0 має порядок $10^{-29} \div 10^{-30}$ Кл·м. Якщо такий діелектрик не перебуває у зовнішньому електричному полі, то внаслідок хаотичного теплового руху молекул вектори їхніх дипольних моментів орієнтовані хаотично (рис.2.7,*а*).

Якщо діелектрик внести в електричне поле, то під дією поля полярні молекули діелектрика намагаються повернутись так, щоб вектори їх дипольних моментів збігались з

напрямком вектора напруженості поля $\vec{E}_0$ (рис.2.5,*в*;рис.2.7,*б*). Але тепловий рух молекул хаотично розкидає диполі і заважає орієнтації всіх векторів $\vec{p}$ (дипольних моментів) вздовж поля. Орієнтація дипольних моментів буде тим повнішою, чим сильніше електричне поле і чим слабший тепловий рух молекул, тобто нижча температура. Цей процес називають ***орієнтаційною поляризацією*** діелектрика.

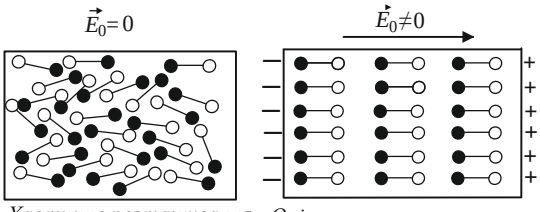


Рис.2.7.

Зауважимо, що молекули полярних діелектриків набувають додаткових, індукованих зовнішнім полем, дипольних моментів (як у випадку неполярних діелектриків) і це теж вносить свій вклад у їхню поляризацію. Проте наведений дипольний момент $p = \beta\varepsilon_o E$ навіть при $E = 10^6$В/м має порядок 10^{-34} Кл·м, тобто в середньому на чотири порядки менший від p_o. Тому електронною поляризацією для полярних діелектриків можна знехтувати.

3). ***Іонні діелектрики – це речовини, молекули яких мають іонну будову.*** Прикладом таких діелектриків є NaCl, NaI, та інші. Іонні кристали є просторовими ґратками з правильним чергуванням іонів різних знаків (рис.2.8). У цих кристалах не можна виділити окремі молекули. Іонні кристали розглядають як систему вміщених одна в одну іонних підграток. У цих діелектриках кожна пара сусідніх різнойменних іонів подібна до диполя. В зовнішньому електричному полі позитивні іони кристалічної ґратки зміщуються по напрямку напруженості поля $\vec{E}$, негативні – в протилежному напрямку. В результаті відбувається певна деформація гратки або, як часто говорять, – зміщення позитивної і негативної підграток, що приводить до виникнення дипольних моментів, тобто до поляризації діелектриків (рис.2.5,*а*). ***Така поляризація називається іонною.***

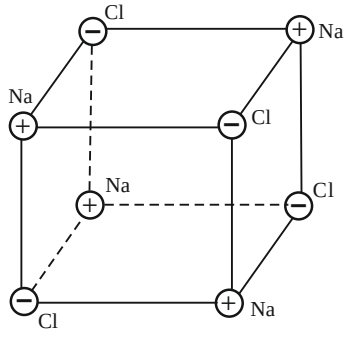


Рис.2.8.

Заряди, які виникають на гранях діелектрика, не вільні, вони зв'язані з атомами та молекулами речовини.

Таким чином, механізм поляризації діелектрика такий: під дією зовнішнього електричного поля молекули речовини набувають дипольні моменти (якщо їх раніше в них не було), які потім орієнтуються за полем. Орієнтація молекул є причиною появи на протилежних поверхнях діелектрика, які перетинаються силовими лініями поля, зарядів протилежних знаків.

Зауваження. Явище поляризації, як правило, супроводжується діелектричними втратами. Так називають потужність, яка витрачається в непровідній речовині при її переполяризації в змінних електричних полях. Ця потужність розсіюється у вигляді тепла, яке

підвищує температуру діелектрика. Діелектричні втрати є важливою характеристикою високочастотних матеріалів. В радіопристроях діелектричні втрати збільшують активний опір електричних кіл і, як наслідок, згасання коливань в контурах. Великі втрати викликають недопустиме нагрівання діелектрика в полі високої частоти, що може призвести до його теплового руйнування.

Явище поляризації діелектриків визначає таке важливе їхнє застосування, як заповнення конденсаторів з метою збільшення ємності останніх.

Поляризацію діелектрика характеризують вектором поляризації – електричним дипольним моментом одиниці об'єму

$$\vec{P} = \sum_{i=1}^{m} \frac{\vec{p}_i}{\Delta V}, \qquad (2.9)$$

де $\vec{p}_i$– дипольний момент i –ої молекули, яка входить в об'єм ΔV; m – число молекул в цьому об'ємі, ΔV – такий малий об'єм, в межах якого можна вважати постійним напруженість електричного поля, густину, температуру, але з іншого боку в ньому міститься достатньо велика для усереднення кількість молекул. Для однорідного діелектрика, який перебуває в однорідному електричному полі, справедлива рівність:

$$\vec{P} = n\vec{p}, \qquad (2.10)$$

де $n = m/\Delta V$ – кількість молекул в одиниці об'єму, $\vec{p}$ – дипольний момент молекули. Так як $\vec{p} = \beta\varepsilon_0\vec{E}$, то $\vec{P} = n\beta\varepsilon_0\vec{E}$. Позначимо $n\beta = \chi$, тоді

$$\vec{P} = \chi\varepsilon_0\vec{E}. \qquad (2.11)$$

Коефіцієнт χ називають *діелектричною сприйнятливістю* речовини або *поляризованістю одиниці об'єму діелектрика*. χ – величина безрозмірна (χ = 80 для води, χ = 25 для спиртів). Діелектрична сприйнятливість макроскопічний параметр, який характеризує здатність структурних одиниць діелектриків до поляризації, тобто до зміщення зарядів, наприклад, в атомах чи молекулах.

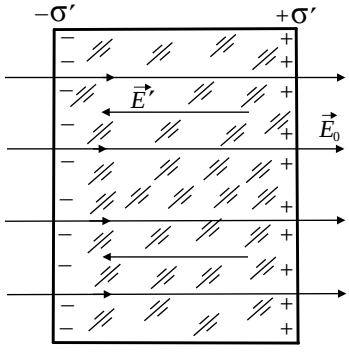


Рис.2.9

Розглянемо однорідний діелектрик, який знаходиться в зовнішньому однорідному електричному полі напруженістю $\vec{E}_0$. Для простоти розгляду будемо вважати, що діелектрик – це прямокутна пластинка і лінії вектора $\vec{E}_0$ перпендикулярні до її поверхні (рис.2.9). Під дією поля діелектрик поляризується, тобто відбувається зміщення зарядів. Внаслідок цього на правій грані діелектрика буде надлишок позитивного заряду з поверхневою густиною $+\sigma'$, на лівій – від'ємного заряду з поверхневою густиною $-\sigma'$. Ці заряди називаються зв'язаними. Зв'язані заряди на гранях пластинки (як дві заряджені площини) створюють електричне поле, напруженість якого $\vec{E}'$ (рис.2.9). Так як вектори $\vec{E}_0$ і $\vec{E}'$ мають протилежні напрямки, то напруженість поля в діелектрику буде:

$$E = E_0 - E' < E_0. \qquad (2.12)$$

Таким чином поляризація діелектрика призводить до послаблення зовнішнього поля $\vec{E}_0$ всередині пластини.

Визначимо поверхневу густину зв'язаних зарядів σ' в загальному випадку, якщо зовнішнє електричне поле $\vec{E}_0$ утворює кут α з нормаллю до грані діелектрика (рис.2.10). Розглянемо циліндр, площі основи якого ΔS знаходяться на гранях пластини, а твірна

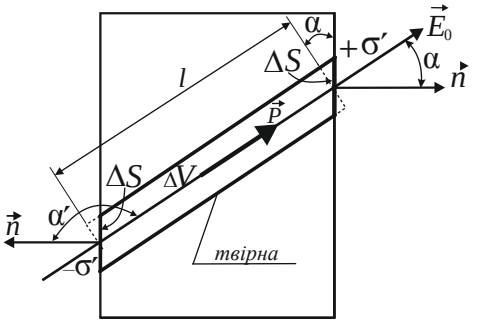


Рис.2.10.

паралельна вектору поляризації $\vec{P}$ і дорівнює якійсь величині *l*. Повний дипольний момент такого циліндричного діелектрика буде дорівнювати

$$P_{\Delta V} = P \cdot \Delta V ,$$

де P - вектор поляризації, $\Delta V = \Delta S \cdot l \cdot \cos\alpha$ – об'єм циліндра. З іншого боку, вважаючи циліндрик за диполь, можемо записати, що

$$P_{\Delta V} = \sigma' \cdot \Delta S \cdot l .$$

Тоді $P \cdot \Delta S \cdot l \cdot \cos\alpha = \sigma' \cdot \Delta S \cdot l$ і

$$\sigma' = P\cos\alpha = P_n = \vec{P} \cdot \vec{n} , \qquad (2.13)$$

де P_n – проекція вектора поляризації на напрямок нормалі до поверхні діелектрика. Для правої поверхні діелектрика на рис.2.10 кут α – гострий, $\cos\alpha > 0$, $P_n > 0$, і, відповідно, σ' для неї позитивна; для лівої поверхні кут α' – тупий, $\cos\alpha' < 0$, $P_n < 0$ і відповідно σ' для неї негативна.

§2.4. Електричне поле всередині діелектрика. Діелектрична проникність та її фізичний зміст. Зв'язок між діелектричною сприйнятливістю та діелектричною проникністю

Розглянемо діелектрик, який знаходиться у зовнішньому електростатичному полі з напруженістю $\vec{E}_0$, яке створене вільними зарядами q_i. Поле всередині діелектрика є суперпозиція зовнішнього поля $\vec{E}_0$ і внутрішнього поля $\vec{E}'$, яке створюється зв'язаними зарядами, тобто

$$\vec{E} = \vec{E}_0 + \vec{E}' . \qquad (2.14)$$

Необхідно уточнити, що під $\vec{E}'$ розуміють деяке усереднене поле зв'язаних зарядів. Властивості вільних і зв'язаних зарядів тотожні. Тому для зв'язаних зарядів теж справедлива теорема Остроградського-Гаусса. Згідно (1.15) можемо записати, що

$$\oint_S \vec{E} d\vec{s} = \oint_S \vec{E}_0 d\vec{s} + \oint_S \vec{E}' d\vec{s} = \frac{1}{\varepsilon_0}(\sum q_i + \sum q'), \qquad (2.15)$$

де $\sum q_i$ – сума вільних зарядів, що охоплюється замкнутою поверхнею *S*, $\sum q'$ – сума зв'язаних зарядів, з якими пов'язана напруженість поля $\vec{E}'$, що теж охоплюється поверхнею *S*. Так як поле $\vec{E}$ всередині діелектрика завжди менше зовнішнього поля $\vec{E}_0$, то потік поля $\vec{E}'$ через замкнуту поверхню *S*:

$$\oint_S \vec{E}' d\vec{s} = \frac{1}{\varepsilon_0}\sum q' < 0.$$

Можна показати, що величина зв'язаних зарядів дорівнює

$$\sum q' = -\oint_s \vec{P}\, d\vec{s}. \qquad (2.16)$$

В результаті (2.15) перепишеться як

$$\oint_s (\varepsilon_0 \vec{E} + \vec{P})\, d\vec{s} = \sum q_i, \qquad (2.17)$$

тому що інтегрування ведеться по одній і тій же замкнутій поверхні *S*.

Для характеристики електричного поля в діелектрику введемо допоміжну величину

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P}, \qquad (2.18)$$

яку будемо називати *вектором електричного зміщення (вектором електричної індукції).*

В результаті рівняння, яке виражає *теорему Остроградського-Гаусса для електричного поля в середовищі,* можна записати у вигляді:

$$\Phi_D = \oint_S \vec{D} \cdot d\vec{S} = \sum q_i, \qquad (2.19)$$

де Φ_D – потік електричного зміщення.

Згідно з цією теоремою ***потік електричного зміщення електричного поля черз довільну замкнуту поверхню дорівнює алгебраїчній сумі вільних зарядів, що охоплюються цією поверхнею.***

Вводячи допоміжну величину, вектор $\vec{D}$, ми позбулися труднощів, які зв'язані з наявністю в (2.15) невідомої величини $\sum q'$. Для вакууму $\vec{P}$=0, тому $\vec{D} = \varepsilon_0 \vec{E}$. Згідно (2.11) вектор поляризації $\vec{P}$ пропорційний до напруженості поля в діелектрику. Отже,

$$\vec{D} = \varepsilon_0 \vec{E} + \chi \varepsilon_0 \vec{E} = \varepsilon \varepsilon_0 \vec{E}, \qquad (2.20)$$

де $\varepsilon = 1 + \chi$. Тоді

$$\vec{D} = \varepsilon \varepsilon_0 \vec{E}. \qquad (2.21)$$

Так же як і χ, ε – безрозмірна величина. Її називають відносною діелектричною проникністю, або просто діелектричною проникністю. Це – важлива константа діелектрика. Для сили взаємодії двох точкових зарядів, занурених в однорідний безмежний діелектрик, закон Кулона (1.2) приймає вид:

$$F = \frac{k q_1 q_2}{\varepsilon r^2}.$$

Необхідно зауважити, що ця формула справедлива лише для рідких і газоподібних діелектриків. Тобто можна сказати, що діелектрична проникність ε показує в скільки раз сила взаємодії двох точкових зарядів в діелектрику менша від цієї ж сили взаємодії зарядів у вакуумі при однакових відстанях між ними.

Далі буде показано, що ε показує в скільки раз напруженість поля в діелектрику менша від напруженості поля поза ним, якщо силові лінії вектора $\vec{E}$ перпендикулярні до поверхні діелектрика.

Відзначимо, що формула (2.21) справедлива для ізотропних діелектриків. Для анізотропних діелектриків напрямки $\vec{D}$ і $\vec{E}$ не співпадають.

Висновки.*1.**Результуюче поле в діелектрику визначає вектор напруженості $\vec{E}$ (формула 2.14) і тому воно залежить від властивостей діелектрика. Вектор $\vec{D}$ від середовища не залежить і описує електростатичне поле, що створюється вільними зарядами. ***Зв'язані заряди, що виникають в діелектрику, викликають перерозподіл вільних зарядів. Тому вектор $\vec{D}$ характеризує електростатичне поле, що створюється вільними зарядами (тобто у вакуумі), але при такому їх розподілі в просторі, який є при наявності діелектрика.

2.Вектор $\vec{D}$ не залежить від властивостей середовища тоді, коли поле створюється зарядженими тілами в однорідному ізотропному середовищі.

3.Лінії вектора $\vec{E}$ можуть починатися і закінчуватись як на вільних зарядах так і на зв'язаних зарядах , а лінії вектора $\vec{D}$ – лише на вільних зарядах. Через області поля, де знаходяться зв'язані заряди, лінії вектора $\vec{D}$ проходять так, ніби зв'язаних зарядів не має.

§2.5.Електричне поле на межі двох діелектриків

На межі двох середовищ з різними діелектричними властивостями вектори електричного поля змінюються за величиною і напрямком. Умови поведінки векторів $\vec{D}$ і $\vec{E}$ на межі поділу середовищ називають граничними. Ці умови отримуються із основних рівнянь поля. Встановимо спочатку граничні умови для нормальної складової вектора $\vec{D}$ (рис.2.11). Нехай лінія MN – межа розділу двох

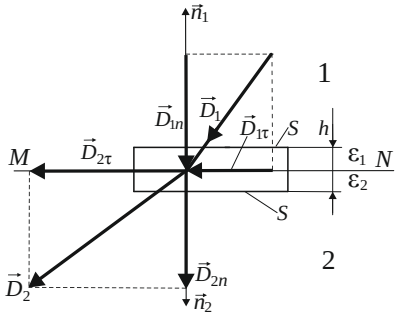


Рис.2.11.

однорідних ізотропних діелектриків 1 і 2, проникності яких ε_1 і ε_2 відповідно. $\vec{D}_1$ – вектор зміщення поля в середовищі 1, а $\vec{D}_2$ – вектор зміщення в середовищі 2. На межі розділу візьмемо уявну циліндричну поверхню, висота якої h, а площа основи S. Будемо вважати площу основи настільки малою, що в її межах поле можна вважати однорідним. $\vec{n}_1$ і $\vec{n}_2$ – нормалі до площини основи в першому і другому середовищах (нормалі направлені назовні поверхні). Використаємо теорему Остроградського-Гаусса (2.19). Якщо вільних зарядів на межі між діелектриками не має, то $\Phi_D = 0$. Потік через основу S в першому середовищі дорівнює:

$$\Phi_{D_{1n}} = \vec{D}_1\vec{n}_1 S = -D_{1n}S,$$

де D_{1n}– проекція вектора $\vec{D}_1$ на нормаль $\vec{n}_1$.

Аналогічно потік через основу S в другому середовищі

$$\Phi_{D_{2n}} = \vec{D}_2\vec{n}_2 S = D_{2n}S,$$

де D_{2n}– проекція вектора $\vec{D}_2$ на нормаль $\vec{n}_2$. Потік через бічну поверхню можна представити у вигляді $\Phi'_{D_n} = \langle D_n\rangle S_{біч}$, де $\langle D_n\rangle$ – значення D_n усереднене по всій бічній поверхні, $S_{біч}$ – бічна поверхня циліндра. Таким чином згідно (2.19) можемо записати:

$$\Phi_D = \Phi_{D_{1n}} + \Phi_{D_{2n}} + \Phi'_{D_n} = 0$$

або

$$-D_{1n}S + D_{2n}S + \langle D_n\rangle S_{біч} = 0. \quad (2.22)$$

Якщо висота циліндра h прямує до нуля, то $S_{біч}$буде прямувати до нуля і із (2.22) отримаємо:

$$D_{1n} = D_{2n}. \quad (2.23)$$

Скориставшись формулою (2.21) можемо записати, що

$$\varepsilon_1\varepsilon_0 E_{1n} = \varepsilon_2\varepsilon_0 E_{2n}.$$

Звідки отримаємо співвідношення:

$$\frac{E_{1n}}{E_{2n}} = \frac{\varepsilon_2}{\varepsilon_1}. \quad (2.24)$$

Виберемо на межі розділу двох діелектриків MN невеликій прямокутний контур Г, довжина якого a, а ширина b. Контур частково проходить в першому діелектрику, частково – у другому (рис.2.12). Межа розділу MN (вісь x) проходить через середини сторін контуру, ширина яких b.

Нехай в діелектриках створено поле, напруженість якого в першому діелектрику дорівнює $\vec{E}_1$, а другому – $\vec{E}_2$. Кожний із векторів $\vec{E}_1$ і $\vec{E}_2$ можна представити у вигляді суми нормальної і тангенціальної складової:

$$\vec{E}_1=\vec{E}_{1n}+\vec{E}_{1\tau};\ \vec{E}_2=\vec{E}_{2n}+\vec{E}_{2\tau}.$$

Тут $\vec{E}_{1\tau}$ і $\vec{E}_{2\tau}$ проекції векторів $\vec{E}_1$ і $\vec{E}_2$ на орт $\vec{\tau}$, який направлений вздовж лінії перетину площини, яка розділяє діелектрики, з площиною, в якій лежать вектори $\vec{E}_1$ і $\vec{E}_2$.

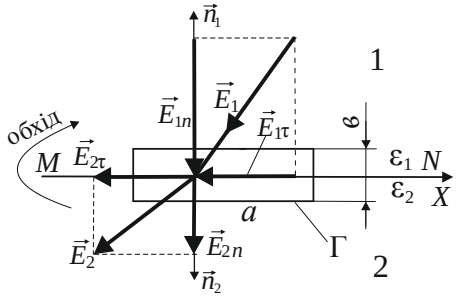


Рис.2.12.

Циркуляція вектора $\vec{E}$ по вибраному нами замкнутому контуру Г згідно теореми про циркуляцію (1.32) повинна дорівнювати нулю. При малих розмірах контуру Г і згідно вказаного на рис.2.12 обходу циркуляція вектора $\vec{E}$ може бути записана у вигляді

$$\oint_\Gamma \vec{E}\,d\vec{l} = -E_{1\tau}a+E_{2\tau}a+\langle E_b\rangle 2b = 0,$$

де $\langle E_b\rangle$ – середнє значення напруженості поля E_b на перпендикулярних до межі $M\,N$ ділянках (довжиною $2b$) контуру Г. Якщо ширина контуру b прямує до нуля, то отримаємо рівність:

$$E_{1\tau} = E_{2\tau}. \qquad (2.25)$$

Замінимо згідно (2.21) проекції вектора $\vec{E}$ проекціями вектора $\vec{D}$, які розділимо на $\varepsilon\varepsilon_0$. Тоді отримаємо співвідношення:

$$\frac{D_{1\tau}}{\varepsilon_1} = \frac{D_{2\tau}}{\varepsilon_2}. \qquad (2.26)$$

Отримані нами результати означають, що при переході через межу розділу двох діелектриків нормальна складова вектора $\vec{D}$ (2.23) і тангенційна складова вектора $\vec{E}$ (2.25) змінюються неперервно. Тангенційна складова вектора $\vec{D}$ (2.26) і нормальна складова вектора $\vec{E}$ (2.24) при переході через границю розділу зазнають розриву. Співвідношення (2.23), (2.24), (2.25) і (2.26) визначають умови, яким повинні задовільняти вектори $\vec{E}$ і $\vec{D}$ на межі двох діелектриків (в тому випадку, якщо на цій межі не має вільних зарядів). Ці співвідношення отримані для електростатичного поля. Однак, вони справедливі і для полів, які змінюються з часом.

Знайдені умови справедливі і для межі діелектрика з вакуумом. В цьому випадку одну із діелектричних проникностей необхідно прийняти за одиницю.

§2.6.Особливі властивості деяких діелектриків

Розглянуті раніше неполярні і полярні діелектрики мають ту особливість, що вони поляризуються тільки при накладанні зовнішнього електричного поля $\vec{E}_0$, а після припинення його дії знову повертаються в неполяризований стан.

Однак у природі існує велика група діелектриків, переважно у твердому кристалічному стані, які виявляють макроскопічну поляризованість, навіть якщо зовнішнього поля не має. Це спонтанно (латинське – spontaneus – самодовільний) поляризовані діелектрики, тобто такі,

поляризація яких викликана не зовнішніми, а внутрішніми причинами. До спонтанно поляризованих діелектриків відносять так звані ***піроелектрики*** і ***сегнетоелектрики***. Спонтанна поляризація є наслідком зміщення "центру ваги" позитивних і негативних зарядів кристалічної ґратки речовини.

Для піроелектриків спостерігається не сама спонтанна поляризація $\vec{P}$ (вона компенсується полями вільних електричних зарядів, що адсорбуються поверхнею із повітря, або компенсується внутрішніми іонами), а її зміна $\Delta\vec{P}$ при швидкій зміні температури ΔT (піроелектричний ефект). Густина поверхневих зарядів, які при цьому виникають,

$$\sigma = p \cdot \Delta T ,$$

де p – піроелектрична стала, величина якої має порядок 10^{-3} Кл/(м^2·К). Відомий піроелектрик – турмалін (боратосилікатний мінерал змінного складу). В ньому при зміні температура на 1℃ виникає електричне поле $E \sim 400$ В/см. Крім турмаліну, до піроелектриків належать сульфат літію, цукор, окремі види кераміки та ін. Характерною особливістю піроелектриків є наявність поляризації в усій області температур їхнього існування. Вони застосовуються як індикатори і приймаючі випромінювання та пристрої теплового бачення.

Сегнетоелектрики

Експериментальні дослідження піроелектриків показали, що вони за своєю поведінкою в зовнішньому електричному полі можуть бути поділені на лінійні і нелінійні діелектрики. Цей поділ визначається характером залежності вектора поляризації $\vec{P}(\vec{E})$. Серед нелінійних твердих діелектриків найбільш поширеними є ***сегнетоелектрики***.

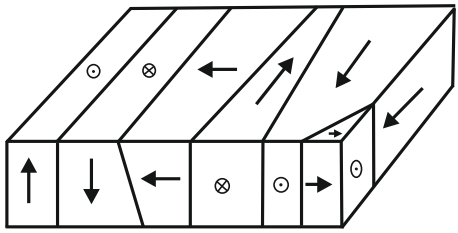


Рис.2.13.Схематичне зображення доменів в сегнетоелектрику ($BaTiO_3$); стрілки і позначення ⊙ і ⊗ вказують на напрямок вектора поляризації $\vec{P}$ домена.

Сегнетоелектрики, кристалічні діелектрики, які володіють в певному інтервалі температур, яка називається полярною фазою, спонтанною поляризацією, яка суттєво змінюється під дією зовнішніх впливів (електричних полів, змін температури, тощо). Сегнетоелектричні властивості були вперше виявлені у кристалів ***сегнетової солі*** $KNaC_4H_4O_6{\cdot}4H_2O$ в 1920р., потім у дигідрофосфата калія $(KDP)KH_2PO_4$). Відомо декілька сотень сегнетоелектриків. Особливість їх в тому, що в них вектор поляризації $\vec{P}$ порівняно легко змінюється під дією електричних полів, пружних деформацій, змін температури та ін. Сегнетоелектрики неоднорідно поляризовані. Вони складаються із ***доменів*** – областей із різними напрямками поляризації (рис.2.13). В результаті сумарний електричний дипольний момент зразка практично відсутній.

Рівноважна доменна структура сегнетоелектрика відповідає мінімуму вільної енергії кристалу. Доменна структура реального кристала визначається природою і характером розподілу його дефектів, а також історією зразка. Число різних доменів, взаємна орієнтація їхньої спонтанної поляризації залежить від ***симетрії кристалу***.

Сегнетоелектриками є кристали, у

яких відсутній центр симетрії. Тому властивості сегнетоелектриків максимально виражені в певному напрямку в кристалі, який називають ***полярною віссю***. Тобто сегнетоелектрики мають різку анізотропію діелектричних властивостей. ***Анізотропія*** означає, що фізичні властивості тіла різні в різних напрямках. Модифікацію кристалів, в яких сегнетоелектрики спонтанно поляризовані, називають ***полярною фазою*** (рис.2.15).

Характеристики деяких сегнетоелектриків

Таблиця 2.1.

Речовина	Хімічна формула	Темпер. фазового переходу T_K,°C	Максимальна спонтанна поляризац. P,мкКл/см^2
Титанат барію	$BaTiO_3$	120	25
Сегнетова сіль	$KNaC_4H_4O_6 \cdot 4H_2O$	$-18 \div -24$	0,25
Дигідрофосфат калію	KH_2PO_4	−150	5,1
Ніобат літію	$LiNB O_3$	1210	50
Титанат вісмуту	$Bi_4Ti_3O_{12}$	675	-

З підвищенням температури вище деякою значення T_K, характерного для кожного сегнетоелектрика, відбувається фазовий перехід, у результаті якого сегнетоелектричні властивості зникають і кристал переходить у звичайний неполярний діелектрик. Такий стан кристалу називають параелектричним, а фазу неполярною. Температуру T_K називають температурою Кюрі. В табл.2.1 приведені деякі характеристики частини сегнетоелектриків. Практично важливим сегнетоелектриком є титанат барію, який використовується для виготовлення конденсаторів з ємністю, яка змінюється в залежності від температури і величини прикладеної напруги.

Залежність діелектричної проникності титану барію від напруженості електричного поля і температури приведені на рис.2.14 і 2.15. Титанат барію має високу механічну міцність, велике значення ε, найбільш просту кристалічну структуру в порівнянні зі всіма відомими сегнетоелектриками, три точки Кюрі (–90 °C, 5 °C і 120 °C).

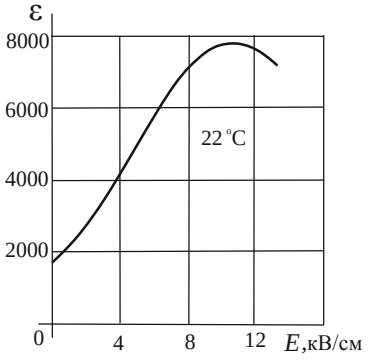


Рис.2.14.Залежність діелектричної проникності титанату барію від напруженості електричного поля.

Дамо якісне пояснення залежностей $\varepsilon(E)$ і $\varepsilon(T)$ для сегнетоелектрика. При внесенні сегнетоелектрика в електричне поле він поляризується. Електрична індукція сегнетоелектрика складається з трьох складових:

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} + \vec{P}_s, \qquad (2.27)$$

де $\vec{E}$ – середнє макроскопічне поле в сегнетоелектрику, $\vec{P}$– середній електричний момент одиниці об'єму, обумовлений електронною і іонною поляризаціями, $\vec{P}_s$ – електричний момент спонтанної поляризації. Вектори $\vec{E}$ і $\vec{P}$ мають однаковий напрямок, а вектор $\vec{P}_s$ в загальному випадку має інший напрямок. Повна діелектрична проникність сегнетоелектрика, яка обумовлена поляризацією, що викликана зовнішнім полем і

спонтанною поляризацією, дорівнює відношенню $\vec{D}/(\varepsilon_0\vec{E})$. Із (2.27) отримаємо:

$$\varepsilon = 1 + \frac{P}{\varepsilon_0 E} + \frac{P_{sE}}{\varepsilon_0 E}, \qquad (2.28)$$

де P_{sE} – проекція вектора $\vec{P}_s$ на напрямок поля $\vec{E}$. Із збільшенням поля $\vec{E}$, повний електричний момент одиниці об'єму (вектор поляризації) може збільшуватись з двох причин: по-перше, внаслідок зростання електричного моменту поляризації зміщення $\vec{P}$, по-друге, внаслідок зростання числа областей спонтанної поляризації (доменів), напрямки електричних моментів яких співпадають з напрямком зовнішнього поля або утворюють з ним гострі кути. Так якісно пояснюється залежність $\varepsilon(E)$.

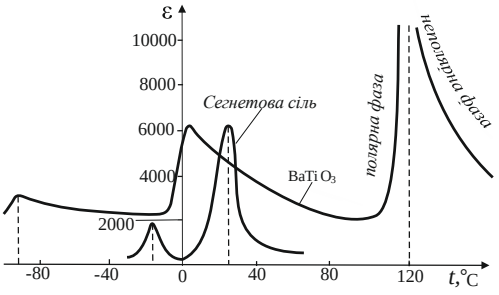


Рис.2.15.Залежність діелектричної проникності сегнетової солі та титанату барію від температури.

Спонтанна поляризація сегнетоелектрика залежить від зміни температури, так як при цьому змінюється інтенсивність теплового руху, який суттєво зменшує її особливо на границях полярної області. В неполярній області сегнетоелектрик поводить себе як звичайний діелектрик, в якому поляризація пропорційна напруженості електричного поля (домени в цій області температур відсутні). Однак, у всіх випадках діелектрична проникність змінюється з температурою. Так вище точки Кюрі ($T > T_K$) ця зміна відбувається за законом Кюрі-Вейса:

$$\varepsilon = \frac{C}{T - T_K},$$

де C і T_K – постійні, з яких T_K називається температурою (точкою) Кюрі. В області точки Кюрі поляризованість сегнетоелектрика аномально велика (рис.2.15). Це пов'язано з тим, що при температурі Кюрі відбувається перебудова структури кристалічної гратки сегнетоелектрика. Титанат барію в неполярній фазі, вище 120°С, має кубічну структуру, нижче 120°С – тетрагональну.

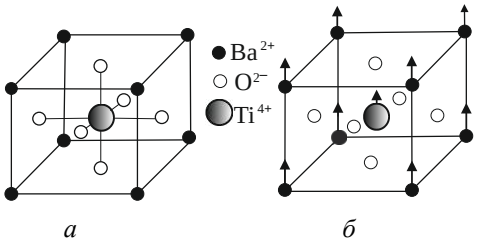


Рис.2.16.

Отже, властивості сегнетоелектриків пояснюються існуванням областей спонтанної поляризації (доменів). Утворення доменів в титанаті барію відбувається таким чином. В неполярній фазі (рис.2.15) (вище точки Кюрі) титан барію має кубічну структуру (рис.2.16,*а*). У вершинах куба – іони Ba^{2+}, в центрах граней – іони O^{2-}, іон Ti^{4+} – в центрі куба. Через наявність центра симетрії дипольний момент елементарної комірки дорівнює нулю і титанат барію в неполярній фазі є звичайним діелектриком з великою діелектричною проникністю.

За температури Кюрі відбувається фазовий перехід, який призводить до деформації кубічної структури (рис.2.16,*б*): іони Ba^{2+} і Ti^{4+} зміщуються відносно іонів O^{2-} і таким чином створюють дипольний момент елементарної комірки, структура якої в

полярній фазі тетрагональна. Близьке розташування окремих диполів один по відношенню до одного обумовлює їх сильну електричну взаємодію між собою. Внаслідок цього ціла група диполів орієнтується в одному напрямку (така орієнтація диполів відповідає мінімуму енергії), тобто утворюються області (домени) з дипольним моментом, який дорівнює сумі дипольних моментів молекул, що входять в цю область. Накладання зовнішнього електричного поля створює переважну орієнтацію векторів поляризації доменів в напрямку поля у зв'язку з чим поляризованність і електрична проникність сегнетоелектрика стають досить великими (рис.2.15).

Підвищення температури в полярній області полегшує орієнтацію векторів поляризації окремих доменів в напрямку поля. При температурі Кюрі крива $\varepsilon(T)$ сегнетоелектрика (рис.2.15) має різко виражений максимум. При підвищенні температури вище точки Кюрі тетрагональна гратка титанату барію перебудовується в правильну кубічну, доменні групи розпадаються і величина ε різко зменшується.

Із аналізу залежності $\varepsilon(T)$ на рис.2.15 витікає, що допускати нагрівання конденсаторів, які містять діелектрик титанат барію, до 120°С не рекомендується, так як сильно змінюється ε, а отже, будуть сильно змінюватися характеристики конденсаторів.

Відзначимо, що сегнетоелектрики використовуються для виготовлення малогабаритних конденсаторів ***великої ємності*** і як нелінійні елементи схем (помножувачів частоти, діелектричних підсилювачів і т. п.).

Для сегнетоелектриків характерне явище *діелектричного гістерезису* (від грецького hysteresis – відставання, запізнення), яке полягає у відставанні змін вектора поляризації $\vec{P}$ від змін вектора напруженості $\vec{E}$ електричного поля (рис.2.17).

Якщо неполяризований сегнетоелектрик помістити в електричне поле і збільшувати напруженість поля E, то викликана цим полем поляризація зростає вздовж кривої OA, яка називається ***основною кривою*** поляризації сегнетоелектрика. E_A – напруженість поля, при якій досягається величина спонтанної поляризації P_s. При зменшенні поля E поляризація зменшується не вздовж кривої OA, а вздовж кривої AB_1, причому в точці B_1

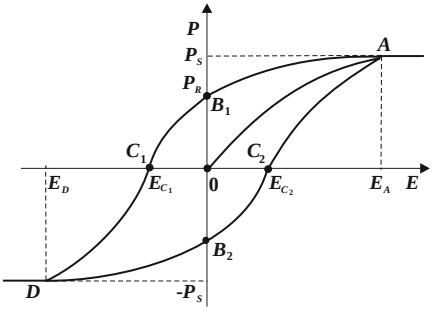


Рис.2.17

зовнішнє поле відсутнє E=0, але сегнетоелектрик поляризований. При зміні напрямку зовнішнього поля вектор поляризації зменшується до нуля в точці C_1, а потім змінивши свій напрямок, зростає за величиною вздовж кривої C_1D. Значення напруженості поля E_{C_1} в точці C_1 називається ***коерцитивною силою***. Наступна зміна зовнішнього поля від E_D до E_A супроводжується зміною вектора поляризації вздовж кривої DB_2C_2A. В точках B_1 і B_2 має місце ***залишкова поляризація*** P_R. Величина напруженості E_{C_2} в точці C_2 теж називається коерцитивною силою. Звужуючи інтервал зміни E, можна отримати нескінченну

множину петель гістерезису (рис.2.17), вкладених одна в одну. Таким чином вектори поляризації $\vec{P}$ і електричного зміщення $\vec{D}$ є багатозначними функціями E. У зв’язку з цим можна говорити про діелектричну сприйнятливість і діелектричну проникність сегнетоелектрика тільки для основної кривої OA (рис.2.17).

При періодичній зміні поляризації сегнетоелектрика, на подолання тертя під час повертання електричних моментів витрачається електрична енергія, яка йде на нагрівання сегнетоелектрика. Площа петлі гістерезису пропорційна до електричної енергії, що перетворюється у внутрішню енергію одиниці об’єму сегнетоелектрика за один цикл.

П’єзоелектрики

Явище виникнення на гранях кристалів зв’язаних електричних зарядів протилежних знаків при їхній механічній деформації називають прямим ***п’єзоелектричним ефектом*** (грецьке piezo – тиснути).

Це явище відкрили і дослідили у 1880 р. французькі фізики Жан і П’єр Кюрі на кристалі кварцу (SiO_2). Дослідження показали, що п’єзоелектричні властивості мають майже 1500 різних кристалічних речовин, у тому числі турмалін, сегнетова сіль, сульфат літія ($Li_2SO_4{\cdot}H_2O$), титанат барію ТВ-1, титанат барію кальцію ТБК – 3, цинкова обманка, хлорат натрію, тростниковий цукор, напівпровідники CdS, InSb, GaAs і багато інших кристалів.

П’єзоелектричний ефект притаманний лише іонним кристалам, які не мають центра симетрії. Якщо кристал має центр симетрії, то при однорідній механічній деформації центр симетрії зберігається, "центри ваги" позитивних і негативних іонів кристала співпадають і, отже, в них не може виникати п’єзополяризація.

Знак зв’язаних зарядів на гранях кристала без центра симетрії при деформаціях однозначно визначається напрямком і характером деформації. Якщо при стисканні на певній грані виникає позитивний заряд, то при розтягу ця грань заряджається негативно. Після припинення деформації кристалічний діелектрик повертається в неполярний стан.

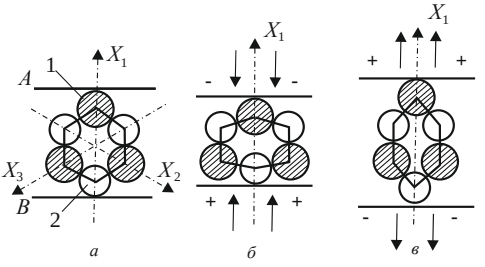


Рис.2.18.Структурна схема кварцу.
Заштриховані кружечки відповідають іонам Si^+, світлі – парі іонів O^-; *а* – недеформований стан; *б* – стиснення вздовж осі X_1; *в* – розтяг вздовж осі X_1. Стрілками показаний напрямок деформації.

Механізм п’єзоефекту пояснимо на прикладі кристалу кварцу, який знайшов широке застосування завдяки своїм механічним і електричним властивостям. Кристалічна гратка кварцу складається із іонів кремнію Si^+ і негативних іонів кисню O^-, має три осі симетрії (на рис.2.18,*а* вони позначені X_1, X_2, X_3), які є полярними осями кристалу.

На рис.2.18,*б* і *в* показані проекції іонів Si^+ і O^-на площину, яка перпендикулярна осі третього порядку X_3. При стиснені вздовж осі X_1 позитивні іони 1(Si^+) і негативні іони 2(O^-) зміщуються в глибину комірки кристалу (рис.2.18,*б*), в результати чого на площинах *A* і *B* появляються заряди. При розтягу (рис.2.18,*в*) на площинах *A* і *B* виникають заряди протилежних знаків.

Максимальний п'єзоелектричний ефект спостерігається тоді, коли деформацію здійснюють у напрямку однієї із полярних осей кристалу. Тому кварцеві пластинки вирізають так, щоб пара площин, що утворюються при розрізанні, була перпендикулярна до однієї із полярних осей.

П'єзоелектричний ефект є оборотним. Накладання різниці потенціалів до граней кристалу (створення всередині напруженості електричного поля) приводить до механічної деформації (розтягу або стискування в залежності від полярності напруги). Якщо подати змінну електричну напругу до граней кристалу, тобто створити змінну поляризацію кристалу, то пластинка буде поперечно розтягуватись і стискуватись вздовж, наприклад, осі X_1 (рис.2.18,*б,в*). Таким чином в пластинці збуджується механічні коливання. Ці коливання максимально інтенсивні, якщо частота змінної прикладеної до пластини напруги збігається із власною частотою коливань пластинки.

Таким чином, ***явище деформації п'єзоелектричного*** кристалу під дією прикладеної до його граней різниці потенціалів електричного поля називають ***оберненим п'єзоелектричним ефектом.***

Використовують п'єзоелектричні матеріали для виготовлення п'єзодатчиків, призначених перетворювати механічні коливання в електричні. Акустичних генераторів для перетворення електричних змін у звукові коливання. Виготовляють адаптери для відтворення звуку, різні прилади для вимірювання частоти і амплітуди вібрації двигунів, турбін, літаків тощо; п'єзоелектричні сейсмографи для реєстрації коливань земної кори, прилади для вимірювання прискорень – акселерометри; п'єзоелектричні манометри (датчики тиску), стабілізатори частоти радіостанцій та інше.

Електрети

Електрети – діелектрики, які довгий час зберігають поляризований стан після припинення зовнішніх дій, які викликали поляризацію, і створюють електричне поле в оточуючому просторі (електричний аналог постійних магнітів).

Якщо речовину, молекули якої мають дипольний момент, розплавити і помістити в сильне електричне поле, то її молекули частково вишикуються по полю. При охолодженні розплаву в електричному полі з подальшим виключенням поля в затверділій речовині поворот молекул затруднений і вони довгий час зберігають переважну орієнтацію (від декількох днів до багатьох років). Перший електрет таким способом був виготовлений в 1922 р. японським фізиком М. Йогучи.

Електрети отримуються із аморфних воску і смол (бджолиний віск, парафін і т.д.), полімерів (поліметилметакрилат, полівінілхлорид та ін.), неорганічних полікристалічних діелектриків (титанати лужноземельних металів, ультрафарфорів та інших керамічних діелектриків), монокристалічних неорганічних діелектриків (лужногалоїдні монокристали, наприклад, LiF, корунд, рутил), скла та ін.

Стабільні електрети отримують нагріванням з наступним охолодженням в сильному електричному полі (*термоелектрети*), освітленням в сильному електричному полі (фотоелектрети), опроміненням радіоактивним випромінюванням (*радіоелектрети*), поляризацією в сильному електричному полі без нагрівання (*електроелектрети*) або в

магнітному полі (*магнітоелектрети*), при застиганні органічних розчинів в електричному полі (кріоелектрети), механічною деформацією полімерів (механоелектрети), тертям (трібоелектрети), дією поля короннного розряду (*короноелектрети*). Всі електрети мають стабільний поверхневий заряд.

Електрети використовують як джерела постійного електричного поля (електретні мікрофони й телефони, вібродатчики, генератори слабих змінних сигналів, електрометри, електростатичні вольтметри та інше), а також як чутливі датчики в пристроях дозиметрії, електричної пам'яті; для виготовлення барометрів, гігрометрів і газових фільтрів, пьєзодатчиків та інше. Фотоелектрети застосовують в електрофотографії.

Контрольні питання

1.Вектор поляризації та його фізичний зміст.

2.Яким чином можна визначити поляризаційні (зв'язані) заряди за допомогою вектора поляризації?

3.Вектор електричного зміщення (індукції) та його фізичний зміст.

4.Теорема Остроградського-Гаусса для електричного поля в діелектриках.

5.Зв'язок між векторами електричного зміщення та вектором напруженості електричного поля.

6.Електричне поле на межі двох діелектриків.

7.Які основні властивості сегнетоелектриків? Яке практичне застосування сегнетоелектриків?

8.Пьєзоелектрики та електрети. Їх практичне застосування.

Розділ 3.Провідники в електричному полі

§3.1.Розподіл зарядів, напруженість та потенціал електричного поля у провіднику

Характерною особливістю провідників є наявність у них вільних носіїв заряду (§2.1). В металах це електрони провідності ("вільні" електрони). Концентрація "вільних" електронів, наприклад міді, є величина n = =8,5·10^{28} м$^{-3}$. Ці електрони можуть рухатись у металах вільно, але не можуть самостійно покидати його поверхню.

Якщо провідник помістити в зовнішнє електростатичне поле $\vec{E}_{зовн}$, то на кожен "вільний" електрон буде діяти сила $\vec{F}_1 = e\vec{E}_{зовн}$

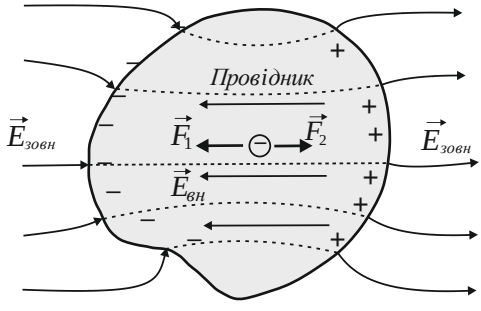


Рис.3.1.

(рис.3.1). Під дією сили $\vec{F}_1$ відбувається переміщення "вільних" електронів і внаслідок цього електричні заряди перерозподіляються: на одній стороні провідника буде надлишок "вільних" електронів, які заряджають її негативно, на іншій виникає їх нестача і ця сторона заряджається позитивно.

Явище перерозподілу вільних носіїв заряду у провіднику під дією зовнішнього електричного поля, внаслідок чого виникає

електризація, називається електростатичною індукцією або електризацією через вплив.

Індуктивні заряди створюють всередині провідника внутрішнє електростатичне поле $\vec{E}_{вн}$, яке діє на "вільні" електрони із силою $\vec{F}_2 = e\vec{E}_{вн}$. Переміщення "вільних" зарядів у провіднику припиняється, коли $\vec{F}_1 + \vec{F}_2 = 0$. Тоді $e(\vec{E}_{зовн} + \vec{E}_{вн}) = 0$. Так як $e \neq 0$, то $\vec{E}_{зовн} + \vec{E}_{вн} = 0$ і $\vec{E}_{вн} = -\vec{E}_{зовн}$. Внутрішнє наведене електричне поле $\vec{E}_{вн}$ буде компенсувати зовнішнє – $\vec{E}_{зовн}$ і рух "вільних" зарядів припиниться. Якщо б всередині провідника залишилось хоч яке-небудь електричне поле, воно змусило би рухатись ще якісь електрони.

Таким чином, ***результуюча напруженість електричного поля всередині провідника дорівнює нулю.***

Тепер розглянемо середину провідника. Оскільки всередині провідника напруженість електричного поля дорівнює нулю, то це значить, що і градієнт потенціалу φ дорівнює нулю: $\vec{E} = \text{grad}\varphi$. Це означає, що φ від точки до точки не змінюється ($\varphi = const$). Таким чином, *будь-який провідник – це еквіпотенціальна область і його поверхня – еквіпотенціальна*.

Якщо в провідному матеріалі електричне поле скрізь дорівнює нулю, то і дивергенція $\vec{E}$ теж дорівнює нулю ($\text{div}\vec{E} = \rho/\varepsilon_0$ формула (1.38)). Тому густина заряду у внутрішній частині провідника дорівнює нулю ($\rho = 0$).

Але якщо всередині провідника не може бути надлишкових зарядів, то як тоді він може бути заряджений? Що ми маємо на увазі, коли говоримо, що провідник "заряджений"? Де ці заряди знаходяться? Мабуть, вони можуть знаходитись на поверхні провідника, де існують сили, які утримують їх. Тож вони не зовсім "вільні" у провідниках.

Необхідно сказати, що спеціально незаряджений провідник є електронейтральним. Сумарний заряд всіх "вільних" електронів в 1м^3 міді Q=8,5·10^{28}·1,6·10^{-19} Кл≅1,36·10^{10} Кл. Цей негативний заряд електронів компенсується позитивним зарядом іонів міді. Якщо провіднику надати певну кількість електронів, то він зарядиться негативно, а якщо забрати в нього частину електронів, то він зарядиться позитивно. Пробійна напруженість електричного поля для повітря $E_i \sim 10^7$ В/м. Це означає, що провідна куля (сфера), радіус якої $r = 1$м, може бути заряджена зарядом $q \cong 4\pi\varepsilon_0 r^2 E_i \cong 10^{-3}$ Кл. Цей заряд $q << Q$. Але величина заряду q є помітною тому, що незаряджений провідник в цілому є електронейтральним. Цей надмірний заряд в будь-якому провіднику знаходиться в тонкому шарі біля поверхні, товщина якого в середньому складає один-два атоми. Для наших сьогоднішніх цілей достатньо правильно буде говорити, що будь-який заряд, який попав на (або в) провідник, збирається на його поверхні. Це зрозуміло ще і з того факту, що однойменні заряди відштовхуються і вони будуть віддалятися один від одного на максимальну відстань у провіднику, тобто на поверхню.

Знайдемо зв'язок між напруженістю поля E біля поверхні зарядженого провідника, де поверхнева густина заряду σ (рис.3.2). Для цього розглянемо замкнену поверхню у вигляді циліндра з основою dS, вісь якого орієнтована вздовж вектора $\vec{D}$ (рис.3.2). Оскільки поверхня провідника еквіпотенціальна, то

вектор $\vec{D}$, як і $\vec{E}$, перпендикулярні до поверхні провідника. Основа циліндра 1 знаходиться зовні провідника, а основа 2 – всередині нього.

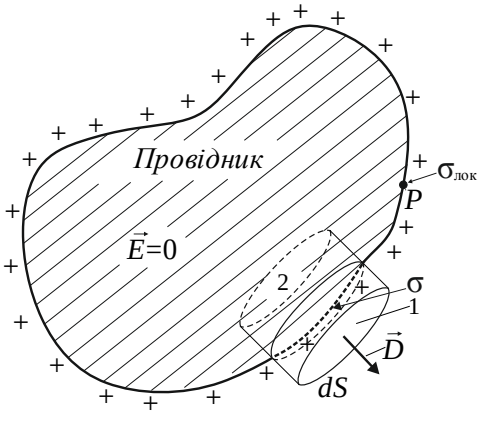


Рис.3.2.

Використаємо теорему Остроградського-Гаусса (2.19). Оскільки поле всередині провідника відсутнє, то потік вектора $\vec{D}$ через замкнуту циліндричну поверхню визначається лише потоком через зовнішню основу циліндра 1:

$$d\Phi_D = D\cdot dS = q = \sigma dS.$$

Звідси

$$D = \sigma = \varepsilon\varepsilon_0 E.$$

Отже, напруженість електричного поля біля поверхні зовні провідника довільної форми дорівнює:

$$E = \sigma/(\varepsilon\varepsilon_0), \qquad (3.1)$$

де ε – відносна діелектрична провідність середовища, в якому знаходиться провідник.

Виникає питання, чому шар зарядів на провіднику створює не таке поле, як шар зарядів сам по собі? Іншими словами, чому у формулі (3.1) (для повітря ε=1) величина напруженості вдвічі більша, ніж у формулі (1.16). Справа в тому, що для того, щоб у провіднику $\vec{E}$ дорівнювала нулю, в ньому обов'язково повинні бути якісь заряди. Безпосередньо близько біля точки P на поверхні (рис.3.2) заряди дійсно створюють поле $E_{лок}=\sigma_{лок}/(2\varepsilon_0)$, як всередині так і зовні провідника. Але всі інші заряди провідника створюють в точці P додаткове поле, яке по величині дорівнює $E_{лок}$. Напруженість сумарного внутрішнього поля дорівнює нулю, а зовнішнього – подвоюється: $2E_{лок}=\sigma_{лок}/\varepsilon_0$, (випадок, коли зовнішнє середовище вакуум (повітря)).

Розподіл зарядів по поверхні провідника

Поверхня зарядженого провідника при рівновазі є еквіпотенціальною, напруженість поля E всередині дорівнює нулю, однак поверхнева густина заряду залежить від кривизни поверхні. Доведемо це на прикладі двох заряджених куль, радіуси яких R_1 і R_2, що з'єднані між собою тонким провідником. Після встановлення рівноваги заряди на кулях розподіляються так, що потенціали куль будуть однаковими (дві кулі з'єднані провідником – це один провідник), тобто $\varphi_1 = \varphi_2$. Оскільки згідно формули (1.19)

$$\varphi_1/(4\pi\varepsilon_0 R_1)$$

і

$$\varphi_2/(4\pi\varepsilon_0 R_2),$$

а

$$q_1 = 4\pi\sigma_1 R_1^2 \text{ і } q_2 = 4\pi\sigma_2 R_2^2,$$

то

$$\frac{\sigma_1}{\sigma_2} = \frac{R_2}{R_1}. \qquad (3.2)$$

Отже, поверхнева густина зарядів на кулях обернено пропорційна їхнім радіусам кривизни, тобто де кривизна поверхні (кривизна кола, радіус якого R, скрізь дорівнює $1/R$) зарядженого провідника більша, там і густина заряду більша. Оскільки біля поверхні зарядженого провідника $E = \sigma/(\varepsilon\varepsilon_0)$, то в

місцях з великою кривизною (вістря) σ і, відповідно, E набувають великих значень.

На рис.3.3 наведено вигляд ліній

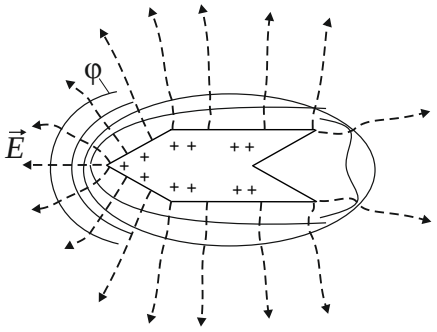


Рис.3.3.

напруженості (штрихові лінії) і еквіпотенційних поверхонь (суцільні лінії) поля позитивно зарядженого провідника циліндричної форми з конічним виступом на одному кінці і конічною западиною на іншому. З рисунка видно, що поблизу вістря і виступів еквіпотенціальні поверхні найсильніше зближені, відповідно і поверхнева густина вільних зарядів на цих ділянках провідника більша, ніж на інших поверхнях тіла. В області конічної западини напруженість поля мінімальна.

Значна величина напруженості поля поблизу гострого виступу на зарядженому провіднику приводить до явища під назвою “електричний вітер”. В атмосферному повітрі завжди є невелика кількість позитивних іонів і вільних електронів, що виникають під впливом космічних променів, випромінювання радіоактивних речовин тощо. В сильному електричному полі поблизу вістря електрони рухаються з великою швидкістю і, стикаючись з молекулами повітря, іонізують їх. Отже, виникає все більша кількість позитивних рухомих зарядів, які утворюють “електричний вітер”. Існування “електричного вітру” супроводжується розряджанням наелектризованого тіла, бо до нього рухаються під дією електричного поля заряди протилежного знаку по відношенню до зарядів наелектризованого тіла. Це приводить стікання електричних зарядів з металевих вістер. Стікання зарядів з вістер використовують для побудови блискавковідводів. У високовольтних лініях електропередач стікання зарядів з загострених частин проводів призводить до значних втрат енергії.

Поле всередині порожнини провідника

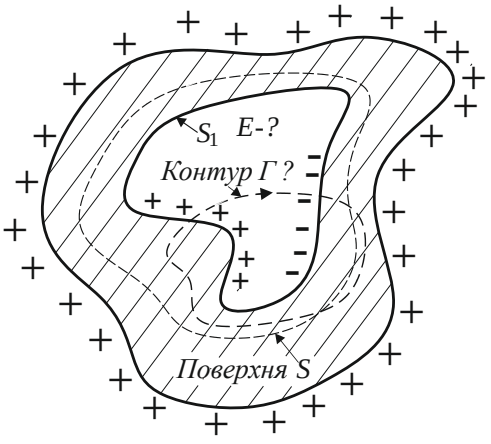


Рис.3.4.

Всередині провідника поля немає, а чи є воно в порожнині? Покажемо, що якщо порожнина пуста, то поля в ній бути не може, яка б не була форма провідника або порожнини (рис.3.4). Розглянемо замкнену поверхню S на рис.3.4, яка оточує собою порожнину, але залишається скрізь в речовині провідника. На поверхні S в кожній її точці напруженість поля дорівнює нулю, так що потік через S дорівнює нулю, і ***сумарний*** заряд всередині S теж повинен дорівнювати нулю. В загальному випадку ми можемо сказати, що на внутрішній поверхні провідника S_1 є однакова кількість позитивного і негативного зарядів, що розташувались в різних місцях цієї поверхні

(рис.3.4). Але однакові за величиною і протилежні за знаком заряди повинні рухатись назустріч одні одним і "знищити" один одного. В цьому ми можемо переконатися, застосувавши закон про циркуляцію вектора $\vec{E}$ (формула (1.32)). Нехай на деяких частинах внутрішньої поверхні з'явилися заряди. Ми знаємо, що десь ще повинна бути така сама кількість протилежних за знаком зарядів. Але будь-які лінії вектора $\vec{E}$ починаються на позитивних зарядах і закінчуються на негативних (розглянемо випадок, коли в порожнині вільні заряди відсутні). Уявімо собі контур *Г*, який перетинає порожнину вздовж силової лінії від якогось позитивного заряду до якогось негативного і повертається до вихідної точки по тілу провідника (рис.3.4). Інтеграл $\int \vec{E}d\vec{l}$ вздовж такої силової лінії в порожнині в межах від позитивного до негативного заряду був би відмінній від нуля, а такий самий інтеграл по шляху контура Г в металі дорівнював би нулю, так як у провіднику $\vec{E}$=0. В результаті ми отримали б, що $\oint_\Gamma \vec{E}d\vec{l} \neq 0$. Але згідно теореми про циркуляцію вектора $\vec{E}$, інтеграл по будь-якому замкнутому контуру Г в електростатичному полі $\oint_\Gamma \vec{E}d\vec{l}$ завжди дорівнює нулю. Значить, всередині порожнини не може бути ніяких полів, що рівнозначно тому, що не може бути ніяких зарядів на внутрішній поверхні провідника S_1.

Таким чином, ми показали, що якщо порожнина повністю оточена провідником, то ніякий статичний розподіл зарядів зовні ніколи не створить поля всередині. Це пояснює принцип "захисту" електронного обладнання, яке поміщають в металічну коробку. В електростатиці (але не у змінних полях) поля по обидві сторони суцільної провідної оболонки повністю не залежать одне від одного.

Властивість зарядів розміщуватись на зовнішній поверхні провідника використовується для створення *електростатичних генераторів*, які призначені для нагромадження великих зарядів і досягнення різниці потенціалів в декілька мільйонів вольт. Винахідником такого генератора є американський фізик Ван-де-Грааф. Використовують ці генератори у високовольтних прискорювачах заряджених частинок.

Теорема єдиності в електростатиці

Розглянемо провідник, якому надали позитивний заряд *q*. Доречно поставити питання: *чи єдиним чином весь цей заряд зможе розподілитись по поверхні провідника*? Відповідь на це питання дає так звана теорема єдиності в електростатиці: *заряд на провіднику розподіляється єдиним чином.* Ця теорема ґрунтується на тому, що заряджений провідник знаходиться в рівноважному стані. В цьому стані електричне поле в об'ємі провідника відсутнє (інакше не зупинилось би перетікання зарядів всередині провідника), провідник стає еквіпотенціальним, надлишкові заряди розподіляються по поверхні провідника, і все це має місце через те, що в металах є практично необмежена кількість вільних носіїв заряду (електрони). Такий єдиностний (або однозначний) розподіл зарядів в провіднику дає можливість ввести поняття про таку властивість провідника, як електроємність.

§3.2. Електроємність відокремленого провідника. Конденсатори

Відокремленим називається провідник, який знаходиться настільки далеко від інших

тіл, що впливом їхніх електричних полів можна знехтувати.

Якщо надати відокремленому провіднику заряд q, то електричне поле, яке створює провідник в будь-якій точці простору визначається суперпозицією полів, що створюються окремими зарядженими ділянками провідника. Кожна така ділянка має заряд $\Delta q_i = \sigma_i \Delta S_i$, де ΔS_i – площа ділянка, σ_i – поверхнева густина заряду на даній ділянці провідника (рис.3.5). Що відбудеться, якщо провіднику надати заряд в n раз більший (або менший), ніж q. Густина заряду в кожній точці провідника згідно ***теореми єдиності*** зміниться в n раз. Напруженість поля $\Delta E_i = k\Delta q_i / r_i^2$, яка створюється кожною ділянкою провідника в будь-якій точці простору,

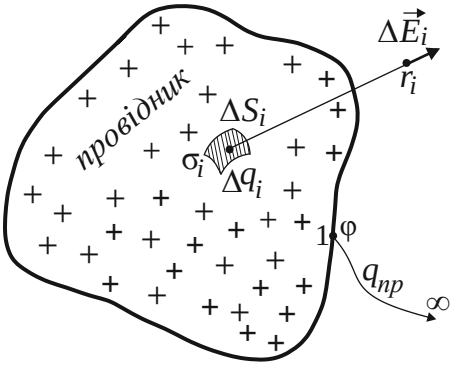


Рис.3.5.

зміниться в n раз, а значить і сумарна напруженість поля в кожній точці зовні провідника зміниться в n раз. Це означає, що елементарна робота, яка здійснюється електричним полем провідника над одиничним позитивним пробним зарядом $q_{пр}$ при його переміщенні на будь-якій ділянці зміниться в n раз. В стільки ж раз зміниться робота електростатичних сил, які діють на одиничний позитивний пробний заряд $q_{пр}$, при переміщенні його від будь-якої точки поверхні провідника (поверхня провідника еквіпотенціальна) на нескінченність. Оскільки робота $A_{1\infty} = q_{пр}\varphi$, де φ потенціал провідника, то це означає, що φ теж зміниться в n раз.

Таким чином, при зміні заряду провідника в n раз в стільки ж раз зміниться потенціал провідника, а значить відношення q/φ залишається незмінним. Величину

$$C = q/\varphi \qquad (3.3)$$

називають електроємністю (або просто ємністю) відокремленого провідника. Чим більший заряд можна надати провіднику при заданому потенціалі, тим більша його ємність. Можна сказати, що *електроємність характеризує здатність провідника накопичувати заряд при даному потенціалі.* Для даного провідника (провідника з ємністю C) існує єдине значення заряду q, який необхідно надати провіднику для того, щоб його потенціал був φ.

Електроємність однозначно визначається формою провідника, його розмірами і електричними властивостями оточуючого середовища.

Якщо розміри провідника *подібно* зменшити в n раз, то і ємність провідника зменшиться в n раз, і навпаки.

Електроємність не залежить ні від матеріалу провідника, ні від його агрегатного стану, ні від форми і розмірів порожнин всередині нього. Це пов'язано з тим, що надлишкові заряди розподіляються тільки на зовнішній поверхні провідника. Електроємність не залежить також від заряду провідника та його потенціалу.

Одиниця ємності – фарада (Ф).

Одна фарада – це ємність такого провідника, потенціал якого змінюється на 1В при наданні йому заряду в 1 Кл (формула 3.3).

Для тіл сферичної форми $\varphi = q/(4\pi\varepsilon_0 R)$. Тому електроємність відокремленого сферичного провідника $C = q/\varphi = 4\pi\varepsilon_0 R$, де R – радіус сфери (кулі). Якщо вважати, що Земля має форму кулі, радіус якої $R_з$ = 6370 км, то електроємність Землі буде

$$C = 4\pi \cdot 8{,}85 \cdot 10^{-12} \cdot 6{,}37 \cdot 10^{6}\text{Ф} =$$

$$= 709 \cdot 10^{-6}\text{Ф} = 709\text{мкФ}.$$

Для того, щоб сферичний провідник мав ємність 1Ф, його радіус повинен бути

$$R = 1/(4\pi\varepsilon_0) \cong 9 \cdot 10^{9}\text{м} = 1411 R_з.$$

На практиці, однак, необхідні пристрої, які мають здатність при малих розмірах і невеликих відносно навколишніх тіл потенціалах нагромаджувати значні за величиною заряди. Для цього використовують *невідокремлені провідники* в пристроях, які називаються *конденсаторами*.

Покажемо, що ємність невідокремленого провідника завжди більша від ємності відокремленого провідника. Нехай провідник 1 (рис.3.6) має заряд q_1 і потенціал його при цьому дорівнює φ_1. Тоді ємність провідника

$$C_1 = q_1/\varphi_1.$$

Допустимо, що досить близько від зарядженого провідника 1 з'явився провідник 2 (рис.3.6). Неважливо, чи провідник 2 заряджений, чи незаряджений. Провідник 2 зазнає впливу електричного поля провідника 1. На поверхні провідника 2, яка ближче до провідника 1, з'являться негативні заряди (протилежні за знаком заряду провідника 1), а на протилежній поверхні – позитивні заряди (рис.3.6). Негативні заряди провідника 2 ближче до провідника 1, ніж його позитивні заряди. Це означає, що негативні заряди провідника 2 впливають сильніше на

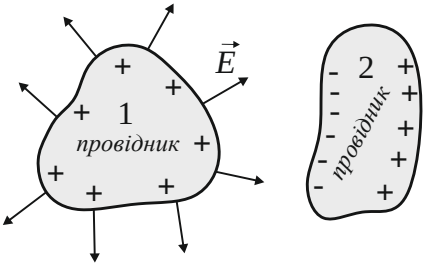


Рис.3.6.

потенціал провідника 1 (понижують φ_1), ніж позитивні заряди провідника 2 (вони підвищують φ_1). Результатом цього впливу є те, що потенціал провідника 1 φ_1 понижується і стає $\varphi_1' < \varphi_1$. Так як заряд q_1 не змінюється, то ємність провідника 1 стає $C_1' = \frac{q_1}{\varphi_1'} > C_1$. Чим ближче знаходиться провідник 2 до провідника 1, тим більша буде його ємність C_1. Зауважимо, що замість провідника 2 може бути і діелектрик. Результатом такої зміни буде теж збільшення ємності провідника 1.

§3.3. Конденсатори

Система, яка зображена на рис.3.6, дає можливість накопичувати (конденсувати) на провіднику 1 при одному і тому ж потенціалі більший заряд, ніж тоді, коли провідник 1 відокремлений. Це дуже важливо для технічного застосування. Вказана система лежить в основі пристроїв, які називаються *конденсаторами*.

Конденсатор – система із двох (або більше) провідників (яких називають обкладками), що розділені діелектриком, товщина якого мала порівняно з лінійними розмірами обкладок. Мала відстань між

обкладками порівняно з їхніми лінійними розмірами дає можливість досягати більшої ємності системи і практично повністю зосередити поле конденсатора між обкладками, тобто, уникнути впливу цього поля на оточуючі тіла і забезпечити незалежність ємності конденсатора від їх оберненого впливу. Оскільки обкладки – провідники, то вони екранують зовнішні поля і, таким чином, усувається і їх вплив на ємність конденсатора.

Конденсатори бувають:

а) плоскі – обкладинками є пара паралельних металічних пластин;

б) сферичні – обкладинками є дві концентричні провідні сфери;

в) циліндричні – обкладинками є два коаксіальні (спільна вісь) провідні циліндри.

Обкладки конденсатора заряджають зарядами однакової величини, але протилежних знаків. В цьому випадку потенціал провідника 1 найбільше понижується зарядженим провідником 2 (рис.3.6). Ємність такого конденсатора визначається як

$$C = \frac{|q|}{\varphi_1 - \varphi_2} = \frac{q}{U}, \qquad (3.4)$$

де $|q|$ – заряд однієї із обкладок конденсатора, як правило позитивно зарядженої, $U = \varphi_1 - \varphi_2$ – різниця потенціалів, або напруга між обкладками.

Плоский конденсатор

Схема плоского конденсатора показана на рис.3.7. Простір між обкладинками (пластинами) заповнений діелектриком з відносною діелектричною проникністю ε. Площа пластин S і, якщо лінійні розміри пластин великі порівняно з відстанню між ними d, то електричне поле між пластинами можна вважити еквівалентним полю між двома нескінченними площинами (див. формулу 1.16), які заряджені різнойменно і поверхневі густини яких дорівнюють $+\sigma$ і $-\sigma$. Крім

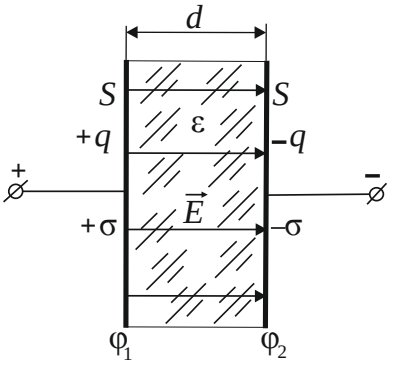


Рис.3.7.

того, відстань d повинна бути настільки малою, щоб порушення однорідності поля поблизу його країв можна було не брати до уваги.

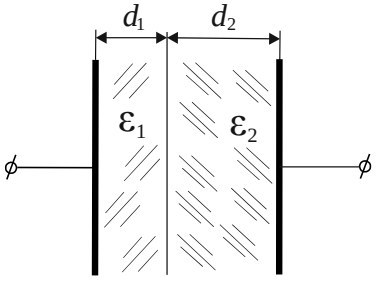


Рис.3.8.

Напруженість електричного поля і різниця потенціалів між обкладками конденсатора в цьому випадку дорівнюють (згідно 1.16):

$$E = \frac{\sigma}{\varepsilon\varepsilon_0} = \frac{q}{\varepsilon\varepsilon_0 S},$$

$$\varphi_1 - \varphi_2 = Ed = \frac{q}{\varepsilon\varepsilon_0 S} d\,.$$

Згідно визначення (3.4) ємність плоского конденсатора буде:

$$C = \frac{q}{\varphi_1 - \varphi_2} = \frac{\varepsilon\varepsilon_0 S}{d}. \qquad (3.5)$$

Ємність конденсатора, який має шаруватий діелектрик (рис.3.8), визначають за формулою:

$$C = \frac{\varepsilon_0 S}{\frac{d_1}{\varepsilon_1} + \frac{d_2}{\varepsilon_2}}. \qquad (3.6)$$

Циліндричний конденсатор

Циліндричний конденсатор – це система із двох коаксіальних циліндрів з радіусами r_1 і r_2 (рис.3.9). Простір між циліндрами заповнений діелектриком з діелектричною проникністю ε. Якщо висота циліндрів значно перевищує їхні радіуси $h>>r_1$ і $h >> r_2$, то можна знехтувати краєвими ефектами і для визначення різниці потенціалів між циліндричними обкладками конденсатора скористатися формулою (1.18), враховуючи присутність діелектрика. Тобто $E = \lambda/(2\pi\varepsilon\varepsilon_0 r)$. Так як $E = -\frac{d\varphi}{dr}$, то

$$-d\varphi = Edr = \lambda dr/(2\pi\varepsilon\varepsilon_0 r).$$

Інтегруючи цей вираз в межах від r_1 до r_2, знаходимо:

$$-\int_{\varphi_1}^{\varphi_2} d\varphi = \frac{\lambda}{2\pi\varepsilon\varepsilon_0}\int_{r_1}^{r_2}\frac{dr}{r};$$

$$-(\varphi_2 - \varphi_1) = \frac{\lambda}{2\pi\varepsilon\varepsilon_0}\ln r\Big|_{r_1}^{r_2};$$

$$\varphi_1 - \varphi_2 = \frac{q}{2\pi\varepsilon\varepsilon_0 h}\ln\frac{r_2}{r_1},$$

де $\lambda = q/h$ – лінійна густина заряду.

Тоді ємність циліндричного конденсатора

$$C = \frac{q}{\varphi_1 - \varphi_2} = \frac{2\pi\varepsilon\varepsilon_0 h}{\ln\frac{r_2}{r_1}}. \qquad (3.7)$$

Якщо зазор d між обкладками конденсатора малий порівняно з r_1 ($d<<r_1$), то $\ln\frac{r_2}{r_1} = \ln\frac{r_1 + d}{r_1} = \ln\left(1 + \frac{d}{r_1}\right)$ можна розкласти в ряд і обмежитись лише членами першого порядку малості (формули наближеного обчислення), а саме:

$$\ln\left(1 + \frac{d}{r_1}\right) \approx \frac{d}{r_1}.$$

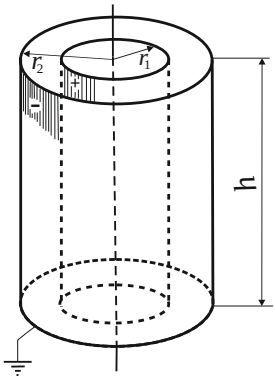


Рис.3.9.

Тоді формула (3.7) набирає вигляду (3.5)

$$C = \frac{2\pi\varepsilon\varepsilon_0 r_1 h}{d} = \frac{\varepsilon\varepsilon_0 S}{d},$$

де $S = 2\pi r_1 h$ – бічна поверхня циліндричної поверхні. Тобто при таких умовах циліндричний конденсатор можна наближено вважати плоским.

Сферичний конденсатор

Обкладки такого конденсатора – це дві концентричні провідні сфери, радіуси яких r_1 і r_2, котрі розділені тонким шаром діелектрика завтовшки $d = r_2 - r_1$ (рис.3.10). Поля поза конденсатором, що створені внутрішньою і

зовнішньою обкладками, взаємно компенсується. Поле між обкладками створюється зарядом q внутрішньої кулі, а заряд $-q$ зовнішньої кулі всередині цієї кулі не створює електричного поля. Напруженість поля між обкладками конденсатора визначається за формулою (1.19). Тоді згідно (1.34) різниця потенціалів між обкладками конденсатора буде: $-\int_{\varphi_1}^{\varphi_2} d\varphi = \frac{q}{4\pi\varepsilon\varepsilon_0}\int_{r_1}^{r_2}\frac{dr}{r^2}$.

Після інтегрування отримаємо:

$$\varphi_1-\varphi_2=\frac{q}{4\pi\varepsilon\varepsilon_0}\left(\frac{1}{r_1}-\frac{1}{r_2}\right).$$

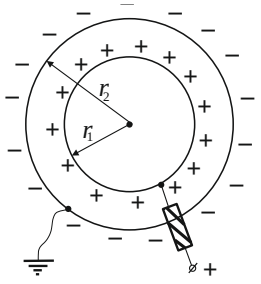


Рис.3.10.

Звідки електроємність сферичного конденсатора буде:

$$C=\frac{q}{\varphi_1-\varphi_2}=\frac{4\pi\varepsilon\varepsilon_0 r_1 r_2}{r_2-r_1}. \quad (3.8)$$

Якщо товщина шару діелектрика d мала, то можна вважати, що $r_2 = r_1 + d \approx r_1$. Тоді

$$C=\frac{4\pi\varepsilon\varepsilon_0 r_1^2}{d}=\frac{\varepsilon\varepsilon_0 4\pi r_1^2}{d}=\frac{\varepsilon\varepsilon_0 S}{d},$$

де $S = 4\pi r_1^2$ – площа обкладки конденсатора. Коли сферичний конденсатор має дуже тонкий шар діелектрика, його електроємність можна обчислювати за формулою ємності плоского конденсатора.

Конденсатор характеризується не лише електроємністю, а й "пробивною напругою" – різницею потенціалів між його обкладками, при якій може відбуватися пробій, тобто електричний розряд через шар діелектрика в конденсаторі. Величина пробивної напруги залежить від властивостей діелектрика, його товщини і форми обкладок.

Для одержання потрібної ємності при відповідній напрузі в електро- і радіотехніці часто конденсатори з'єднують між собою в батареї. З'єднання може бути паралельним, послідовним.

При паралельному з'єднанні (рис.3.11) напруга U на кожному із конденсаторів однакова. Тоді заряди на кожному із конденсаторів $q_1 = C_1U; q_2 = C_2U;\ldots, \quad q_n = C_nU$.

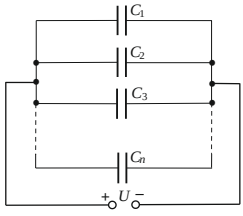


Рис.3.11

Заряд батареї конденсаторів

$$q = q_1 + q_2 + \cdots + q_n = (C_1 + C_2 + \cdots + C_n)U.$$

Електрична ємність батареї

$$C=\frac{q}{U}=C_1+C_2+\cdots+C_n. \quad (3.9)$$

Тобто при паралельному з'єднанні конденсаторів результуюча ємність батареї дорівнює сумі ємностей окремих конденсаторів.

При послідовному з'єднанні негативно заряджену обкладку першого конденсатора з'єднують з позитивно зарядженою обкладинкою другого і т. д. (рис.3.12).

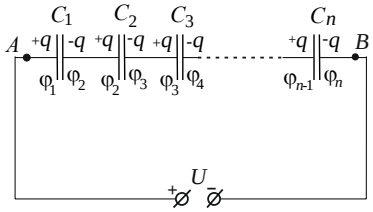


Рис.3.12.

Якщо на батарею подають напругу $U=\varphi_A-\varphi_B$, то до такої різниці потенціалів заряджаються тільки крайні обкладки першого і останнього конденсатора, причому

$$U=\varphi_A-\varphi_B=(\varphi_1-\varphi_2)+(\varphi_2-\varphi_3)+\cdots+(\varphi_{n-1}-\varphi_n), \quad (3.10)$$

де $\varphi_1,\varphi_2,\ldots,\varphi_n$ – потенціали пластин конденсаторів (рис.3.12).

Провідник, що з'єднує негативно заряджену обкладку першого та позитивну заряджену обкладку другого конденсатора, можна розглядати разом з обкладками як один провідник (одне провідне тіло). Вільні заряди цього тіла під дією поля прикладеної напруги перерозподіляються так, що на одній обкладці з'являються негативний заряд $-q$, а на іншій – позитивний $+q$. Тому можна зробити висновок, що заряди на всіх послідовно з'єднаних конденсаторах, незалежно від їхньої ємності, однакові і дорівнюють заряду всієї батареї, тобто

$$q=q_1=q_2=\ldots=q_n.$$

Ємність батареї послідовно з'єднаних конденсаторів

$$C=\frac{q}{\varphi_A-\varphi_B}.$$

Звідки

$$\varphi_A-\varphi_B=\frac{q}{C}.$$

Для першого конденсатора $\varphi_1-\varphi_2=\frac{q}{C_1}$, для другого $\varphi_2-\varphi_3=\frac{q}{C_2}$ і т. д. Тоді (3.10) перепишеться так:

$$\frac{q}{C}=\frac{q}{C_1}+\frac{q}{C_2}+\ldots+\frac{q}{C_n}.$$

Після скорочення на q отримуємо формулу для обчислення електроємності батареї послідовно з'єднаних конденсаторів:

$$\frac{1}{C}=\frac{1}{C_1}+\frac{1}{C_2}+\ldots+\frac{1}{C_n}. \quad (3.11)$$

Якщо в батарею з'єднують n конденсаторів з однаковою ємністю C_o, то при паралельному їх з'єднанні $C=nC_o$, а при послідовному – $C=C_o/n$. При послідовному з'єднанні конденсаторів результуюча ємність батареї завжди менша від найменшої ємності конденсатора, що входить до батареї.

§3.4 Енергія зарядженого відокремленого провідника та конденсатора. Енергія електростатичного поля

Щоб зарядити провідник, необхідно виконувати роботу проти кулонівських сил електричного відштовхування між однойменними зарядженими частинками. Заряджання провідника можна здійснювати різними способами. Нехай на провідник із нескінченності послідовно переносяться заряди величиною dq. При накопиченні заряду на провіднику наступне перенесення на нього зарядів dq все більш утруднюватиметься внаслідок зростання кулонівських сил відштовхування між однойменними зарядами. Якщо в якийсь момент часу потенціал поверхні провідника дорівнює φ, то при перенесенні із

нескінченності на провідник заряду dq необхідно виконати роботу (формула (1.26))

$$dA = dq(\varphi_\infty - \varphi) = -\varphi dq,$$

де $\varphi_\infty = 0.$

Оскільки потенціал провідника зв'язаний з його зарядом і ємністю співвідношенням (3.3), то

$$dA = -\frac{1}{C} q dq.$$

Повна робота при зміні заряду провідника від 0 до q буде:

$$A = \int_0^q dA = -\frac{1}{C}\int_0^q q dq = -\frac{q^2}{2C}.$$

Ця робота йде на надання енергії зарядженому провіднику, тобто

$$W = -A = \frac{q^2}{2C} = \frac{C\varphi^2}{2} = \frac{q\varphi}{2}. \qquad (3.12)$$

За допомогою формули (3.12) можна знайти енергію зарядженого конденсатора, як енергію двох заряджених обкладок (провідників).

$$W = \frac{1}{2}(+q)\varphi_1 + \frac{1}{2}(-q)\varphi_2 = \frac{1}{2}q(\varphi_1 - \varphi_2) = \frac{1}{2}qU.$$

Остаточно можемо записати, що енергія зарядженого конденсатора запишеться так із урахуванням (3.4):

$$W = \frac{qU}{2} = \frac{CU^2}{2} = \frac{q^2}{2C}, \qquad (3.13)$$

де $U = \varphi_1 - \varphi_2$ – різниця потенціалів між обкладками конденсатора.

Формулу (3.13) перепишемо в дещо іншому вигляді. Будемо вважати електричне поле всередині конденсатора однорідним, тобто знехтуємо краєвими ефектами. Тоді $U = Ed$, $C = \varepsilon\varepsilon_0 S/d$ і вираз (3.13) набирає вигляду:

$$W = \frac{1}{2}\frac{\varepsilon\varepsilon_0 S}{d}E^2 d^2 = \frac{\varepsilon\varepsilon_0 E^2}{2}V, \qquad (3.14)$$

де $V = Sd$ – об'єм, обмежений пластинами (обкладками) конденсатора. Тому можемо вважати, що відношення

$$\omega = \frac{W}{V} = \frac{\varepsilon\varepsilon_0 E^2}{2} \qquad (3.15)$$

є енергія, що приходиться на одиницю об'єму конденсатора, тобто об'ємна густина енергії електричного поля. Якщо поле неоднорідне, тобто $E = E(x, y, z),$ то енергія електричного поля в елементарному об'ємі поля буде:

$$dW = \omega(x, y, z)dV = \frac{\varepsilon\varepsilon_0 E^2}{2}dV.$$

Тоді формула (3.14) може бути записана як

$$W = \tfrac{1}{2}\int_V \varepsilon\varepsilon_0 E^2 dV. \qquad (3.16)$$

З математичної точки зору співвідношення (3.14) і (3.16) є іншими формами запису співвідношення (3.13). Однак формальне математичне перетворення рівнянь часто відкриває можливості зовсім нового фізичного інтегрування виражених ними співвідношень. Так формула (3.16) виражає електричну енергію у вигляді нескінченної суми доданків, кожний із яких дорівнює $\frac{1}{2}\varepsilon\varepsilon_0 E^2 dV$ і належить певному об'єму dV. Це відповідає уявленням теорії поля (теорії близькодії), де вважається, що енергія подібно до речовини, розподілена у просторі з об'ємною густиною (3.15), яку ми можемо записати ще й так:

$$\omega = \frac{\varepsilon\varepsilon_0 E^2}{2} = \frac{ED}{2} = \frac{D^2}{2\varepsilon\varepsilon_0}. \qquad (3.17)$$

Для ізотропних діелектриків напрямки векторів $\vec{E}$ і $\vec{D}$ співпадають. Тому $\omega = \vec{E}\vec{D}/2$.

Враховуючи співвідношення (2.18), отримаємо для ω наступний вираз:

$$\omega = \frac{\vec{E}(\varepsilon_0\vec{E} + \vec{P})}{2} = \frac{\varepsilon_0 E^2}{2} + \frac{\vec{E}\vec{P}}{2}. \qquad (3.18)$$

Перший доданок у (3.18) відповідає густині енергії поля $\vec{E}$ у вакуумі. Другий доданок – це густина енергії, яка витрачається на поляризацію діелектрика.

Формула (3.13) відповідає положенню теорії далекодії, де W розглядається як потенціальна енергія заряджених тіл, що притягуються або відштовхуються один від одного. Тобто у концепції далекодії вважається, що електрична енергія локалізована на заряджених провідниках.

Якому з цих двох уявлень про локалізацію електричної енергії надати перевагу? В межах електростатики принципово неможливо виконати жодного досліду, який дав би можливість зробити вибір одного з цих уявлень. Це пояснюється тим, що в електростатиці електричне поле невіддільне від зарядів, джерел поля. Величиною і розміщенням зарядів однозначно визначається електростатичне поле. Навпаки, знаючи напруженість електростатичного поля в будь-яких точках простору, можна однозначно вказати розподіл зарядів. Зовсім інша ситуація у разі змінних полів. Змінні електромагнітні поля можуть існувати самостійно, незалежно від джерел, які їх збудили. Електромагнітне поле при цьому існуватиме у вигляді електромагнітних хвиль, які, поширюючись у просторі, переносять енергію. Ця енергія не може бути потенціальною енергію взаємодіючих зарядів, оскільки їх немає.

Якщо електростатику розглядати як частину електродинаміки, то необхідно надавати перевагу теорії поля з її уявленнями про локалізацію електричної енергії у просторі.

Контрольні питання

1.Основні властивості провідників в електричному полі.

2.Напруженість електричного поля біля поверхні зарядженого провідника.

3.Які властивості провідника дають можливість ввести поняття його ємності?

4.Конденсатори. Їх властивості та вимоги до конструкцій. Типи конденсаторів.

5.Енергія зарядженого провідника та конденсатора. Об’ємна густина енергії та загальна формула для розрахунку енергії електричного поля.

6.Де локалізована енергія електричного поля?

Розділ 4. Постійний електричний струм

§4.1.Електричний струм та його види. Струм провідності та умови його існування

Електричний струм – упорядкований (направлений) рух електрично заряджених частинок або заряджених макроскопічних тіл.

За напрямок струму приймають напрямок руху позитивно заряджених частинок; якщо струм створюється негативно зарядженими частинками (наприклад електронами), то напрямком струму вважають протилежний напрямок руху частинок. Розрізняють такі електричні струми: 1) провідності; 2) конвекційний; 3) дифузійний.

1.*Струм провідності – це направлене переміщення вільних електричних зарядів, які знаходяться в провідному середовищі (твердому, рідкому чи газоподібному) відносно інших частинок, що входять до складу цього середовища, внаслідок дії на вільні заряди електричного поля.* Про наявність електричного струму *в провідниках* можна судити по дії, які він викликає: нагрівання провідників, зміна хімічного складу, *створення*

магнітного поля. Магнітна дія струму проявляється у всіх без винятку провідниках; в *надпровідниках* не відбувається виділення теплоти, а хімічна дія струму спостерігається переважно в електролітах. Магнітне поле створюється не тільки струмом провідності або конвекційним струмом, але і змінним електричним полем в діелектриках і вакуумі. Величина, яка пропорційна швидкості зміни електричного поля з часом, Дж. Максвелл назвав струмом зміщення (розглядається в подальшому). Струм зміщення входить в рівняння Максвелла на рівних правах із струмом, який обумовлений рухом зарядів. Тому повний електричний струм, який дорівнює сумі струму провідності і струму зміщення, визначає створюване ним магнітне поле.

В металах електричний струм створюють "вільні" електрони, в напівпровідниках – негативно заряджені електрони і позитивно заряджені дірки, в провідних розчинах (електролітах) – позитивно і негативно заряджені іони, в газах – електрони, позитивно і негативно заряджені іони. Струмом провідності є і електричний струм у вакуумі (наприклад, рух електронів в електронних лампах, електронно-променевих трубках).

2.*Конвекційний струм* – рух макроскопічних заряджених тіл як цілого (наприклад, заряджених крапель дощу).

3.*Дифузійний струм* – направлений рух електрично заряджених частинок внаслідок їх дифузії *без дії на них електричного поля*. В напівпровідниках можна створити в певному місці надлишкову концентрацію носіїв заряду порівняно із сусіднім місцем. Утворюється градієнт концентрації (швидкість зміни концентрації носіїв заряду з відстанню), що і призводить до направленого руху заряджених частинок.

В даній главі мова буде йти про струм провідності. Для появи і існування струму провідності необхідно, щоб виконувались дві умови:

Перша – наявність у даному середовищі вільних електричних зарядів – носіїв струму.

Друга – на електричні заряди необхідно, щоб діяла сила $\vec{F} = q\vec{E}$. Отже, в даному середовищі повинно існувати електричне поле, енергія якого витрачалась би на переміщення електричних зарядів. Зрозуміло, що *провідник (метал) зі струмом вже перестає бути еквіпотенціальним.* Порушити еквіпотенціальність провідника можна за допомогою пристрою (*джерела енергії)*, який називається джерелом *електрорушійної сили* (в подальшому скорочено е.р.с.).

До характеристик струму провідності можна віднести середню швидкість направленого руху носіїв заряду, яку ще називають швидкістю дрейфу $\vec{\upsilon}_{др}$. Відомо, що носії заряду і у відсутності, і при наявності електричного поля у провідниках приймають участь в хаотичному тепловому русі, середня швидкість якого є набагато більшою, ніж $\vec{\upsilon}_{др}$. В силу рівноправності всіх напрямків теплового руху і великої кількості носіїв заряду хаотичний рух не може створити електричного струму. Струм провідності виникає при створенні у провіднику електричного поля і накладанні на хаотичний рух направленого руху, який обумовлений дією поля. Якщо $\vec{\upsilon}_T$ – вектор швидкості теплового (хаотичного) руху, а $\vec{\upsilon}$ – вектор швидкості направленого руху

конкретного носія заряду, то усереднення сумарної швидкості носія заряду $\vec{v}_T + \vec{v}$ по всім носіям заряду дає, що

$$< \vec{v}_T + \vec{v} > = < \vec{v}_T > + < \vec{v} > = < \vec{v} > = \vec{v}_{др},$$

де середнє значення *вектора* швидкості теплового руху дорівнює нулю, а середнє значення *вектора швидкості направленого* руху є вектор дрейфової швидкості.

Основною кількісною характеристикою електричного струму є *сила струму I, скалярна величина, яка вимірюється кількістю електрики (величиною заряду), що проходить через поперечний переріз провідника за одиницю часу.* Якщо за нескінченно малий проміжок часу dt через поперечний переріз провідника проходить заряд величиною dq, то сила струму.

$$I = \frac{dq}{dt}. \qquad (4.1)$$

Співвідношення (4.1) виражає миттєве значення сили струму. Якщо сила струму і його напрям з часом не змінюється, то такий струм називають *стаціонарним* або *постійним*. Силу постійного струму визначають співвідношенням

$$I = \frac{q}{t}, \qquad (4.2)$$

де $q-$ електричний заряд, який проходить через поперечний переріз провідника за час $t.$ Сила струму в СІ вимірюється в амперах (А). *Одиниця сили струму ампер є основною в системі* СІ.

У загальному випадку електричний струм може бути розподілений по перерізу провідника, через який він проходить, нерівномірно. Тому для характеристики розподілу електричного струму по перерізу провідника вводять поняття *густини струму***.**

Під густиною струму у кожній даній точці провідника треба розуміти границю, до якої прямує відношення сили струму ΔI до елемента площі перерізу ΔS, розташованого перпендикулярно до напрямку струму, при $\Delta S \to 0$, тобто

$$j = \lim_{\Delta S \to 0} \frac{\Delta I}{\Delta S} = \frac{dI}{dS}. \qquad (4.3)$$

Якщо електричний струм розподілений рівномірно по перерізу провідника, через який він проходить, то густина струму

$$j = \frac{I}{S}, \qquad (4.4)$$

де $S-$ площа поперечного перерізу.

Одиницею густини струму в СІ є ампер на метр квадратний (А/м^2).

Припустімо, що у провіднику заряд окремого носія заряду дорівнює заряду електрона e, їхня концентрація – n і середня швидкість направленого (дрейфуючого) руху $\vec{v}_{др}$, то згідно з визначенням густини струму

$$\vec{j} = en\vec{v}_{др}. \qquad (4.5)$$

Швидкість $\vec{v}_{др}$ характеризує рух носіїв заряду у даній точці провідника, тому вектор $\vec{j}$ визначає електричний струм у тій самій точці.

Якщо струм створюється рухом позитивних і негативних зарядів, то у загальному випадку вираз для вектора густини струму записують так:

$$\vec{j} = e^+ n^+ \vec{v}^+_{др} + e^- n^- \vec{v}^-_{др}, \qquad (4.6)$$

де n^+ і n^- – концентрації позитивних і негативних носіїв заряду, e^+ і e^-, $\vec{v}^+_{др}$ і $\vec{v}^-_{др}$ – їхні заряди і дрейфові швидкості, відповідно.

§4.2.Закон Ома та Джоуля–Ленца в інтегральній та диференціальній формах

Нехай в провіднику створене електричне поле $\vec{E}$, внаслідок того, що торець 1 провідника має потенціал φ_1, а торець 2 - потенціал φ_2, причому $\varphi_2 < \varphi_1$ (рис. 4.1).

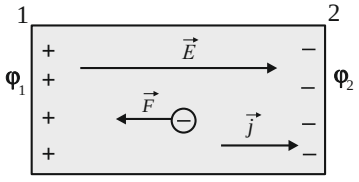


Рис.4.1.

Будемо вважати, що це електричне поле з напруженістю $\vec{E}$ стаціонарне (незмінне з часом). На вільні заряди діє сила $\vec{F} = e\vec{E}$, яка зумовлює стаціонарний рух електричних зарядів провідника – постійний струм. Зрозуміло, що це можливо, коли різниця потенціалів $\varphi_1 - \varphi_2$ підтримується постійною. Цю різницю потенціалів називають напругою і позначають буквою U, тобто $\varphi_1 - \varphi_2 = U$.

Стаціонарне електричне поле за своїми фізичними властивостями не відрізняється від електростатичного поля, оскільки воно є полем електричних зарядів, розміщених у певних ділянках кола струму з незмінною в часі густиною.

Якщо струм постійний, то густина електричних зарядів у будь-якій точці простору не змінюється з часом, хоч має місце потік електричних зарядів. Тут на місце тих зарядів, що переходять в інше місце, надходять нові заряди. Такі заряди, як підтверджує дослід, створюють в просторі таке ж кулонівське електричне поле, яке б створили нерухомі заряди тієї ж густини. З цього випливає, що електричне поле стаціонарних струмів також є потенціальним. Проте не слід вважати, що електричне поле постійних струмів і кулонівське поле зарядів одне й те саме. Між ними є істотні відмінності. При рівновазі зарядів, напруженість кулонівського поля у провіднику дорівнює нулю. Електричне поле постійних струмів є також кулонівським полем, але заряди, що його збуджують, знаходяться у русі, тому що поле постійних струмів існує і всередині провідника. При його відсутності у провіднику струму не було б. Силові лінії електростатичного поля завжди перпендикулярні до поверхні провідника, коли в провіднику струм відсутній. За наявності стаціонарного струму у провіднику його поверхня не є еквіпотенціальною і лінії напруженості поля, напрямок яких співпадає з напрямком струму, можуть бути направлені під гострим кутом до поверхні провідника.

Оскільки поле постійних струмів має потенціальний характер, то можна говорити про потенціал окремих точок кола постійного струму і про різницю потенціалів між ними.

Німецький фізик Г. Ом у 1827р. експериментально довів, що сила струму I в провіднику прямо пропорційна напрузі U на його кінця

$$I = kU, \qquad (4.7)$$

де k – коефіцієнт пропорційності, який залежить від розмірів провідника, матеріалу, з якого він виготовлений, та його фізичного стану. Його називають ***електричною провідністю провідника.*** На практиці частіше користуються оберненою до електропровідності величиною, яку Ом назвав опором провідника R:

$$R = 1/k.$$

Тоді формулу (4.7) можна записати так:

$$I = \frac{U}{R}. \qquad (4.8)$$

Опір провідників в СІ вимірюють в омах (Ом), провідність - в сименсах (См).

На основі (4.8) визначають одиницю опору в 1 Ом. Це опір такого провідника, в якому при напрузі між його кінцями 1В існує струм силою 1А. *Електричний опір* металів пов'язаний з розсіюванням електронів провідності на теплових коливаннях кристалічної решітки і її структурних неоднорідностях.

Формула (4.8) виражає в *інтегральній формі закон Ома для однорідної ділянки кола. Ділянку кола називають однорідною, якщо на ній немає джерел струму*. Закон Ома справджується з великою точністю для металевих провідників та електролітів. Відхилення від закону Ома становить близько 1% для струмів дуже великої густини – кількох мільйонів амперів на 1 см^2. Закон порушується і в тих випадках, коли в провідниках не вистачає носіїв електричних зарядів і настає насичення струму (несамостійний розряд у газах), а також при досить високих напругах та ін. Залежність $I = f(U)$ називають вольт-амперною характеристико (ВАХ) провідника.

Зауважимо, що опір провідника в першому наближенні не залежить від величини струму, який проходить по провіднику, а залежить від його геометричних розмірів, форми та від матеріалу провідника і його стану. Для однорідних провідників циліндричної форми

$$R = \rho \frac{l}{S}, \qquad (4.9)$$

де l – довжина провідника; S – площа його поперечного перерізу; ρ – коефіцієнт пропорційності, який залежить від матеріалу та фізичного стану провідника. Його називають *питомим опором* матеріалу. Одиницею питомого опору в СІ є Ом на метр (Ом·м). В електротехніці часто довжину провідників вимірюють у метрах, а площу поперечного перерізу – в квадратних міліметрах. Тоді питомий опір доцільно подавати в Ом·мм^2/м.

Питомий опір провідників значною мірою залежить від домішок, а також від способу їх виготовлення. Найменший питомий опір мають срібло і мідь. На практиці для передавання електричної енергії частіше всього використовують дріт із міді або алюмінію. Для виготовлення реостатів, котушок опору використовують сплави з великим питомим опором (ніхром, нікелін). Як уже зазначалося, питомий опір провідників залежить не тільки від хімічного складу речовини, але й від його стану, зокрема від температури. Залежність питомого опору однорідної речовини від температури характеризують температурним коефіцієнтом опору

$$\alpha = \frac{1}{\rho}\frac{d\rho}{dT}. \qquad (4.10)$$

Температурний коефіцієнт опору числово дорівнює відносній зміні опору провідника при зміні його температури на 1К. Величина α для різних речовин різна. Для всіх чистих металів температурний коефіцієнт опору близький до 1/273 К$^{-1}$. Температурний коефіцієнт опору дещо змінюється при зміні температури даної речовини. Деякі сплави мають досить малі значення температурного коефіцієнта опору α. Так, для константану α=0,1·10^{-4} К$^{-1}$. Дріт з таких сплавів використовують для виготовлення еталонів опору, котушок вимірювальних місткових схем, тощо.

Як показують досліди, для кожного хімічного металевого провідника існує певний

інтервал температур, в якому залежність питомого опору від температури має лінійний характер, тобто

$$\rho = \rho_0(1+\alpha t), \qquad (4.11)$$

де ρ_0 – питомий опір, що відповідає початку температурного інтервалу. Здебільшого за початок цього інтервалу беруть температуру 0℃. Залежність опору металів від температури покладено в основу роботи приладів для вимірювання температури (термометри опору), приладів для вимірювання енергії випромінювання (болометри) та автоматичних пристроїв для термостатування, терморегулювання будь-яких процесів з великою точністю.

Користуючись поняттям густини струму, отримаємо формулу *закону Ома в диференціальній формі*, яка встановлює зв'язок між характеристиками струму в певній точці провідника. Виділимо подумки в однорідному провіднику прямий циліндр довжиною Δl з площею основи ΔS так, щоб його твірні збігалися з напрямом ліній струму (рис.4.2). Вважатимемо, що різниця потенціалів у перерізах 1 і 2 $\Delta\varphi = \varphi_1 - \varphi_2$, $\varphi_1 > \varphi_2$, а опір виділеного циліндра $R = \rho\dfrac{\Delta l}{\Delta S}$. Згідно закону Ома (4.8) для цього елемента провідника струм

$$I = \frac{\Delta\varphi}{R} = \frac{\Delta\varphi}{\rho\dfrac{\Delta l}{\Delta S}} = \frac{1}{\rho}\frac{\Delta\varphi}{\Delta l}\Delta S.$$

Звідси отримаємо, що

$$j = \frac{I}{\Delta S} = \frac{1}{\rho}\frac{\Delta\varphi}{\Delta l}. \qquad (4.12)$$

Позначимо $1/\rho = \sigma$. Величину σ називають ***питомою провідністю*** (у СІ вимірюється в Ом$^{-1}$·м$^{-1}$. Оскільки $\lim_{\Delta l\to 0}\frac{\Delta\varphi}{\Delta l} = E$ – напруженість електричного поля в даній точці провідника, то вираз (4.12) перепишемо так:

$$j = \sigma E. \qquad (4.13)$$

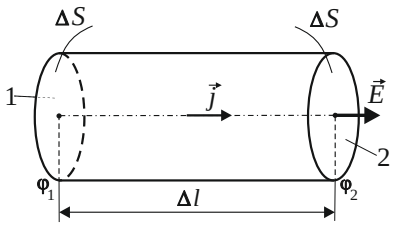


Рис.4.2.

Оскільки напрями $\vec{E}$ і $\vec{j}$ збігаються, то рівняння (4.13) можна записати у векторній формі:

$$\vec{j} = \sigma\vec{E}. \qquad (4.14)$$

Співвідношення (4.13) і (4.14) виражають *закон Ома у диференціальній формі для однорідної ділянки кола.*

Закон Ома в диференціальній формі зв'язує густину струму в кожній точці всередині провідника з напруженістю електричного поля в тій самій точці. Закон (4.14) справедливий і для змінних струмів і є одним із основних рівнянь електродинаміки (рівнянь Максвелла).

Оскільки направлений рух носіїв заряду створюється електричним полем у провіднику, то можна вважати, що дрейфова швидкість $\vec{\upsilon}_{др}$ носіїв заряду пропорційна до напруженості поля $\vec{E}$ в провіднику:

$$\vec{\upsilon}_{др} \sim \vec{E}, \text{ або } \vec{\upsilon}_{др} = u\vec{E}, \qquad (4.15)$$

де u – рухливість носіїв заряду.

Рухливість носіїв заряду числово дорівнює швидкості напрямленого руху зарядів, якого вони набувають під дією електричного поля у провіднику, напруженість якого дорівнює 1 В/м.

Тоді формулу (4.5) для густини струму можна записати у вигляді

$$\vec{j} = enu\vec{E} . \qquad (4.16)$$

Порівнюючи (4.16) і (4.14), отримаємо, що

$$\sigma = enu. \qquad (4.17)$$

Отже, питома електропровідність провідників (металів) прямо пропорційна до концентрації вільних електронів та їх рухливості.

Робота електричного струму

Якщо по ділянці провідника з опором R, на кінцях якої напруга U, проходить струм I (рис.4.1) то через його поперечний переріз за час dt буде перенесено електричний заряд $dq = Idt$. За визначенням, напруга $U = \varphi_1 - \varphi_2$ між довільними точками провідника дорівнює роботі, яка виконується при перенесенні одиниці заряду між цими точками провідника. При перенесенні заряду dq виконується робота

$$dA = Udq = IUdt .$$

У випадку постійного струму

$$A = IUt . \qquad (4.18)$$

На основі закону Ома (4.8) можна записати

$$A = I^2Rt = \frac{U^2}{R}t . \qquad (4.19)$$

Закон Джоуля-Ленца

Якщо постійний струм проходить по нерухомому провіднику, в якому відсутні хімічні реакції, робота струму перетворюється у внутрішню енергію провідника і він нагрівається. Вивчаючи теплову дію електричного струму, англійський вчений Д. Джоуль і російський Е. Ленц відкрили закон, який формулюється так: кількість теплоти Q, що виділяється в провіднику при проходженні в ньому постійного струму, прямо пропорційна квадрату сили струму I, опору провідника R, і часу проходження струму, тобто

$$Q = I^2Rt . \qquad (4.20)$$

Формула (4.20) виражає закон Джоуля-Ленца в інтегральній формі.

Оскільки кількість теплоти Q і робота A є кількісними мірами перетворення енергії електричного струму у внутрішню енергію, то всі формули роботи можуть бути використані для визначення кількості теплоти.

Запишемо закон Джоуля-Ленца у диференціальній формі через локальні характеристики провідника і електричного поля в ньому. Користуючись співвідношеннями

$$I = jS = \sigma ES \text{ і } \sigma = \frac{1}{\rho},$$

вираз (4.20) запишемо так:

$$Q = \sigma^2E^2S^2 \cdot \rho\frac{l}{S}t = \sigma E^2Slt = \sigma E^2Vt \text{ де}$$

$V = lS -$ об'єм провідника.

Кількість теплоти, що виділяється в одиниці об'єму провідника за одиницю часу (питома теплова потужність),

$$Q_{пит} = \frac{Q}{Vt} = \sigma E^2 = \rho j^2 . \qquad (4.21)$$

На основі закону Ома в диференціальній формі (4.14) перепишемо (4.21) так:

$$Q_{пит} = \vec{j}\vec{E} . \qquad (4.22)$$

Співвідношення (4.21) і (4.22) виражають закон Джоуля-Ленца у диференціальній формі.

§4.3. Електрорушійна сила

Для того, щоб в провіднику існував струм (рис.4.1), необхідно підтримувати в ньому електричне поле, тобто підтримувати різницю потенціалів $\varphi_1 - \varphi_2$. Якщо цього не робити, то електрони, рухаючись до торця 1, будуть понижувати його потенціал φ_1, підвищувати потенціал торця 2 і з часом різниця $\varphi_1 - \varphi_2$ стане дорівнювати нулю, поле всередині провідника зникне і струм перестане існувати. Щоб підтримувати сталою різницю потенціалів $\varphi_1 - \varphi_2 \neq 0$, необхідно електрони від торця 1 переносити до торця 2 за напрямком напруженості поля $\vec{E}$. Іншими словами, електрони необхідно переміщувати від позитивних зарядів до негативних, тобто виконувати роботу переміщення, витрачаючи при цьому енергію. Таке переміщення можна здійснювати тільки за допомогою *сторонніх сил*. Ці сили можуть бути обумовленні хімічними процесами в гальванічних елементах і акумуляторах; дифузією носіїв струму в неоднорідному середовищі; електричними полями, які створюються магнітними полями, що змінюються з часом. Сторонні сили виникають в генераторах за рахунок механічної енергії обертання ротора генератора. Сторонні сили не можуть бути кулонівськими силами. Кулонівські сили є потенціальними, і робота таких сил по переміщенню зарядів по замкнутому контуру (колу) дорівнює нулю. Під дією кулонівських сил заряди можуть тільки перерозподілятись поверхнею провідника, яка зразу ж стає еквіпотенціальною. Отже, на певних ділянках кола (ділянка 1-2 на рис.4.3) сили, що діють з боку електричного поля $\vec{E}$ на носії заряду, не є єдиною причиною струму.

Сторонні сили, які діють на заряди, Место для формулы.разом із силами поля $\vec{E}$ зумовлюють постійний струм. Щоб врахувати дію сторонніх сил, вводять поняття вектора *напруженості поля сторонніх сил* $\vec{E}_{ст}$, який за визначенням числово дорівнює силі, яка діє на одиничний позитивний заряд.

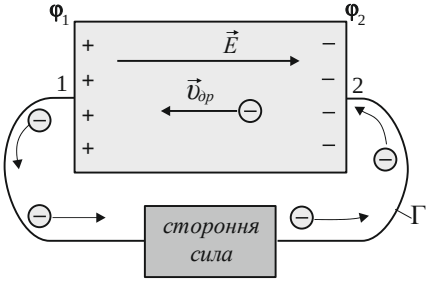


Рис.4.3.

Сторонні сили, переміщуючи електричні заряди, виконують роботу. Тому *фізична величина, що числово дорівнює роботі, яка виконується сторонніми силами під час переміщення одиничного позитивного заряду, називається електрорушійною силою (е.р.с), що діє в колі*:

$$\mathcal{E} = A_{ст}/q. \qquad (4.23)$$

Сторонняя сила $\vec{F}_{ст}$, що діє на заряд q, дорівнює

$$\vec{F}_{ст} = q\vec{E}_{ст}.$$

На електричних схемах сторонню силу зображують так, як показано на рис.4.4.

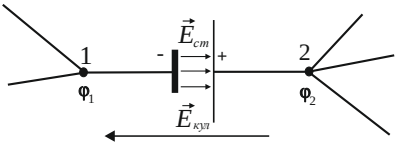


Рис.4.4.

Напруженість поля сторонніх сил $\vec{E}_{ст}$ діє в межах пристрою (джерела струму) і направлена від "мінуса" до "плюса".

Робота сторонніх сил по переміщенню в напрямі струму заряду q вздовж усього замкнутого кола Г (рис.4.3) дорівнює

$$A_{\text{о}} = \oint_{\Gamma} \vec{F}_{ст} d\vec{l} = q \oint_{\Gamma} \vec{E}_{ст} d\vec{l} \ .$$

Тоді

$$\varepsilon = \frac{A_{\text{о}}}{q} = \oint_{\Gamma} \vec{E}_{ст} d\vec{l} \ . \qquad (4.24)$$

Е.р.с., що діє в замкнутому колі, визначається циркуляцією вектора напруженості сторонніх сил $\vec{E}_{ст}$.

Е.р.с., яка діє на якійсь ділянці кола 1-2 (рис.4.4), дорівнює

$$\varepsilon = \int_{1}^{2} \vec{E}_{ст} d\vec{l}. \qquad (4.25)$$

Результуюча сила, що діє в колі на заряд:

$$\vec{F} = \vec{F}_{ст} + \vec{F}_{кул} = q(\vec{E}_{ст} + \vec{E}_{кул}). \qquad (4.26)$$

Робота, яка виконується результуючою силою над зарядом q на ділянці кола 1-2 (рис.4.5) дорівнює

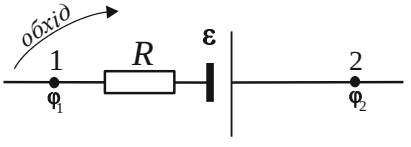


Рис.4.5.

$$A_{12} = \int_1^2 \vec{F}\, d\vec{l} = q \int_1^2 (\vec{E}_{ст} + \vec{E}_{кул}) d\vec{l} = \\ = q \int_1^2 \vec{E}_{ст}\, d\vec{l} + q \int_1^2 \vec{E}_{кул}\, d\vec{l} = \\ = q\varepsilon + q(\varphi_1 - \varphi_2). \qquad (4.27)$$

При цьому враховано співвідношення (1.34).

Напругою U_{12} на ділянці кола 1-2 називається фізична величина, що визначається роботою, яка виконується сумарним полем електростатичних і сторонніх сил при переміщенні одиничного позитивного заряду на даній ділянці кола.

Отже

$$U_{12} = \frac{A_{12}}{q} = \varphi_1 - \varphi_2 + \varepsilon. \qquad (4.28)$$

Поняття напруги є узагальненим поняттям різниці потенціалів: напруга на кінцях ділянки кола дорівнює різниці потенціалів в тому випадку, якщо на цій ділянці не прикладена е.р.с.

Ділянка кола, на якій на носії струму діють сторонні сили, називається неоднорідною. Ділянка кола, на якій не діють сторонні сили, називається однорідною. Для однорідної ділянки кола

$$U_{12} = \varphi_1 - \varphi_2. \qquad (4.29)$$

На неоднорідній ділянці кола на носії струму діє сила (4.26). Тому дрейфова швидкість $\vec{v}_{др}$ буде пропорційна силі (4.26). Відповідно густина струму буде пропорційна до суми напруженостей $\vec{E}_{кул} + \vec{E}_{ст}$:

$$\vec{j} = \sigma(\vec{E}_{кул} + \vec{E}_{ст}) = \frac{1}{\rho}(\vec{E}_{кул} + \vec{E}_{ст}). \qquad (4.30)$$

Цей вираз є *узагальненим законом Ома в диференціальній формі* для неоднорідної ділянки кола. З закону (4.30) можна отримати і інтегральну форму закону. Для цього помножимо рівняння (4.30) скалярно на елемент $d\vec{l}$ осі провідника, взятому у напрямі вектора $\vec{j}$, і проінтегруємо від перерізу 1 до перерізу 2 нерозгалуженої ділянки кола (рис.4.5):

$$\int_1^2 \frac{\vec{j} d\vec{l}}{\sigma} = \int_1^2 \vec{E}_{кул} d\vec{l} + \int_1^2 \vec{E}_{ст} d\vec{l} \ . \qquad (4.31)$$

Оскільки

$$\frac{\vec{j} d\vec{l}}{\sigma} = \frac{j dl \cos 0^\circ}{\sigma} = I \frac{\rho dl}{S} = IdR \, ,$$

де $I-$ сила струму, $S-$ площа поперечного перерізу провідника, dR – опір елемента провідника, довжина якого dl, $\vec{j} \parallel d\vec{l}$ і струм I однаковий в усіх перерізах кола. Тоді

$$\int_1^2 \frac{\vec{j}d\vec{l}}{\sigma} = \int_1^2 IdR = I\int_1^2 dR = IR_{12}, \quad (4.32)$$

де R_{12} – опір ділянці кола $1-2$.

Враховуючи співвідношення (1.34), (4.25) і (4.32), рівняння (4.31) перепишемо так:

$$IR_{12} = \varphi_1 - \varphi_2 + \mathcal{E}_{12}, \quad (4.33)$$

де $\mathcal{E}_{12}$ позначено сумарну е.р.с., яка діє на ділянці кола $1-2$.

Поділивши (4.33) на R_{12} отримаємо математичний вираз закону Ома для неоднорідної ділянки кола:

$$I = \frac{\varphi_1 - \varphi_2 + \mathcal{E}_{12}}{R_{12}}. \quad (4.34)$$

При користуванні законом (4.34) необхідно користуватись певними правилами:

1.Вибрати довільно обхід на ділянці кола $1-2$ (рис.4.5).

2. Вибрати довільно напрям струму на цій ділянці. Якщо напрямок струму і напрям обходу співпадають, то алгебраїчну величину струму треба приймати, як позитивну величину ($I>0$). Якщо напрям струму і напрям обходу на ділянці 1-2 протилежні, то ($I<0$).

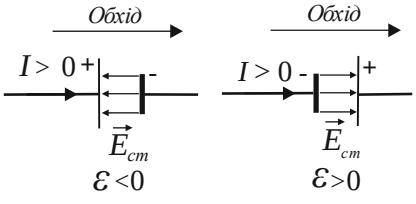


Рис.4.6.

3.Е.р.с. $\mathcal{E}$ є теж величина алгебраїчна. Вважається, що дія $\mathcal{E}$ має напрям, що співпадає з напрямком поля $\vec{E}_{cm}$ (рис.4.6). Якщо напрям $\vec{E}_{cm}$ співпадає із напрямком обходу на ділянці $1-2$, то $\mathcal{E}>0$. Якщо напрям поля $\vec{E}_{cm}$ і напрям обходу протилежні, то $\mathcal{E}<0$ (рис. 4.6).

Якщо електричне коло замкнене, то точки 1 и 2 ділянки кола (рис. 4.5) співпадають. Це означає, що $\varphi_1 = \varphi_2$ і $R_{12} = R$, де R- загальний опір кола. Тому закон Ома для замкненого кола має такий вигляд:

$$I = \frac{\mathcal{E}}{R}, \quad (4.35)$$

де $\mathcal{E}$ – алгебраїчна сума всіх е.р.с., які діють у цьому колі.

Гальванічний елемент

У звичайних батареях (акумуляторах) переміщення носіїв заряду в області, де електричне поле протилежне струму, відбувається за допомогою хімічної енергії. За рахунок цієї енергії позитивний заряд може рухатись до місця з більш високим електричним потенціалом, якщо він при цьому приймає участь в хімічній реакції, яка дає йому більше енергії, ніж потрібно для того, щоб подолати силу кулонівського відштовхування. За рахунок цієї енергії негативні заряди рухаються до місця з більш меншим електричним потенціалом, якщо вони беруть участь в хімічній реакції.

Щоб розглянути, як цей механізм працює, дослідимо це на прикладі нормального елемента Вестона, який має велике значення в лабораторії, тому що є еталоном при точних вимірюваннях напруги (рис.4.7). Він складається із Н подібної скляної посудини, яка наповнена розчином сульфату кадмію, $CdSO_4$.

У дно кожного плеча впаяні зовнішні виходи, які здійснюють контакти із

внутрішніми електродами. Лівим внутрішнім електродом є рідка чиста ртуть (4), а правим – та ж ртуть, в якій розчинений металічний кадмій (2). (В ртуті розчиняється багато металів: такий розчин називається амальгамою). Зліва над поверхнею ртуті знаходиться декілька кристалів сульфату ртуті, Hg_2SO_4 (3) – сполука, яка дуже слабо розчиняється у воді. Між зовнішніми виходами існує різниця потенціалів причому лівий вихід позитивний по відношенню до правого. (Абсолютна величина потенціалу не має значення; тут важлива тільки різниця).

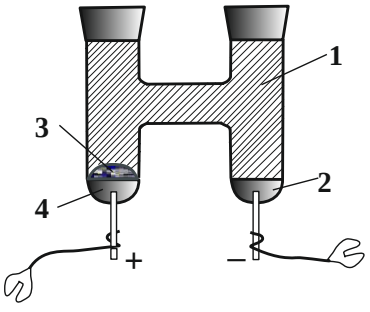


Рис.4.7. 1-водний розчин сульфату кадмію $Cd^{++}SO_4^{--}$ (електроліт); 2- амальгама кадмію (кадмій, розчинений у ртуті); 3-сульфат ртуті (Hg_2SO_4) (дуже слабо розчинний); 4-чиста ртуть. Це "ненасичений" елемент.

Ось як працює такий елемент: частина іонів елементу кадмію переходить із амальгами (2) у водний розчин (1) (рис.4.8,*б*) , причому кожний іон кадмію залишає після себе два електрони і амальгамний електрод (2) набуває значний негативний заряд. Однак цей процес зупиняється, як тільки в електроді накопичується стільки електронів, що їхня сила притягання іонів Cd^{++} перешкоджає подальшому їхньому виходу у водний розчин (1). Якщо тепер з'єднати виводи елемента провідним опором, то електрони потечуть від негативного електрода до позитивного. Це дозволить новій порції іонів Cd^{++} перейти в розчин (1), а залишені ними електрони просто

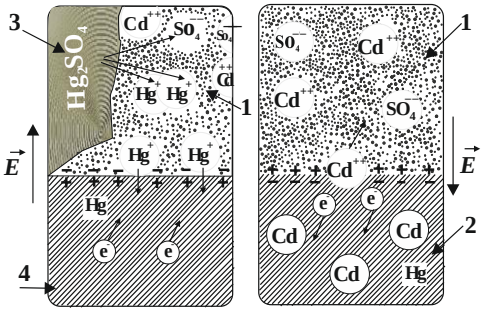
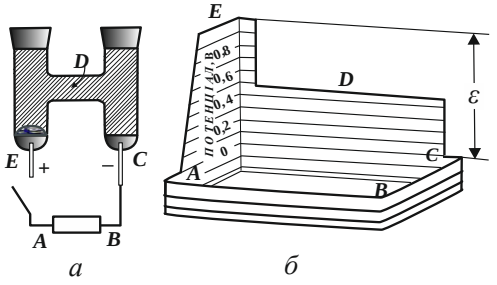


Рис. 4.8.Тут схематично показано, що відбувається на поверхні позитивного електрода (*а*) і на поверхні негативного електрода (*б*), коли елемент Вестона підключають до зовнішнього навантаження .

поповнять негативний заряд цього електрода. В колі буде протікати стаціонарний струм, який супроводжується міграцією іонів, які замикають коло всередині водного розчину (1). Тим часом дещо відбувається і на другому електроді (рис.4.8,*а*). Іони ртуті Hg^+ покидають розчин (3), зустрічаються з електронами, які приходять зовні, і стають нейтральними атомами ртуті. В розчині нові іони Hg^+ виникають при розчиненні Hg_2SO_4, причому одночасно в електроліт (1) добавляються нові сульфатні іони Hg^+. Тим часом, як показано на рис.4.8,*б*, неперервно йде дисоціація атомів кадмію, які поступають в електроліт (1) уже у вигляді іонів Cd^{++}.

В результаті, по суті, відбувається вилучення електронів від атомів кадмію і приєднання їх до іонів ртуті. Хімік сказав би, що кадмій окислюється, а ртуть відновлюється. *Елемент працює тому, що цей обмін енергетично вигідний***.** Відносна сила зв'язку електронів в структурі атома кадмію і атома ртуті така, що, виражаючись фігурально,

намагання атомів ртуті приєднати електрони переважає бажання атомів кадмію утримувати їх. Зауважимо, що на кожній із поверхонь розділу іони Cd^{++} (рис.4.8,*б*) і іони Hg^{+} (рис.4.8,*а*) рухаються проти сил $q\vec{E}$ електричного поля.

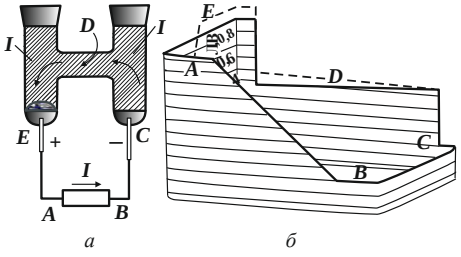


а *б*

Рис.4.9. Розподіл в розімкнутому колі (*а*) потенціалу (*б*). Потенціал правого затискача *C* прийнятий за нульове значення.

Розглянемо тепер зміну електричного потенціалу у всій системі як при протіканні струму, так і при його відсутності.

На рис.4.9,*б* по вертикалі відкладено зміну потенціалу вздовж кола, яке розімкнуте в одній точці. Різниця потенціалів на затискачах при розімкнутому колі є е.р.с. елемента, яка позначена як $\mathcal{E}$. Електричне поле дорівнює градієнту потенціалу з оберненим знаком. Як і в будь-якому електростатичному полі лінійний інтеграл $\oint_\Gamma \vec{E}d\vec{l} = 0$. На рис.4.10,*б* показаний розподіл потенціалу в тому випадку, коли по зовнішньому опору протікає струм. Всередині електроліту є електричне поле, яке направлене в ту ж сторону, що і струм. Розчин сульфату кадмію веде себе як звичайний омічний опір. Тепер різниця потенціалів на затискачах менша $\mathcal{E}$ із-за наявності внутрішнього спаду напруги на електроліті, а також, можливо, із-за додаткового опору перехідних шарів. Лінійний інтеграл $\oint_\Gamma \vec{E}d\vec{l}$ по всьому колу, як і раніше, дорівнює нулю. Нехай в результаті протікання струму через кожний переріз кола пройшов заряд q, тоді $\mathcal{E}q$ є енергія, яка розсіяна у зовнішньому колі і всередині елемента. Ця енергія отримана із хімічної енергії елемента, яка є джерелом існування сторонньої сили $\vec{E}_{ст}$.



а *б*

Рис.4.10. Розподіл потенціалу (*б*) у випадку, коли по зовнішньому опору протікає струм (*а*). Зверніть увагу на спад потенціалу всередені електроліту. Цей розгляд не придатний для елемента Вестона, який використовується як еталон. Еталонний елемент застосовується в умовах, коли протікає дуже малий струм.

Ланцюжок реакцій в елементі оборотний. Це означає, що якщо включити в коло назустріч елементу інше джерело з більшою е.р.с., то струм потече в оберненому напрямку і описані процеси будуть протікати в оберненому порядку. Саме це відбувається при зарядці батареї акумуляторів.

Е.р.с. елемента залежить від властивостей атомів. Її значення близьке до 1В, тому що енергія зв'язку зовнішніх електронів в атомі лежить в діапазоні декількох електрон-вольт, а в е.р.с. проявляється, по суті, різниця цих енергій зв'язку.

Сам елемент Вестона використовується не як джерело електричної енергії, а швидше, як еталон різниці потенціалів. Тому ситуація, яка показана на рис.4.10, коли протікає досить значний струм і вихідна напруга падає приблизно на 10 % , є грубим зловживанням до такого елементу.

Е.р.с. елемента Вестона відтворювана у високій степені. Якщо водний розчин (1) (рис.4.7) насичений надлишком сульфату кадмію в області обох електродів, то е.р.с. при

20℃ дорівнює (1,018800-1,019600)В з похибкою до 0,002 % і стабільністю (за 1 рік до 20 мкВ). При цьому допускається протікання струму (0,5-10 мкА). Використовуючи, як еталон, елемент Вестона і відповідний потенціометр, можна надійно вимірювати різницю потенціалів з точністю до 10^{-6}В.

§4.4.Розрахунок розгалужених електричних кіл за допомогою правил Кірхгофа

За допомогою закону збереження заряду і закону Ома (4.33) для неоднорідної ділянки кола можна розрахувати будь-яке складне електричне коло. Такий розрахунок значно полегшується, якщо застосувати два правила Кірхгофа.

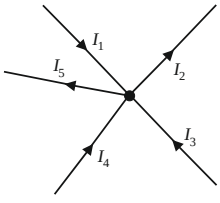


Рис.4.11.

Перше правило Кірхгофа встановлює взаємозв'язок між струмами в провідниках, що сходяться у вузлі. ***Вузлом*** розгалуженого кола називають точку, в якій сходяться три і більше провідників (рис.4.11). Нехай постійні струми I_1, I_3, I_4 входять у вузол, а струми I_2, I_5 виходять із нього. Так як у вузлах заряди не накопичуються, то кількість заряду, що приносять струми I_1, I_3, I_4 до вузла дорівнює кількості заряду, що виносять струми I_2, I_5 з вузла. Тому на основі *закону збереження заряду* можна записати, що

$$I_1 + I_3 + I_4 - I_2 - I_5 = 0, \quad (4.36)$$

або

$$-I_1 - I_3 - I_4 + I_2 + I_5 = 0. \quad (4.37)$$

Рівняння (4.36) або (4.37) виражають перше правили Кірхгофа. Для зручності користування цим рівнянням струмам, що входять у вузол, надають один знак, а струмам, що виходять – протилежний. Так, якщо вважати струм, що входить у вузол, додатним, а струм, що виходить із вузла, – від'ємним, то перше правило Кірхгофа можна сформулювати так: ***алгебраїчна сума усіх сил струмів, що сходяться в будь-якому вузлі розгалуженого кола, дорівнює нулю:***

$$\sum_{i=1}^{m} I_i = 0. \quad (4.38)$$

Друге правило Кірхгофа. Виділимо в розгалуженому колі простий замкнутий контур, наприклад *ABCDA* (рис.4.12).

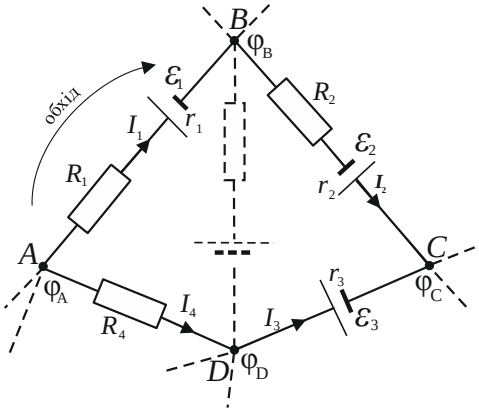


Рис.4.12.

Застосуємо закон Ома (4.33) до кожної ділянки кола цього контуру:

$$I_1R_1 + I_1r_1 = \varphi_A - \varphi_B + \mathcal{E}_1;$$

$$I_2R_2 + I_2r_2 = \varphi_B - \varphi_C + \mathcal{E}_2;$$

$$I_3r_3 = \varphi_C - \varphi_D + \mathcal{E}_3;$$

$$I_4R_4 = \varphi_D - \varphi_A.$$

Додамо почленно ліві і праві частини цих рівностей. Тоді отримаємо, що

$$I_1R_1+I_1r_1+I_2R_2+I_2r_2+I_3r_3+ +I_4R_4=\mathcal{E}_1+\mathcal{E}_2+\mathcal{E}_3. \quad (4.39)$$

Аналогічні співвідношення отримуються для будь-якого простого замкнутого контуру. Узагальнюючі співвідношення (4.39), запишемо, що

$$\sum IR +\sum Ir = \sum \mathcal{E}. \quad (4.40)$$

Співвідношення (4.40) виражає ***друге правило Кірхгофа: у будь-якому простому замкнутому контурі, довільно обраному у розгалуженому електричному колі, алгебраїчна сума добутків сил струмів на опори відповідних ділянок цього контуру дорівнює алгебраїчній сумі е.р.с., що діють у контурі.***

У сумі (4.40) знаки струмів і е.р.с. $\mathcal{E}$ (див.правила на стор.68) беруть згідно вибраного обходу по контуру (рис.4.12). Обхід для всіх замкнутих контурів розгалуженого кола треба здійснювати в одному напрямку. Найчастіше обхід контура вибирають за стрілкою годинника.

Якщо при розв'язуванні системи рівнянь, складених на основі (4.38) і (4.40), сили струмів виявились додатними, то це означає, що напрямок їх обрано правильно. У противному разі дійсний напрямок струму протилежний до обраного нами.

Контрольні питання

1.Визначення струму. Умови існування струму провідності.

Характеристики струму.

2.Причини існування електричного опору металів.

3.Визначення електрорушійної сили сторонніх сил. В чому різниця між е.р.с. та різницею потенціалів?

4.Закон Ома в інтегральній та диференціальній формах для неоднорідної ділянки кола.

5.Закон Ома в інтегральній та диференціальній формах для однорідної ділянки кола.

6.Закон Ома для замкнутого кола.

7.Вектор густини струму. Дрейфова швидкість носіїв заряду та її фізичний зміст.

8.Перше правило Кірхгофа. На основі якого фізичного закону воно встановлене?

9.Друге правило Кірхгофа. На основі якого фізичного закону воно базується?

10.Закон Ленца-Джоуля в інтегральній та диференціальній формах.

Розділ 5.Магнітне поле

§5.1.Взаємодія струмів. Закон Ампера. Магнітне поле як форма існування матерії

Нерухомі точкові заряди (рис.5.1) взаємодіють за законом Кулона (формула 1.2). Ця взаємодія, за теоріє близькодії, здійснюється через електростатичне поле, яке існує навколо нерухомих заряджених частинок. Разом з тим, на основі закону Кулона не можна пояснити характеру взаємодії тих же зарядів, які рухаються .

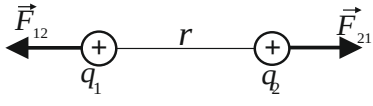


Рис.5.1.

Два позитивні нерухомі заряди відштовхуються з силою Кулона (рис.5.1). З початком руху зарядів виникає нова якість взаємодії (рис.5.2). Експериментально

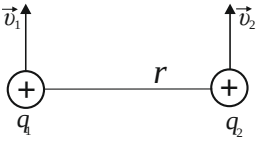


Рис.5.2.

досліджувати взаємодію двох зарядів, які

рухаються, складно. Але коли рухається багато зарядів, то експериментальні дослідження реальні. Направлений рух зарядів – це струми.

Французький фізик Ампер в 1820р. відкрив, що два тонких прямолінійних металічних провідники, якими протікає електричний струм, взаємно притягуються (рис.5.3,*а*), якщо струми мають однаковий напрямок (позитивні заряди рухаються в одному напрямку (рис.5.2) хоч між ними діє кулонівська сила відштовхування). Якщо ж струми мають протилежні напрямки (позитивний і негативний заряди рухаються в одному напрямку і між ними діє кулонівська сила притягання), то провідники відштовхуються (рис.5.3,*б*).

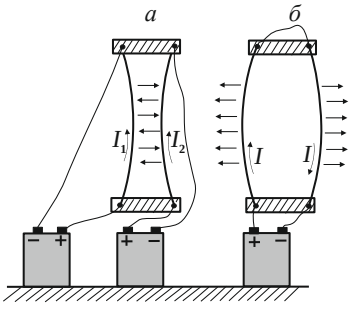


Рис.5.3.

Зауважимо, що електрична взаємодія самих проводів дорівнює нулю, так як в цілому вони електронейтральні.

Дослідним шляхом Ампер встановив закон взаємодії двох тонких прямолінійних паралельних провідників, якщо відстань між ними r набагато менша їх довжини:

$$F_l = 2kI_1I_2 / r, \qquad (5.1)$$

де I_1 і I_2 – сила струму, відповідно, в першому і другому провідниках, F_l – сила їхньої взаємодії, що приходиться на одиницю довжини проводу, $2k$ – коефіцієнт пропорційності.

На основі формули (5.1) встановлюється одиниця сили струму в СІ – (ампер) – це сила постійного струму, який протікаючи двома паралельними прямолінійними дуже тонкими провідниками нескінченої довжини, які розташовані на відстані $r = 1$м один від одного у вакуумі, визвав би між цими провідниками силу, що дорівнювала б $2\cdot10^{-7}$ ньютон на кожен метр довжини.

В раціональному виді формулу (5.1) записують так:

$$F_l = \frac{\mu_0}{4\pi} \cdot \frac{2I_1I_2}{r}, \qquad (5.2)$$

де μ_0 – так звана ***магнітна стала***. Якщо I_1=I_2 =1А, r = 1 м, F_l =2·10^{-7}Н/м, то із (5.2) отримаємо, що

$$\mu_0 = 4\pi \cdot 10^{-7} \text{ Гн/м}. \qquad (5.3)$$

Взаємодія струмів здійснюється через поле, яке називають магнітним. Магнітне поле отримало свою назву у зв'язку з тим, що Ерстед (практично одночасно з відкриттям Ампера) відкрив дію електричного струму на магнітну стрілку (рис.5.4).

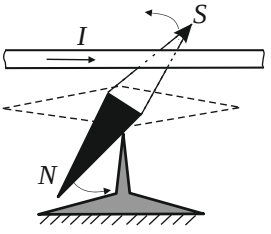


Рис.5.4.

Магнітна стрілка розміщувалась під проводом. При пропусканні по провіднику струму, стрілка орієнтувалась перпендикулярно до проводу. При зміні напрямку струму стрілка поверталась на 180°. Те саме мало місце при переносі стрілки із положення під проводом в положення стрілки

над проводом. При віддаленні магнітної стрілки від провідника орієнтуюча дія струму зменшувалась.

У цих дослідах було зроблено декілька фундаментальних відкриттів. *По-перше*, показано, що електричний струм створює магнітне поле (на магнітну стрілку може діяти тільки магнітне поле), *по-друге*, що величина магнітного поля в різних точках простору навколо провідника зі струмом різна (зменшується з відстанню від провідника), і, *по-третє*, що магнітне поле має напрям у просторі, тобто є полем векторним. Після відкриття Ерстеда почалися інтенсивні дослідження магнітних полів і струмів (рис.5.5).

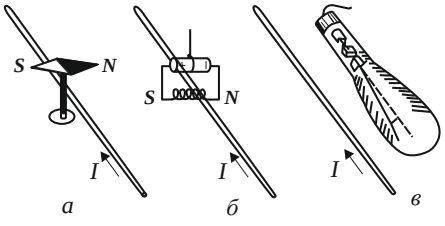


Рис.5.5.

Дія магнітного поля провідника зі струмом на магнітну стрілку (рис. 5.5,*а*) нічим не відрізняється від його дії на котушку, по якій протікає струм від батареї (рис 5.5,*б*), або дії на рухомий пучок електронів в електронно-променевій трубці (рис.5.5,*в*). Останні досліди свідчать не тільки про те, що магнітне поле створюється рухомими електричними зарядами, але й про те, що магнітне поле, в свою чергу, діє лише на рухомі електричні заряди і цією дією інтегрально визначається взаємодія провідників зі струмом, або провідника зі струмом і магнітної стрілки, або, взагалі, магнітна взаємодія. Отже, навколо будь-якого рухомого заряду, чи то буде електрон, іон або заряджене тіло, крім електричного поля, існує також і магнітне поле. Електричне поле діє як на рухомі, так і на нерухомі електричні заряди. Магнітне поле діє лише на рухомі електричні заряди.

§5.2.Взаємодія елементів струму

У дослідах Ампера насамперед було встановлено, що сила взаємодії двох провідників пропорціональна величині струму в кожному із провідників. Далі досліди показали, що коли провід зі струмом зігнути, як показано на рис.5.6,*а*, то він не чинить магнітної дії. І навпаки, такий провідник не зазнає дії сили з боку інших провідників зі струмом. Магнітна дія не спостерігається і тоді, коли одну частину проводу (і притому довільно) обмотати навколо другої (рис.5.6,*б*).

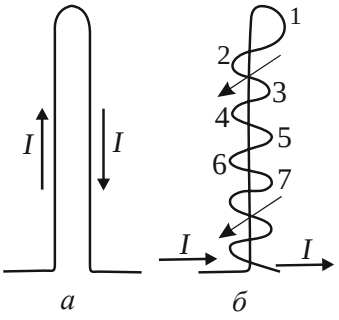


Рис.5.6.Зігнуті провідники *а* і *б* не чинять магнітної дії.

З цих результатів випливає висновок, що які-небудь елементи провідника $d\vec{l}_1, d\vec{l}_2$ і $d\vec{l}_3$ разом чинять таку саму магнітну дію, як один елемент $d\vec{l}$ (рис.5.7), що замикає ці відрізки. Зокрема, дія зігнутих відрізків 12 і 23 провідника *б* (рис.5.6) така, як коли замість них був би прямолінійний відрізок, що сполучає точки 1 і 3, дія 34 і 45 дорівнює дії 35 і т.д. , і тому дія всього провідника *б* така сама, як і провідника a (рис.5.6), тобто дорівнює нулю.

Із сказаного випливає, що магнітна дія нескінченно малого відрізка проводу залежить від добутку $Id\vec{l}$, де I – величина струму , а $d\vec{l}$ – вектор, що має довжину відрізка dl і направлений вздовж струму. ***Цей добуток називають елементом струму.***

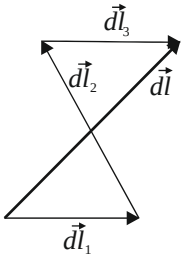


Рис.5.7.До поняття елемента струму.

Поняття елемента струму в законах магнітної взаємодії відіграє те саме значення, що й поняття точкового заряду в законах електричної взаємодії. Сила $d\vec{F}_{12}$, з якою діє який-небудь елемент струму $I_1 d\vec{l}_1$ на інший елемент струму $I_2 d\vec{l}_2$, пропорціональна величині кожного з елементів струму, тобто їх добутку:

$$dF_{12} \sim I_1 d\vec{l}_1 \cdot I_2 d\vec{l}_2 .$$

Ця сила залежить також від відстані r_{12} між елементами струму. Ампер припустив, що вона обернено пропорціональна відстані в деякій степені n і з дослідів зробив висновок, що $n = 2$:

$$dF_{12} \sim 1/r_{12}^2 .$$

Величина і напрям сили взаємодії двох елементів струму залежить ще від їх взаємного ***орієнтування***. Визначення цієї залежності було дуже складним і потребувало ряду досліджень різних учених. Розглянемо спочатку випадок, коли відрізки провідників (елементи струму) $d\vec{l}_1$ і $d\vec{l}_2$ лежать в одній площині, і позначимо через $\vec{r}_{12}$ радіус-вектор, проведений від першого відрізка до другого

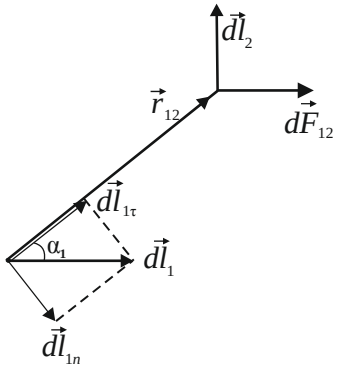


Рис.5.8.Взаємодія двох елементів струму, які лежать в одній площині.

(рис.5.8). Згідно зі сказаним вище, дію відрізка $d\vec{l}_1$ можна замінити дією двох його складових $d\vec{l}_{1\tau}$ і $d\vec{l}_{1n}$, одна з яких направлена вздовж радіуса-вектора $\vec{r}_{12}$, друга – перпендикулярна до нього.

З даних досліду зроблено висновок, що відрізок $d\vec{l}_{1\tau}$ не чинить ніякої дії на $d\vec{l}_2$ при будь-якому орієнтуванні останнього. Тому сила, яка діє на $d\vec{l}_2$, пропорційна $dl_{1n} = dl_1 \sin\alpha_1$, де α_1 – кут між векторами $d\vec{l}_1$ і $\vec{r}_{12}$:

$$dF_{12} \sim \sin\alpha_1 = \sin(d\vec{l}_1, \vec{r}_{12}) .$$

У загальному випадку відрізки провідників $d\vec{l}_1$ і $d\vec{l}_2$ можуть не лежати в одній площині. Тоді відрізок $d\vec{l}_2$ можна розкласти на дві складові, одна з яких, $d\vec{l}_{2\tau}$ лежить у площині П, що містить $d\vec{l}_1$ і $\vec{r}_{12}$ (рис.5.9), а друга, $d\vec{l}_{2n}$, перпендикулярна до

цієї площини. Досліди ведуть до висновку, що сила, яка діє на $d\vec{l}_{2n}$, завжди дорівнює нулю. Тому сила, яка діє на $d\vec{l}_2$, пропорціональна складовій $dl_{2\tau} = dl_2 \sin\alpha_2$, де α_2 – кут між вектором $d\vec{l}_2$ і нормаллю $\vec{n}$ до площини П:

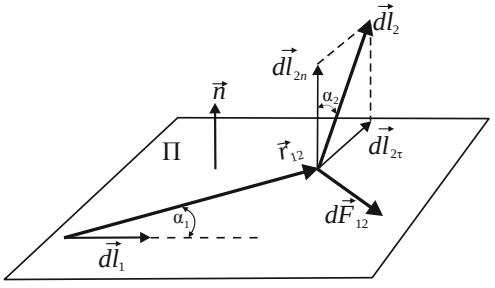


Рис.5.9.Взаємодія двох елементів струму в загальному випадку.

$$dF_{12} \sim \sin\alpha_2 = \sin(d\vec{l}_2, \vec{n}).$$

Об'єднуючи всі знайдені дослідні результати в одну формулу, знаходимо:

$$dF_{12} = k\frac{I_1 dl_1 \cdot I_2 dl_2 \cdot \sin\alpha_1 \cdot \sin\alpha_2}{r^2_{12}}. \quad (5.4)$$

Тут k означає коефіцієнт пропорціональності, який залежить від вибору системи одиниць (в СІ $k = \mu_0/(4\pi)$).

Формула (5.4) виражає величину сили dF_{12}. Напрям цієї сили визначається таким правилом: сила, яка діє на елемент струму $I_2 d\vec{l}_2$, перпендикулярна до цього елемента і лежить у площині, що містить елемент струму $I_1 d\vec{l}_1$ і радіус – вектор $\vec{r}_{12}$ (рис.5.9). При цьому напрям сили визначається ще ***правилом правого свердлика***: якщо обертати свердлик з правою нарізкою так, щоб його рукоятка переміщувалась від елемента струму $I_2 d\vec{l}_2$, до нормалі $\vec{n}$, то поступальний рух свердлика збігається з напрямом $d\vec{F}_{12}$. А напрям нормалі $\vec{n}$ також визначається цим правилом: він збігається з поступальним рухом свердлика, якщо його обертати від елемента струму $I_1 d\vec{l}_1$ до радіус – вектора $\vec{r}_{12}$ (рис.5.9). Величину і напрям сили магнітної взаємодії можна подати однією формулою, якщо скористатися поняттям векторного добутку двох векторів. Векторний добуток $d\vec{l}_1 \times \vec{r}_{12}$ направлений уздовж нормалі $\vec{n}$ до площини П, а векторний добуток цього вектора на вектор $d\vec{l}_2$, тобто подвійний векторний добуток $d\vec{l}_2 \times d\vec{l}_1 \times \vec{r}_{12}$ лежить в самій площині П і перпендикулярний до $d\vec{l}_2$, тобто направлений так само, як і сила $d\vec{F}_{12}$. Тому ми можемо написати, що

$$d\vec{F}_{12} = \frac{\mu_0}{4\pi}\frac{I_2 d\vec{l}_2 \times I_1 d\vec{l}_1 \times \vec{r}_{12}}{r^3_{12}}. \quad (5.5)$$

Закон Ампера, виражений формулами (5.4) і (5.5), є основним у вченні про магнетизм і відграє те саме значення, що і закон Кулона в електростатиці. Користуючись цим законом, можна обчислити силу взаємодії між елементами струму, на які можна розбити будь-які замкнені провідники, а підсумовуючи потім ці сили для всіх елементів струму, можна обчислити силу взаємодії між замкненими контурами скінченних розмірів будь-якої форми.

На останок зазначимо, що закон Ампера у формі (5.4) і (5.5), на перший погляд не задовільняє третій закон Ньютона. Так, наприклад, якщо розглянути два елементи струму 1 і 2, зображенні на рис. 5.10, то закон Ампера (5.5) дає, що сила F_{12}, яка діє на

елемент 2, дорівнює нулю; але сила F_{21}, що діє на елемент 1, не дорівнює нулю. Такий результат вийшов тому, що на досліді можна

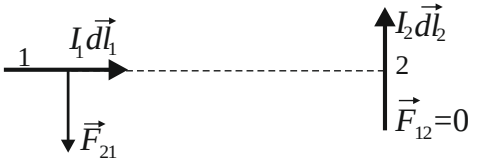


Рис.5.10.Два елементи струму.

дослідити тільки взаємодію замкнених контурів скінченної величини і не можна реалізувати взаємодію окремих елементів струму. Тому з досліду можна вивести закон взаємодії елементів струму тільки з точністю до деякого доданка, який перетворюється в нуль при підсумуванні по замкненому контуру. Такий доданок опущено в (5.4) і (5.5), і це є причиною позірного порушення третього закону Ньютона. Однак цей доданок не відграє ніякого значення, бо закон Ампера ми завжди застосовуємо тільки до замкнених контурів, а в цьому випадку він завжди дає результати, які узгоджуються з третім законом Ньютона.

Магнітне поле

При поясненні магнітної взаємодії струмів ми стикаємося з тими самими питаннями, що й при поясненні електричної взаємодії зарядів. І тут можна запитати, чому виникають сили, які діють на контур зі струмом у присутності іншого контура, і як ці сили передаються від одного провідника до іншого? Чи відбуваються які-небудь зміни в просторі навколо проводу зі струмом, коли іншого проводу немає і магнітна взаємодія не проявляється?

З тих самих причин, які викладені в розділі 1, сучасна фізика спростовує можливість далекодії в магнітних явищах, так само як і в електричних. Причину виникнення сил магнітної взаємодії ми бачимо в появі навколо провідників зі струмом *магнітного поля.* Далі ми побачимо, що магнітне поле є носієм ряду фізичних властивостей. Основна властивість магнітного поля полягає в тому, що на провідники зі струмом, які перебувають у ньому, діють сили.

Магнітне поле виникає навколо проводу зі струмом завжди, навіть при відсутності інших провідників, коли магнітна взаємодія не спостерігається. І в цьому випадку в просторі, який оточує провідник зі струмом, відбуваються певні фізичні зміни. Основне завдання дослідження магнітних явищ полягає у вивченні властивостей магнітного поля і законів, яким воно підпорядковується.

Термін «Магнітне поле» ввів в 1845 році англійський фізик М. Фарадей, який вважав, що як електрична, так і магнітна взаємодія здійснюється через єдине матеріальне (електромагнітне) поле. Класична теорія електромагнітного поля була створена англійським фізиком Дж. Максвеллом (1873р.), квантова теорія – в 20 роках 20 століття.

§5.3.Індукція магнітного поля. Закон Біо-Савара-Лапласа

Кількісною характеристикою магнітного поля є спеціальна фізична величина – індукція магнітного поля, яку ми визначаємо аналогічно до напруженості електричного поля (розділ 1, формули (1.5) і (1.6)). Основний закон магнітної взаємодії струмів є закон Ампера (5.5), який виражає силу взаємодії двох елементів струму. У ньому ми можемо виділити частину

$$d\vec{B}_1 = \frac{\mu_0}{4\pi}\frac{I_1 d\vec{l}_1 \times \vec{r}_{12}}{r^3_{12}}, \qquad (5.6)$$

яка не містить другого елемента струму $I_2 d\vec{l}_2$. Ця величина $d\vec{B}_1$ залежить тільки від елемента струму $I_1 d\vec{l}_1$ і положення тієї точки, де перебуває елемент струму $I_2 d\vec{l}_2$, і тому характеризує магнітне поле елемента струму $I_1 d\vec{l}_1$ в точці перебування елемента струму $I_2 d\vec{l}_2$. Вона і дістала назву ***індукція магнітного поля.*** Вводячи поняття індукції магнітного поля в ***закон Ампера*** (5.5), ми можемо подати силу, яка діє на елемент струму $I_2 d\vec{l}_2$, у вигляді:

$$d\vec{F}_{12} = I_2 d\vec{l}_2 \times d\vec{B}_1 . \qquad (5.7)$$

Отже, ми приходимо до такого визначення індукції магнітного поля: *індукція магнітного поля* $d\vec{B}$*, яка створюється в даній точці елементом струму* $Id\vec{l}$ *, дорівнює*

$$d\vec{B} = \frac{\mu_0}{4\pi} \frac{Id\vec{l} \times \vec{r}}{r^3}, \qquad (5.8)$$

де $\vec{r}$ – радіус-вектор, проведений з елемента провідника $d\vec{l}$ в розглядувану точку. Через те, що векторний добуток двох векторів є також вектором, то очевидно, що ***індукція магнітного поля є вектор.***

З (5.8) випливає, що *величина* індукції магнітного поля в точці, віддаленій на відстань r від елемента струму, дорівнює:

$$dB = \frac{\mu_0}{4\pi} \frac{Idl \sin\alpha}{r^2}, \qquad (5.9)$$

де α – кут між $d\vec{l}$ і радіусом вектором $\vec{r}$, проведеним від елемента струму в розглядувану точку (рис.5.11).

Напрям вектора індукція магнітного поля перпендикулярний до $d\vec{l}$ і $\vec{r}$, тобто перпендикулярний до площини, що містить обидва ці вектори. Цей напрям підлягає правилу правого свердлика: напрям магнітного поля збігається з напрямом руху кінця рукоятки свердлика з правою нарізкою, що рухається поступально в напрямі струму. Так, наприклад, якщо струм проходить вертикально

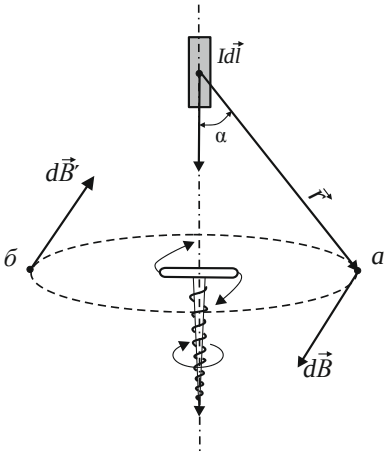


Рис.5.11.Індукція магнітного поля елемента струму.

зверху вниз (рис.5.11), то правий свердлик треба обертати за стрілкою годинника (якщо дивитися зверху), а кінець його рукоятки, що перебуває в точці a , рухатиметься при цьому від рисунка до читача; так буде напрямлене і магнітне поле в точці a. У точці $б$ воно направлене протилежно від читача за рисунок. Формули (5.8) і (5.9), які виражають індукцію магнітного поля елемента струму, називають ***законом Біо-Савара-Лапласа.*** Закон Біо-Савара-Лапласа є одним із основних експериментальних законів електромагнітних явищ і він, подібно до закону Кулона, лежить в основі класичної електродинаміки. Цей закон дає змогу розраховувати індукцію магнітних полів струмів.

Для магнітного поля, як і для електричного, справедливий ***принцип накладання, або суперпозиції:*** якщо є кілька елементів струму або кілька контурів зі струмом, кожний з яких створює магнітне

поле, то при одночасній дії всіх струмів індукція $\vec{B}$ результуючого поля дорівнює векторній сумі індукцій окремих полів:

$$\vec{B} = \vec{B}_1 + \vec{B}_2 + \vec{B}_3 + ... = \sum \vec{B}_k \cdot \quad (5.10)$$

Користуючись виразом (5.8) або (5.9) для індукції магнітного поля елемента струму і принципом накладання (5.10), можна обчислити магнітне поле будь-якого провідника зі струмом.

Необхідно зазначити, що таку характеристику магнітного поля, як індукція $\vec{B}$, часто вводять таким чином.

Розглянемо замкнений плоский контур зі струмом, розміри якого малі порівняно з відстанню до струмів, що утворюють поле. За позитивний напрямок нормалі $\vec{n}$ приймається напрямок поступального руху свердлика, рукоятка якого обертається за напрямком струму в контурі (рис.5.12).

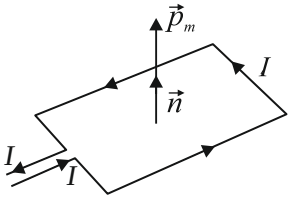


Рис.5.12.

Контур зі струмом характеризується ***магнітним моментом*** $\vec{p}_m$, величина якого дорівнює добутку сили струму I, що протікає у контурі, на площу поверхні контуру S:

$$\vec{p}_m = IS\vec{n},$$

де $\vec{n}$ одиничний вектор нормалі до поверхні рамки. Напрямок вектора $\vec{p}_m$ збігається з напрямком позитивної нормалі рамки.

Контур зі струмом в магнітному полі повертається, набуваючи рівноважного положення і його позитивна нормаль розміщується вздовж осі магнітної стрілки в напрямку від її магнітного полюса (рис.5.13). Контуром зі струмом можна скористатись і для кількісного опису магнітного поля. На контур в магнітному полі діє пара сил. Обертальний момент пари сил M залежить від властивостей контуру:

$$M \sim p_m .$$

Якщо контур зі струмом повернути на 90° від рівноважного положення, то на нього буде діяти максимальний обертальний момент M_{max}.

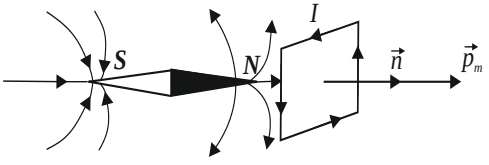


Рис.5.13.

Якщо в дане місце магнітного поля поміщати контури з різними магнітними моментами, то на них діятимуть різні обертальні моменти, але відношення M_{max}/p_m для всіх контурів однакове і служить кількісною характеристикою магнітного поля:

$$B = M_{max}/p_m .$$

Магнітна індукція $\vec{B}$ в даному місці магнітного поля визначається максимальним обертальним моментом, що діє на контур з одиничним магнітним моментом.

Одиниця магнітної індукції – тесла: 1 Тл – магнітна індукція такого магнітного поля, в якому на рамку з магнітним моментом 1А·м^2, діє максимальний момент сили 1 Н·м.

За напрямок магнітної індукції приймається напрямок магнітного моменту

контуру $\vec{p}_m$, який знаходиться в рівноважному положенні у цьому полі.

Для графічного зображення магнітних полів зручно користувались лініями магнітної індукції.

Лініями магнітної індукції називають такі лінії, дотичні до яких в кожній точці збігаються з напрямком вектора $\vec{B}$ в цих точках поля.

Напрямок лінії індукції магнітного поля струму визначається за правилом свердлика: якщо вкручувати свердлик за напрямком руху струму в провіднику, то напрямок руху його рукоятки покаже напрям ліній магнітної індукції.

Лінії магнітної індукції можна спостерігати за допомогою дрібних металевих ошурків, які в магнітному полі поводять себе, як маленькі магнітні стрілки. Вигляд лінії магнітної індукції простих магнітних полів показаний на рис.5.14.

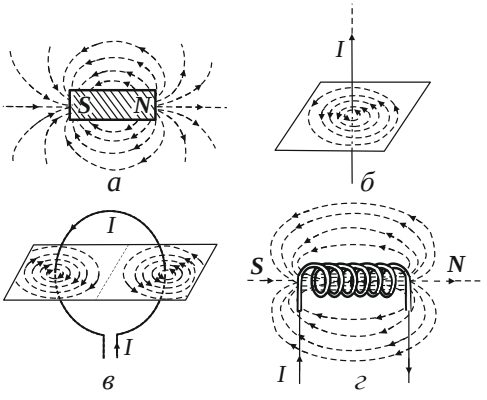


Рис.5.14.

Лінії магнітної індукції охоплюють провідник зі струмом, який створює поле. Поблизу провідника лінії лежать в площині, що перпендикулярна до провідника (рис.5.14,*б*). Лінії індукції магнітного поля не можуть обриватися ні в одній точці поля, тобто вони завжди замкнені. ***Лінії індукції постійного магніту*** (рис.5.14,*а*) ***виходять із його північного полюса N і входять у південний полюс S.***

Магнітне поле соленоїда, тобто довгої котушки зі струмом (рис.5.14,*г*), подібне до магнітного поля штабового магніту. Північний полюс магніту збігається з тим кінцем соленоїда, з якого струм у витках тече проти годинникової стрілки. Магнітне поле колового струму (рис.5.14,*в)*, який є одним витком соленоїда, подібне на поле дуже короткого штабового магніту, що розташований в центрі витка, так щоб його вісь була перпендикулярна до площини.

§5.4.Застосування закону Біо-Савара--Лапласа до розрахунку магнітних полів

1.Магнітне поле прямолінійного провідника зі струмом

Знайдемо індукцію магнітного поля, яка створюється прямим проводом зі струмом в точці *A* (рис.5.15), що знаходиться на відстані r від осі проводу. Поділимо провідник на нескінченно малі елементи Idl (наприклад елемент *BC*). За формулою (5.9) знайдемо величину індукції $d\vec{B}$, яка створюється елементом $Id\vec{l}$ в точці *A:*

$$dB=\frac{\mu_0}{4\pi}\frac{Idl\sin\alpha}{AC^2}.$$

На основі формули (5.8) неважко переконатись, що для заданого напрямку струму I всі елементарні значення індукції магнітного поля напрямлені в точці *A* в один бік по одній прямій перпендикулярно до площини рисунка від нас. Тому результуюча індукція магнітного поля всіх елементів струму в точці *A* за принципом суперпозиції (5.10) визначиться так:

$$dB = \frac{\mu_0}{4\pi}\int_l \frac{Idl\sin\alpha}{AC^2}, \qquad (5.11)$$

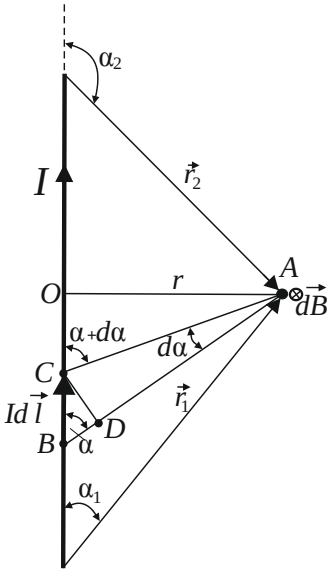


Рис.5.15.

де інтегрування ведеться по всій довжині проводу. Щоб виконати інтегрування перейдемо до однієї змінної α. Будемо вважати, що в трикутнику *ABC* кут *CAB* дорівнює $d\alpha$, який є нескінченно малим. Тоді із трикутника *ACD* можемо знайти, що $CD = AC\cdot d\alpha$. Із трикутника *BCD* знайдемо, що $BC=dl=\frac{CD}{\sin\alpha}=\frac{ACd\alpha}{\sin\alpha}$. Із трикутника *ACO* знаходимо, що

$AC = OA/\sin(\alpha+d\alpha) = r/\sin(\alpha+d\alpha)$.

$\mathrm{Sin}(\alpha+d\alpha) = \sin\alpha\cdot\cos(d\alpha) +$

$+\cos\alpha\cdot\sin d\alpha \cong \sin\alpha$, тому що $\cos(d\alpha)\approx 1$, а $\sin d\alpha \approx 0$. Тоді

$$BC = dl = \frac{rd\alpha}{\sin^2\alpha}.$$

Тепер вираз (5.11) перепишемо так:

$$B = \frac{\mu_0 I}{4\pi r}\int_{\alpha_1}^{\alpha_2}\sin\alpha d\alpha.$$

Звідки після інтегрування отримаємо, що

$$B = \frac{\mu_0 I}{4\pi r}(\cos\alpha_1 - \cos\alpha_2), \qquad (5.12)$$

де α_1 і α_2 – кути, що утворюються радіус-векторами $\vec{r}_1$ і $\vec{r}_2$, проведеними в точку *A* відповідно із початку і кінця провідника, з напрямком струму. У випадку нескінченно довгого прямого провідника $\alpha_1 \to 0$, $\alpha_2 \to \pi$ і формула (5.12) прийме вид:

$$B = \frac{\mu_0 I}{2\pi r}, \qquad (5.13)$$

де $2\pi r$ – довжина силової лінії, радіус якої *r*.

2.Поле колового витка зі струмом

Нитка зі струмом у вигляді кільця, радіус якого *R*, показані на рис.5.16,*а*. Можемо здогадуватись без всяких розрахунків, що магнітне поле повинно мати вигляд, як показано на рис.5.16,*б*, де нарисовані декілька силових ліній в площині *ZOX*. Поле має обертальну симетрію відносно осі *Z*, а самі силові лінії симетричні відносно площини колового витка (*YOX*). Дуже близько до нитки поле нагадує поле довгого прямого проводу.

Поле на осі *Z* вирахуємо, використовуючи співвідношення (5.8). Кожний елемент кільця $Id\vec{l}$ створює в точці *A* вектор $d\vec{B}$, який перпендикулярний до $\vec{r}$ і до $Id\vec{l}$. Розглянемо діаметрально протилежний до $Id\vec{l}$ елемент струму $Id\vec{l}_1$, який в точці *A* створює вектор $d\vec{B}_1$ (рис.5.16,*а*). Проекції векторів $d\vec{B}$ і $d\vec{B}_1$ на будь-яку вісь, яка перпендикулярна до осі Z,

однакові за величиною і протилежні за напрямком. Це означає, що складатись будуть Z компоненти

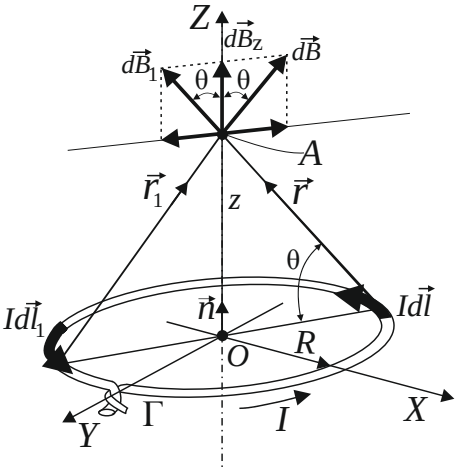


Рис. 5.16,*а*.

вектора $d\vec{B}$. Згідно (5.9)

$$dB = \frac{\mu_0 Idl\sin(\pi/2)}{4\pi r^2},$$

а

$$dB_z = \frac{\mu_0}{4\pi}\frac{Idl}{r^2}\cos\theta = \frac{\mu_0}{4\pi}\frac{IRdl}{r^3},$$

де враховано, що $\cos\theta = R/r$.

Інтегруючи по всьому кільцю, отримаємо:

$$B_z = \frac{\mu_0 IR}{4\pi r^3}\oint_\Gamma dl = \frac{\mu_0 IR}{4\pi r^3}2\pi R = \\ = \frac{\mu_0 IR^2}{2r^3} = \frac{\mu_0 IR^2}{2(R^2+z^2)^{3/2}}. \qquad (5.14)$$

При цьому враховано, що $r = \sqrt{R^2+z^2}$, де z – координата точки *A* на осі *Z*. Якщо врахувати, що магнітний момент колового витка зі струмом $\vec{p}_m = \pi R^2 I\vec{n}$, то формулу (5.14) перепишемо так:

$$\vec{B}_z = \frac{\mu_0 \vec{p}_m}{2\pi(R^2+z^2)^{3/2}}. \qquad (5.15)$$

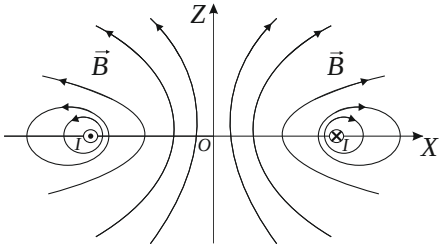


Рис.5.16,*б*.

В центрі кільця $z = 0$ і величина індукції поля згідно (5.14) буде:

$$B_0 = \frac{\mu_0 I}{2R}. \qquad (5.16)$$

3. Магнітне поле соленоїда

Циліндрична провідна котушка, яка показана на рис.5.17,*а*, називається соленоїдом. Будемо вважати, що провід розподілений щільно і рівномірно, так, що число витків на одиницю довжини вздовж циліндра є величиною постійною, яка дорівнює n. Таким чином, в дійсності струм протікає по спіралі, але якщо витків багато і вони розташовані щільно один до одного, то ми можемо цим знехтувати і розглядати соленоїд як сукупність кілець зі струмом. Тоді ми можемо використати співвідношення (5.14) за основу для розрахунку індукції в любій точки на осі циліндра z, наприклад в точці *A*. Розглянемо спочатку вклад кільця зі струмом, яке розташоване між радіусами, що проведені із точки *A* і утворюють з віссю *Z* кути від α до $\alpha + d\alpha$ (рис.5.17,*б*). Ширина цього кільця dz. Із трикутника *ABC*,

знайдемо, що $AC=BC/\sin\alpha = R/\sin\alpha$. Тоді

$$dz = \frac{AC \cdot d\alpha}{\sin\alpha} = \frac{Rd\alpha}{\sin^2\alpha}.$$

Сила струму, який протікає кільцем, ширина якого dz і містить *dN* витків, буде:

$$dI = IdN = Indz = \frac{nIRd\alpha}{\sin^2\alpha}.$$

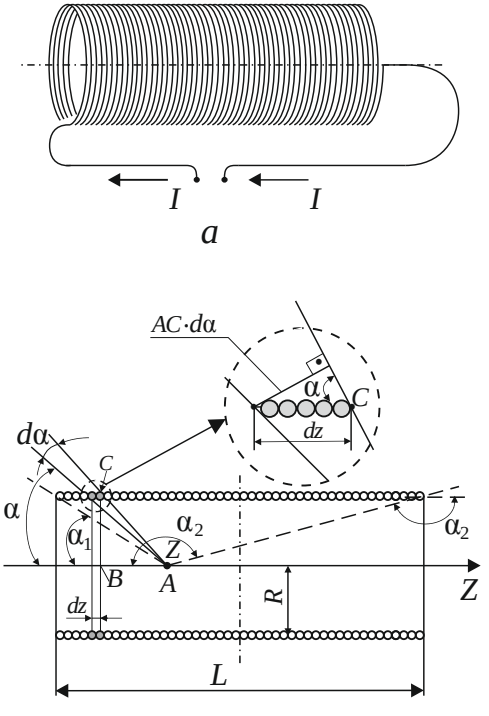
Так як $r = AC = R/\sin\alpha$, $AB=z = AC\cos\alpha = R\cos\alpha/\sin\alpha$, то за формулою (5.14) отримаємо:

$$dB_z = \frac{\mu_0 R^2 dI}{2R^3}\sin^3\alpha = \frac{\mu_0 nI}{2}\sin\alpha d\alpha.$$

Інтегрування в межах від α_1 до α_2 дає

$$B_z = \frac{\mu_0 nI}{2}\int_{\alpha_1}^{\alpha_2}\sin\alpha d\alpha = \frac{\mu_0 nI}{2}(\cos\alpha_1 - \cos\alpha_2), \quad (5.17)$$

де α_1 і α_2 кути показані на рис.5.17,*б*.



а

б

Рис.5.17.

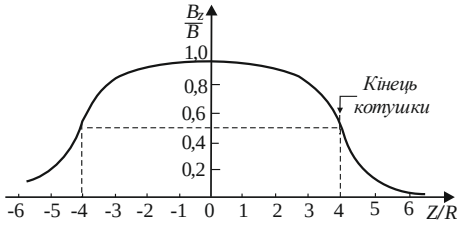


Рис.5.18.

Скористаємось рівнянням (5.17) і побудуємо графік залежності індукції в точці *A* від її положення на осі відносно середини соленоїда (рис.5.18). Приймаємо, що довжина котушки L в чотири рази більша її діаметра 2*R*. По осі ординат відкладено відношення величини індукції B_z до величини поля в котушці нескінченно великої довжини з тією ж

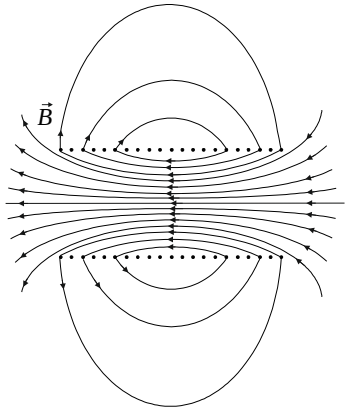


Рис.5.19.Силові лінії магнітного поля соленоїда. Точками позначений переріз проводів котушки.

кількістю витків n на одиницю довжини і з тією ж силою струму в кожному витку. Для нескінченно довгої котушки $\alpha_1 \to 0$, $\alpha_2 \to \pi$ і формула (5.17) прийме вид:

$$B = \mu_0 nI. \qquad (5.18)$$

В центрі котушки «чотири до одного» індукція поля досить близька до величини , що визначається формулою (5.18), і залишається майже сталою, поки не наблизимось до одного із кінців котушки. На рис.5.19 зображені лінії індукції магнітного поля всередині і зовні котушки з такими пропорціями. Зауважимо, що деякі силові лінії проходять через обмотку. Циліндричний шар струму є поверхнею розриву для магнітного поля.

4. Магнітне поле електричного заряду, який рухається

За законом Біо-Савара-Лапласа (5.8) магнітна індукція пропорційна силі струму, тобто кількості елементарних зарядів, які проходять за одиницю часу через переріз

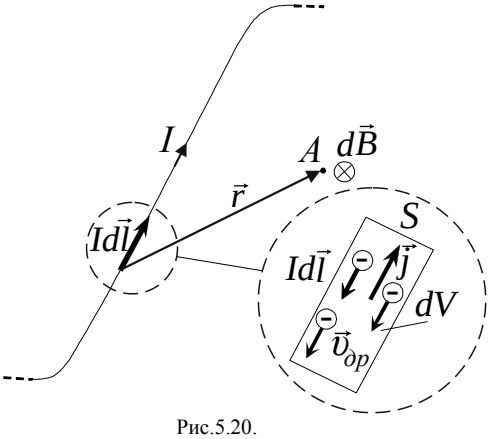


Рис.5.20.

провідника. Знайдемо індукцію магнітного поля, яка зв'язана з одним елементарним зарядом, який рухається. Для цього в законі (5.8) елемент струму $Id\vec{l}$ (рис.5.20) перепишемо так:

$$Id\vec{l} = jSd\vec{l} = \vec{j}Sdl = en\vec{v}_{др}dV = e\vec{v}_{др}dN,$$

де $\vec{j}$ – вектор густини струму, n – концентрація носів струму, $\vec{v}_{др}$ – швидкість їх направленого руху (дрейфова швидкість), S – переріз проводу, dN – кількість носіїв заряду в об'ємі елемента струму $Id\vec{l}$.

Тоді закон (5.8) перепишеться так:

$$d\vec{B} = \frac{\mu_0}{4\pi}\frac{e\vec{v}_{др}\times\vec{r}}{r^3}dN.$$

Так як всі носії заряду (електрони) в середньому мають однакову направлену швидкість руху, то індукція магнітного поля, що створюється одним елементарним зарядом, буде:

$$\vec{B}_e = \frac{d\vec{B}}{dN} = \frac{\mu_0}{4\pi}\frac{e\vec{v}_{др}\times\vec{r}}{r^3},$$

або в загальному випадку

$$\vec{B}_q = \frac{\mu_0}{4\pi}\frac{q\vec{v}\times\vec{r}}{r^3}, \qquad 5.19)$$

де q – заряд, що рухається із швидкістю $\vec{v}$.

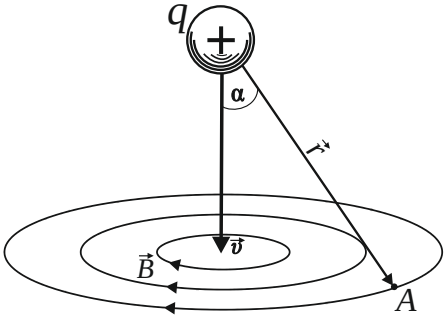


Рис.5.21.

Із формули (5.19) витікає, що вектор $\vec{B}$ в кожній точці *A* направлений перпендикулярно до площини, яка проходить через вектори $\vec{v}$ і $\vec{r}$ (рис.5.21). Напрямок вектора $\vec{B}$ для руху позитивно заряджених частинок найпростіше визначати правилом «лівої руки». Чотири пальці лівої руки розташувати по напрямку вектора $\vec{v}$, в долоню направити вектор $\vec{r}$ і великий палець покаже напрямок вектора $\vec{B}$.

Визначення напрямку $\vec{B}$ для негативно заряджених частинок здійснюється за правилом «правої руки».

§5.5.Циркуляція вектора індукції магнітного поля або закон повного струму

В електростатиці було встановлено, що робота при переміщенні одиничного пробного заряду в електричному полі нерухомого заряду не залежить від форми шляху. Тобто робота при переміщенні одиничного пробного заряду в електричному полі по довільному замкнутому контуру дорівнює нулю. Такі поля називають потенціальними.

Математично умова потенціальності поля записувалась у вигляді рівності (1.32). Ця умова вказувала на те, що силові лінії електростатичного поля незамкнені: початок вони беруть на позитивних зарядах і закінчуються на негативних або починаються на нескінченності і прямують у нескінченність.

До сьогоднішнього дня магнітних зарядів в природі не виявлено. Магнітні поля створюються рухомими електричними зарядами, тобто струмами. Спостереження спектрів магнітних полів (рис.5.14) показують, що лінії індукції магнітного поля замкненні. Поля, які мають замкнені силові лінії, називаються ***вихровими*** полями. Це відповідає тому, що в природі не має магнітних зарядів.

Обчислимо циркуляцію вектора $\vec{B}$ по замкнутому контуру Г. За визначенням циркуляція (формула 1.32) дорівнює інтегралу

$$\oint_{\Gamma} \vec{B}d\vec{l} \ . \qquad (5.20)$$

Найпростіше обчислити цей інтеграл у випадку поля прямого нескінченно довгого струму. Нехай замкнутий контур Г лежить в площині, яка перпендикулярна струму (рис.5.22); струм перпендикулярний до площини рисунка і направлений на нас. В кожній точці A контуру Г вектор $\vec{B}$, направлений по дотичній до кола, радіус якого $OA=r$. Коло лежить в площині рисунка, а центр O знаходиться на осі струму. Тому

$$\vec{B}d\vec{l} = Bdl\cos\alpha = Bdl\cos(\pi-\beta) = \\ = -Bdl\cos\beta = -Bdl_B$$

($dl_B = dl\cdot\cos\beta -$ проекція елемента $d\vec{l}$ на напрямок вектора $\vec{B}$). Із рисунка 5.22 видно, що $dl_B = rd\varphi$, де r – відстань від проводу зі струмом до dl, $d\varphi -$ кут, на який повертається радіальна пряма OA при переміщенні вздовж контуру на відрізок dl.

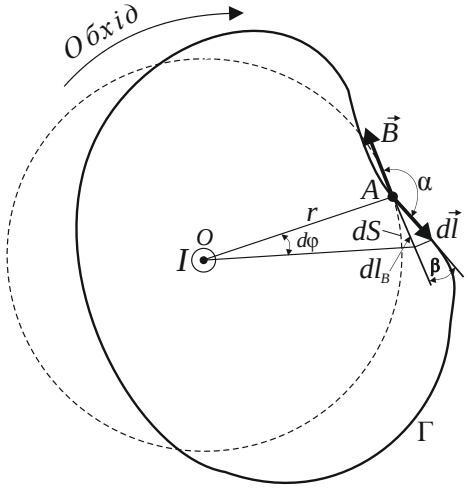


Рис.5.22.

Таким чином, із врахуванням (5.13), отримаємо:

$$\vec{B}d\vec{l} = -Bdl_B = -Brd\varphi = -\frac{\mu_0 I}{2\pi}d\varphi .$$

Тоді (5.20) запишеться так:

$$\oint_{\Gamma} \vec{B}\,d\vec{l} = -\frac{\mu_0}{2\pi}\oint_{\Gamma} d\varphi . \qquad (5.21)$$

При обході по контуру, який охоплює струм, радіальна пряма *OA* (рис.5.22) весь час повертається в одному напрямку, тому $\oint_\Gamma d\varphi = 2\pi$. Зовсім інша ситуація, коли контур Г не охоплює струм (рис.5.23). В цьому випадку радіальна пряма *OA'* повертається за обходом спочатку проти годинникової стрілки (ділянка 1-2), а потім в протилежному напрямку (ділянка 2-1, пряма *OA''*), внаслідок чого $\oint_\Gamma d\varphi = 0$. Враховуючи цей результат можемо (5.21) записати так:

$$\oint_\Gamma \vec{B}\, d\vec{l} = -\mu_0 I, \qquad (5.22)$$

де I струм, що охоплюється контуром Г.

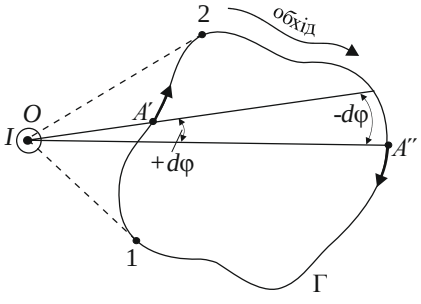


Рис.5.23.

Знак у виразі (5.22) залежить від напрямку обходу по контуру Г і напрямку струму. Якщо напрямок обходу утворює з напрямком струму лівогвинтову систему (як на рис.5.22), то $\oint_\Gamma \vec{B}d\vec{l} < 0$, в іншому випадку $\oint_\Gamma \vec{B}d\vec{l} > 0$. Знак циркуляції вектора $\vec{B}$ можна врахувати, вважаючи струм I алгебраїчною величиною. Струм I вважається позитивним, якщо його напрямок зв'язаний з обходом по контуру Г правилом правого свердлика; струм протилежного напрямку буде негативним.

Припустимо, що деякий контур Г охоплює декілька проводів зі струмом (рис.5.24). Згідно принципу суперпозиції (5.10)

$$\oint_\Gamma \vec{B}d\vec{l} = \oint_\Gamma (\sum_k \vec{B}_k)d\vec{l} = \sum_k \oint_\Gamma \vec{B}_k d\vec{l}\,.$$

Кожен із інтегралів в цій сумі дорівнює $\mu_0 I_k$, якщо струм I_k охоплюється контуром. Кожний струм враховується стільки разів, скільки раз він охоплюється контуром.

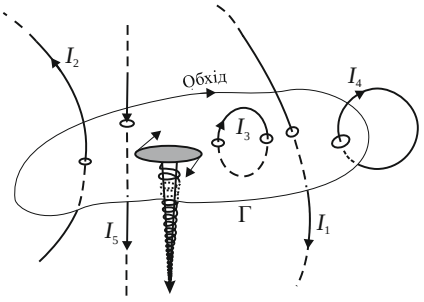


Рис.5.24.

Наприклад, на рис.5.25 струм охоплюється контуром два рази.

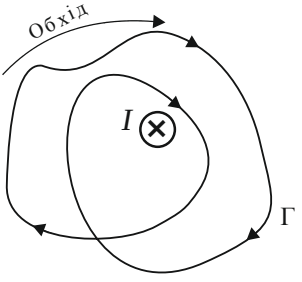


Рис.5.25.

Значить

$$\oint_\Gamma \vec{B}d\vec{l} = \mu_0 \sum_k I_k\,. \qquad (5.23)$$

(Нагадаємо, що I_k – алгебраїчна величина). Для струмів на рис.5.24 можемо записати, що

$$\sum_k I_k = I_1 - I_2 - I_4 + I_5\,.$$

Закон, який виражається рівністю (5.23) називається ***теоремою про циркуляцію***

вектора індукції магнітного поля, або законом повного струму для струмів провідності: циркуляція вектора $\vec{B}$ ***по довільному замкнутому контуру дорівнює добутку магнітної сталої на алгебраїчну суму струмів, що охоплюються цим контуром.***

Отриманий вираз закону ***повного струму справедливий лише*** для постійного магнітного поля у вакуумі.

§5.6.Теорема Остроградського-Гаусса для магнітного поля

Доведемо, що магнітний потік вектора індукції $\vec{B}$ через довільну замкнуту поверхню S дорівнює нулю. Потік вектора $\vec{B}$ через поверхню S визначається аналогічно як і потік вектора напруженості електричного поля $\vec{E}$ (формули (1.10) і (1.11). Для магнітних полів, які створюються рухомими зарядами, справедливий принцип суперпозиції. Крім того, потік геометричної суми декількох векторів $\vec{B}_k$ через любу поверхню дорівнює алгебраїчної сумі потоків цих векторів через ту ж поверхню. Тому при розрахунку магнітного потоку достатньо обмежитись окремим випадком, коли поле $\vec{B}$ створюється окремим точковим зарядом, який рівномірно рухається із швидкістю $\vec{\upsilon}$. Для простоти будемо вважати, що заряд рівномірно рухається перпендикулярно до площини рисунка 5.26. Згідно формули (5.19) і рис.5.21 магнітними силовими лініями будуть коаксіальні кола (мають спільний центр), площини яких паралельні площині рисунка, а центри розташовані на прямій, вздовж якої рухається заряд. Візьмемо нескінченно тонку кільцеву трубку *ABCD*, яка утворена магнітними силовими лініями. Через осьову симетрію магнітний потік через поперечний переріз трубки буде залишатись сталим по всій її довжині. Трубка перетне замкнену поверхню S парне число раз, наприклад два. Магнітні потоки через площадки $d\vec{S}_1$ і $d\vec{S}_2$, які трубка вирізає на поверхні S, однакові за величиною, але протилежні за знаком (нормаль $\vec{n}_1$ протилежна до напрямку індукції $\vec{B}$, а нормаль $\vec{n}_2$ співпадає з напрямком вектора $\vec{B}$). Сума потоків $d\Phi_1 = \vec{B}_1 d\vec{S}_1$ і $d\Phi_2 = \vec{B}_2 d\vec{S}_2$ через такі

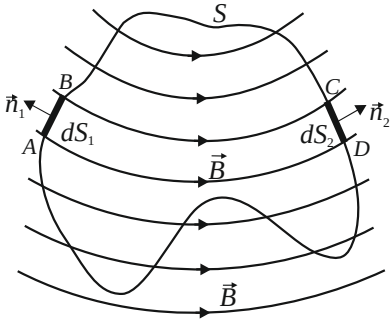


Рис.5.26.

площадки дорівнює нулю. Це справедливо і в тому випадку, коли трубка перетинає поверхню S довільне число раз, оскільки це число завжди парне. Але весь простір можна розбити на нескінченно тонкі кільцеві трубки, і кожна із них не буде вносити ніякого вкладу в магнітний потік через замкнену поверхню S. Сумарний магнітний потік через таку поверхню буде дорівнювати нулю. Таким чином, теорема доведена, тобто

$$\oint_S \vec{B} d\vec{S} = 0 . \qquad (5.24)$$

Рівняння (5.24) відображає той факт, що в природі не має джерел магнітного поля у вигляді магнітних зарядів, подібних до електричних, а силові лінії магнітного поля

постійних струмів замкнуті, і, отже, їх потік через довільну замкнуту поверхню дорівнює нулю. Таке поле називається ***вихровим*** або ***соленоїдальним***.

§5.7.Магнітне поле довгого соленоїда, тороїда та прямого проводу

Соленоїд. Якщо довжина котушки $L >> 2R$ (§5.4), де R – радіус перерізу котушки, то таку котушку можна вважати довгим (нескінченно довгим) соленоїдом. Магнітне поле $\vec{B}$ всередині такого соленоїда є однорідним (рис.5.18;5.19) і визначається за формулою (5.18). Магнітне поле $\vec{B}'$ зовні соленоїда теж будемо вважати однорідним. Отримаємо формулу (5.18) за допомогою закону повного струму (5.23). Спочатку покажемо, що для довгого соленоїда індукція магнітного поля зовні соленоїда дорівнює нулю. Для цього візьмемо площину, перпендикулярну до осі соленоїда (рис.5.27).

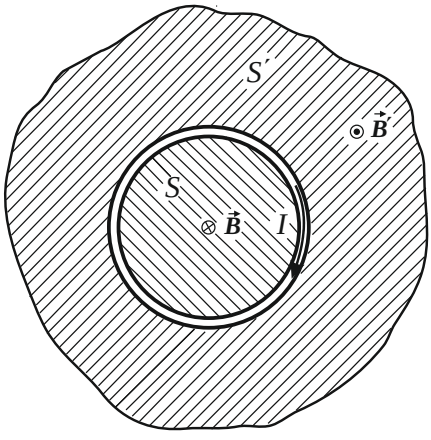


Рис.5.27.

Внаслідок замкнутості силових ліній вектора $\vec{B}$ магнітні потоки через внутрішню частину S цієї площини і через зовнішню частину S' повинні бути однаковими за величиною і протилежними за знаком, щоб виконувалось теорема (5.24).

Так як поля однорідні і вектори $\vec{B}$ і $\vec{B}'$ перпендикулярні до площини, то кожен із потоків дорівнює добутку BS і $B'S'$. Таким чином, отримуємо співвідношення

$$BS = B'S'.$$

Ліва частина цієї рівності кінцева величина. В правій частині цієї\ рівності площина S' може бути нескінченно великою, що можливо, коли $B' = 0$. Так ми довели, що зовні нескінченно довгого соленоїда магнітна індукція дорівнює нулю.

Розглянемо соленоїд в площині, яка містить вісь соленоїда (рис.5.28). Для використання закону (5.23) виберемо замкнутий контур Г 1-2-3-4-1, як показано на рис.5.28.

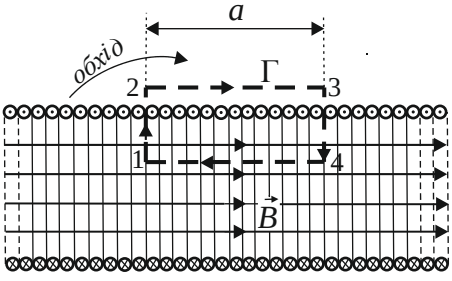


Рис.5.28.

Ширина контура дорівнює a, обхід по контуру за годинниковою стрілкою. Тоді циркуляція вектора $\vec{B}$ по такому контуру буде дорівнювати $-Ba$, а алгебраїчна сума струмів, що охоплюється цим контуром, буде дорівнювати $-naI$, де n – густина намотки соленоїда. Згідно (5.23) отримаємо рівність:

$$-Ba = -\mu_0 nI .$$

Звідки

$$B = \mu_0 nI , \qquad (5.18')$$

що співпадає з формулою (5.18).

Тороїдом називають кільцеву котушку, яку можна отримати із соленоїда зігнувши його так, як показано на рис.5.29. Лінії магнітної індукції поля тороїда повинні мати форму кіл, центри яких лежать на прямій, що проходить через центр тороїда *O* і перпендикулярна до площини рисунка. За законом повного струму (5.23) циркуляція вектора $\vec{B}$ вздовж кола, радіус якого r (рис.5.29), буде дорівнювати:

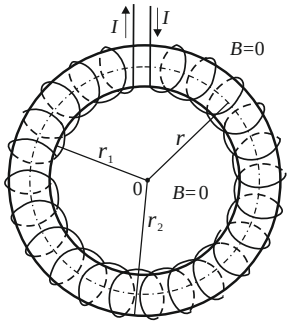


Рис.5.29.

$$\oint_{\Gamma} \vec{B}d\vec{l} = \oint_{\Gamma} Bdl = B\oint_{\Gamma} dl = B2\pi r\,. \quad (5.25)$$

Позначимо кількість витків обмотки тороїда N, а струм в обмотці – I. Якщо $r < r_1$, то коло радіусом r не охоплює провідників зі струмом і у формулі (5.23) $\sum_k I_k = 0$. Це означає, що індукція $B = 0$. Якщо $r > r_2$, то коло радіусом r охоплює $2N$ провідників зі струмом I. В N із них струм протікає в одному напрямку, а в інших N провідниках – у зворотньому напрямку. Тому знову $\sum_k I_k = 0$ і *B*=0. Отже поза тороїдом магнітного поля немає.

Тобто магнітне поле тороїда локалізується всередині об'єму тороїда (об'єму, який охоплюється витками тороїда і для якого $r_1 < r < r_2$). Коло радіусом r, яке лежить всередині тороїда, охоплює N провідників, струми в яких дорівнюють I і однаково напрямлені. Тому $\sum_k I_k = NI$.

Згідно (5.23) і (5.25) запишемо, що $2\pi rB = \mu_0 NI$. Звідки отримаємо:

$$B = \mu_0 \frac{NI}{2\pi r}. \quad (5.26)$$

Магнітна індукція поля всередині тороїда зменшується із збільшенням відстані від його центра *O*:

$$B_{max} = \mu_0 \frac{NI}{2\pi r_1},\ B_{min} = \mu_0 \frac{NI}{2\pi r_2}.$$

Індукція на основній лінії тороїда $r = (r_1 + r_2)/2$ дорівнює:

$$B = \mu_0 \frac{N}{2\pi r} I = \mu_0 nI, \quad (5.27)$$

де n – кількість витків на одиницю довжини середньої лінії тороїда. Бачимо, що формули для індукції поля для нескінченно довгого соленоїда (5.18), (5.18') і для тороїда формула (5.27) однакові. Тому дуже часто використовується таке поняття як тонкий тороїд. Тороїд вважається тонким, коли діаметр його перерізу $2(r_2 - r_1)$ є набагато менший довжини середньої лінії $2\pi r$. Для такого тороїда, як і для нескінченно довгого соленоїда, магнітне поле всередині витків вважається однорідним, а індукція розраховується за формулою (5.27).

Прямий провід. Розглянемо розрахунок магнітного поля довгого прямого проводу в точці, яка лежить поза проводом на відстані r від його осі (рис.5.30). Використаємо для обчислення закон повного струму (5.23). За контур Г зручно взяти силову лінію (коло), радіус якої r. Площина кола перпендикулярна до струму і має центр на осі струму. Оскільки $\vec{B}$

і $d\vec{l}$ мають однаковий напрямок і $|\vec{B}|$ є однаковою в кожній точці контуру Г, то

$$\oint_{\Gamma} \vec{B}d\vec{l} = \oint_{\Gamma} Bdl = B\oint_{\Gamma} dl = B2\pi r \cdot$$

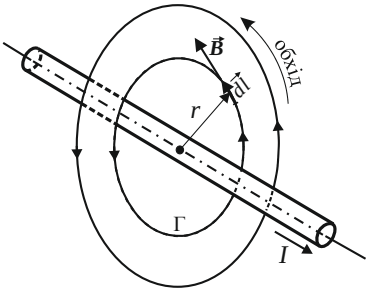


Рис.5.30.

Тоді згідно (5.23)

$$2\pi rB = \mu_0 I .$$

Звідки

$$B = \mu_0 \frac{I}{2\pi r} . \qquad (5.28)$$

Цей результат ми отримали вже раніше (див.формулу 5.13). Бачимо, що розрахунок за допомогою закону (5.23) є простіший, ніж безпосереднє сумування полів окремих елементів струму. Обчислимо тепер індукцію поля в якій небудь точці *A* ***всередині*** проводу, яка лежить на відстані r від осі проводу (рис.5.31,*а*). Замкнутий контур Г виберемо знову у вигляді кола, що проходить через точку *A*, а центр лежить на осі проводу. Тоді за формулою (5.23), як і в попередньому випадку можемо записати, що

$$\oint_{\Gamma} \vec{B}d\vec{l} = 2\pi rB = \mu_0 \sum_k I_k = \mu_0 \pi r^2 j ,$$

де $j = I/(\pi R^2)$ – густина струму (стала у всіх точках провідника). Звідси отримаємо, що

$$B = \frac{\mu_0 I}{2\pi R^2} r. \qquad (5.29)$$

Тут I – повна величина сили струму через весь переріз проводу , R – радіус проводу.

Отже індукція магнітного поля всередині проводу збільшується з відстанню від осі за лінійним законом (формула (5.29)), а в зовнішньому просторі зменшується за гіперболічним законом (формула (5.28)). Цю залежність зображено графічно на рис.5.31,*б*.

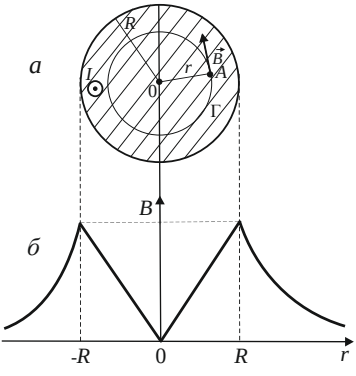


Рис.5.31.

§5.8.Контур зі струмом у магнітному полі

З'ясуємо, як поводить себе контур Г зі струмом в магнітному полі. Розглянемо випадок, коли поле однорідне ($\vec{B} = \text{const}$) (рис.5.32). Нехай контур Г орієнтований так, що позитивна нормаль $\vec{n}$ перпендикулярна до вектора $\vec{B}$ (рис.5.32). Розіб'ємо площу контура на вузькі паралельні вектору $\vec{B}$ смужки, ширина яких dz. Зліва і справа смужка обмежена елементами контуру $d\vec{l}_1$ і $d\vec{l}_2$. Згідно закону Ампера (5.7) на елемент струму $Id\vec{l}_1$ буде діяти сила $d\vec{F}_1 = Id\vec{l}_1 \times \vec{B}$, а на $Id\vec{l}_2$ –сила $d\vec{F}_2 = Id\vec{l}_2 \times \vec{B}$. Сила $d\vec{F}_1$ направлена за рисунок (від нас). Модуль цієї сили дорівнює $dF_1 = IBdl_1 \sin\alpha_1 = IBdz.$. На елемент струму $Id\vec{l}_2$ діє сила $d\vec{F}_2$, яка направлена на нас. Модуль цієї сили дорівнює:

$$dF_2 = IBdl_2 \sin\alpha_2 = IBdl_2 \sin(\frac{\pi}{2}+\beta) =$$

$$= IBdl_2 \cos\beta = IBdz.$$

Тому $d\vec{F}_1 = -d\vec{F}_2$.

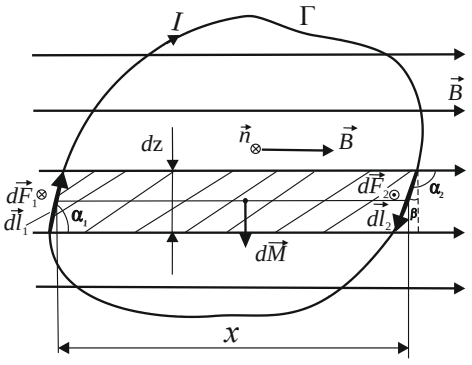


Рис.5.32.

Отриманий нами результат означає, що сили, які прикладені до протилежних елементів контуру $d\vec{l}_1$ і $d\vec{l}_2$ утворюють пару, обертальний момент якої дорівнює

$$dM = F_1 \cdot x = IBxdz = IBdS,$$

де dS – площа смужки , x – середня лінія, а dz – висота трапеції (смужки).

Якщо обертати свердлик за напрямками дії сил $d\vec{F}_1$ і $d\vec{F}_2$, то він буде рухатись в напрямку вектора $d\vec{M}$. Це означає, що вектор $d\vec{M}$ буде перпендикулярний до векторів $\vec{n}$ і $\vec{B}$. Тобто можемо записати, що

$$d\vec{M} = I\vec{n} \times \vec{B}dS .$$

Просумуємо цей вираз по всіх смужках і отримаємо обертальний момент, який діє на контур:

$$\vec{M} = \int_S I\,\vec{n} \times \vec{B}dS = I\,\vec{n} \times \vec{B}\int_S dS = I\,\vec{n} \times \vec{B}S . \quad (5.30)$$

При отриманні формули (5.30) враховано, що поле однорідне, тому добуток $\vec{n} \times \vec{B}$ для всіх смужок є однаковим і може бути винесений за знак інтеграла. Величина S у виразі (5.30) є площа контуру Г.

Вираз (5.30) можна представити ще і у такому виді

$$\vec{M} = IS\vec{n} \times \vec{B}.$$

Якщо врахувати, що $\vec{p}_m = IS\vec{n}$ (§5.3) є дипольний магнітний момент контуру зі струмом, то

$$\vec{M} = \vec{p}_m \times \vec{B}. \quad (5.31)$$

Модуль вектора $\vec{M}$ дорівнює

$$M = p_m B \sin\alpha, \quad (5.32)$$

де α – кут між векторами $\vec{p}_m$ і $\vec{B}$.

Із механіки відомо, щоб кут α збільшити на $d\alpha$ необхідно виконати роботу проти сил, які діють на контур в магнітному полі,

$$dA = Md\alpha = p_m B \sin\alpha d\alpha . \quad (5.33)$$

Повертаючись в початкове положення, контур може повернути затрачену на його поворот механічну роботу, виконавши її над яким-небудь тілом. Значить, робота (5.33) йде на збільшення потенціальної енергії $W_{p_{мех}}$, яку має контур зі струмом в магнітному полі,

$$dW_{p_{мех}} = p_m B \sin\alpha d\alpha.$$

Інтегруючи цей вираз, знаходимо, що

$$W_{p_{мех}} = \int dW_{p_{мех}} = p_m B \int \sin\alpha d\alpha$$

$$= -\,p_m B \cos\alpha + \text{const}.$$

Якщо прийняти, що const = 0, то

$$W_{p_{мех}} = -\,p_m B \cos\alpha = -\vec{p}_m \vec{B}. \quad (5.34)$$

Якщо магнітний момент $\vec{p}_m$ контуру паралельний до напрямку зовнішнього поля $\vec{B}$ $(\alpha = 0)$, то $M = 0$, а $W_{p_{min}} = -p_m B$ і це положення контуру відповідає стійкій рівновазі. Якщо $\vec{p}_m$ антипаралельний $\vec{B}(\alpha = \pi)$, то

$M = 0$, а $W_{p\,max} = p_m B$ і контур Г перебуває в нестійкій рівновазі.

Розглянемо тепер виток зі струмом у неоднорідному магнітному полі (рис.5.33). При цьому вважатимемо для простоти, що магнітне поле в центрі витка перпендикулярне до площини витка. Сили $d\vec{F}$, які діють на окремі ділянки витка, будуть, як і раніше, перпендикулярні до струму і до магнітного поля.

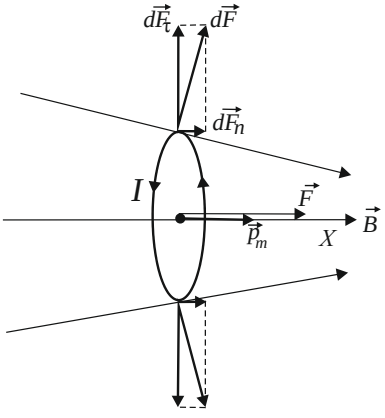


Рис.5.33.

Однак тепер, через те що силові лінії вже не паралельні, ці сили утворюватимуть деякий кут з площиною витка. Складові цих сил $d\vec{F}_\tau$, паралельні витку, створять зусилля, які розтягують або стискають виток. А складові $d\vec{F}_n$, які перпендикулярні до площини витка, в сумі дадуть результуючу силу $\vec{F}$, яка прагне переміщати виток у магнітному полі.

Застосовуючи правило лівої руки, легко бачити, що коли момент струму $\vec{p}_m$ паралельний магнітному полю (як зображено на рисунку), то виток втягуватиметься в область сильнішого поля. А коли вектор $\vec{p}_m$ антипаралельний полю, то виток виштовхуватиметься і переміщатиметься в область слабшого поля.

Легко зрозуміти, що в загальному випадку неоднорідного поля, індукція $\vec{B}$ якого не перпендикулярна до площини витка, діятимуть і пара сил, яка прагне повернути виток, і сила, яка спричинює поступальне переміщення витка.

З допомогою виразу (5.34) для енергії контуру в магнітному полі легко знайти вираз для сили $\vec{F}$. Якщо орієнтація магнітного моменту $\vec{p}_m$ по відношенню до поля (вектора $\vec{B}$) залишається незмінною $\alpha = \text{const}$, то $W_{p_{мех}}$ буде залежати тільки від x (через $B(x)$). Із механіки відомо, що

$$F_x = -\frac{\partial W_{p_{мех}}}{\partial x} = p_m \frac{\partial B}{\partial x} \cos\alpha.$$

Якщо вважати, що в інших напрямках поле змінюється слабо, то проекціями сили на інші осі можна знехтувати і вважати, що $F = F_x$.

Тому

$$F = p_m \frac{\partial B}{\partial x} \cos\alpha. \qquad (5.35)$$

Формула (5.35) повністю відображає поведінку контуру зі струмом в магнітному полі. Наприклад: Якщо $\vec{p}_m$ і $\vec{B}$ мають один напрямок ($\alpha = 0$), сила $F>0$, тобто направлена в сторону зростання $\vec{B}$ ($\frac{\partial B}{\partial x} > 0$); для випадку $\frac{\partial B}{\partial x} < 0$ знак і напрямок сили зміниться на протилежний, але сила все одно буде втягувати контур в область сильного поля. Якщо $\vec{p}_m$ і $\vec{B}$ антипаралельні ($\alpha = \pi$), сила негативна, тобто направлена в сторону зменшення $\vec{B}$.

Підсумовуючи сказане вище, ми бачимо, що й сили, які діють на контур зі струмом у магнітному полі, і магнітне поле, яке

створюється контурами, визначаються їх магнітними моментами.

Виникнення пари сил, яка діє на рамку зі струмом у магнітному полі, використовують для побудови дуже поширеного типу електровимірювальних приладів – ***магнітоелектричних приладів.*** Вони мають легку рамку з дротиною, закріплену на наріжному підвісі тієї чи іншої конструкції, здатну повертатися в магнітному полі. Зрозуміло, що повертання рамки зі струмом в магнітному полі, знайшло своє використання в побудові електродвигунів.

§5.9.Робота при переміщенні провідника зі струмом в магнітному полі

Розглянемо контур зі струмом , який утворений нерухомими проводами і перемичкою *KL=l*, яка може ковзати по них (рис.5.34). Нехай контур знаходиться у зовнішньому магнітному полі, яке будемо

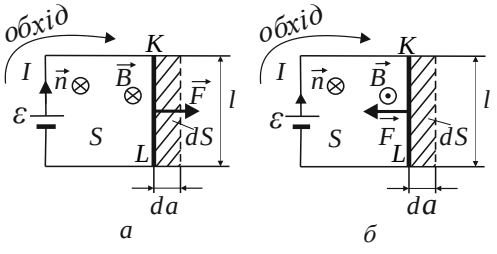


Рис.5.34.

вважати однорідним і перпендикулярним до площини контуру. На перемичку *KL* зі струмом в магнітному полі буде діяти сила Ампера (5.7), яка для нашого випадку, показаного на рис.5.34, дорівнює:

$$F = IlB.$$

При переміщенні перемички вправо на відстань da ця сила виконує позитивну роботу

$$dA = Fda = IBlda = IBdS, \quad (5.36)$$

де $dS = lda -$ заштрихована площа на рис.5.34,*а*.

З'ясуємо, як змінюється при переміщенні перемички *KL* потік магнітної індукції через площу контуру. У загальному випадку потік вектора індукції через площу контура зі струмом дорівнює

$$\Phi = \int \vec{B}\,\vec{n}dS,$$

де $\vec{n}$ – позитивна нормаль, тобто нормаль, яка утворює з напрямком струму в контурі (з обходом по контуру) правогвинтову систему (§5.3). Тоді у випадку, показаному на рис. 5.34,*а*, потік буде позитивним і дорівнювати BS ($S -$ площа контуру. При переміщенні перемички вправо площа контуру отримає позитивний приріст dS. В результаті потік також отримає позитивний приріст $d\Phi = BdS$. Тому вираз (5.36) можна представити у вигляді

$$dA = Id\Phi. \quad (5.37)$$

При напрямку магнітного поля на нас (рис.5.34,*б*) сила, яка діє на перемичку, буде направлена вліво. Тому при переміщенні перемички вправо на da магнітна сила виконує негативну роботу

$$dA = -IBlda = -IBdS. \quad (5.38)$$

В цьому випадку потік через контур дорівнює $\vec{B}\vec{n}S = BS\cos\pi = -BS$. При зміні площі контуру на dS потік отримує приріст $d\Phi = \vec{B}\vec{n}dS = BdS\cos\pi = -BdS$. Значить, вираз (5.38) також можна записати у вигляді (5.37).

Величину $d\Phi$ у виразі (5.37) можна трактувати як потік через площу, яку описує перемичка при своєму русі. Тому можна сказати, що робота, яку виконує магнітна сила над ділянкою контуру зі струмом, дорівнює добутку сили струму на величину магнітного

потоку через поверхню, яку описує ця ділянка при своєму русі.

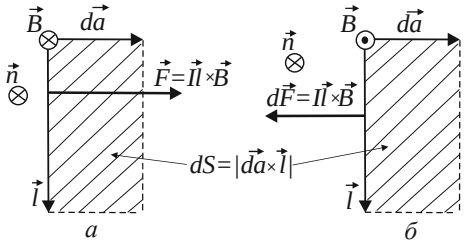


Рис.5.35.

Формули (5.36) і (5.38) можна об'єднати в один векторний вираз. Представимо перемичку як вектор $\vec{l}$, що має напрямок струму (рис.5.35). Незалежно від напрямку вектора $\vec{B}$ (від нас або на нас) силу Ампера, яка діє на перемичку, можна представити у вигляді $\vec{F} = I\vec{l} \times \vec{B}$.

При переміщенні перемички на $d\vec{a}$ сила виконує роботу

$$dA = \vec{F}d\vec{a} = I\vec{l} \times \vec{B}d\vec{a}.$$

Виконаємо в змішаному добутку векторів циклічну перестановку множників. В результаті отримаємо , що

$$dA = I\vec{B}d\vec{a} \times \vec{l}. \qquad (5.39)$$

Із рис.5.35 видно, що векторний добуток $d\vec{a} \times \vec{l}$ дорівнює за величиною площадці dS, яку описує перемичка при її русі, і має напрямок позитивної нормалі $\vec{n}$. Тобто $d\vec{a} \times \vec{l} = \vec{n}dS$. Тому (5.39) перепишеться так:

$$dA = I\vec{B}\vec{n}dS. \qquad (5.40)$$

У випадку, що показаний на рис.5.35,*а*, $\vec{B}\vec{n} = B,$ і ми приходимо до формули (5.36). У випадку, показаному на рис.5.35,*б*, $\vec{B}\vec{n} = -B$ і ми приходимо до формули (5.38).

Вираз $\vec{B}\vec{n}dS$ визначає приріст магнітного потоку через контур, обумовлений переміщенням перемички. Таким чином формулу (5.40) можна записати у вигляді (5.37). Однак формула (5.40) має перевагу перед (5.37), так як із неї «автоматично» отримується знак $d\Phi$, а значить і знак dA.

Розглянемо жорсткий довільної форми контур Г зі струмом, що знаходиться в довільному магнітному полі. Знайдемо роботу, яка виконується при довільному нескінченно малому переміщенні контуру. Нехай елемент контуру $d\vec{l}$, переміщається на $d\vec{a}$ (рис.5.36).

При цьому над ним магнітна сила виконує роботу

$$dA_{ел} = Id\vec{l} \times \vec{B}d\vec{a}. \qquad (5.41)$$

Тут $\vec{B}$ – магнітна індукція в тому місці, де знаходиться елемент контуру $d\vec{l}$.

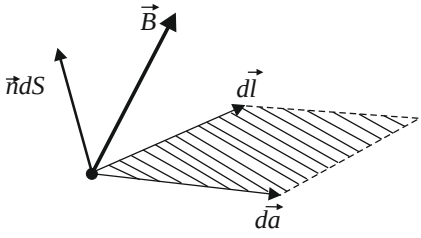


Рис.5.36.

Здійснивши в (5.41) циклічну перестановку множників, отримаємо

$$dA_{ел} = I\vec{B}d\vec{a} \times d\vec{l}. \qquad (5.42)$$

Векторний добуток $d\vec{a} \times d\vec{l}$ дорівнює по модулю площі паралелограма, побудованого на векторах $d\vec{a}$ і $d\vec{l}$, тобто площі dS, яку описує елемент при його переміщенні на $d\vec{a}$. Тому

$$\vec{B}d\vec{a} \times d\vec{l} = \vec{B}\vec{n}dS = d\Phi_{ел}, \qquad (5.43)$$

де $\vec{n}$ – позитивна нормаль до площадки dS, $d\Phi_{ел}$ – приріст магнітного потоку через контур, обумовлений переміщенням елемента контура

$d\vec{l}$. Прийнявши до уваги рівність (5.43), напишемо (5.42) у вигляді

$$dA_{ел} = I\, d\Phi_{ел}. \qquad (5.44)$$

Просумувавши вираз (5.44) по всіх елементах контуру, отримаємо вираз для роботи магнітних сил при довільному нескінченно малому переміщенні контуру Г:

$$dA = \oint_{\Gamma} dA_{ел} = \oint_{\Gamma} I d\Phi_{ел} = I \oint_{\Gamma} d\Phi_{ел} = I d\Phi, \qquad (5.45)$$

де $d\Phi$– повний приріст потоку через весь контур Г при його довільному нескінченно малому переміщенні.

Щоб знайти роботу, яка виконується при кінцевому довільному перемещенні контуру з початкового положення в кінцеве, проінтегруємо вираз (5.45):

$$A_{12} = \int_{\Phi_1}^{\Phi_2} dA = I \int_{\Phi_1}^{\Phi_2} d\Phi = I(\Phi_2 - \Phi_1), \qquad (5.46)$$

де Φ_1 і Φ_2– значення магнітного потоку через контур Г в початковому і кінцевому положеннях. Таким чином, робота, яку виконують магнітні сили над контуром, дорівнює добутку сили струму в контурі на приріст магнітного потоку через контур.

Відмітимо, що робота (5.46) виконується не за рахунок енергії зовнішнього магнітного поля, а за рахунок джерела, яке підтримує незмінним струм в контурі.

§5.10.Дія електричного і магнітного полів на рухомий електричний заряд

1.*Сила Лоренца.* Виникнення сили Ампера (5.7), що діє на елемент провідника зі струмом у магнітному полі, можна пояснити так. При проходженні струму носії заряду в провіднику рухаються напрямлено. Тому магнітне поле відхиляє їх в один бік. При цьому вони стикаються з кристалічною решіткою металу і передають їй певний імпульс, якого набули під дією магнітного поля. Макроскопічним результатом елементарних процесів зіткнення окремих носіїв заряду з кристалічною граткою провідника є виникнення сили Ампера.

Магнітне поле діє на вільні електрони в металі і за відсутності електричною струму в провіднику. Та електрони в цьому випадку рухаються тільки хаотично і сумарний імпульс, який вони надають кристалічній решітці провідника, дорівнює нулю і провідник залишається нерухомим.

Для обчислення сили, що діє на рухомий заряд в магнітному полі, використаємо формулу (5.7) для випадку магнітного поля з індукцією $\vec{B}$:

$$d\vec{F} = Id\vec{l} \times \vec{B}.$$

Елемент струму, як і в §5.4, можна представити так:

$$Id\vec{l} = e\vec{v}_{др} dN.$$

Оскільки електрони рухаються направлено, створюючи струм, то сила, яка припадає на один електрон дорівнює:

$$\vec{F}_л = \frac{d\vec{F}}{dN} = e\vec{v}_{др} \times \vec{B}.$$

У загальному випадку

$$\vec{F}_л = \frac{d\vec{F}}{dN} = e\vec{v} \times \vec{B}. \qquad (5.47)$$

Напрямок сили Лоренца визначається за правилом векторного добутку або правилом лівої руки: якщо долоню лівої руки розмістити так, щоб в неї входив вектор $\vec{B}$, чотири витягнуті пальці спрямувати вздовж вектора швидкості $\vec{v}$ руху позитивних зарядів, то відігнутий на 90° великий палець покаже напрямок сили, що діє на позитивний заряд. На негативний заряд сила діє в протилежному напрямку (рис.5.37).

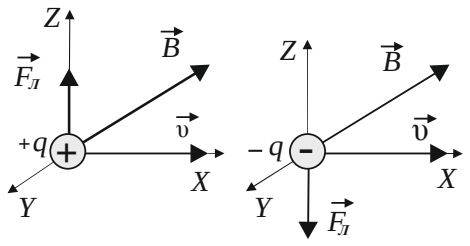


Рис.5.37.

Отже, магнітне поле не діє на електричні заряди, які не рухаються.

Сила Лоренца завжди перпендикулярна до швидкості руху зарядженої частинки, тому вона змінює лише напрямок цієї швидкості, не змінюючи її модуля. Тому сила Лоренца роботи не виконує і кінетична енергія частинки при русі в магнітному полі не змінюється.

Якщо на рухомий електричний заряд, крім магнітного поля з індукцією $\vec{B}$, діє і електричне поле з напруженістю $\vec{E}$, то результуюча сила $\vec{F}$, яка прикладена до заряду:

$$\vec{F} = q\vec{E} + q\vec{\upsilon} \times \vec{B}. \qquad (5.48)$$

Це формула Лоренца.

Якщо заряджена частинка рухається в магнітному полі зі швидкістю $\vec{\upsilon}$ вздовж ліній магнітної індукції або в протилежний бік до напрямку магнітної індукції, то $\alpha = 0$, або $\alpha = \pi$. У такому разі $F_л = 0$, магнітне поле на частинку не діє і вона рухається рівномірно і прямолінійно.

Якщо заряджена частинка рухається в магнітному полі зі швидкістю $\vec{\upsilon}$ перпендикулярно до вектора $\vec{B}$, то сила Лоренца є стала за модулем і нормальна до траєкторії частинки. Частинка рухатиметься по колу, бо сила Лоренца за другим законом Ньютона буде створювати доцентрове прискорення. Отже,

$$q\upsilon B = \frac{m\upsilon^2}{r}. \text{ Звідси } r = \frac{m}{q}\frac{\upsilon}{B},$$

де r – радіус кола.

Використавши зв'язок $\upsilon = \omega r$, знайдемо циклічну частоту ω та період обертання T частинки навколо ліній індукції в магнітному полі:

$$\omega = \frac{\upsilon}{r} = \frac{q}{m}B, T = \frac{2\pi r}{\upsilon} = \frac{2\pi m}{Bq}. \qquad (5.49)$$

Період обертання частинки в однорідному магнітному полі не залежить від її швидкості (при $\upsilon << c$, де c – швидкість світла у вакуумі). На цьому ґрунтується дія циклічних прискорювачів заряджених частинок. Якщо швидкість зарядженої частинки $\vec{\upsilon}$ напрямлена під кутом α до вектора $\vec{B}$ (рис.5.38), то її рух можна подати у вигляді суперпозиції:

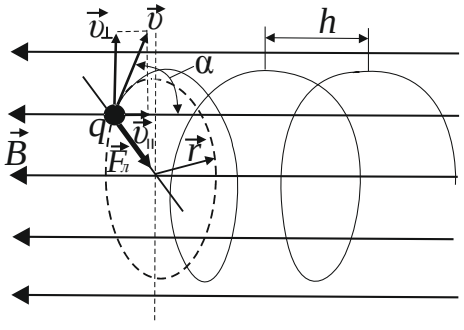


Рис.5.38.

1)рівномірного прямолінійного руху вздовж поля зі швидкістю $\upsilon_{||} = \upsilon\cos\alpha$;

2)рівномірного руху зі швидкістю $\upsilon_{\perp} = \upsilon\sin\alpha$ вздовж кола, площина якого перпендикулярна до поля $\vec{B}$. Радіус кола $r = \frac{m\upsilon\sin\alpha}{qB}$. В результаті складання обох

рухів виникає рух вздовж спіралі, вісь якої паралельна до магнітного поля. Крок гвинтової лінії

$$h = \upsilon_{||} T = \upsilon T \cos\alpha = \frac{2\pi m \upsilon \cos\alpha}{qB}.$$

Напрямок, в якому закручується спіраль залежить від знака заряду частинки.

2. Ефект Холла

В 1879 р. Е.Холл здійснив наступний експеримент. Він пропускав електричний струм I через золоту пластинку у вигляді паралелепіпеда і вимірював різницю потенціалів $\Delta\varphi$ між точками C і D на верхній і нижній гранях (рис.5.39).

Ці точки лежать у одному і тому поперечному перерізі пластинки. Тому виявилось, що $\Delta\varphi = 0$. Коли пластинку зі струмом Холл помістив в однорідне магнітне поле, лінії магнітної індукції якого перпендикулярні до бічних граней пластинки, то було вставлено, що різниця потенціалів $\Delta\varphi_H = \varphi_C - \varphi_D \neq 0$ і вона прямо пропорційна

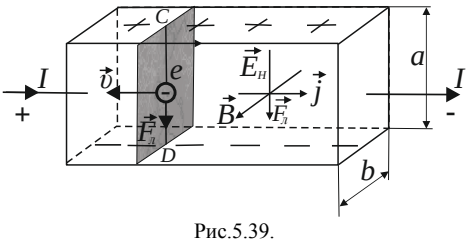


Рис.5.39.

силі струму *I*, магнітній індукції поля *B* і обернено пропорційна ширині пластинки, тобто

$$\Delta\varphi_H = \varphi_C - \varphi_D = R_H \frac{IB}{b}, \qquad (5.50)$$

де R_H – коефіцієнт пропорційності названий ***сталою Холла***. А явище, яке експериментально виявив Холл, дістало назву ***ефекту Холла*** .

Наступні дослідження показали, що ефект Холла спостерігається в усіх провідниках і напівпровідниках. Зміна напрямку струму або напрямку магнітного поля на протилежний викликає зміну знаку різниці потенціалів $\Delta\varphi$.

Розглянемо, як можна пояснити ефект Холла. Помістимо металеву пластинку зі струмом густиною $\vec{j}$ в магнітне поле з індукцію $\vec{B}$, лінії індукції якого перпендикулярні до $\vec{j}$ (рис.5.39). На електрони діє сила Лоренца, яка направлена вниз і дорівнює $F_л = Bev$. Тому на нижній грані пластинки збиратиметься нескомпенсований негативний заряд і вона заряджатиметься негативно, а на верхній грані пластинки виникатиме нестача негативних зарядів і вона заряджатиметься позитивно. Внаслідок цього між гранями пластинки виникає додаткове поперечне електричне поле, направлене згори вниз. Розглянемо момент динамічної рівноваги, коли сили $F_л$ і F_H з якими діють на електрони магнітне поле та холлівське електричне поле, стануть рівними (ця рівність настає вже через 10^{-12} с після замикання кола):

$$F_л = F_H, \quad Bev = eE_H.$$

Звідси

$$E_H = \upsilon B.$$

Невідому швидкість υ напрямленого руху виразимо через густину струму j:

$$j = en\upsilon, \upsilon = j/(e\,n),$$

де $n-$ концентрація електронів в металевій пластинці, $\upsilon -$ їх швидкість направленого руху.

Тоді

$$E_H = \frac{Bj}{ne}.$$

Помножимо ліву частину виразу на ab, а праву на $S = ab$:

$$E_H ab = \frac{Bj}{ne} S.$$

Оскільки $E_H a = \Delta\varphi_H$, а $jS = I$, то

$$\Delta\varphi_H = \frac{1}{en}\frac{BI}{b}. \qquad (5.51)$$

Порівнюючи вираз (5.51) з виразом (5.50), який отриманий на основі експерименту, отримуємо, що стала Холла R_H обернено пропорційна до добутку заряду електрона на їх концентрацію n:

$$R_H = \frac{1}{ne}. \qquad (5.52)$$

За виміряними значеннями сталої Холла можна:

1) *визначити концентрацію носіїв струму,* якщо характер провідності і заряд q носіїв струму відомі, а саме

$$n = 1/(qR_H).$$

Так, для одновалентних металів виявилося, що концентрація електронів провідності збігається з концентрацією атомів.

Оскільки питома електропровідність речовини обчислюють за формулою (§4.2) $\sigma = neu$, де u – рухливість носіїв заряду, то $\sigma = \frac{1}{R_H} u$ і $u = R_H \sigma$, тобто добуток $R_H \sigma$, визначає рухливість носіїв заряду.

2) *зробити висновок про природу провідності,* оскільки знак сталої Холла збігається із знаком заряду q.

3.Визначення питомого заряду електрона e/m_e методом магнетрона

Для експериментального визначення питомого заряду методом магнетрона використовують електронні лампи спеціальної конструкції. Особливістю цих ламп є те, що

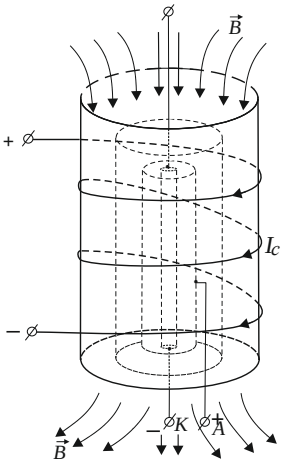


Рис.5.40.

катод і анод у них виготовлені у вигляді двох співвісних циліндричних поверхонь. Лампу (магнетрон) поміщають всередину соленоїда так, щоб його магнітне поле було напрямлене вздовж осі лампи (рис.5.40). Розглянемо характер руху електронів у лампі. Якщо струму в соленоїді немає $(\vec{B} = 0)$, електрони рухаються від катода до анода в радіальному напрямі вздовж силових ліній електричного поля між катодом і анодом (рис.5.41,*а*). При заданому значенні анодної напруги між анодом

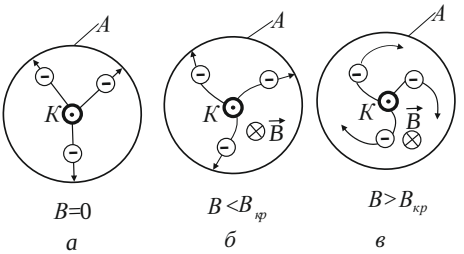


Рис.5.41.

і катодом установлюється анодний струм I_a=const. Пропустимо струм, сила якого I_c, витками соленоїда і створимо, таким чином, однорідне магнітне поле вздовж катода. На рухомі електрони діятиме сила Лоренца $F_л$, яка перпендикулярна до напрямку швидкості руху електронів. Траєкторія електронів

викривляється (рис.5.41,*б*), але всі вони будуть потрапляти на анод і, отже, анодний струм I_a не змінюється. З ростом I_c збільшується $\vec{B}$ і кривизна траєкторії руху електронів також зростатиме. При певному, критичному, значенні індукції поля електрони почнуть рухатись по колах (рис.5.41,*в*) і не будуть потрапляти на анод. Анодний струм I_a різко спадає.

Критичні значення сили струму соленоїда $I_{кр}$, при яких різко зменшується анодний струм, визначають за так званими скидовими характеристиками лампи, під якими розуміють графіки залежності анодного струму I_a від струму в соленоїді I_c при заданих значеннях анодної напруги (рис.5.42). При цьому $B_{кр} = \mu_0 n I_{кр}$, де $n=N/l = 1/d$ – густина намотки (N–загальна кількість витків, l–довжина, d – діаметр дроту соленоїда).

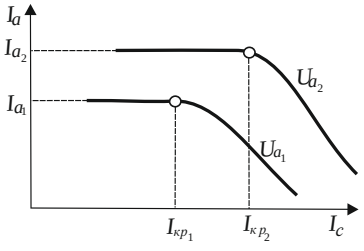


Рис.5.42.

Робочу формулу для визначення e/m_e виводять, виходячи з диференціального рівняння руху електрона у схрещених електричному полі $\vec{E}$ між анодом і катодом і магнітному полі $\vec{B}$ соленоїда:

$$m_e \frac{d\vec{v}}{dt} = e\vec{E} + e\vec{v} \times \vec{B}.$$

Розв'язок цього рівняння для $B = B_{кр}$ дає, що

$$\frac{e}{m_e} = \frac{8U_a}{B_{кр}^2 R^2 (1-\frac{r^2}{R^2})^2}, \qquad (5.53)$$

де R і r радіуси відповідно анода і катода. Питомий заряд електрона, за сучасними вимірюваннями, становить $1{,}76 \cdot 10^{11}$Кл/кг. Зауважимо лише, що цей метод називають методом магнетрона тому, що конфігурація електричного і магнітного полів, в яких рухаються електрони, нагадує конфігурацію схрещених полів у магнетронах – багаторезонаторних приладах для генерування електромагнітних коливань в області надвисоких частот.

Контрольні питання

1.Що є джерелом існування магнітного поля?

2.Що є характеристикою магнітного поля у вакуумі?

3.В яких фізичних одиницях вимірюється індукція магнітного поля?

4.Пояснити формулу для індукції магнітного поля електричного заряду, який рухається.

5.Сформулювати та пояснити закон Біо-Савара-Лапласа.

6.Пояснити суть принципу суперпозиції магнітних полів.

7.Сформулювати і пояснити теорему про циркуляцію вектора індукції магнітного поля для вакууму та теорему Остроградского-Гаусса для магнітного поля. На який характер магнітного поля вони вказують?

8.Магнітне поле прямолінійного провідника та колового витка зі струмом.

9.Магнітне поле соленоїда та тороїда.

10.Сила Ампера, сила Лоренца.

11Де використовується поведінка контура зі струмом в магнітному полі?

12.В яких дослідах вивчається рух заряджених частинок в схрещених електричному і магнітному полях?

Розділ 6.Постійне магнітне поле в речовині

§6.1.Намагнічування середовищ

Досі ми розглядали магнітне поле у вакуумі. Якщо провідники зі струмом розміщені не у вакуумі, а в іншому середовищі, то магнітне поле змінюється.

Це означає, що речовини в магнітному полі намагнічуються, тобто самі стають джерелами магнітного поля. Результуюче магнітне поле в середовищі є сумою полів, які створюються провідниками зі струмом $\vec{B}_0$ і намагніченим середовищем $\vec{B}'$, і тому обидва поля в сумі дають результуюче поле

$$\vec{B} = \vec{B}_0 + \vec{B}'. \qquad (6.1)$$

Речовини, які здатні намагнічуватись, називаються магнетиками. Згідно гіпотези Ампера, причина намагнічування полягає в тому, що в усіх речовинах існують найдрібніші електричні струми, які замикаються в межах кожного атома (молекулярні струми, рис.6.1). Кожний такий струм має магнітний момент і створює в оточуючому просторі магнітне поле. У відсутності зовнішнього магнітного поля молекулярні струми орієнтовані хаотично внаслідок чого їхнє сумарне магнітне поле дорівнює нулю. Під дією зовнішнього поля магнітні моменти молекул орієнтуються по полю (§5.8), внаслідок чого магнетик намагнічується – його сумарний магнітний момент стає відмінним від нуля. Магнітні поля окремих молекулярних струмів в цьому випадку вже не компенсують один одного і виникає деяке середнє поле $\vec{B}'$. Тому намагнічений магнетик можна представити як систему найдрібніших орієнтованих струмів (рис.6.1).

У §5.8 ми бачили, що всі магнітні дії замкнених струмів визначаються їх магнітним моментом:

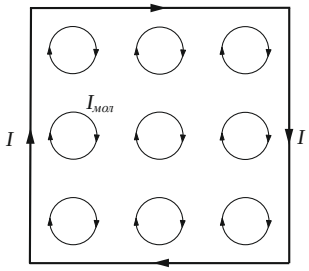


Рис.6.1.

$$\vec{p}_m = IS\vec{n},$$

де $I-$ сила струму, $S-$ площа, яка обігається струмом, $\vec{n}-$ одиничний вектор нормалі до площини витка зі струмом, який з напрямком струму утворює правогвинтову систему. Кожний молекулярний струм у магнетику має певний магнітний момент, а отже, і магнетик у цілому при намагнічування набуває магнітного моменту, який дорівнює векторній сумі моментів усіх молекулярних струмів. Тому магнітний стан речовини можна цілком охарактеризувати магнітним моментом його одиниці об'єму. Ця величина називається *вектором намагнічування* і позначається буквою $\vec{J}$. Якщо магнетик намагнічений неоднорідно, то намагніченість в даній точці визначається наступним виразом

$$\vec{J} = \frac{\sum_{\Delta V} \vec{p}_m}{\Delta V}, \qquad (6.2)$$

де ΔV –фізично нескінченно малий об'єм, взятий біля точки, яку розглядаємо, $\vec{p}_m$ – магнітний момент окремої молекули. Сумування виконується по всіх молекулах, що знаходяться в об'ємі ΔV (порівняйте із формулою (2.9)).

Вектор намагнічування є основною величиною, яка характеризує магнітний стан речовини. Знаючи вектор намагнічування в

кожній точці якого-небудь тіла, можна визначити і магнітне поле, яке створюється намагніченим тілом, що розглядається. Задача значно спрощується, якщо вектор намагнічування однаковий в усіх точках магнетика (*однорідне* намагнічування).У цьому випадку при додаванні молекулярних струмів прилеглі їхні відрізки, що мають протилежні напрями струмів, взаємно компенсуються і залишаються тільки відрізки струмів, які прилягають до поверхні магнетика. Тому дія всіх молекулярних струмів буде така сама, як дія деякого поверхневого струму, який обтікає намагнічений магнетик (рис.6.1). У цьому розумінні можна сказати, що при всуванні в соленоїд залізного осердя на поверхні осердя ніби виникають додаткові невидимі ампер-витки, які додаються до ампер-витків намагніченої котушки (рис.6.2).

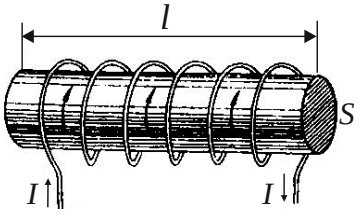


Рис.6.2.Поверхневі струми в намагніченому циліндрі.

Величина зазначеного поверхневого струму визначається значенням намагнічування $\vec{J}$. Розглянемо в однорідному магнітному полі (довгий соленоїд) досить довгий циліндричний стрижень (рис.6.2) і позначимо через i_1 лінійну густину поверхневого струму, тобто величину струму на одиницю довжини стрижня. Тоді повна величина поверхневого струму стрижня є $i_1 l$, де l – довжина стрижня. Якщо S – площа перерізу стрижня, то величина його магнітного момента дорівнює:

$$i_1 lS = i_1 V$$

($V -$ об'єм стрижня). З іншого боку, за визначенням намагнічування J, цей самий момент дорівнює JV. Прирівнюючи обидва вирази, ми знаходимо:

$$J = i_1. \qquad (6.3)$$

При однорідному намагнічуванні значення намагнічення J дає безпосередньо лінійну густину поверхневого струму магнетика.

Поле $\vec{B}'$ так само, як і поле $\vec{B}_0$, не має джерел (магнітних зарядів). Тому потік результуючого поля (6.1) через замкнуту поверхню дорівнює нулю:

$$\oint_S \vec{B} d\vec{S} = \oint_S \vec{B}_0 d\vec{S} + \oint_S \vec{B}' d\vec{S} = 0. \quad (6.4)$$

Таким чином, формула (5.24) справедлива не тільки для поля у вакуумі, але і для поля у речовині.

§6.2.Магнітомеханічні явища

Природа молекулярних струмів стала зрозумілою після того, як дослідами Резерфорда було встановлено, що атоми всіх речовин складаються із позитивно зарядженого ядра і негативно заряджених електронів, які рухаються навколо нього.

Рух електронів в атомах підпорядковується квантовим законам, тому поняття траєкторії до руху електронів в атомі не застосовується. Однак, магнетизм речовини вдається пояснити, користуючись простою моделлю Бора, згідно з якою електрони в атомах рухаються по стаціонарних кругових орбітах. Рух електрона по одній з таки орбіт (рис.6.3) еквівалентний коловому струму, тому він має ***орбітальний магнітний момент*** $\vec{p}_m = IS\vec{n}$, модуль якого $p_m = IS$, де S – площа орбіти електрона: $S = \pi r^2$. Вектор $\vec{p}_m$ напрямлений в той самий бік, що й індукція

магнітного поля в центрі колового струму I. Кількість обертів електрона за секунду $\nu = \frac{\upsilon}{2\pi r}$. Тоді сила струму $I = e\nu = \frac{e\upsilon}{2\pi r}$ і

$$p_m = \frac{e\upsilon}{2\pi r}\pi r^2 = \frac{e\upsilon r}{2}. \qquad (6.4)$$

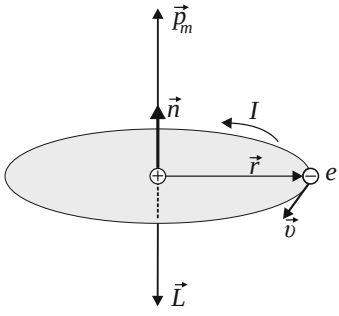


Рис.6.3.

З іншого боку, кожний електрон, що рівномірно рухається по орбіті, має ***орбітальний механічний момент імпульсу***, який дорівнює $\vec{L} = \vec{r} \times m_e\vec{v}$, і цей вектор направлений перпендикулярно до площини руху електрона. Так як $\vec{L} \perp \vec{v}$, то величина

$$L = m_e\upsilon r. \qquad (6.5)$$

Ця формула нерелятивістська, але для атома вона достатньо справедлива, оскільки для електрона на борівській орбіті відношення υ/c (c – швидкість світла у вакуумі) в загальному випадку дорівнює по порядку величини $e^2/(4\pi\varepsilon_0\hbar c) = 1/137$, або близько 1% ($\hbar$ – стала Планка).

Тоді справедливе відношення:

$$\frac{p_m}{L} = \frac{e}{2m_e}.$$

Вектори $\vec{p}_m$ і $\vec{L}$ напрямлені у взаємно протилежні сторони. Тому

$$\vec{p}_m = -\frac{e}{2m_e}\vec{L} \text{ і } \vec{p}_m = -g\vec{L}. \quad (6.6)$$

де $g = \frac{e}{2m_e}$ – гіромагнітне відношення орбітальних моментів.

Гіромагнітне відношення, яке експериментально визначили Ейнштейн і де-Гааз для електрона, виявилось в два рази більшим від гіромагнітного відношення орбітальних моментів.

Щоб пояснити результат експерименту, припустили, що електрон крім орбітальних моментів $\vec{p}_m$ і $\vec{L}$, має ***власний механічний момент імпульсу*** $\vec{L}_s$, що називається ***спіном електрона*** (детальніше розглядається в розділі 17). Спін є невід'ємною властивістю електрона, подібно до того, як заряд e і маса m_e. Спін електрона проявляється в багатьох експериментальних фактах. Спіну електрона $\vec{L}_s$ відповідає власний (спіновий) магнітний момент $\vec{p}_{ms}$, який дорівнює :

$$\vec{p}_{ms} = -g_s\vec{L}_s = -\frac{e}{m_e}\vec{L}_s, \quad (6.7)$$

де $g_s = \frac{e}{m_e}$ – гіромагнітне відношення спінових моментів.

Спін мають не тільки електрони, але й інші елементарні частинки. Максимальне значення проекції спіна на виділений в просторі напрямок, наприклад, на напрямок зовнішнього поля дорівнює цілому, або напівцілому значенню величини $\hbar$, яка дорівнює сталій Планка h, поділеній на 2π. Для електрона $L_s = \frac{1}{2}\hbar$, тому говорять, що спін електрона дорівнює $1/2$. Таким чином, $\hbar$ є природною одиницею моменту імпульсу, подібно тому як елементарний заряд e є природною одиницею заряду.

Згідно (6.7) спіновий магнітний момент електрона дорівнює

$$p_{ms} = -\frac{e}{m_e}L_s = -\frac{e\hbar}{2m_e}. \qquad (6.8)$$

Величину

$$\mu_Б = \frac{e\hbar}{2m_e} = 0{,}927 \cdot 10^{-23} \text{ Дж/Тл}$$

називають ***магнетоном Бора.***

Магнітний момент електрона складається з його орбітального і спінового магнітного моментів. Відповідно магнітний момент атома $\vec{p}_{ma}$ дорівнює сумі магнітних моментів електронів, що входять до складу атома, і магнітного моменту ядра, який значно менший від магнітного моменту електронів. Тому магнітним моментом ядра, як правило, нехтують. Отже,

$$\vec{p}_{ma} = \sum_{i=1}^{Z} \vec{p}_{mi} + \sum_{i=1}^{Z} \vec{p}_{msi},$$

де Z – атомний номер хімічного елемента.

За відсутності зовнішнього магнітного поля магнітні моменти орієнтовані хаотично, внаслідок чого сумарний магнітний момент речовини дорівнює нулю і результуюча магнітна індукція поля, яке створене мікроскопічними струмами, дорівнює нулю. Під дією зовнішнього магнітного поля магнітні моменти молекул орієнтуються в певному напрямку і виникає внутрішнє магнітне поле, сумарний магнітний момент відмінний від нуля і це значить, що речовина намагнічується.

Таким чином, магнетизм речовин обумовлений трьома причинами:

1) орбітальним рухом електронів навколо атомних ядер;

2) спіном електрона;

3) магнітним моментом атомних ядер.

Важкі атомні ядра рухаються значно повільніше легких електронів. Тому магнітні моменти атомних ядер в тисячі раз менші орбітальних і спінових магнітних моментів електронів.

§6.3.Закон повного струму для магнітного поля в речовині. Напруженість магнітного поля. Магнітна сприйнятливість і проникність

Запишемо закон повного струму, згідно формули (5.23), для результуючого магнітного поля (6.1) в речовині:

$$\oint_{\Gamma} \vec{B} d\vec{l} = \oint_{\Gamma} \vec{B}_0 d\vec{l} + \oint_{\Gamma} \vec{B}' d\vec{l} = \mu_0 (\sum_i I_i + \sum I_{мол}), \quad (6.9)$$

де $\sum_i I_i$ – алгебраїчна сума макрострумів (струмів провідності), $\sum I_{мол}$ – алгебраїчна сума мікрострумів (молекулярних струмів), що охоплюються заданим контуром Г. З контуром Г зв'язана деяка поверхня S (рис.6.4). Тому

$$\sum I_{мол} = \int_S \vec{j}_{мол} \cdot d\vec{S}, \quad (6.10)$$

де $\vec{j}_{мол}$ – густина молекулярних струмів, яка залежить від значень вектора $\vec{B}$, який в свою

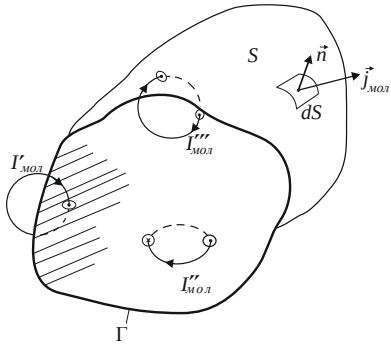


Рис.6.4.

чергу залежить від $\vec{j}_{мол}$. Щоб обійти цю трудність, знайдемо таку допоміжну фізичну величину, циркуляція якої визначається лише силою макроскопічних струмів. Щоб встановити вид цієї допоміжної величини, виразимо густину молекулярних струмів $\vec{j}_{мол}$ через намагніченість магнетика. Для цього вирахуємо суму (6.10). В алгебраїчну суму (6.10) входять тільки ті молекулярні струми, які

є “нанизані” на контур Г (рис.6.4). Струми, які не “нанизаються” на контур Г, або не перетинають цю поверхню зовсім, або перетинають цю поверхню два рази – один раз в одному напрямку, другий раз в іншому (струм $I''_{мол}$ на рис.6.4) не дають вкладу в алгебраїчну суму струмів (6.10).

Із рис.6.5 видно, що елемент контура $d\vec{l}$, який утворює з напрямком намагніченості $\vec{J}$ кут α, "нанизує" на себе ті молекулярні струми, центри яких попадають всередину косого циліндра з об’ємом $S_{мол} \cdot dl \cdot \cos\alpha$ ($S_{мол}$ – площа, яка охоплюється одним молекулярним струмом). Якщо n – число молекул в одиниці об’єму магнетика, то сумарний струм, який охоплюється елементом контура dl, дорівнює $I_{мол} \cdot nS_{мол}dl \cdot \cos\alpha$. Добуток $I_{мол} \cdot S_{мол} = p_m$, де p_m – магнітний момент окремого молекулярного струму.

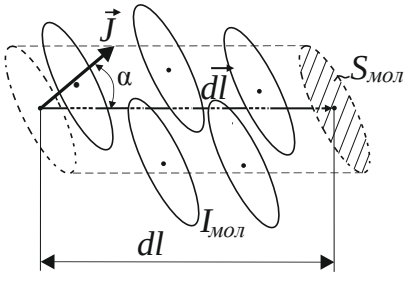


Рис.6.5.

Значить вираз $I_{мол}S_{мол} \cdot n$ є магнітний момент одиниці об’єму, тобто модуль вектора $\vec{J}$, а $I_{мол} \cdot S_{мол} \cdot n\cos\alpha$ – проекція вектора $\vec{J}$ на напрямок елемента $d\vec{l}$. Таким чином, сумарний молекулярний струм, що охоплюється елементом контура $d\vec{l}$ дорівнює $\vec{J}d\vec{l}$, а сума молекулярних струмів (6.10), які охоплюються всім контуром Г, дорівнює:

$$\int_S \vec{j}_{мол} \cdot d\vec{S} = \oint_\Gamma \vec{J}d\vec{l}. \qquad (6.11)$$

Із врахуванням (6.11), (6.9) можна переписати так:

$$\oint_\Gamma \left(\frac{\vec{B}}{\mu_0} - \vec{J}\right) d\vec{l} = \sum_i I_i. \qquad (6.12)$$

Вираз в дужках (6.12)

$$\vec{H} = \frac{\vec{B}}{\mu_0} - \vec{J} \qquad (6.13)$$

і є шукана допоміжна величина, циркуляція якої визначається лише макроскопічними струмами і (6.12) перепишеться так:

$$\oint_\Gamma \vec{H}\, d\vec{l} = \sum_i I_i. \qquad (6.14)$$

Вектор $\vec{H}$ *називається напруженістю магнітного поля.*

Отримана формула (6.14) виражає теорему про циркуляцію вектора $\vec{H}$ (закон повного струму для магнітного поля в середовищі): *циркуляція вектора напруженості магнітного поля по деякому замкнутому контуру дорівнює алгебраїчній сумі макроскопічних струмів, що охоплюються цим контуром.*

Напруженість магнітного поля $\vec{H}$ є аналогом електричного зміщення $\vec{D}$ (2.18).

У вакуумі $\vec{J} = 0$, тому $\vec{H} = \vec{B}/\mu_0$ і

$$\vec{B} = \mu_0\vec{H}. \qquad (6.15)$$

Як показують досліди, в несильних магнітних полях намагніченість $\vec{J}$ прямо пропорційна до напруженість поля $\vec{H}$, що викликає намагнічення, тобто

$$\vec{J} = \chi\vec{H}, \qquad (6.16)$$

де χ – магнітна сприйнятливість речовини. Тоді (6.13) із врахуванням (6.16) перепишеться так:

$$\vec{H} = \frac{\vec{B}}{\mu_0(1+\chi)}. \qquad (6.17)$$

Безрозмірна величина

$$\mu = 1 + \chi \qquad (6.18)$$

називається *відносною магнітною проникністю* речовини.

На відміну від діелектричної сприятливості χ (формула(2.11)), яка може мати тільки позитивні значення (поляризованість $\vec{P}$ в ізотропному діелектрику завжди направлена по полю $\vec{E}$), магнітна сприятливість χ буває як позитивна, так і негативна. Тому магнітна проникність μ може бути як більша одиниці, так і менша одиниці.

Із врахуванням (6.18) формула (6.17) прийме вигляд

$$\vec{B} = \mu\mu_0\vec{H}. \qquad (6.19)$$

Таким чином, напруженість магнітного поля $\vec{H}$ є вектор, який має той же напрямок, що і вектор $\vec{B}$, але в $\mu\mu_0$ раз меншу за модулем величину (в анізотропних середовищах вектори $\vec{H}$ і $\vec{B}$, взагалі кажучи, не співпадають за напрямком).

§6.4.Магнітне поле на межі двох магнетиків

Біля поверхні розділу двох магнетиків вектори $\vec{B}$ і $\vec{H}$ повинні задовольняти граничним умовам, які отримуються із співвідношень:

$$\oint_S \vec{B}d\vec{S} = 0, \qquad \oint_\Gamma \vec{H}\,d\vec{l} = \sum_i I_i. \qquad (6.20)$$

(див.формули (5.24) і (6.14)). Розглянемо стаціонарні поля. Візьмемо на межі розділу двох магнетиків з проникностями μ_1 і μ_2 уявну циліндричну поверхню, висота якої h, а площі основ $S_1 = S_2 = S$, які розташовані по різні сторони поверхні розділу (рис.6.6). Потік вектора $\vec{B}$ через цю циліндричну поверхню

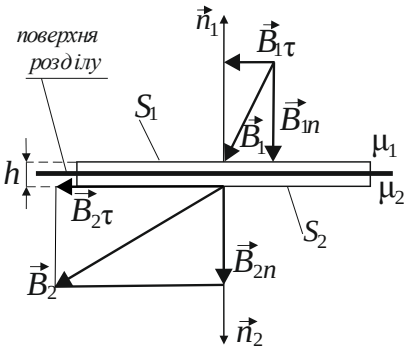


Рис.6.6.

дорівнює:

$$\Phi_B = \vec{B}_1\vec{n}_1S_1 + \vec{B}_2\vec{n}_2S_2 + < B > S_{бічна}$$

або

$$\Phi_B = -B_{1n}S_1 + B_{2n}S_2 + < B > S_{бічна},$$

де $<B>$ −деяке усереднене значення магнітної індукції на межі розділу середовищ, B_{1n} і B_{2n} − проекції векторів $\vec{B}_1$ і $\vec{B}_2$ на відповідні нормалі $\vec{n}_1$ і $\vec{n}_2$ до поверхні циліндра (рис.6.6). Оскільки потік вектора $\vec{B}$ через будь-яку замкнуту поверхню дорівнює нулю, то при умові $h \to 0$ ($S_{бічна} \to 0$), прийдемо до співвідношення

$$B_{1n} = B_{2n}. \qquad (6.21)$$

Замінивши, згідно (6.19), складові вектора $\vec{B}$ відповідними складовими вектора $\vec{H}$, помноженими на $\mu_0\mu$, отримаємо співвідношення

$$\mu_0\mu_1H_{1n} = \mu_0\mu_2H_{2n},$$

із якого витікає, що

$$\frac{H_{1n}}{H_{2n}} = \frac{\mu_2}{\mu_1}. \qquad (6.22)$$

Тепер візьмемо на межі магнетиків прямокутний контур Г (рис.6.7) і вирахуємо для нього циркуляцію $\vec{H}$. При малих розмірах контуру Г циркуляцію можна представити у вигляді: $\oint_\Gamma \vec{H}\,d\vec{l} = \int_1^2 \vec{H}_1\,d\vec{l} + \int_2^3 \langle\vec{H}\rangle\,d\vec{l} +$

$$+\int_3^4 \vec{H}_2\, d\vec{l} + \int_4^1 \langle \vec{H} \rangle\, d\vec{l},$$

або

$$\oint_\Gamma \vec{H}\, d\vec{l} = -H_{1\tau} a + 2\langle H_l \rangle b + H_{2\tau} a,$$

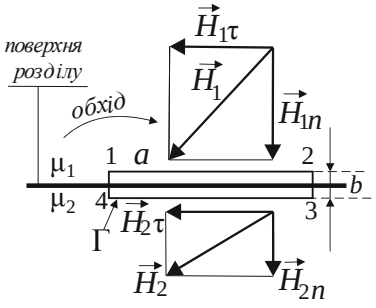


Рис.6.7.

де $\langle H_l \rangle$ – середнє значення H_l на перпендикулярних до поверхні розділу ділянках контуру, $H_{1\tau} = \vec{H}_1 \cdot \vec{\tau}_1$, $H_{2\tau} = \vec{H}_2 \cdot \vec{\tau}_2$, $\vec{\tau}_1$ і $\vec{\tau}_2$ – одиничні вектори дотичні до межі розділу магнетиків. $H_{1\tau}$ і $H_{2\tau}$ так звані тангенційні складові векторів $\vec{H}_1$ і $\vec{H}_2$. Якщо на межі розділу не течуть макроскопічні струми, то згідно (6.20) при умові, що $b \to 0$, отримаємо:

$$\oint_\Gamma \vec{H}\, d\vec{l} = -H_{1\tau} a + H_{2\tau} a = 0.$$

Звідки отримаємо співвідношення:

$$H_{1\tau} = H_{2\tau} \qquad (6.23)$$

(порівняйте з (2.25)).

Замінимо складові $\vec{H}$ у (6.23) відповідними складовими $\vec{B}$, згідно (6.19), поділеними на $\mu_0\mu$, отримаємо співвідношення

$$\frac{B_{1\tau}}{\mu_0\mu_1} = \frac{B_{2\tau}}{\mu_0\mu_2},$$

із якого витікає, що

$$\frac{B_{1\tau}}{B_{2\tau}} = \frac{\mu_1}{\mu_2}. \qquad (6.24)$$

Із співвідношень (6.21) і (6.23) можна зробити висновок, що при переході через межу розділу двох магнетиків нормальна складова вектора $\vec{B}$ і тангенційна складова вектора $\vec{H}$

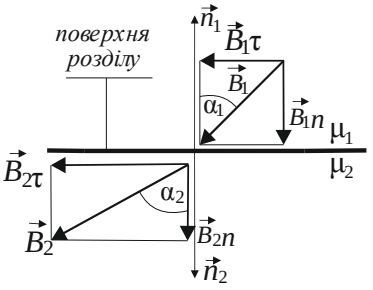


Рис.6.8.

змінюються неперервно. Із (6.22) і (6.24) витікає, що тангенційна складова вектора $\vec{B}$ і нормальна складова вектора $\vec{H}$ при переході через межу розділу двох магнетиків зазнають розриву. Таким чином, при переході через межу розділу двох середовищ вектор $\vec{B}$ поводить себе аналогічно вектору $\vec{D}$, а вектор $\vec{H}$ – аналогічно вектору $\vec{E}$.

На рис.6.8 показано поведінку вектора $\vec{B}$ при перетині межі розділу двох магнетиків ($\mu_2 > \mu_1$). Позначимо кути між лініями $\vec{B}$ і нормаллю до поверхні розділу відповідно α_1 і α_2 (рис.6.8). Відношення тангенсів цих кутів дорівнює

$$\frac{\mathrm{tg}\alpha_1}{\mathrm{tg}\alpha_2} = \frac{B_{1\tau}/B_{1n}}{B_{2\tau}/B_{1n}}.$$

Звідки, з урахуванням (6.21) і (6.24), отримаємо закон заломлення ліній магнітного індукції.

$$\frac{\mathrm{tg}\alpha_1}{\mathrm{tg}\alpha_2} = \frac{\mu_1}{\mu_2}. \qquad (6.25)$$

При переході в магнетик з більшою μ ($\mu_2 > \mu_1$) лінії магнітної індукції відхиляються від нормалі до поверхні розділу двох середовищ. Це призводить до згущення ліній. Згущування ліній $\vec{B}$ у речовині з більшою

магнітною проникністю дає можливість формувати магнітні пучки, тобто надавати їм необхідну форму і напрямок. Для того, щоб здійснити *магнітний захист* деякого об'єму, його оточують залізним екраном. Із рис.6.9 видно, що згущення ліній магнітної індукції в товщі екрану призводить до ослаблення поля всередині.

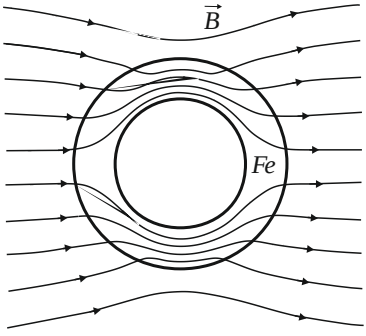


Рис.6.9.

§6.5.Класифікація магнетиків

Формула (6.16) визначає магнітну сприйнятливість χ одиниці об'єму речовини. Замість цієї сприйнятливості часто використовують сприйнятливість для одного кіломоля речовини, яку називають кіломолярною $\chi_{км}$ (для хімічно простих речовин кілоатомною $\chi_{кат}$).Очевидно, що $\chi_{км}(\chi_{кат})$ має розмірність $м^3/кмоль$, $(м^3/кат)$.

В залежності від знаку і величини магнітної сприйнятливості всі магнетики поділяються на три групи:

1) Діамагнетики, у яких χ від'ємна і мала за абсолютною величиною $(\chi_{км} \sim -(10^{-8} \div 10^{-7}) м^3/кмоль)$. До таких речовини відносяться: вуглекислота, вода, срібло, вісмут, мідь, цинк, золото, аргон та інші. Діамагнетики послаблюють зовнішне магнітне поле. Від'ємне значення χ в діамагнетиках означає, що в цих речовинах вектор намагнічення направлений протилежно зовнішньому (намагнічуючому) полю.

2) парамагнетики, це речовини у яких χ теж невелика, але позитивна величина $(\chi_{км} \sim 10^{-7} \div 10^{-6} м^3/кмоль)$. Такими речовинами є: кисень, алюміній, літій, натрій, калій, магній, кальцій, титан, вольфрам, платина, уран, плутоній, хлористе залізо $(FeCl_3)$ та інші.

3) феромагнетики це речовини , у яких χ додатня і досягає дуже великих значень $(\chi_{км} \sim 10^3\ м^3/кмоль)$. Крім того, на відміну від діа- і парамагнетиків, для яких χ не залежить від H, сприйнятливість феромагнетиків є функція напруженості магнітного поля. Феромагнетиками є хімічні елементи Fe, Co, Ni (так звані $3d$-метали) і рідкоземельні метали Gb, Tb, Dy, Ho, Er та інші. Феромагнетиками є теж багаточисельні металічні бінарні і більші складні (багатокомпонентні) сплави вище згаданих металів між собою та з іншими неферомагнітними системами, наприклад: Fe_3Al, Ni_3Mn, $MnPt_3$, $ZrZn_2$, $ZnCMn_3$, $AlCMn_3$ та інші.

§6.6.Діа - і парамагнетизм

Діамагнетизм (від грецького $dia-$ префікс, який тут означає розходження (силових ліній), і магнетизм), властивість речовини намагнічуватись назустріч напрямку зовнішнього магнітного поля, яке діє на неї. Діамагнетизм властивий всім речовинам.

Розглянемо вплив зовнішнього магнітною поля $\vec{B}$ на рух електронів в атомах речовини.

Нехай навколо ядра атома рівномірно рухається по коловій орбіті електрон. Якщо магнітне поле відсутнє, то кулонівська сила є

доцентровою і згідно другого закону Ньютона

$$\frac{Ze^2}{4\pi\varepsilon_0 r^2} = \frac{m_e \upsilon_o{}^2}{r} = m_e \omega_0^2 r, \qquad (6.26)$$

де υ_o – лінійна швидкість електрона на орбіті, ω_0 – його кутова швидкість, r – радіус орбіти.

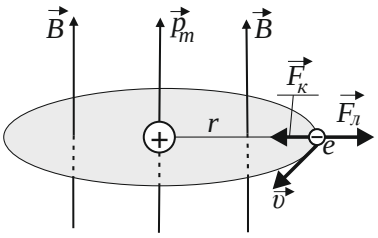


Рис.6.10.

В однорідному магнітному полі, лінії індукції якого перпендикулярні до площини орбіти електрона, на кожен електрон діятиме ще й сила Лоренца

$$F_л = e\upsilon B = e\omega r B,$$

де ω – кутова швидкість при наявності магнітного поля (рис.6.10).

Тепер рівняння Ньютона запишемо так:

$$\frac{Ze^2}{4\pi\varepsilon_0 r^2} - e\omega r B = m_e \omega^2 r$$

або із урахуванням (6.26)

$$m_e r(\omega^2 - \omega_0^2) = -e\omega r B. \qquad (6.27)$$

Із (6.27) отримаємо, що

$$(\omega - \omega_0)(\omega + \omega_0) = -\frac{e}{m_e} B\omega.$$

Якщо позначити $\omega - \omega_0 = \omega_L$, а $\omega + \omega_0 \approx 2\omega$, то

$$\omega_L = -\frac{eB}{2m_e}. \qquad (6.28)$$

Частоту (6.28) називають частотою *лармової прецесії* або просто *ларморовою частотою*.

Отже, магнітне поле викликає зміну кутової швидкості обертання електрона по орбіті. У випадку, коли $\vec{B}$ не перпендикулярне площині орбіти, то дія поля полягає у збудженні прецесії орбіти навколо поля (рис.6.11).

Магнітний момент електрона $\vec{p}_m$ описує конус навколо $\vec{B}$. В цьому випадку кутова швидкість ω_L такої прецесії теж виражається формулою (6.28). Кутова швидкість ω_L є однаковою для всіх електронів атома і не залежить ні від кута нахилу орбіти φ відносно вектора $\vec{B}$, ні від радіуса орбіти або швидкості електрона.

***Теорема Лармора**: єдиним наслідком впливу магнітного поля на орбіту електрона* в *атомі є прецесія орбіти і вектора* $\vec{p}_m$ з *кутовою швидкістю* ω_L *навколо осі, яка проходить через ядро атома і паралельна вектору* $\vec{B}$ *індукції магнітного поля.*

Прецесія орбіти обумовлює додатковий рух електрона навколо напрямку поля. Якщо б відстань r електрона від осі, що проходить через центр орбіти, не змінювалась, то додатковий рух електрона відбувався б по колу, радіус якого r.

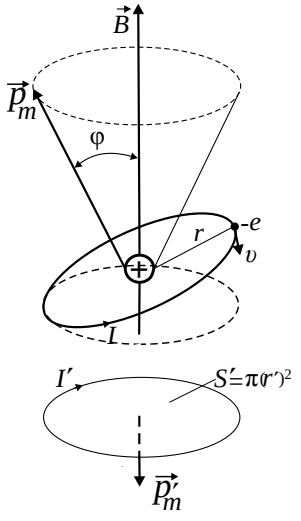


Рис.6.11.

Йому відповідав би коловий струм

$$I' = e\frac{\omega_L}{2\pi} = \frac{e^2 B}{4\pi m_e},$$

магнітний момент якого дорівнює

$$p'_m = I'S' = \frac{e^2}{4\pi m_e} S'B$$

і напрямлений в бік, протилежний до вектора $\vec{B}$ (рис.6.11):

$$\vec{p}'_m = -\frac{e^2}{4\pi m_e} S'\vec{B}. \qquad (6.29)$$

Магнітний момент $\vec{p}'_m$ називається індукованим магнітним моментом. Розрахунок показує, що $S' = \pi(r')^2 = \frac{2}{3}\pi r^2$.

В результаті отримуємо середнє значення індукованого магнітного моменту одного електрона:

$$\langle p'_m \rangle = -\frac{e^2 B}{6m_e} r^2. \qquad (6.30)$$

При отриманні виразу для $\langle p'_m \rangle$ передбачалось, що електрони рухаються по колових орбітах. В загальному випадку, наприклад, для еліптичної орбіти, потрібно замість r^2 брати $\langle r^2 \rangle$, тобто середній квадрат відстані електрона від ядра.

Індукований магнітний момент атома дорівнює сумі індукованих моментів електронів, що входять до складу атома:

$$p'_{ma} = \sum \langle p'_m \rangle = -\frac{e^2 B}{6m_e} \sum_{i=1}^{Z} \langle r_i^2 \rangle, \quad (6.31)$$

де Z – атомний момент хімічного елемента (число електронів в атомі дорівнює Z).

Таким чином, під дією зовнішнього магнітного поля відбувається прецесія електронних орбіт з однаковою для всіх електронів кутовою швидкістю (6.28). Обумовлений прецесією додатковий рух електронів приводить до виникнення індукованого магнітного момента атома (6.31), який направлений проти зовнішнього поля. Ларморова прецесія виникає у всіх без винятку речовин. Однак в тих випадках, коли атоми мають самі по собі магнітні моменти, магнітне поле не тільки індукує моменти (6.31), але і орієнтує магнітні моменти атомів по напрямку зовнішнього поля. При цьому виникає позитивний (тобто направлений по полю) магнітний момент, який буває значно більший, ніж індукований від'ємний момент (6.31). Це означає, що результуючий момент стає позитивний і речовина поводить себе як парамагнетик.

Тому діамагнетиками називають речовини, магнітні моменти атомів або молекул яких дорівнюють нулю, коли немає зовнішнього магнітного поля.

Коли внести діамагнетик у зовнішнє магнітне поле, у кожному його атомі індукується магнітний момент (6.31), який направлений протилежно до вектора $\vec{B}$ магнітної індукції поля.

Якщо (6.31) помножити на число Авогадро N_A, то отримаємо магнітний момент одного кіломоля ізотропного діамагнетика J. Поділивши J на напруженість магнітного поля H отримаємо кіломолярну сприйнятливість $\chi_{км}$. Магнітна проникність діамагнетиків практично дорівнює одиниці, тому $B = \mu_0 H$.

$$\chi_{км} = -\frac{J}{H} = -\frac{\mu_0 J}{B} = \\ = -\frac{\mu_0 N_A e^2}{6m_e} \sum_{i=1}^{Z} \langle r_i^2 \rangle. \qquad (6.32)$$

Зазначимо, що строга квантово-механічна теорія дає такий самий вираз для $\chi_{км}$, як (6.32).

Підставивши в (6.32) числові позначення для μ_0, N_A, e і m_e, отримаємо, що

$$\chi_{км} = -3{,}55 \cdot 10^{12} \sum_{i=1}^{Z} \langle r_i^2 \rangle. \quad (6.33)$$

Радіуси електронних орбіт мають величину порядка 10^{-10} м. Значить, кіломолярна діамагнітна сприйнятливість буде порядка $10^{-8} - 10^{-7}$ м^3/кмоль, що добре узгоджується з експериментальними даними.

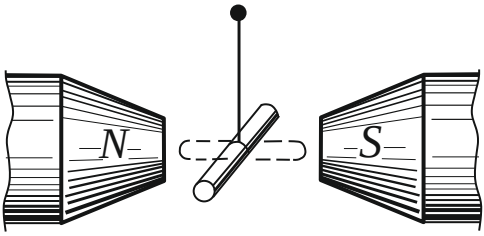


Рис.6.12.Діамагнітна паличка вісмуту в магнітному полі.

Стрижень з діамагнітного матеріалу намагнічується в напрямку, протилежному до напрямку зовнішнього магнітного поля. Тому в неоднорідному магнітному полі діамагнетик виштовхується в область слабшого поля і встановлюється так, щоб його вісь була перпендикулярна до вектора магнітної індукції $\vec{B}$ поля (рис.6.12).

Якщо векторна сума орбітальних магнітних моментів усіх електронів атома або молекули не дорівнює нулю, то атом загалом має деякий магнітний момент $\vec{p}_{ma}$. *Такі атоми (молекули) називаються парамагнітними, а речовини, що складаються з них – парамагнетиками.*

Парамагнетизм (від грецького para- біля, рядом і магнетизм) – властивість речовин (парамагнетиків), які розміщені у зовнішньому магнітному полі, намагнічуватись (набувати магнітний момент) в напрямку, який співпадає з напрямком цього поля.

За відсутності зовнішнього магнітного поля парамагнетик ненамагнічений, оскільки внаслідок теплового руху власні магнітні моменти атомів орієнтовані хаотично ($J = 0$).

Розглянемо, що станеться при внесенні парамагнетика в однорідне магнітне поле, магнітна індукція якого дорівнює $\vec{B}$.

Кожен електрон атома парамагнітного тіла бере участь у двох рухах: орбітальному і прецесійному. Згідно з теоремою Лармора, всі магнітні моменти електронів атома і результуючий магнітний момент атома прецесують навколо напрямку $\vec{B}$ з однаковою кутовою швидкістю ω_L.

Тепловий рух атомів парамагнетика і їх зіткнення спричинюють поступове згасання прецесії магнітних моментів атомів і, як наслідок, зменшення кутів між напрямами $\vec{p}_{ma}$ і $\vec{B}$ (рис.6.11). Отже, незважаючи на утворюване тепловим рухом «розкидання» атомів, цей рух водночас сприяє переважній орієнтації магнітних моментів атомів у напрямку зовнішнього магнітного поля, оскільки саме по собі магнітне поле може спричинювати лише прецесію $\vec{p}_{ma}$ навколо напрямку $\vec{B}$.

Магнітний момент p_{ma} окремого атома має величину $\sim 10^{-23}$Дж/Тл, але сукупна дія магнітних моментів усіх атомів, що містяться в одиниці об'єму речовини приводить до ефекту намагнічування, що значно перевищує діамагнітний ефект. Тому парамагнетик намагнічується "за полем", тобто в напрямку магнітної індукції $\vec{B}$ зовнішнього магнітного поля.

Класичну теорію парамагнетизму розвинув П. Ланжевен, який розглянув задачу про поведінку молекулярних струмів в однорідному магнітному полі.

Орієнтуюча дія магнітного поля на атом залежить від магнітного моменту атома і від магнітної індукції $\vec{B}$ поля. *«Розкидаюча»* дія теплового руху визначається величиною kT, пропорційною до середньої теплової енергії однієї частинки. Виявилося. що результуюча дія двох протилежних факторів залежить від відношення

$$\alpha = \frac{p_{ma}B}{kT},$$

де $p_{ma}B$ – енергія магнітного момента p_{ma} в магнітному полі *B*.

Ланжевен знайшов залежність намагніченості J парамагнетики від параметра α:

$$J = np_{ma}L(\alpha),$$

де n-концентрація атомів парамагнетика, $L(\alpha)$ – класична функція Ланжевена, яка має вигляд:

$$L(\alpha) = \left\{\frac{e^{\alpha} + e^{-\alpha}}{e^{\alpha} - e^{-\alpha}} - \frac{1}{\alpha}\right\}.$$

Якщо $\alpha \ll 1$, то $L(\alpha) \sim \frac{\alpha}{3}$; якщо $\alpha \gg 1$, тоді $L(\alpha) \to 1$.

При $T \approx 300K$ умова $\alpha = 1$ може справджуватись лише в дуже сильних магнітних полях.

В дуже сильних магнітних полях або при дуже низьких температурах, коли більша частина векторів $\vec{p}_{ma}$ «орієнтована» вздовж напряму індукції магнітного поля, спостерігається *явище насиченості*: намагніченість J не залежить від $\vec{B}$.

Далеко від області насиченості $L(\alpha) \sim \frac{\alpha}{3}$, тоді

$$\vec{J} = \frac{np_{ma}^2}{3kT}\vec{B} = \chi\frac{\vec{B}}{\mu_0}.$$

Магнітна сприйнятливість парамагнітних речовин:

$$\chi = \frac{\mu_0 np_{ma}^2}{3kT}.$$

Замінивши n числом Авогадро N_A, отримаємо вираз для кіломолярної сприйнятливості

$$\chi_{км} = \frac{\mu_0 N_A p_{ma}^2}{3kT}. \qquad (6.34)$$

Отже $\chi_{км} > 0$. Якщо прийняти, що $p_{ma} \approx (0{,}2 \div 0{,}4)\mu_Б$, тоді значення $\chi_{км} \sim 10^{-7} \div 10^{-6}$ м^3/кмоль.

Макроскопічно парамагнетизм виявляється в тому, що парамагнетики втягуються в неоднорідне магнітне поле, а в однорідному полі парамагнітний стрижень орієнтується паралельно до ліній індукції магнітного поля (рис.6.13). При нагріванні парамагнетика, який внесений у зовнішнє магнітне поле, тепловий рух атомів зростатиме і руйнуватиме ту орієнтацію елементарних магнітних моментів

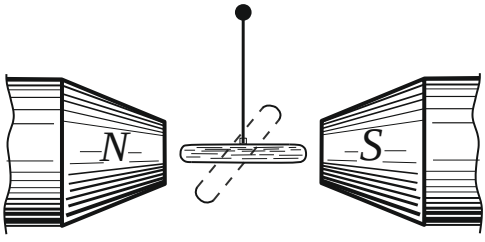


Рис.6.13.Ампула з парамагнітним розчином хлористого заліза в магнітному полі.

частини атомів, яка встановилася під дією зовнішнього поля. Отже, магнітна сприйнятливість парамагнетиків χ, як величина, що характеризує з макроскопічного погляду магнітні властивості речовини, повинна залежати від температури T:

$$\chi_{км} = C/T, \qquad (6.35)$$

де $C = \frac{\mu_0 N_A P_{ma}^2}{3k}$ – стала Кюрі.

Співвідношення (6.35) виражає ***закон Кюрі:*** *магнітна сприйнятливість парамагнетиків обернено пропорційна до його термодинамічної температури.* Квантова теорія парамагнетизму враховує той факт, що можливі тільки дискретні орієнтації магнітного момента атома $\vec{p}_{ma}$ відносно індукції поля $\vec{B}$. Вона дає для $\chi_{км}$ вираз аналогічний (6.34).

§6.7.Феромагнетизм

Речовини, які здатні добре намагнічуватися,називаються феромагнетиками (від лат. ferrum – залізо). Їх намагніченість J велика і в 10^{10} число раз перевищує намагніченість діа – і парамагнетиків, які належать до категорії слабомагнітних речовин.

Магнітна проникність більшості феромагнетиків за звичайних температур вимірюється багатьма сотнями і тисячами одиниць, а в деяких спеціально приготовлених і оброблених феромагнетиків вона досягає мільйона.

Феромагнетики, крім здатності добре намагнічуватися, мають ряд властивостей, які істотно відрізняють їх від діа– і парамагнетиків.

Крива намагнічування. Характерною особливістю феромагнетиків є складна нелінійна залежність між індукцією B і напруженістю поля H. Ця залежність була встановлена в класичних працях О.Г. Столєтова на прикладі м'якого (відпаленого) заліза.

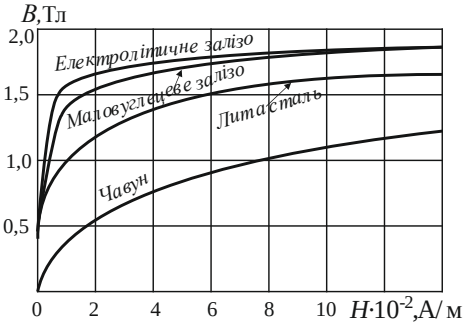


Рис.6.14.Залежність магнітної індукції від напруженості магнітного поля.

Залежність індукції B від напруженості магнітного поля H у феромагнетиках має вигляд, зображений на рис.6.14. Індукція спочатку швидко збільшується, але в міру намагнічування магнетика її зростання сповільнюється.

За значеннями індукції B і поля H можна визначити намагнічення магнетика $J = B/\mu_0 - H$ (магнітний момент одиниці об'єму). Характер залежності J від H для феромагнетиків зображено на рис.6.15. Намагнічення J, подібно до індукції, спочатку швидко зростає, але потім настає *магнітне насичення*, за якого намагнічення досягає деякого максимального значення $J_{нас}$ і практично перестає залежати від напруженості поля.

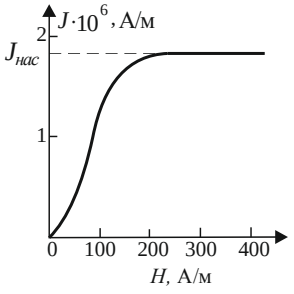


Рис.6.15.Крива намагнічення феромагнетика.

Внаслідок нелінійної залежності B від H магнітна проникність $\mu = B/(\mu_0 H)$ *залежить від напруженості магнітного поля.* Крива залежності μ від H (рис.6.16) зростає із збільшенням поля від початкового значення $\mu_п$ до деякої максимальної величини μ_m, але потім, після проходження через максимум, зменшується і асимптотично прямує до значення, дуже близького до одиниці.

Магнітна сприйнятливість феромагнетика $\chi = J/H$ також не є сталою і залежить від напруженості поля. Цю залежність зображено на рис.6.16. Магнітна прийнятливість χ, як і проникність μ, має максимум і при великих полях асимптотично прямує до значення, близького до нуля.

Залежність χ від H часто називають *кривою Столєтова* на честь О. Г. Столєтова, який вперше встановив особливості цієї кривої на прикладі м'якого заліза і вказав на наукову і технічну важливість цієї залежності.

Особливості намагнічування феромагнетиків, які розглядаються, показують, що використання феромагнетиків для одержання потужних магнітних полів дуже ефективне в областях намагнічування, які далекі від насичення. А у випадку дуже потужних полів настає магнітне насичення і магнітна проникність значно спадає.

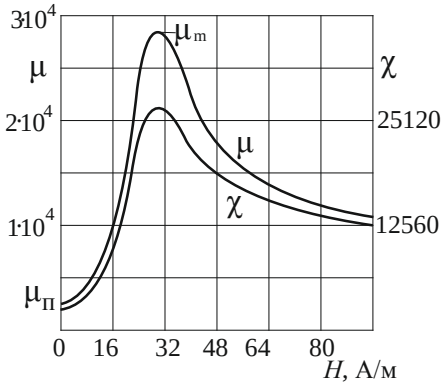


Рис.6.16.Залежність магнітної проникності μ і магнітної сприйнятливості χ від напруженості поля для заліза "армко".

Тому феромагнетики застосовують для отримання полів, напруженість яких не перевищує $\sim 10^6$А/м. А якщо необхідне ще потужніше поле, то застосовувати феромагнетики практично не ефективно. Так, наприклад, в дослідах П.Л.Капіци поля з напруженістю до $3 \cdot 10^7$А/м були добуті тільки за допомогою котушок зі струмом без феромагнетиків.

Анізотропія намагнічування. Усі феромагнетики в магнітному розумінні є *анізотропними*. Однак якщо феромагнетик має дрібнокристалічну структуру і окремі кристалики в ньому розміщені цілком безладно, то ця анізотропія не проявляється і його намагнічення не залежить від напряму поля. А коли феромагнетик є єдиним кристалом, то вигляд кривої намагнічення буде різний і залежатиме від напряму намагнічуючого поля відносно осей кристала. На рис.6.17 зображено елементарну комірку кристала заліза (центрований куб) і показано кристалографічні напрями, які позначені символами [100] (ребро куба), [110] (діагональ грані) і [111] (просторова діагональ). Для кожного феромагнетика існує напрям, в якому

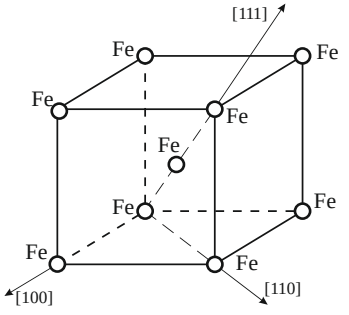


Рис.6.17.Елементарна кристалічна комірка заліза і її основні кристалографічні напрямки. [100] є напрямом легкого намагнічування, а [111] – напрямом важкого намагнічування.

намагнічення (при даному полі) найбільше (*напрям легкого намагнічування* [100]), і напрям, в якому намагнічення найменше (*напрям важкого намагнічування* [111] (рис.6.17)).

Гістерезис. Припустимо, що ми намагнічуємо спочатку ненамагнічений феромагнетик і, помістивши його всередину намагнічуючої котушки, збільшуємо магнітне поле всередині магнетика від нуля до деякого значення H_1 (рис.6.18). Значення індукції в магнетику визначатиметься кривою індукції $01A$ і зобразиться відрізком ординати $0B_1$. Якщо тепер знову зменшувати магнітне поле, то зменшення індукції зображатиметься вже не відрізком кривої індукції $1O$, а кривою $1B'$, і коли поле знову дорівнюватиме нулю, індукція

не дорівнюватиме нулю, а виражатиметься відрізком OB'. Феромагнетик у цьому стані є *постійним магнітом.* Далі, якщо змінити напрям струму в намагнічуючій котушці і перемагнічувати зразок у зворотному напрямі, то зміна індукції описуватиметься відрізком кривої $B'2$. При наступній зміні поля в зворотному напрямі індукція змінюватиметься згідно з кривою $2B''1$. При циклічному перемагнічуванні феромагнетика зміна індукції в ньому зображатиметься петлеподібною замкненою кривою $1B'2B''1$.

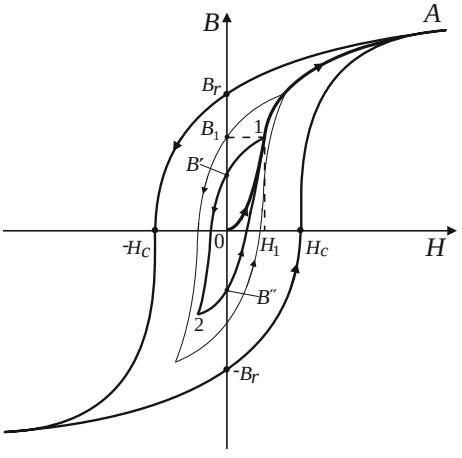


Рис.6.18.Магнітний гістерезис.

Ми бачимо, що значення індукції у феромагнетику визначається не тільки існуючим магнітним полем, але ще залежить від попередніх станів намагнічування, причому відбувається своєрідне *відставання* зміни індукції від зміни напруженості поля. Це явище називається магнітним *гістерезисом,* а зазначена вище петлеподібна крива залежності B від H при циклічному перемагнічуванні називається *петлею гістерезису*. Магнітний гістерезис подібний до діелектричного гістерезису в сегнетоелектриках (§2.6).

З кривих (рис.6.18) видно, що при усуненні намагнічуючого поля феромагнетик зберігає залишкове намагнічення, причому всередині магнетика існує деяка залишкова індукція. При збільшенні амплітуди намагнічуючого поля індукція прямує до граничного значення B_r при H=0 (рис.6.18). Щоб "знищити" це залишкове намагнічення, всередині феромагнетика слід створити певне поле, напрямлене протилежно початковому намагнічуючому полю, яке зображене відрізком OH_c. Це поле називають *затримуючою*, або *коерцитивною,* силою феромагнетика. Коерцитивна сила характеризує властивість феромагнетика зберігати намагніченість. Матеріали з великою коерцитивною напруженістю дають широку петлю гістерезису і називаються «твердими» магнітними матеріалами. З них виготовляють постійні магніти. «М'які» магнітні матеріали мають малу коерцитивну силу і дають вузьку петлю гістерезису. Ці матеріали використовують для виготовлення осердь трасформаторів.

З цим пов’язаний практичний спосіб, який використовують для *розмагнічування* феромагнетиків. Для цього феромагнетик поміщають всередину котушки, яка живиться

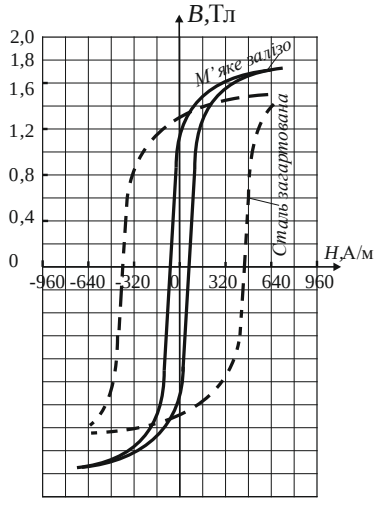


Рис.6.19.Криві гістерезису м’якого заліза і загартованої сталі.

змінним струмом, і величину струму поступово зменшують до нуля. При цьому феромагнетик піддається багаторазовим циклічним перемагнічуванням, що відповідає різним петлям гістерезису, які поступово зменшуючись, стягуються до точки 0 (рис.6.18), де намагнічування дорівнює нулю. Гістерезис залежить значною мірою від складу феромагнетика і від його обробки. Для чистого «м'якого» заліза, тобто відпаленого і потім повільно охолодженого, гістерезис виражений дуже слабко і петля гістерезису дуже вузька. Але в загартованої сталі гістерезис значний (рис.6.19).

Температура Кюрі. Здатність пара– і феромагнетиків намагнічуватися різна за різних температур, тобто їх магнітна сприйнятливість залежить від температури. Вона зменшується із збільшенням температури. Навпаки, магнітна сприйнятливість діамагнетиків практично не залежить від температури.

Для багатьох *парамагнітних* речовин зміна χ з температурою підлягає закону, встановленому Кюрі (6.35). Магнітна сприятливість таких речовин монотонно змінюється із зміною температури. Такі речовини називаються *нормальними парамагнетиками*.

Залежність магнітної сприйнятливості від температури для феромагнетиків має складніший характер. При підвищенні температури здатність феромагнетиків намагнічуватися зменшується. При цьому зменшуються значення їх магнітної сприйнятливості і проникності при будь-якому значенні магнітного поля, ослаблюється гістерезис і зменшується намагнічення насичення $J_{нас}$. При деякій температурі T_K, яка називається *температурою Кюрі*, феромагнітні властивості зникають зовсім. Температура Кюрі різна для різних феромагнетиків; її значення для деяких речовин подані в таблиці.

Речовина	T_K, °C
Гадоліній	17
Залізо	770
Кобальт	1150
Нікель	360
30% -ний пермалой (сплав 70% Fe, 30% Ni)	70
78% -ний пермалой	550

При температурах вищих температури Кюрі, феромагнетик перетворюється на парамагнетик. Залежність магнітної сприйнятливості χ від температури для таких парамагнетиків підлягає закону Кюрі - Вейсса, який має вигляд:

$$\chi = \frac{C}{T - T_K}. \qquad (6.36)$$

Тут C – стала, яка залежить від роду речовини, а T_K – температура Кюрі.

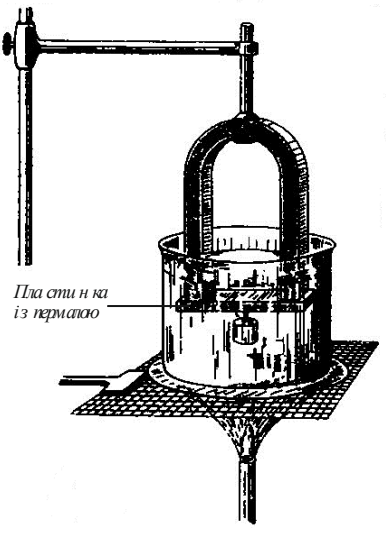


Рис.6.20.При намагнічуванні пермалоєвої пластинки вище температури Кюрі вона перестає притягуватися до магніту.

Як видно з таблиці, температура Кюрі для деяких сплавів дуже низька, наприклад для 30–процентного пермалою — близько 70°С. Тому, коли пермалоєву пластинку піднести до

полюсів постійного магніту за кімнатної температури, то вона притягується до магніту і втримує деякий вантаж. Але коли магніт з пластинкою помістити у воду, нагріту до температури 80—90°С, то пластинка втрачає феромагнітні властивості, її намагнічення практично зникає і вона під дією вантажу відривається від магніту (рис.6.20).

Робота при намагнічування

Розрахунок дає, що робота dw, яка потрібна для збільшення індукції на малу величину dB в одиниці об'єму магнетика, дорівнює

$$dw = HdB. \qquad (6.37)$$

Оскільки робота намагнічування залежить від процесів всередині магнетика, то вираз (6.37) буде справедливий, очевидно, для магнетика будь якої форми.

Розглянемо спочатку магнетик без гістерезису, для якого пряма і зворотна вітки кривої намагнічення збігаються (рис.6.21).

Елементарна робота, яка потрібна для нескінченно малого збільшення намагнічення, на цьому графіку виражається згідно з (6.37) величиною площі 12341. Повна робота, яка затрачається при збільшенні індукції від нуля до заданого значення B_1, дорівнює площі $0AB_10$, обмеженої кривою намагнічення і віссю B (заштрихована площа). Ця робота виконується джерелом при намагнічуванні магнетика. При розмагнічуванні магнетика в коло джерела повертається енергія, яка запасена в магнітному полі. Вона, як і раніше, виражається величиною площі, обмеженої кривою намагнічення і віссю *B*. Якщо гістерезису немає, то обидві вітки кривої збігаються і при розмагнічуванні повертається та сама робота, яка була затрачена при намагнічуванні.

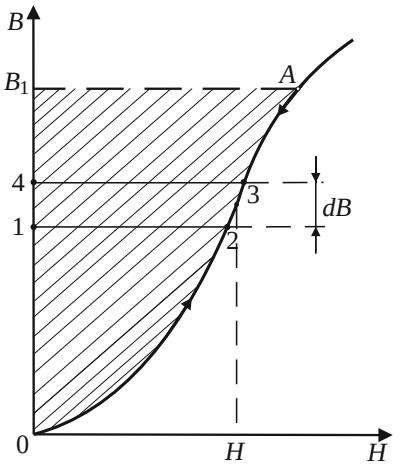


Рис.6.21.Робота намагнічування магнетика без гістерезису.

Інакше буде для магнетика з гістерезисом.У цьому разі при збільшенні індукції від значення B_1 до деякого іншого значення B_2 (рис.6.22) потрібна буде робота, яка дорівнює площі, яка обмежена віткою кривої намагнічення 1, тобто площі $B_1a1bB_2B_1$, а при розмагнічуванні до вихідного стану робота, яка повертається, дорівнюватиме площі $B_2b2aB_1B_2$, яка має меншу величину. Тому при повному циклі перемагнічування в кожну одиницю об'єму магнетика вводиться енергія

$$w = S, \qquad (6.38)$$

де S−площа петлі гістерезису, яка має розмірність одиниць роботи (енергії). Ця енергія витрачається на виконання роботи проти коерцитивних сил у магнетику і в кінцевому підсумку перетворюється в тепло. Тому феромагнетики при циклічному перемагнічуванні нагріваються і тим більше, чим сильніше виражений гістерезис.

В існуванні тепла гістерезису можна впевнитися за допомогою простого досліду. Якщо всередину дротяної котушки, яка живиться технічним змінним струмом, опустити стальний стержень, то він виконуватиме 50

циклів перемагнічування за секунду і вже через 1-2хв досліду значно нагрівається.

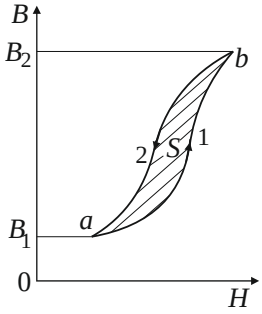


Рис.6.22.Робота при циклічному перемагнічуванні пропорціональна площі петлі гістерезису.

Тепло гістерезису завжди враховують при розрахунку різних електричних пристроїв змінного струму, які містять феромагнетики, що зазнають періодичного перемагнічування. Такі, наприклад, залізні осердя трансформаторів і обертові залізні якорі динамомашини. Наявність гістерезису в них веде до некорисної витрати частини енергії на тепло гістерезису і зменшує коефіцієнт корисної дії установок. Тому для таких пристроїв і застосовують спеціальні сорти м'якого заліза (трансформаторне залізо), в яких гістерезис виражений слабко.

§6.8.Магнітні матеріали. Ферити

У сучасній електротехніці феромагнетики відіграють значну роль. Використовуючи феромагнетики, ми примушуємо елементарні струми брати участь в утворенні магнітного поля і, так би мовити, «безкоштовно» збільшуємо магнітне поле в сотні і тисячі разів порівняно з полем одних намагнічуючих котушок. Залежно від призначення феромагнетиків до них ставлять різні вимоги. Так, наприклад, для застосування в трансформаторах найважливішими вимогами є висока магнітна проникність і слабкий гістерезис («магнітно-м'які» матеріали), а для виготовлення постійних магнітів — залишкове намагнічення і велика коерцитивна сила. Як феромагнітні матеріали тепер широко застосовують залізо і його сплави з іншими елементами. Підбираючи склад сплавів і варіюючи їх обробку, можна дістати різні феромагнітні матеріали, які відрізняються надзвичайною різноманітністю магнітних властивостей. У таблицях подано магнітні характеристики деяких речовин, які застосовуються в сучасній техніці.

Наведені дані показують, що у виготовленні магнітних матеріалів досягнуто значних успіхів. Деякі сплави (алніко, магніко) відзначаються дуже високими значеннями коерцитивної сили і залишкової індукції і тому дають змогу виготовляти виняткові за якістю

Матеріали магніто - м'які	Склад (%), останнє залізо і домішки	Відносна початкова проникність	Відносна максим. проникність	Коерцитивна сила, А/м	Залишкова індукція,Тл
Залізо	3,3 Si	600	10000	16	2,00
Залізо чисте	0,05 (дом.)	10000	200000	4	2,15
Залізо технічне	0,20 (дом.)	150	5000	80	2,15
Сталь м'яка	0,2 C	120	2000	140	2,12
Пермалой	78,5 Ni	8000	100000	4	1,08
Нікель	99 Ni	110	600	56	0,61
Ферит	Ni-Zn	1000	2000	8	0,25

Матеріали магніто - жорсткі	Склад(%), останнє залізо	Коерцитивна сила, А/м	Залишкова індукція,Тл
Сталь вуглецева	0,9C, 1Mn	4000	1,00
Сталь вольфрамова	0,4C, 6W	5200	1,05
Сталь кобальтова	1,0C, 3Co, 4Cr, 0,4Mn	6400	1,00
Алніко	19Ni, 10Al, 18Co, 3Cu	52000	0,90
Платина – залізо	78Pt	120000	0,60
Платина - кобальт	77Pt, 23Co	320000	0,50
Магніко	13,5Ni, 9Al, 24Co, 3Cu	56000	1,30

постійні магніти, які широко застосовують і тепер у магнітоелектричних вимірювальних приладах і інших пристроях, в яких потрібне потужне постійне магнітне поле.

Дуже цікаві магнітні властивості рідкоземельних елементів. Серед цієї групи елементів і їх хімічних сполук є діамагнетики, парамагнетики і феромагнетики. З них найсильнішим феромагнетиком є гадоліній. Він намагнічується до насичення в порівняно слабких полях і його намагнічення насичення більше, ніж у чистого заліза.

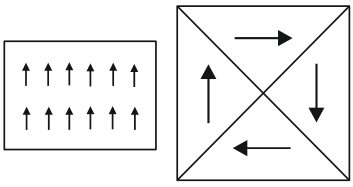
Дослідження магнітних сплавів і хімічних сполук привели і до важливих, і до цікавих результатів, які відкривають нові технічні можливості застосування магнітних матеріалів. Так, наприклад, було виявлено, що деякі сплави з неферомагнітних елементів при певному співвідношенні між компонентами мають великий феромагнетизм. Такими є сплави марганець — вісмут, марганець — сурма, хром—телур, тощо. Інше важливе досягнення полягає в добуванні і дослідженні *феритів*, які є феромагнітними хімічними сполуками типу $MeO \cdot Fe_2O_3$, де Me — один (або суміш) двовалентних катіонів Mn, Co, Ni, Cu, Mg, Zn, Cd, Fe^{+2}. На відміну від заліза та інших феромагнітних металів, ферити є магнітними напівпровідниками і мають великий питомий електричний опір порядку $1 \div 10^4$ Ом · м. Цим і зумовлена велике технічне значення феритів. Феромагнітні метали не можна використовувати в радіотехніці високих частот внаслідок їх великої електропровідності і великих втрат на вихрові струми, які виникають при цьому. Ферити цього недоліку не мають і дають змогу по-новому розв'язувати ряд завдань радіотехніки.

§6.9.Пояснення феромагнетизму

Сучасна теорія феромагнетиків спиралася в своєму розвитку на такі основні дослідні дані. *По-перше*, ми бачили в §6.7, що в деяких феромагнетиків можна змінювати намагнічення від початкового нульового значення до величезного значення насичення під дією мізерного намагнічуючого поля. Це є характерним для феромагнетиків і різко відрізняє їх від парамагнетиків. Зазначимо для порівняння, що нормальна парамагнітна сіль, наприклад $FeS0_4$, при кімнатній температурі під дією поля в 10 А/м збільшує своє намагнічення приблизно на 10^{-9} Тл, тобто на величину, у сотні мільйонів раз меншу, ніж деякі магнітно–м'які феромагнітні сплави.

Друга особливість стосується величини магнітного момента атомів феромагнітних речовин. Прямі досліди показують, що магнітні моменти атомів феромагнітних речовин мають той самий порядок величини, що й атоми парамагнетиків, і вимірюються небагатьма магнетонами Бора. Звідси випливає, що феромагнетизм не можна пояснити за допомогою теорії, подібної до теорії парамагнетизму, і, що феромагнітні властивості не зумовлені наявністю магнітного момента атома в цілому.

Третій важливий дослідний факт зв'язаний з величиною гіромагнітного відношення (§6.2). У феромагнетиках воно приблизно в два рази більше, ніж очікуване теоретично значення для електронних орбіт, і відповідає відношенню власних магнітного і механічного моментів електрона (формула (6.8)). Ця обставина вказує, що намагнічування феромагнетиків зумовлюється дуже значним орієнтуванням *власних магнітних моментів електронів* (електронних спінів), але не



а *б*

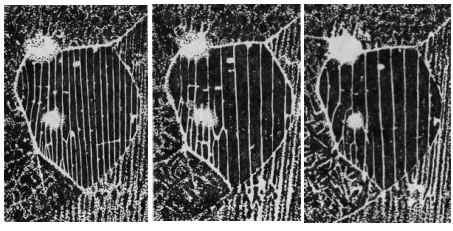
Рис.6.23.Схематичне зображення:
а – розміщення електронних спінів при спонтанному намагнічуванні (низька температура); *б* – напрямки намагнічення в окремих доменах.

магнітних моментів атомів в цілому. За сучасними уявленнями, які надійно обґрунтовані на досліді, суть феромагнетизму полягає в тому, що значне орієнтування елементарних магнітних моментів виникає у феромагнетиках *незалежно від зовнішнього магнітного поля,* бо феромагнетик намагнічений до насичення, яке відповідає даній температурі, уже без усякого поля (рис.6.23,*а*).

Наявність такого *самодовільного*, або *спонтанного, намагнічування* є найхарактернішою властивістю феромагнетиків. Спонтанне намагнічування виникає через те, що на елементарні магнітні моменти діють сили молекулярної взаємодії не магнітного походження. Ці уявлення були висловлені в працях П. Вейсса в 1907 р.

Дослід показує, що феромагнетики при відсутності зовнішнього поля можуть бути і не намагнічені. Для пояснення цієї позірної суперечності Вейсс зробив другу основну гіпотезу, за якою феромагнетик розбивається на велику кількість малих (але макроскопічних) областей, або *доменів* (визначення поняття "домен" приводиться в §2.3). Кожна з цих областей за температур, нижчих за температуру Кюрі, намагнічена значною мірою, але напрями намагнічування в окремих доменах різні, а саме такі, що повний магнітний момент феромагнетика дорівнює нулю (рис.6.23,*б).*

Питання про фізичну причину спонтанного намагнічування принципово розв'язали в 1928 р. Френкель, а потім Гайзенберг, які показали, що значне орієнтування електронних спінів спричинюється так званими силами *обмінної взаємодії.* Наявність цього нового класу сил, який не пояснюється в класичній фізиці, було з'ясовано тільки з розвитком квантової фізики атома. Існування доменів у феромагнетиках тепер доведено різними дослідами. Найпряміший метод полягає в добуванні так званих *порошкових фігур*. Якщо на добре відполіровану поверхню феромагнетика помістити шар рідини, в якій є завислі найдрібніші крупинки феромагнітного порошку (наприклад, Fe_2O_3), то ці крупинки осідатимуть переважно на ті місця, поблизу яких магнітне поле неоднорідне. Але саме поблизу меж доменів і виникають неоднорідності поля і тому осілий порошок позначить межі областей спонтанного



а *б* *в*

Рис.6.24.Домени в деформованому кременистому залізі: *а*−без поля; *б*−магнітне поле перпендикулярне до площини рисунка; *в*−магнітне поле тієї самої напруженості, але протилежного напряму.

намагнічування. На рис.6.24 наведені фотографії порошкових фігур, які видно в мікроскоп при невеликому збільшенні. Особливо слід зазначити, що домени справді спостерігаються і за відсутності зовнішнього

магнітного поля. Порошкові фігури були використані в численних дослідженнях феромагнетиків і дали змогу визначити розміри, форму і розміщення доменів, а також зміни доменів у зовнішньому магнітному полі.

Розглянемо тепер докладніше, чому у феромагнетиках виникають численні домени. Для простоти вважатимемо, що феромагнетик дуже анізотропний і що його вісь легкого намагнічування перпендикулярна до поверхні зразка. На рис.6.25,*а* зображено феромагнетик, який складається з одного домена. У цьому випадку в зовнішньому просторі виникає магнітне поле, яке містить у собі певну енергію. На рис.6.25,*б* є два домени з протилежним напрямом намагнічування.

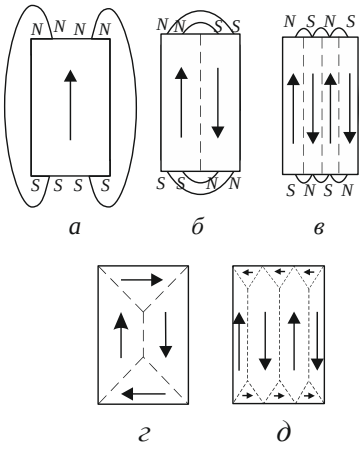


Рис.6.25.Пояснення виникнення доменів.

Зовнішнє магнітне поле тут зменшується із збільшенням відстані швидше, ніж у випадку *а*, і енергія, яка є в полі, менша. У випадку рис.6.25,*в* магнітне поле практично існує тільки в безпосередній близькості від поверхні магнетика і енергія поля ще швидше зменшується. На рис.6.25,*г* зображено такий випадок, коли в зовнішньому просторі магнітного поля немає зовсім. Тут є «замикаючі» домени у формі тригранних призм, бічні поверхні яких скрізь утворюють кут 45° з вектором намагнічення. Внаслідок цього магнітний потік проходить виключно всередині феромагнетика, він замикається граничними доменами, чим і зумовлена їх назва — «замикаючі домени». Стан *г* енергетично вигідніший, ніж попередні стани, і тому феромагнетик, який перебуває, наприклад, у стані *а*, прагнутиме перейти в стан *г*. Нарешті, на рис.6.24,*д* зображено цілу сукупність доменів разом з доменами, що їх замикають, у якій також немає зовнішнього потоку. Така форма доменів справді спостерігається на досліді. Отже, феромагнетик на домени розбивається тому, що при утворенні доменних структур енергія феромагнетика зменшується.

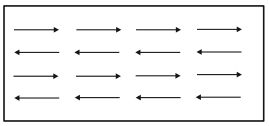
Вище ми розглядали феромагнетик за відсутності зовнішнього магнітного поля і враховували тільки його власну магнітну енергію. Якби були ще й інші джерела енергії, наприклад механічні напруги або зовнішне магнітне поле, то форма доменів визначалася б повною енергією системи. Тому за наявності механічних напруг і зовнішнього поля доменна структура змінюється. При виникненні спонтанного намагнічування (орієнтування електронних спінів) магнетик *деформується.* Якщо при температурі вищій, ніж температура Кюрі, вирізати з монокристала феромагнетика кулю, то при охолодженні нижче температури Кюрі куля перетвориться в еліпсоїд. Форма і розміри доменів змінюються і при перемагнічуванні. Тому феромагнетик в цілому не намагнічений і при намагнічуванні деформується.

Явище деформації при намагнічуванні відкрив у середині минулого століття Джоуль; воно називається *магнітострикцією*. Деформації, які виникають при цьому, дуже малі: відносні видовження зразка $\Delta l/l$ в полях ~

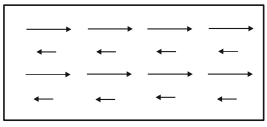
10^5 А/м звичайно мають порядок $10^{-5} \div 10^{-6}$. Магнітострикцію використовують, як і п'єзоелектричний ефект (§2.6) для побудови потужних випромінювачів ультразвукових хвиль та інших цілей.

Сили обмінної взаємодії спричинюють у феромагнетиках *паралельне* орієнтування електронних спінів. Однак обмінні сили залежать від структури тіла і тому характер орієнтування спінів, який спричинюється ними, може бути різний. Виявляється, що існують речовини, в яких також виникає значне орієнтування електронних спінів, але, на відміну від феромагнетиків, електронні спіни орієнтовані в них *попарно антипаралельно.* У найпростішому випадку електронні спіни утворюють ніби дві просторові підрешітки, вставлені одна в одну і намагнічені в протилежних напрямах (рис.6.26,*а*).

Речовини, в яких намагнічення обох підрешіток однакове за величиною, називаються *антиферомагнетиками.* Їх існування теоретично завбачив Л. Д. Ландау в 1933 р. Антиферомагнетиками є деякі сполуки марганцю (MnO, MnS), хрому (NiCr, Cr_20_3), ванадію (VO_2) тощо. Такі речовини за низьких температурах мають дуже малу магнітну сприйнятливість. При підвищенні температури точна попарна антипаралельність електронних спінів порушується і магнітна сприйнятливість збільшується. При деякій температурі *(антиферомагнітна точка Кюрі*, або *точка Нєєля)* області спонтанного орієнтування електронних спінів руйнуються і антиферомагнетик перетворюється в парамагнетик. При дальшому підвищенні температури магнітна сприйнятливість, як у всякого парамагнетика, зменшується, а отже, при антиферомагнітній температурі Кюрі магнітна сприйнятливість має максимум.



а



б

Рис.6.26.Характер намагнічування в антиферомагнетиках (*а*) і в феритах (*б*) (схематично).

Якщо намагнічення обох підрешіток не однакове за величиною, то виникає нескомпенсований антиферомагнетизм і речовина може набути значного магнітного моменту. Такий характер намагнічування спостерігається у *феритах* (рис.6.26,*б*).

§6.10.Процеси намагнічування феромагнетика

Одним із основних завдань теорії феромагнетизму є пояснення технічної кривої намагнічування, тобто залежність намагнічення J від напруженості поля H, або індукції B всередині феромагнетика від H. Розглянемо полікристалічний звичайний кусок заліза. Всередині нього є велика кількість маленьких кристаликів, кристалічні осі яких направлені в різні сторони. Ці кристалики – це не те ж саме, що домени.

Раніше ми вважали, що домени є частиною одного кристалу, а в куску заліза, як видно з рис.6.27, містяться багато різних кристалів з різною орієнтацією намагнічення. В кожному із цих кристалів, взагалі кажучи, міститься декілька доменів. Коли до куска полікристалічного матеріалу прикласти слабе магнітне поле, доменні бар'єри в кристаликах

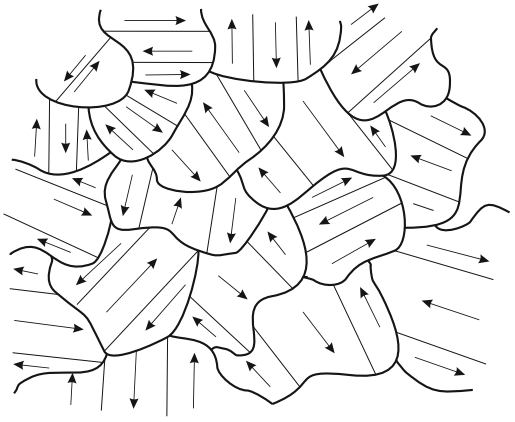


Рис.6.27.Мікроструктура ненамагніченого полікристалічного феромагнітного матеріалу. Кожний кристалик має напрямок легкого намагнічування і розбивається на домени, які звичайно спонтанно намагнічені в цьому напрямку.

починають зміщуватись і домени, напрямок намагніченості яких співпадає з напрямком легкого намагнічування, ростуть все більше і більше. До тих пір поки поле залишається дуже малим, цей ріст оборотний; якщо поле виключити, намагніченість знову повернеться до нуля. Ця ділянка кривої намагнічування позначена на рис.6.28 буквою a.

Для більших полів в області, яка на рис.6.28 позначена буквою b, все стає більш складним. В кожному маленькому кристалику матеріалу є напруги і дислокації (порушення кристалічної гратки), домішки і дефекти. І при всіх полях, за винятком дуже слабких, стінки доменів при своєму русі наштовхуються на них.

Між доменною стінкою і дислокацією (або домішкою) виникає взаємодія. В результаті, коли стінка наштовхується на перешкоду, вона ніби приклеюється і тримається так, поки поле не досягне певної величини. Коли поле дещо зросте, стінка раптово зривається. Таким чином, рух доменної стінки є не плавним, як в ідеальному кристалі; вона рухається скачкоподібно, зупиняючись на мить. Якщо розглянуть криву намагнічування в мікроскопічному масштабі, то побачили б щось подібне як на вставці рис.6.28. Найваждивіше в тому, що ці скачки намагнічування викликають втрату енергії. Перш за все, коли стінка домена проскакує накінець перешкоду, вона дуже швидко рухається до наступної. Швидкий рух спричиняє за собою і швидку зміну магнітного поля, що в свою чергу створює в кристалі вихрові струми (див.наступний розділ). Останні витрачають енергію на нагрівання металу.

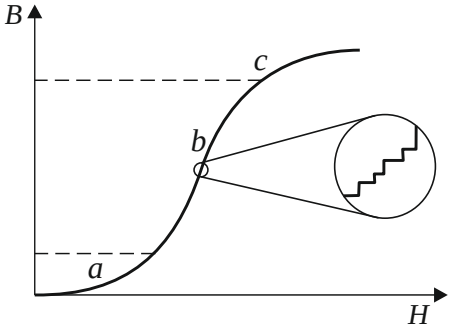


Рис.6.28.Крива намагнічування полікристалічного заліза.

Інший ефект в тому, що, коли домен раптово змінюється, частина кристалів із-за магнітострикції змінює свої розміри. Кожний раптовий зсув доменної стінки створює невелику звукову хвилю, яка теж виносить енергію. Через ці ефекти ця частина кривої намагнічування є необоротною, має місце втрата енергії. Це і є причиною гістерезисного ефекту, або рух скачками вперед – одне, а рух назад – вже інше, і при русі в обидва кінці втрачається енергія.

Зрештою, при достатньо сильних полях, коли всі доменні стінки зсунуті і намагніченість кожного кристалика направлена

по найближчій до поля осі легкого намагнічування, залишаються ще деякі кристалики, у яких вісь легкого намагнічування далека від напрямку зовнішнього магнітного поля. Щоб повернути ці магнітні моменти необхідне ще додаткове поле. Таким чином, в сильних полях саме в області, яка позначена на рис.6.28 буквою c, намагніченість зростає повільно, але гладко. Намагнічення не відразу досягає свого насичення, оскільки в цій останній частині кривої відбувається доорієнтація атомних магнітиків в сильному полі. Тепер зрозуміло, чому крива намагнічення полікристалічного матеріалу має вигляд, показаний на рис.6.28. Такий характер залежності *B(H)* дає можливість пояснити залежність $\mu(H)$, який приведений на рис.6.16. Для цього на рис.6.29,*а* зображено ще раз основна крива намагнічення. Проведемо з початку координат пряму лінію, яка проходить через довільну точку кривої *B(H)*. Тангенс кута нахилу цієї прямої пропорційний відношенню *B/H,* тобто магнітній проникності (див.формулу (6.19)) для відповідного значення напруженості поля. При збільшеності H від нуля кут нахилу (а значить і μ) спочатку зростає. В точці 2 він досягає максимуму (пряма 0-2 є дотичною до кривої *B(H)),* а потім зменшується. На рис.6.29,*б* приведений графік залежності μ від H. Із рисунка видно, що максимальне значення проникності досягається дещо раніше, ніж насичення. При необмеженому зростанні H проникність асимптотично наближається до одиниці. Із формул (6.13) і (6.19) легко отримати вираз для μ, а саме

$$\mu = 1 + J/H. \qquad (6.37)$$

Оскільки намагнічення J не може перевищувати $J_{нас}$ (рис.6.15), то при $H \to \infty, \mu \to 1$.

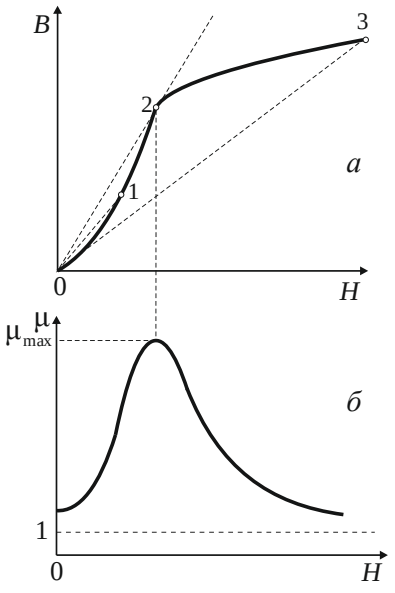


Рис.6.29.

Контрольні питання

1.Які речовини називаються магнетиками?

2.За допомогою яких фізичних величин проводять класифікацію магнетиків?

3.Провести класифікацію магнетиків.

4.Дати визначення вектора намагнічення речовини.

5.Які існують характеристики магнітного поля в речовині?

6.Який існує зв'язок між вектором напруженості і вектором індукції магнітного поля в феромагнетиках?

7.Схематично зобразити залежність індукції магнітного поля від напруженості магнітного поля у феромагнетиках. Пояснити цю залежність.

8.Записати співвідношення між складовими векторів B і H на межі двох магнетиків.

9.Який фізичний зміст магнітної проникності феромагнетика?

10.Чому значення магнітної проникності для різних феромагнетиків не приводяться в довідникових таблицях?

11.Дати визначення коерцитивної сили та залишкової намагніченості для феромагнетика.

12.Які феромагнетики мають назву магніто-м'яких та магніто-жорстких ?

13.Де в техніці зв'язку застосовують феромагнетики?

Розділ 7.Електромагнітна індукція

§7.1.Явище електромагнітної індукції

У попередніх розділах вивчались властивості статичних, незмінних в часі електричних та магнітних полів. Переходячи в область динамічних електромагнітних явищ, ми зустрічаємось з цілим рядом якісно нових властивостей полів та їхніх закономірностей. В основі електродинаміки змінних полів лежить фундаментальний закон електромагнітної індукції, відкритий дослідно у 1831р. видатним англійським вченим М. Фарадеем. Суть цього закону полягає у виникненні струму в довільному замкненому провідному контурі при зміні магнітного потоку, що пронизує цей контур.

Відкриття зв'язку між електрикою та магнетизмом стало визначною подією в історії розвитку електромагнетизму, оскільки ці явища раніше вважались абсолютно взаємонезалежними. Спочатку у 1820р. Х. Ерстед відкрив, що навколо провідників зі струмом виникає магнітне поле, а трохи пізніше в цьому ж році А. Ампер встановив, що на провідники зі струмом у магнітному полі діє сила. Результати цих досліджень стали основою припущення, що й електричне поле можна отримати за допомогою магнітного. Для перевірки цієї гіпотези були проведені різноманітні експерименти. Один із них полягає у спробі виявити струм у замкненому провіднику, розміщеному біля іншого провідника, по якому проходить постійний електричний струм. У своїх ранніх дослідах Фарадей з подивом виявив, що постійний струм не впливає на контур, який близько до нього лежав. Він виготовляв різні провідні котушки, намотуючи два проводи так, щоб їхні витки розташовувались дуже щільно один до одного, але були розділені ізолюючим шаром тканини або паперу. Одна з котушок складала із гальванометром замкнене коло. Через іншу котушку пропускався великий струм від батареї. На розчарування Фарадея, стрілка гальванометра не відхилялась. Але в одному з дослідів він помітив дуже слабкий рух стрілки гальванометра при включені струму, та такий самий рух при виключені. Продовжуючи ці досліди він скоро встановив, що струми в інших провідниках наводяться не постійним, а змінним струмом. Однією з блискучих експериментальних знахідок в дослідах Фарадея того періоду була заміна гальванометра (він зрозумів, що гальванометр був недостатньо чутливим приладом для виявлення короткочасних імпульсів струму) простою маленькою котушкою, в яку помістив ненамагнічену стальну стрілку (рис.7.1).

Фарадей виявив, що стрілка намагнічувалась імпульсом індукованого

струму при включенні первинного струму і що напрямок намагнічування був протилежним, якщо його причиною було виключення струму у первинному контурі. Так було відкрито явище електромагнітної індукції.

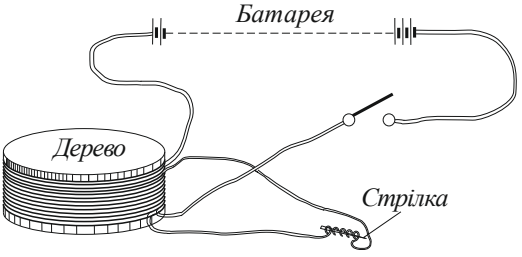


Рис.7.1.

Розглянемо деякі досліди, що ілюструють *електромагнітну* індукцію. Для цього скористаємось двома дротяними котушками 1 і 2 (рис.7.2), одну з яких (1)

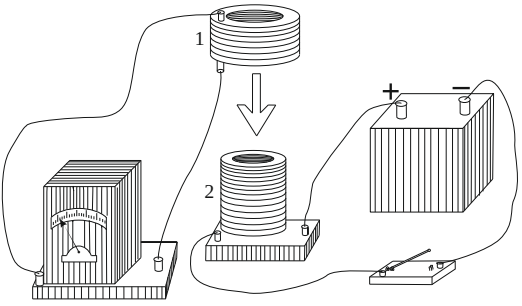


Рис.7.2

можна надівати на іншу (2). Сполучимо котушку 1 з гальванометром, а котушку 2 з джерелом струму. Якщо котушка 1 нерухома відносно котушки 2 (тобто відносно магнітного поля), то в колі котушки 1 не буде струму, яким би сильним не будо магнітне поле котушки 2. Почнемо тепер переміщувати котушку 1. Ми побачимо, що гальванометр покаже появу струму. Цей струм існує тільки при русі котушки, і він тим більший, чим швидше рухається котушка. Струм у колі котушки 1 виникає і при зближені котушок, і при віддаленні їх, однак струми в обох випадках мають протилежні напрями. Якщо залишити котушку 1 нерухомою, а рухати котушку зі струмом 2, то гальванометр теж показує струм при русі котушки.

Ми могли б залишити обидві котушки 1 та 2 нерухомими, але змінювати величину струму в котушці 2 за допомогою реостата. Тоді при всякій зміні величини струму (тобто магнітного поля котушки 2) у колі котушки 1 виникає струм. Напрям струму в котушці 1 при підсиленні магнітного поля протилежний напряму струму при ослабленні поля; струм у котушці 1 не виникає, якщо магнітне поле залишається постійним. Ці досліди показують, що причиною появи індукційного струму є *зміна магнітного поля.*

Те, як створюється ця зміна значення не має. У досліді, зображеному на рис.7.3 обидві котушки 1 та 2 нерухомі, а в котушку 2 ми всуваємо, або висуваємо з неї залізне

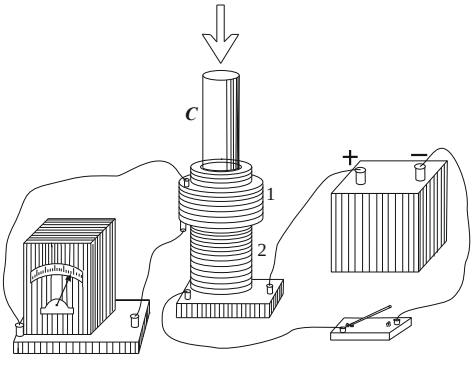


Рис.7.3.

осереддя ***C***. При всуванні осердя воно намагнічується, і магнітне поле підсилюється; при висуванні осердя поле зменшується. Струм у колі котушки 1 проходить тільки при русі осердя. Змінне магнітне поле ми можемо створити і рухом постійного магніту. Якщо зовсім забрати котушку зі струмом 2 і всувати

(або висувати) в котушку 1 постійний магніт, то гальванометр також показує струм. Цей струм виникає і тоді, коли рухається котушка 1, а магніт перебуває у спокої. Спостерігаючи напрями струму при зближенні магніту і котушки і при віддаленні їх, можна впевнитися, що вони протилежні, як і в попередніх дослідах.

Результати своїх численних дослідів Фарадей подав у такій наочній формі. Зображатимемо магнітне поле за допомогою ліній магнітної індукції. Тоді величина магнітної індукції поля характеризуватиметься густотою ліній індукції. Уявімо собі тепер, що замкнений провідник рухається в магнітному полі і переходить в область сильнішого поля. Тоді кількість ліній індукції, які охоплює провідник, збільшиться. Навпаки, при русі провідника в область слабшого поля число ліній індукції, які охоплює провідник зменшиться. Але магнітне поле є поле *вихрове* (§5.6), і його лінії індукції завжди замкнені. Внаслідок цього лінії індукції поля зчеплені з дротяним контуром, як ланки ланцюга (рис.7.4).

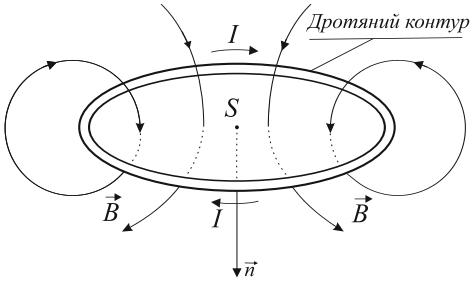


Рис.7.4.

Тому всяка зміна кількості ліній індукції, які охоплює контур, може статися тільки внаслідок *перетину* ними дротяного контуру. Точно так само, коли провідник перебуває у спокої, але змінюється величина магнітної індукції. То при підсиленні поля густота ліній індукції збільшуватиметься і вони стягуватимуться одна до одної, а при ослабленні поля – розходитимуться одна від одної. І в цьому випадку провідник перетне деяке число ліній індукції. Тому Фарадей зробив висновок, що *індукційний струм виникає в провіднику тоді, коли провідник, або яка-небудь його частина перетинає лінії магнітної індукції.* Можемо зробити такий висновок, що виникнення індукційного струму завжди пов’язане із зміною магнітного потоку через поверхню, яку охоплює провідник. Ці зміни можуть відбуватися з різних причин, зокрема через:

– переміщення постійного магніту, відносно нерухомого провідника;

– переміщення контуру, відносно нерухомого магніту;

– замикання та розмикання струму в обмотці нерухомого електромагніту, розміщеного поблизу провідника;

– відносне переміщенні контуру і електромагніту;

– зміну магнітної індукції поля електромагніту (виймання осердя при сталому струмі в обмотці або зміну струму реостатом);

– зміну комутатором напрямку струму в обмотці електромагніту;

– постійний рух контуру в неоднорідному магнітному полі;

– обертальний рух контуру в однорідному магнітному полі.

Відкриття електромагнітної індукції мало величезне наукове і технічне значення. Це явище показало, що можна не тільки створити магнітне поле за допомогою струмів, а й, навпаки, мати електричні струми за допомогою магнітного поля. Цим було остаточно встановлено взаємний зв'язок між електричними і магнітними явищами. Завдяки

явищу електромагнітної індукції можна перетворювати механічну енергію в електричну, а також передавати електричну енергію з одного кола в інше.

Технічне значення цього відкриття полягає в тому, що на явищі електромагнітної індукції ґрунтується дія всіх потужних джерел струму (промислових генераторів постійного та змінного струму, трансформаторів тощо), без яких сучасний розвиток електротехніки був би неможливим.

§7.2 Закон Ленца

Фарадей встановив, що напрямок індукційного струму в провіднику залежить від характеру зміни (збільшення, чи зменшення) магнітного потоку ($\Delta\Phi>0$, чи $\Delta\Phi<0$) через його контур. Якщо при внесенні постійного магніту в котушку стрілка гальванометра відхиляється в один бік, то при вийманні магніту вона відхиляється в інший бік.

Загальне правило, за допомогою якого можна визначити напрямок індукційного струму в замкненому провіднику, сформулював Е.Х. Ленц:

Індукційний струм у замкненому провіднику завжди має такий напрямок, що створений ним струмом власний магнітний потік протидіє тим змінам зовнішнього магнітного току, які збуджують індукційний струм.

Стисло закон Ленца можна подати так: *індукційний струм у всіх випадках напрямлений так, що його дія протилежна дії причини, яка зумовила виникнення цього струму.* Пояснимо закон Ленца на прикладі. Зрівноважимо прямий магніт (рис.7.5) так, щоб він міг переміщуватися з малим тертям у вертикальному напрямі. Розмістимо під магнітом дротяну котушку, так щоб магніт міг вільно входити в неї і сполучимо котушку з гальванометром. Швидко всунемо тепер магніт у котушку, переміщаючи його рукою згори вниз. При цьому в колі котушки виникне індукційний струм і гальванометр дасть відкид.

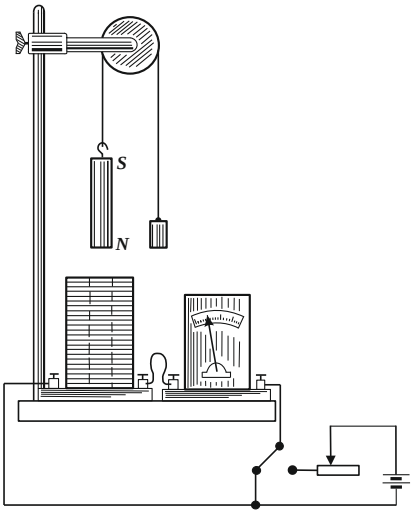


Рис.7.5.Демонстрування закону Ленца

Запам’ятавши напрям індукційного струму (напрям відхилення стрілки), розімкнемо коло і, залишивши магніт всередині котушки, приєднаємо кінці проводів до батареї елементів (гальванометр перед цим треба зашунтувати, бо струм від батареї буде значно більший, ніж індукційний). Якщо струм в котушці має той самий напрям, який мав індукційний струм, то магніт *виштовхується* з котушки.

Закон Ленца справджується і в тих випадках, коли провідники нерухомі, а змінюється магнітне поле (величина стуму). У цьому випадку індукційні струми завжди спричиняють таке поле, яке прагне протидіяти змінам зовнішнього поля, які спричинили ці струми. Так, наприклад, якщо в досліді (рис.7.2) котушки нерухомі, то при вмиканні струму в котушці 2 (його наростанні) напрям

струму в котушці 1 буде протилежний (індукційний струм прагне послабити наростаюче поле котушки 2), а при вимиканні струму в котушці 2 (його зменшенні) струм у котушці 1 буде напрямлений так само, як і в котушці 2, прагнучи підтримати магнітне поле, яке послаблюється.

Закон Ленца випливає із закону збереження енергії. Справді, індукційні струми, як і всякі електричні струми, виконують певну роботу. Це означає, що при русі замкненого провідника в магнітному полі повинна бути виконана додаткова робота зовнішніх сил. Ця робота і виникає тому, що індукційні струми, взаємодіючи з магнітним полем, спричиняють сили, напрямлені протилежно рухові, тобто перешкоджають йому переміщатись.

§7.3.Основний закон електромагнітної індукції

Індукційний струм I_i у замкненому провіднику, опір якого R, виникає під дією е.р.с. $\mathcal{E}_i$, яку можна виразити за законом Ома:

$$\mathcal{E}_i = I_i R.$$

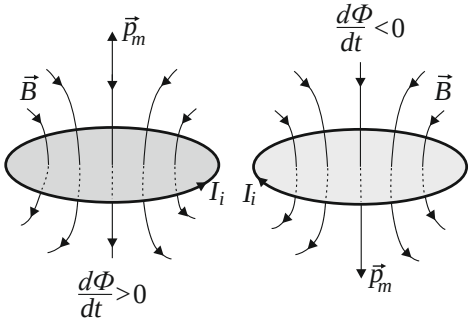
Оскільки та сама е.р.с. у провідниках з різним опором створює неоднакові струми, то для кількісної характеристики явища електромагнітної індукції зручніше користуватися величиною $\mathcal{E}_i$, а не силою індукційного струму I_i.

Внаслідок численних дослідів Фарадей встановив основний кількісний закон електромагнітної індукції. Однак ми розглянемо спочатку інше формулювання цього закону, яке дав згодом Максвелл.

У §7.1 ми бачили, що індукційний струм, а отже, і е.р.с. індукції, виникають тільки тоді, коли провідник *перетинає* лінії магнітної індукції, тобто коли повне число ліній індукції, що проходять через площу, обмежену дротяним контуром, змінюється (рис.7.4). Але певне число ліній магнітної індукції, що проходять через яку-небудь поверхню, створює магнітний потік через цю поверхню. Звідси можна зробити висновок, що причина виникнення е.р.с. електромагнітної індукції є зміна *магнітного потоку*. Аналізуючи результати дослідів Фарадея, Максвелл встановив, що у всіх випадках е.р.с. *електромагнітної індукції пропорційна швидкості зміни магнітного потоку через площу, обмежену контуром.* Або, інакше:

$$\mathcal{E}_i = k d\Phi/dt, \qquad (7.1)$$

де $\mathcal{E}_i$ – е.р.с. індукції , Φ– магнітний потік, t – час, а k – множник пропорційності, який залежить від вибору системи одиниць. Е.р.с. електромагнітної індукції вважають позитивною, якщо магнітний момент $\vec{p}_m$ відповідного їй індукційного струму утворює гострий кут з лініями магнітної індукції того поля, яке наводить цей струм (рис.7.6).



а *б*

Рис.7.6.

Тоді у випадку, зображеного на рис.7.6,*а* $\mathcal{E}_i<0$, а для зображеного на рис.7.6,*б* $\mathcal{E}_i>0$. В системі СІ k= 1 і

$$\mathcal{E}_i = -d\Phi/dt. \qquad (7.2)$$

Знак " – " є математичним виразом правила Ленца. Ця формула, яка об'єднує

закони Фарадея і Ленца, є математичним виразом основного закону електромагнітної індукції, яку дав Максвелл: *електрорушійна сила електромагнітної індукції в замкненому контурі числово дорівнює і протилежна за знаком швидкості зміні магнітного потоку крізь поверхню, обмежену контуром*. Якщо е.р.с. індукції виникає при зміні магнітного потоку, який пронизує котушку з *N* витків, то її величина буде відповідно в *N* разів більшою, ніж для одного витка, тобто:

$$\mathcal{E}_i = -Nd\Phi/dt = -d\psi/dt, \qquad (7.3)$$

де $\Psi = N\Phi$ – повний магнітний потік через котушку, що носить назву *потокозчеплення*.

На основі закону електромагнітної індукції можна дати визначення одиниці магнітного потоку вебера: 1Вб – це такий магнітний потік, при зменшенні якого до нуля, протягом 1с в колі, яке він пронизував, виникає е.р.с. індукції в 1В.

Формула (7.2) виражає закон електромагнітної індукції в загальній формі. Вона застосовна як до нерухомих контурів, так і до рухомих провідників у магнітному полі. Похідна від магнітного потоку по часу, що входить до неї, у загальному випадку, складається з двох частин, одна з яких зумовлена зміною магнітної індукції з часом, а друга – рухом контура відносно магнітного поля (або його деформацією).

§7.4.Пояснення закону електромагнітної індукції

1.Виведення закону електромагнітної індукції із закону збереження енергії.

Німецький фізик Гельмгольц показав, що основний закон електромагнітної індукції – закон Фарадея – є наслідком закону збереження енергії.

Помістимо в однорідне магнітне поле з індукцією $\vec{B}$ провідник *CD* завдовжки *l* з опором *R*, який є ділянкою замкненого кола *CDKLC* (рис.7.7). Опором з'єднувальних провідників *DK, KL* і *LC* нехтуємо. Провідник *DC* рухається без тертя по ділянках *DK* і *LC*.Джерело постійного струму з е.р.с. $\mathcal{E}$ і внутрішнім опором *r* створює в замкненому колі за законом Ома струм *I:* $I = \frac{\mathcal{E}}{R+r}$. При цьому на провідник діє сила Ампера (5.7)

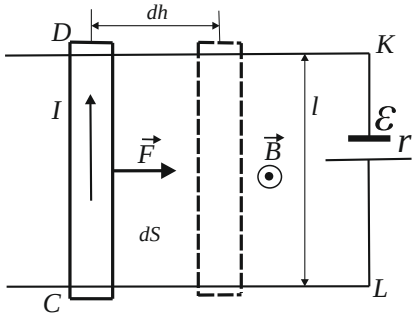


Рис.7.7.

$F = BIl$, внаслідок чого провідник переміщуватиметься прискорено і за час dt пройде шлях dh. При проходженні струму I в замкненому колі *CDKLC* протягом часу dt відбуваються такі процеси:

а) виконується робота по подоланню електричного опору $R+r$ кола:

$$dA_1 = I^2(R+r)dt;$$

б) виконується механічна робота при переміщенні провідника на шляху dh:

$$dA_2 = Fdh = IBld = IBdS = Id\Phi.$$

За законом збереження енергії сума робіт dA_1 і dA_2 повинна дорівнювати повній роботі джерела струму

$$dA = \mathcal{E}dq = \mathcal{E}Idt,$$

тобто

$$I^2(R+r)dt + Id\Phi = \mathcal{E}Idt.$$

Звідси:

$$I = \frac{\mathcal{E} - \frac{d\Phi}{dt}}{R+r}.$$

Ця формула визначає закон Ома для замкненого кола, яке поміщене в магнітне поле. Роль діючої е.р.с. в такому колі відіграє вираз:

$$\mathcal{E} - \frac{d\Phi}{dt} \quad \text{або} \quad \mathcal{E} + \left(-\frac{d\Phi}{dt}\right).$$

Додатковий член $(-\frac{d\Phi}{dt})$, який відрізняється від нуля тільки у разі зміни магнітного потоку через поверхню dS повинен мати розмірність е.р.с.. Він і визначає е.р.с. індукції $\mathcal{E}_i$: $\mathcal{E}_i = -\frac{d\Phi}{dt}$.

2.Пояснення закону електромагнітної індукції за допомогою сили Лоренца

Розглянемо контур з рухомою перемичкою 1-2 (рис.7.8). Розмістимо його в однорідне магнітне поле, яке перпендикулярне до площини контуру Г і направлене за рисунок.

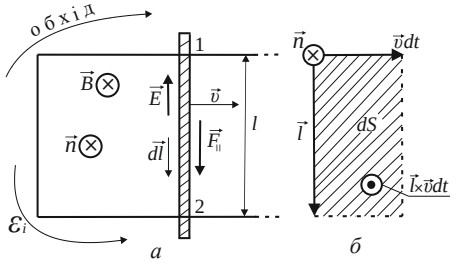


Рис.7.8.

Приведемо перемичку в рух із швидкістю $\vec{v}$. З тією ж швидкістю відносно поля стануть переміщуватись і носії струму в перемичці – електрони. В результаті на кожний електрон почне діяти направлена вздовж перемички магнітна сила

$$\vec{F}_{\parallel} = -e\vec{v} \times \vec{B}. \qquad (7.4)$$

(див.§5.10, формула (5.47)); заряд електрона дорівнює $-e$). Дія цієї сили еквівалентна дії на електрон, електричного поля $\vec{E}$ (рис.7.8,*а*), напруженість якого

$$\vec{E} = \vec{v} \times \vec{B}. \qquad (7.5)$$

Це поле не електростатичного походження. Згідно (4.25) його циркуляція по контуру Г дає величину е.р.с., яка індукується в контурі:

$$\mathcal{E}_i = \oint_\Gamma \vec{E}d\vec{l} = \oint_\Gamma \vec{v}\times\vec{B}d\vec{l} = \int_1^2 \vec{v}\times\vec{B}d\vec{l} \qquad (7.6)$$

(підінтегральна функція відмінна від нуля тільки на ділянці перемички 1-2, яка рухається).

Будемо вважати $\mathcal{E}_i$ позитивною в тому випадку, коли її напрямок утворює з напрямком нормалі $\vec{n}$ до контуру Г правогвинтову систему.

Виберемо обхід по контуру Г за годинниковою стрілкою, тоді нормаль $\vec{n}$ буде мати напрямок , як показано на (рис.7.8 *а*). При розрахунку циркуляції необхідно обходити контур за годинниковою стрілкою і відповідно вибрати напрямок вектора $d\vec{l}$. Якщо в (7.6) винести постійний вектор $\vec{v} \times \vec{B}$ за знак інтегралу, то отримаємо:

$$\mathcal{E}_i = \vec{v} \times \vec{B} \int_1^2 d\vec{l} = \vec{v} \times \vec{B} \cdot \vec{l},$$

де $\vec{l}$ – вектор, який показаний на (рис.7.8,*б*.). Здійснимо в отриманому виразі циклічну перестановку векторів (множників), після чого помножимо і розділимо його на dt:

$$\mathcal{E}_i = \vec{v} \times \vec{B} \cdot \vec{l} = \vec{B}\frac{\vec{l}\times\vec{v}dt}{dt}. \qquad (7.7)$$

Із (рис.7.8,б) видно, що $\vec{l} \times \vec{v}\, dt = -\vec{n}dS$, де dS – приріст площі контуру Г за час dt. За визначенням магнітного потоку вираз

$$\vec{B}d\vec{S} = \vec{B}\vec{n}dS$$

є потік через площадку dS, тобто приріст потоку $d\Phi$ через контур Г.

Тоді

$$\vec{B}\cdot\vec{l}\times\vec{v}\,dt=-\vec{B}\vec{n}dS=-d\Phi.$$

Із врахуванням цього виразу (7.7) перепишеться так:

$$\mathcal{E}_i=-\frac{d\Phi}{dt}. \qquad (7.7')$$

Таким чином, ми отримали закон (7.2) за допомогою сили Лоренца. $\frac{d\Phi}{dt}$ і $\mathcal{E}_i$ мають протилежні знаки. Знак потоку і знак $\mathcal{E}_i$ зв'язані з вибором напрямку обходу по контуру Г (напрямком нормалі). На рис.7.8,*а* знак $\frac{d\Phi}{dt}$ позитивний, а знак $\mathcal{E}_i$ – негативний. Якби ми вибрали обхід проти годинникової стрілки (нормаль направлена на нас), то знак $\frac{d\Phi}{dt}$ був би негативний, а знак $\mathcal{E}_i$ – позитивний.

Зауважимо, що в міркуваннях, які привели нас до формули (7.7'), роль сторонніх сил, які підтримували струм в контурі, грають магнітні сили. Робота цих сил над одиничним позитивним зарядом, яка дорівнює за визначенням е.р.с., відмінна від нуля. Ця обставина начебто суперечить твердженню про те, що магнітна сила роботи над зарядами не виконує. Це протиріччя зникає, якщо врахувати, що сила (7.4) є не повна магнітна сила, що діє на електрон, а лише її складова, що діє паралельно проводу і обумовлена швидкістю $\vec{v}$ (сила $\vec{F}_{\parallel}$ рис.7.9).

Під дією сили $\vec{F}_{\parallel}$ електрон починає рухатись вздовж проводу із швидкістю $\vec{v}_{\parallel}$, в результаті чого виникає перпендикулярна до проводу інша складова магнітної сили:

$$\vec{F}_{\perp}=-e\vec{v}_{\parallel}\times\vec{B}$$

(ця складова не вносить вкладу в циркуляцію, оскільки перпендикулярна до $d\vec{l}$). Повна магнітна сила, яка діє на електрон

$$\vec{F}=\vec{F}_{\parallel}+\vec{F}_{\perp}.$$

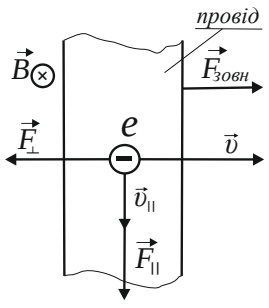


Рис.7.9.

Робота цієї сили над електроном за час dt

$$dA=\vec{F}_{\parallel}\vec{v}_{\parallel}dt+\vec{F}_{\perp}\vec{v}dt=F_{\parallel}v_{\parallel}dt-F_{\perp}vdt=$$
$$=-evv_{\parallel}Bdt+evv_{\parallel}Bdt=0.$$

При цьому враховано вирази для сил $\vec{F}_{\parallel}$ і $\vec{F}_{\perp}$, а також те, що напрямок векторів $\vec{F}_{\parallel}$ і $\vec{v}_{\parallel}$ однакові, а векторів $\vec{F}_{\perp}$ і $\vec{v}$ – протилежні.

Таким чином, робота повної магнітної сили над зарядом дорівнює нулю.

Наведене нами пояснення виникнення е.р.с. індукції належить до випадку, коли магнітне поле постійне, а змінюється геометрія контуру. Однак магнітний потік через контур може змінюватись також внаслідок зміни $\vec{B}$. В цьому випадку виникнення е.р.с. є принципово іншим і це головне у відкритті Фарадея.

3. Пояснення закону електромагнітної індукції за допомогою вихрового електричного поля

Розглянемо випадок електромагнітної індукції, коли провідний контур, в якому індуктується струм, нерухомий, а зміна магнітного потоку обумовлена зміною магнітного поля. Виникнення індукційного струму свідчить про те, що зміна магнітного

поля викликає появу в контурі сторонніх сил, які діють на носії заряду. Ці сторонні сили не пов'язані ні з хімічними, ні з тепловими процесами в проводі; вони не можуть бути і магнітними силами, тому що такі сили роботи над зарядами не виконують. Залишається стверджувати, що індукційний струм обумовлений електричним полем, яке виникає в проводі. Позначимо напруженість цього поля $\vec{E}_в$ (індекс $в$ означає – вихрове). Згідно (4.25), циркуляція цього поля по замкнутому контуру Г дорівнює е.р.с. що діє в цьому контурі, тобто:

$$\mathcal{E}_i = \oint_Г \vec{E}_в \, d\vec{l}. \qquad (7.8)$$

Якщо враховувати закон (7.2) і вираз $\Phi = \int_S \vec{B} d\vec{S}$, то (7.8) можна переписати так:

$$\oint_Г \vec{E}_в \, d\vec{l} = -\frac{d}{dt}\int_S \vec{B} d\vec{S},$$

де інтеграл в правій частині береться по довільній поверхні *S*, яка опирається на контур Г (рис.7.10).

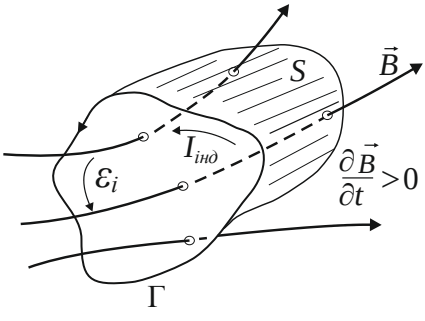


Рис.7.10.

Оскільки контур і поверхня нерухомі, операцію диференціювання по часу і інтегрування по поверхні можна поміняти місцями:

$$\oint_Г \vec{E}_в \, d\vec{l} = -\int_S \frac{\partial \vec{B}}{\partial t} d\vec{S}. \qquad (7.9)$$

У зв'язку з тим, що вектор $\vec{B}$ залежить в загальному випадку від часу і координат, ми записали під знаком інтеграла символ частинної похідної по часу $\frac{\partial}{\partial t}$ (інтеграл $\int_S \vec{B} d\vec{S}$ є функція тільки від часу). Формула (7.9) виражає закон електромагнітної індукції (7.2) записаний в іншій формі.

Аналізуючи явище електромагнітної індукції, Максвелл зробив висновок, що причина виникнення е.р.с. індукції полягає у виникненні *електричного поля* (рис.7.11). *При*

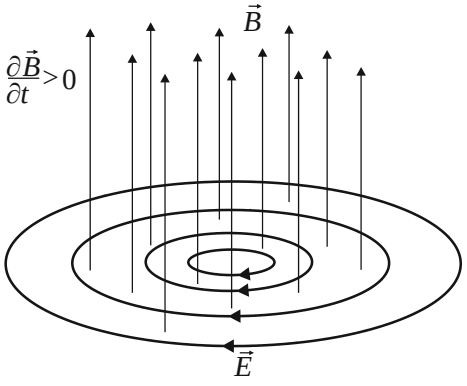


Рис.7.11.При зміні магнітного поля виникає вихрове електричне поле. Зображений на рисунку напрям $\vec{E}$ відповідає зростанню $\vec{B}$.

цьому провідники відіграють другорядну роль і є тільки своєрідним приладом, який виявляє це поле. Під дією поля електрони провідності в дротині приходять у рух; коли дротина замкнута, у ній виникає індукційний струм. Істотна особливість цього явища та, що електричне поле, яке виникає, не є електростатичним. Силові лінії електростатичного поля завжди розімкнені. Вони починаються і закінчуються на електричних зарядах, відповідно до цього циркуляція вектора $\vec{E}$ в електростатичному полі завжди дорівнює нулю. Через це електростатичне поле не може підтримувати замкнутий рух зарядів і отже, не може

призвести до виникнення електрорушійної сили. Навпаки, електричне поле, яке виникає при електромагнітній індукції, має замкнені силові лінії, тобто є вихровим полем. Таке поле спричинює в дротині рух електронів по замкнутих траєкторіях і веде до виникнення електрорушійної сили; при цьому сторонніми силами є сили вихрового електричного поля. Циркуляція вектора $\vec{E}$ по замкнутому контуру в цьому полі не дорівнює нулю; її значення між двома якими-небудь точками вже не визначається тільки положенням цих точок, як це було у випадку електростатичного поля, але залежить ще від форми контура (провідника), який сполучає ці точки. Таким чином, електричне поле може бути як потенціальним, так і вихровим. В загальному випадку, електричне поле може складатись із поля $\vec{E}_q$, яке створюється зарядами, і поля $\vec{E}_в$, яке обумовлене змінним магнітним полем. Сумарне поле $\vec{E} = \vec{E}_q + \vec{E}_в$, але

$$\oint_\Gamma \vec{E}_q d\vec{l} = 0.$$

Тому співвідношення (7.9) можемо переписати так:

$$\oint_\Gamma \vec{E} d\vec{l} = -\int_S \frac{d\vec{B}}{dt} d\vec{S}\,. \qquad (7.10)$$

Це рівняння є одним із основних в електромагнітній теорії Максвелла.

Існування взаємозв'язку між електричним і магнітним полями (що виражається наприклад рівнянням (7.10)) є причиною того, що окремий розгляд електричного і магнітного полів має тільки відносний зміст. Дійсно, електричне поле створюється системою нерухомих зарядів. Однак, якщо заряди нерухомі відносно деякої інерціальної системи відліку, то відносно інших інерціальних систем, ці заряди рухаються, а отже породжують не лише електричне, а й магнітне поле. Нерухомий провід з постійним струмом створює в кожній точці простору постійне магнітне поле. Однак відносно інших інерціальних систем цей провід знаходиться в русі. Тому магнітне поле в будь-якій точці з даними координатами *x,y,z,* яке він створює, буде змінюватись з часом, і, отже, породжувати вихрове електричне поле. Таким чином, поле яке відносно деякої системи відліку є «чисто» електричним, або «чисто» магнітним, відносно інших систем відліку буде являти собою сукупність електричного і магнітного полів, які утворюють єдине електромагнітне поле.

§7.5 Самоіндукція. Індуктивність

Явище електромагнітної індукції спостерігається в усіх випадках, коли змінюється магнітний потік, який пронизує контур. Зокрема, цей потік може створюватися струмом, який проходить в самому розглядуваному контурі. Тому при всякій зміні величини струму в якому-небудь контурі в ньому виникає е.р.с. індукції, яка спричиняє додатковий струм у контурі. Це явище називається *самоіндукцією,* а додаткові струми, які спричинюються е.р.с. самоіндукції – *екстраструмами* самоіндукції.

На рис.7.12 подано схему досліду для спостереження екстраструмів. Котушка, яка має кілька тисяч витків дроту, увімкнена в коло, що містить батарею $\mathcal{E}$, реостат *R* і вимикач *K*. Паралельно до котушки приєднаний "грубий" гальванометр Г. При замкненому ключі струм батареї розгалужуються: його частина *I* проходить через котушку, а частина I_1 — через гальванометр. Якщо розімкнути ключ, то магнітний потік у котушці зникатиме і в ній виникне екстраструм самоіндукції I_s *(ек-*

страструм розмикання). Згідно з законом Ленца, він перешкоджатиме спаданню магнітного потоку, тобто буде напрямлений у котушці так само, як і спадний струм. Цей екстраструм проходить повністю через гальванометр (ключ *K* розімкнутий), де його напрям протилежний початковому струму I_1. Тому гальванометр дає відкид у зворотний бік. При замиканні ключа (установленні струму) в котушці також виникає екстраструм *(екстраструм замикання).* Його напрям у

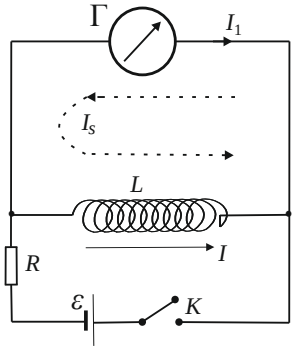


Рис.7.12.Спостереження екстраструму самоіндукції.

котушці протилежний зростаючому струму батареї. Однак у випадку замикання екстраструм розподіляється між батареєю і гальванометром і, крім того, його напрям у гальванометрі — такий самий, як і в наростаючого струму батареї. Тому екстраструм замикання помітно значно гірше.

Якщо помістити в котушку залізне осердя, то екстраструми значно збільшуються. У цьому разі гальванометр можна замінити невеликою лампочкою розжарювання. При розмиканні ключа лампочка дає яскравий спалах.

Розглянемо, від чого залежить е.р.с. самоіндукції. Нехай котушка *L* (соленоїд) має *N* витків, по яких проходить струм *I* від зовнішнього джерела. Цей струм створює в соленоїді магнітне поле, яке зосереджене в його об'ємі і це поле можна вважати однорідним. Нехай весь об'єм магнітного поля соленоїда заповнений однорідною речовиною з відносною магнітною проникністю μ. Якщо змінювати струм в соленоїді, то власний магнітний потік *Ф*, який пронизує його поперечний переріз, теж змінюватиметься. Магнітний потік крізь площу *S*, що обмежена одним витком, дорівнює:

$$\Phi = BS = \mu\mu_0 nSI = \mu\mu_0 \frac{NS}{l} I,$$

де l – довжина соленоїда, $B = \mu\mu_0 H$,

$$H = nI = \frac{N}{l} I.$$

За законом Фарадея:

$$\mathcal{E}_s = -N\frac{d\Phi}{dt} = -\frac{d}{dt}(\mu\mu_0 \frac{N^2 S}{l} I).$$

Припустимо, що сили, які деформують соленоїд, дуже малі і тому l і S залишаються сталими. Якщо всередині соленоїда знаходиться феромагнітне осердя, то магнітний потік *Ф* змінюється не лише внаслідок зміни сили струму *I,* але і від зміни μ. Позначимо:

$$L = \mu\mu_0 \frac{N^2 S}{l}. \qquad (7.11)$$

Тоді:

$$\mathcal{E}_s = -\frac{d}{dt}(LI) = -L\frac{dI}{dt} - I\frac{dL}{dt}.$$

Якщо магнітна проникність середовища μ не змінюється, то L=const і

$$\mathcal{E}_s = -L\frac{dI}{dt}. \qquad (7.12)$$

Е.р.с. самоіндукції прямо пропорційна швидкості зміни сили струму в контурі. Параметр *L* характеризує індивідуальні властивості контуру. Його називають *коефіцієнтом індуктивності,* або просто *індуктивністю* контуру. Індуктивність контуру залежить від форми контуру, його розмірів та відносної магнітної проникності середовища, в якому він знаходиться.

З виразів (7.3) і (7.12) можемо записати, що

$$LdI = d\psi_s,$$

де ψ_s – потокозчеплення самоіндукції.

Після інтегрування ця рівність має вигляд:

$$LI = \psi_s.$$

Отже, індуктивністю замкненого контуру називається скалярна величина L. яка дорівнює відношенню потокозчеплення самоіндукції контуру ψ_s до сили струму I в цьому контурі:

$$L = \frac{\psi_s}{I}. \qquad (7.13)$$

Одиниця індуктивності — генрі (Гн): 1Гн – індуктивність такого контуру, потокозчеплення самоіндукції якого при струмі 1А дорівнює 1Вб.

Формулу (7.12) можна використати для провідників довільної форми, якщо знайти потокозчеплення $\psi_s = LI$, яке зв'язане з цим провідником при силі струму *I*. У виразі для е.р.с. самоіндукції знак "–" зумовлений правилом Ленца і показує, що наявність індуктивності контура приводить до сповільнення зміни струму в ньому. Тобто індуктивність контуру є мірою його інертності відносно зміни струму.

Знайдемо закон зміни струму *I* в електричному колі, показаному на рис.7.12, у випадку відсутності гальванометра під час вмикання і вимикання з нього джерела е.р.с. $\mathcal{E}$. Сила струму *I* в колі за законом Ома дорівнює:

$$I = \frac{\mathcal{E} + \mathcal{E}_s}{R} = \frac{\mathcal{E} - L\frac{dI}{dt}}{R},$$

якщо зовнішнє магнітне поле постійне.

Розділимо змінні в цьому рівнянні:

$$IR = \mathcal{E} - L\frac{dI}{dt}, \quad \frac{dI}{dt} = \frac{\mathcal{E} - IR}{L},$$

$$\frac{dI}{\mathcal{E} - IR} = \frac{dt}{R}, \qquad -\frac{d(\mathcal{E} - IR)}{\mathcal{E} - IR} = \frac{dt}{R}.$$

Вважаючи $\mathcal{E}$, *R*, *L* сталими, проінтегруємо останній вираз:

$$\int \frac{d(\mathcal{E} - IR)}{\mathcal{E} - IR} = -\frac{1}{R}\int dt,$$

$$\ln(\mathcal{E} - IR) = -\frac{t}{R} + \ln C,$$

де C – довільна стала інтегрування.

Потенціюючи дане рівняння, отримаємо:

$$\mathcal{E} - IR = Ce^{-\frac{R}{L}t}.$$

В початковий момент часу t = 0, сила струму дорівнює I_0. Тоді: $\mathcal{E} - I_0R = C$.

В результаті:

$$\mathcal{E} - IR = (\mathcal{E} - I_0R)e^{-\frac{R}{L}t}.$$

Звідси знайдемо вираз для сили струму:

$$IR = I_0Re^{-\frac{R}{L}t} + \mathcal{E} - \mathcal{E}e^{-\frac{R}{L}t}$$

і

$$I = I_0e^{-\frac{R}{L}t} + \frac{\mathcal{E}}{R}(1 - e^{-\frac{R}{L}t}).$$

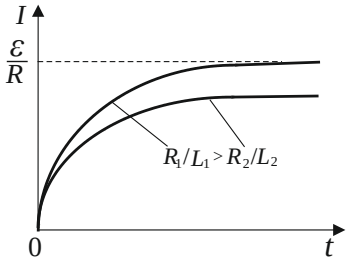


Рис.7.13.

Для випадку вмикання джерела е.р.с. $\mathcal{E}$, коли початковий струм $I_0 = 0$, отримуємо вираз для екстраструму замикання:

$$I = \frac{\mathcal{E}}{R}(1 - e^{-\frac{R}{L}t}). \qquad (7.14)$$

Струм у колі поступово збільшується від нуля при $t = 0$, до значення $\mathcal{E}/R$, яке відповідає величині постійного струму (рис.7.13). Зростання струму відбувається тим швидше, чим більше відношення *R/L*.

Для випадку вимкнення джерела ($\mathcal{E}$=0)

$$I = I_0e^{-\frac{R}{L}t}. \qquad (7.15)$$

Струм у колі поступово зменшується від початкового значення I_0 до нуля за експоненціальним законом. Струм зменшується тим швидше, чим більше значення *R/L*, тобто чим більший опір кола і чим менша його індуктивність (рис.7.14).

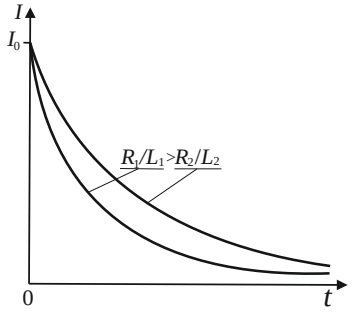


Рис. 7.14.

§7.6. Взаємна індукція

Розглянемо тепер два контури з струмом, наприклад два колові витки 1 і 2 (рис.7.15). Частина ліній індукції поля, яка створюється контуром 1, проходитиме через контур 2, тобто буде зчеплена з цим контуром. І навпаки, певна кількість ліній індукції поля, створюваного контуром 2, буде зчеплена з контуром 1. У цьому випадку ми говоримо, що між обома контурами існує *магнітний зв'язок*. Індукція поля контура 1 пропорціональна величині струму I_1 в цьому контурі. Тому магнітний потік $\Phi_{1,2}$ через контур 2, який створюється контуром 1, також пропорціональний струму I_1:

$$\Phi_{12} = L_{12} I_1. \qquad (7.16)$$

Коефіцієнт $L_{1,2}$ називається *коефіцієнтом, взаємної індукції* контурів 1 і 2. Він, очевидно, дорівнює магнітному потоку через контур 2, що створюється контуром 1 при величині струму в ньому, яка дорівнює одиниці. З порівняння (7.16) з (7.13) видно, що розмірність L_{12} – така ж, як і розмірність

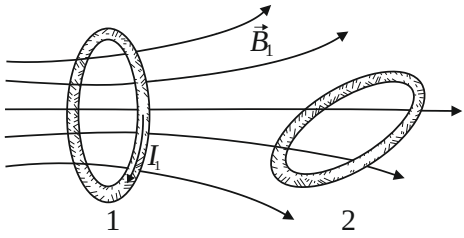


Рис.7.15.Магнітний зв'язок двох контурів.

індуктивності, і тому коефіцієнт взаємної індукції вимірюється тими самими одиницями, що й індуктивність. Так само, коли в контурі 2 є струм деякої величини I_2, то він створює магнітний потік Φ_{21} через контур 1, причому

$$\Phi_{21} = L_{21} I_2. \qquad (7.16, a)$$

Тут L_{21} є коефіцієнт взаємної індукції контурів 2 і 1.

Можна показати, що для будь-яких двох контурів коефіцієнти взаємної індукції завжди однакові:

$$L_{12} = L_{21}. \qquad (7.17)$$

Наявність магнітного зв'язку між контурами проявляється в тому, що при всякій зміні величини струму в одному з контурів у другому контурі виникає е.р.с. індукції. За основним законом електромагнітної індукції (§7.3) маємо:

$$\mathcal{E}_{i2} = -\frac{d\Phi_{12}}{dt} = -L_{12}\frac{dI_1}{dt}, \qquad \mathcal{E}_{i1} = -\frac{d\Phi_{21}}{dt} = -L_{12}\frac{dI_2}{dt}, \qquad (7.18)$$

де $\mathcal{E}_{i2}$ — е.р.с. індукції, яка виникає в контурі 2, а $\mathcal{E}_{i1}$ — е.р.с. індукції в контурі 1.

Коефіцієнт взаємної індукції залежить від форми і розмірів контурів, і від їх взаємного розміщення. Він залежить також від властивостей навколишнього середовища.

Розглянемо простий приклад

визначення коефіцієнта взаємної індукції. Нехай є дві тороїдні одношарові котушки 1 і 2, які щільно прилягають одна до одної (рис.7.16). У цьому випадку всі лінії індукції поля, які створюються однією котушкою, проходять і через другу котушку. Індукція магнітного поля котушки 1 дорівнює:

$$B_1 = \mu_0 n_1 I_1 = \mu_0 \frac{N_1}{l} I_1,$$

де $N_1 -$ кількість витків котушки 1, $l -$ довжина тороїда по середній лінії. Це поле створює крізь один виток котушки 2 магнітний потік

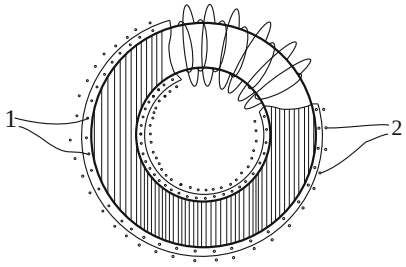


Рис.7.16.Дві обмотки з магнітним зв'язком.

$$B_1 S = \frac{\mu_0 N_1 I_1}{l} S,$$

де S – площа перерізу котушок. Повний потік крізь усі N_2 витків котушки 2 буде:

$$\Phi_{12} = \frac{\mu_0 N_1 N_2 S}{l} I_1 \cdot$$

Звідки для коефіцієнта взаємної індукції дістаємо вираз:

$$L_{12} = \frac{\mu_0 N_1 N_2 S}{l} \cdot \qquad (7.19)$$

Якщо обчислити магнітний потік, який створюється котушкою 2 крізь котушку 1 , то дістанемо:

$$B_2 = \mu_0 \frac{N_2 I_2}{l}, \qquad \Phi_{2,1} = \frac{\mu_0 N_2 N_1 S}{l} I_2 \cdot$$

Звідси ми знаходимо для коефіцієнта взаємної індукції L_{21} попередній вираз (7.19), згідно з формулою (7.17).

Коли всередині котушок є осердя з речовини з магнітною проникністю μ, магнітний потік збільшується в μ раз і коефіцієнт взаємної індукції буде в μ раз більший.

§7.7. Власна енергія струму

Щоб збільшити струм в електричному колі, необхідно виконати деяку роботу. Цю роботу виконує джерело струму, увімкнене в коло. Навпаки, при всякому зменшенні струму в колі вивільнюється деяка енергія і джерело струму виконує меншу роботу, ніж при постійному струмі. Розглянемо це питання докладніше.

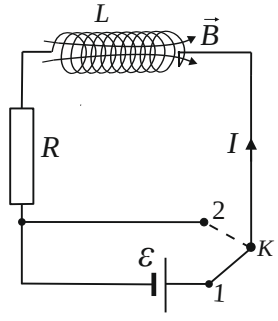


Рис.7.17.

Повернемось до кола, зображеного на рис.7.17, що містить індуктивність L і має повний опір R, і припустимо спочатку, що в ньому є постійний струм, який встановився. Величина цього струму визначається електрорушійною силою джерела струму $\mathcal{E}$ і опором кола:

$$I = \frac{\mathcal{E}}{R} \cdot$$

За час dt в колі виділиться тепло Ленца — Джоуля $RI^2 dt$. Оскільки $RI = \mathcal{E}$, то

$$RI^2 dt = \mathcal{E} I dt = \mathcal{E} dq = dA.$$

Але права частина цієї рівності виражає роботу джерела струму за той самий час dt. Ми бачимо, що робота джерела точно дорівнює теплу Ленца — Джоуля, а отже, щоб підтримати постійне магнітне поле котушки, ніякої роботи виконувати не потрібно.

Уявимо собі тепер, що струм у колі збільшується зі швидкістю dI/dt. У цьому разі в колі виникає ще е. р. с. самоіндукції $\mathcal{E}_s$, яка спричинить екстраструм

$$\delta I = \frac{|\mathcal{E}_s|}{R} = \frac{L}{R}\frac{dI}{dt}.$$

Він направлений проти струму I і тому повна величина струму в колі буде $(I - \delta I)$.

У подальшому ми вважатимемо, що процес збільшення струму відбувається дуже повільно ($\delta I \ll I$) і при розрахунках враховуватимемо тільки члени першого порядку малості.

У розглядуваному процесі за час dt виділиться тепло Ленца—Джоуля $dQ' = R(I - \delta I)^2 dt$. Воно менше порівняно з випадком постійного струму на величину

$$dQ' - dQ = R(I - \delta I)^2 dt - RI^2 dt = \\ = -2RI\delta I dt = -2LI dI.$$

Тут dI є збільшення струму в колі за час dt.

За цей самий час джерело струму виконає роботу $dA' = \mathcal{E}(I - \delta I)dt$ і тому батарея $\mathcal{E}$ виконає меншу роботу:

$$dA' - dA = \mathcal{E}(I\text{-}\delta I)dt - \mathcal{E}I dt = -\mathcal{E}\delta I dt = \\ = -LI dI.$$

Отже, у випадку наростаючого струму робота джерела струму більша за кількість тепла, яке виділилося (в алгебраїчному розумінні). Надлишок роботи джерела знайдемо так:

$$(dA' - dA) - (dQ' - dQ) = LI dI.$$

Тому $dW = LI dI$ і є та робота, яка потрібна для збільшення величини струму в колі від значення I до $I + dI$. Повна робота, яка потрібна для встановлення струму I дорівнює:

$$W = L\int_0^I I dI = \frac{1}{2}LI^2. \qquad 7.20)$$

При вимиканні джерела струму робота W виділяється в колі; її виконують екстраструми розмикання. Тому вираз (7.20) дає енергію, яку запасає контур зі струмом. Вона називається *власною енергією струму*. Саме ця енергія і проявлялася в досліді §7.5 (рис.7.12) у вигляді відкиду стрілки гальванометра і спалаху лампочки від екстраструмів розмикання.

Корисно зіставити вирази для власної енергії контура зі струмом і енергії зарядженого конденсатора (§3.4): $q^2/(2C)$ (q– заряд конденсатора, C – його ємність). Енергія конденсатора пропорціональна квадрату *заряду,* а енергія струму пропорціональна квадрату *величини струму,* тобто залежить від *швидкості* руху зарядів.

У механіці є два види енергії: потенціальна і кінетична. Потенціальна енергія стисненої пружини дорівнює $kx^2/2$, де x — зміщення кінця пружини, а k — її пружність (жорсткість), а кінетична енергія рухомого тіла є $mv^2/2$, де m — маса тіла, а v — його швидкість. Розвиваючи аналогію між електричними і механічними явищами, ми бачимо, що енергія конденсатора відповідає потенціальній енергії в механіці, а власна енергія струму — кінетичній енергії. При цьому величина $1/C$, обернена ємності, аналогічна пружності (жорсткості) пружини k, а індуктивність L — масі тіла m.

§7.8.Енергія магнітного поля

Будь-який електричний струм завжди оточений магнітним полем. Тому можна запитати, де саме зосереджена, локалізована, власна енергія струму: всередині проводів, де рухаються електричні заряди, чи в магнітному полі, тобто в середовищі, яке оточують струми?

Відповідь на це запитання може дати тільки дослід. Однак, поки ми маємо справу з постійними струмами, такі досліди неможливі, бо в цьому випадку струми завжди оточені магнітним полем і, навпаки, магнітне поле завжди супроводиться струмам, які його підтримують. Це стосується і постійних магнітів, які також мають замкнені молекулярні струми (див. розд.6).

Відповісти на це запитання можна тоді, коли дослідити змінні магнітні поля або електромагнітні хвилі. В електромагнітних хвилях ми маємо магнітні поля, які змінюються в просторі і часі і здатні існувати без струмів, які їх підтримують. Оскільки електромагнітні хвилі містять у собі і переносять певну енергію, то звідси ми робимо висновок, що *енергія зосереджена в магнітному полі.*

Знайдемо величину енергії, яка міститься в одиниці об'єму магнітного поля. Розглянемо замкнену тороїдну котушку. Її індуктивність (формула 7.11) є

$$L = \mu\mu_0 \frac{N^2 S}{l},$$

де μ – магнітна проникність осердя тороїда, а μ_0 – магнітна стала вакууму. Підставимо цей вираз у (7.20) і отримаємо:

$$W = \frac{1}{2}\mu\mu_0 \frac{N^2 S}{l} I^2.$$

Але врахувавши співвідношення (5.18) та (6.19) отримаємо, що:

$$H = nI = \frac{N}{l} I$$

є напруженість поля всередині котушки. Тому

$$W = \frac{1}{2}\mu\mu_0 H^2 V,$$

де $V = Sl$ – об'єм котушки. Ми бачимо, що енергія однорідного магнітного поля пропорціональна об'єму V, який займає поле. Тому енергія одиниці об'єму поля, або *об'ємна густина енергії* магнітного поля дорівнює:

$$\omega = \frac{\mu\mu_0 H^2}{2}.$$

Оскільки індукція магнітного поля (формула (6.19) $B = \mu\mu_0 H$, то

$$\omega = \frac{BH}{2} = \frac{B^2}{2\mu\mu_0} = \frac{\mu\mu_0 H^2}{2}. \quad (7.21)$$

Якщо магнітне поле неоднорідне, то його можна розбити на нескінченно малі елементи об'єму dV у кожному з яких поле можна вважати однорідним. Енергія, яка є в елементі об'єму dV буде ωdV. Повна енергія будь-якого магнітного поля дорівнює:

$$W = \int_V \omega dV = \frac{1}{2}\int_V BH dV, \quad (7.22)$$

де інтегрування поширюється на весь об'єм V зайнятий магнітним полем.

§ 7.9.Взаємна енергія двох струмів

Обчислимо енергію магнітного поля, створеного двома контурами із струмом. Розглянемо найпростіший випадок двох тороїдних котушок (рис.7.16), в яких магнітне поле однорідне, що перебувають у вакуумі ($\mu = 1$). Напруженість сумарного поля всередині котушок:

$$H = H_1 \pm H_2 = \frac{N_1 I_1 \pm N_2 I_2}{l}.$$

Вибір знака " +" або "−" залежить від того, однакові чи протилежні напрями струмів в обох

котушках. В одиниці об'єму цього поля є енергія $\frac{1}{2}\mu_0 H^2$, а повна енергія, яка міститься в усьому об'ємі поля $V = Sl$, дорівнює:

$$W=\frac{1}{2}\mu_0 H^2 Sl=\frac{\mu_0 S}{2l}\cdot$$
$$\cdot[(N_1 I_1)^2+(N_2 I_2)^2\pm 2(N_1 I_1)(N_2 I_2)].$$

Скориставшись виразами для індуктивностей L_1 і L_2 обох котушок (7.11) і їхнім коефіцієнтом взаємної індукції L_{12} (7.19), знайдений результат можна подати в такому вигляді:

$$W = \frac{1}{2}L_1 {I_1}^2 + \frac{1}{2}L_2 {I_2}^2 \pm L_{12} I_1 I_2. \quad (7.23)$$

Перший член цієї формули дає власну енергію струму I_1 (7.20), а другий доданок виражає власну енергію струму I_2. Формула показує, що енергія двох струмів у контурах, які мають магнітний зв'язок, відрізняється від суми власних енергій струмів на величину

$$W_{12} = \pm L_{12} I_1 I_2. \quad (7.24)$$

Вона називається *взаємною енергією* двох струмів.

Зі сказаного видно, що причина виникнення взаємної енергії струмів полягає в тому, що в полі кількох контурів зі струмом додаються напруженості поля (принцип накладання, або суперпозиції полів), тоді як енергія магнітного поля пропорціональна квадрату напруженості, а квадрат суми не дорівнює сумі квадратів. Уявімо собі, що ми спочатку створюємо струм I_1 в контурі 1, а струм $I_2 = 0$. Для цього потрібна буде робота $\frac{1}{2}L_1 {I_1}^2$. Створюючи потім струм у контурі 2, ми виконаємо роботу, яка згідно (7.23) дорівнює $\frac{1}{2}L_2 {I_2}^2 \pm L_{12} I_1 I_2$. Якби ми спочатку створили струм I_2 (а струм I_1 дорівнював би нулю), то для цього потрібна була б робота $\frac{1}{2}L_2 {I_2}^2$. Але тоді наступне створення струму I_1 потребувало б роботи $\frac{1}{2}L_1 {I_1}^2 \pm L_{12} I_1 I_2$. Ми бачимо, що робота утворення струму в контурі в присутності другого контура зі струмом не дорівнює роботі утворення того самого струму у відокремленому від інших контурів. Вона відрізняється на величину взаємної енергії обох струмів W_{12}.

Ми дістали формули (7.23) і (7.24), розглядаючи окремий випадок тороїдних котушок. Однак легко впевнитися, що цей результат справедливий для будь-яких контурів довільної форми, які перебувають у будь-якому середовищі.

§7.10.Вихрові струми

Якщо в змінному магнітному полі розміщений який-небудь масивний провідник, то вихрове електричне поле спричинює в цьому провіднику індукційний струм. Густина цього струму в якій-небудь точці провідника за законом Ома дорівнює $\vec{j} = \sigma\vec{E}$. Оскільки лінії напруженості $\vec{E}$ замкнені, то й лінії струму також замикаються всередині провідника, внаслідок чого такі струми називаються *вихровими струмами.*

Вихрові струми спричинюють нагрівання провідників. Якщо всередині котушки зі змінним струмом помістити провідне тіло, наприклад металевий диск, орієнтований перпендикулярно до осі котушки, то диск можна розжарити до високої температури і розплавити.

Нагрівання провідників вихровими струмами застосовують в індукційних металургійних печах для плавлення металів і приготування їхніх сплавів, в лабораторній практиці для розжарювання металів, а також для приготування їжі та інших цілей. Вихрові

струми виникають також під час руху масивних провідників у магнітному полі. Взаємодіючи з магнітним полем, вихрові струми спричинюють виникнення сил, що діють на рухоме провідне тіло, які за законом Ленца завжди *протидіють* руху. Ця гальмівна дія дає змогу також легко виявити вихрові струми. На рис.7.18 зображено маятник з масивною сочевицею з червоної міді, яка при коливаннях маятника проходить між полюсами електромагніту. При вимкненому електромагніті маятник коливається завдяки великій масі з дуже малим згасанням, але при вмиканні електромагніту різко зупиняється.

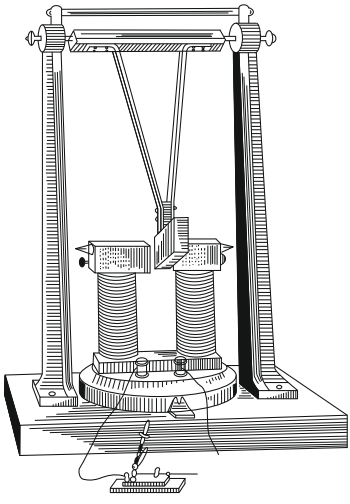


Рис.7.18.Металевий маятник гальмується в магнітному полі внаслідок утворення вихрових струмів.

Якщо мідній монеті надати можливість вільно падати без початкової швидкості між полюсами потужного електромагніту, вона рухатиметься майже рівномірно з дуже малою швидкістю порядку 1см/с так, як тоді, коли рух відбувається в дуже в'язкому середовищі. Якщо мідний диск розмістити поблизу постійного магніту і привести магніт в обертання, то диск почне обертатися в той самий бік, що й магніт (рис.7.19).

Взаємодія вихрових струмів з магнітним полем підлягає третьому закону Ньютона. Тому, якщо в останньому досліді замість магніту обертати диск, то магніт також приходить в обертання. Сили, які виникають від вихрових струмів і діють на рухомі провідники в магнітному полі використовують в багатьох вимірювальних приладах (напр. електричні лічильники, електромагнітне заспокоювання вимірювальних приладів, тахометри, тощо).

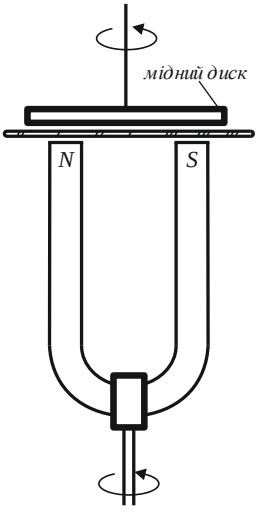


Рис.7.19.Мідний диск, підвішений поблизу обертового магніту, приходить в обертання.

У деяких технічних пристроях вихрові струми відіграють шкідливу роль. Так, наприклад, у залізних осердях трансформаторів і обертових частинах електричних генераторів вихрові струми спричинюють некорисне нагрівання і зменшують к.к.д. цих пристроїв. Для ослаблення вихрових струмів такі деталі виготовляють з тонких листів, розділених найтоншими шарами ізолятора, щоб ізолюючі шари перетинали можливі лінії вихрових

струмів. Суть цього прийому пояснюється дослідом, який зображено на рис.7.20.

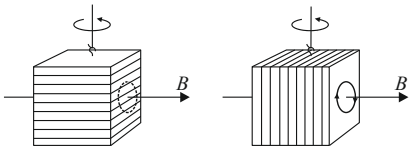


Рис.7.20.Ослаблення вихрових струмів. У випадку *а* вихрові струми значно слабші, ніж у випадку *б*.

Підвісимо кубик, набраний з тонких металевих пластинок, на нитці, помістимо його між полюсами електромагніту і, закрутивши попередньо нитку, спостерігатимемо обертання кубика при розкручуванні нитки. Якщо підвісити кубик у положення *а*, то ізолюючі шари перетинатимуть лінії вихрових струмів (одну з них зображено пунктиром). У цьому випадку вихрові струми виникають тільки в межах кожної з пластинок, їх дія дуже слабка і кубик вільно обертається з великою кутовою швидкістю. А якщо кубик підвісити в положення, як показано на рис.7.20,*б*, то ізолюючі шари будуть паралельні лініям вихрових струмів і не перешкоджатимуть їх утворенню. Кубик поводитиме себе так, ніби ізолюючих шарів не має зовсім, і повертатиметься дуже повільно внаслідок гальмівної дії вихрових струмів.

§7.11.Витіснення змінного струму (скін-ефект)

Якщо в однорідному провіднику з сталим поперечним перерізом є постійний струм, то густина струму однакова в різних точках перерізу провідника. Інше спостерігається при змінному струмі. У цьому випадку густина струму неоднакова за перерізом: вона найбільша на поверхні і найменша на осі провідника. Ця нерівномірність тим більша, чим товстіший провідник і чим більша частота змінного струму, і при дуже великих частотах струм практично існує тільки в тонкому поверхневому шарі, внаслідок чого явище і називається ***скін-ефектом*** (від англ. skin — шкірочка, поверхневий шар).

Це цікаве явище також пояснюється виникненням вихрового електричного поля електромагнітної індукції. Розглянемо провідник зі змінним струмом, і нехай в якийсь момент часу струм *I* має напрям, показаний на рис.7.21. Цей струм створює всередині провідника магнітне поле *H*, силові лінії якого лежать у площині, перпендикулярній до осі провідника.

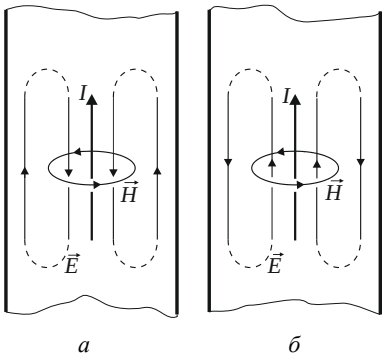


Рис.7.21.Пояснення витіснення змінного струму на поверхню провідника: *H*— магнітне поле струму; *E* — вихрове електричне поле, яке ним спричинюється. *а* – струм зростає, *б* – струм спадає.

Припустимо, що струм *I* підсилюється. Тоді зростаюче магнітне поле *H* спричинить виникнення вихрового електричного поля *E* (рис.7.21,*а),* яке біля поверхні провідника напрямлене однаково зі струмом *I*, а на осі провідника — протилежно струму. Отже, це поле підсилюватиме струм на поверхні і ослаблюватиме його на осі.

Припустимо тепер, що струм *I* зменшується. У цьому випадку послаблююче

магнітне поле *H* спричинить електричне поле *E,* яке буде напрямлене протилежно порівняно з першим випадком (рис.7.21,*б*), тобто буде біля поверхні протилежне струму, а на осі — збігатиметься зі струмом. В обох випадках, і при зростанні, і при послабленні струму *I,* вихрове електричне поле на осі провідника перешкоджатиме, а на поверхні сприятиме змінам струму *I*, а отже на осі провідника змінний струм буде змінюватись слабо, а на поверхні сильно.

Закон розподілу густини струму по перерізу провідника особливо простий, коли провідник має форму плоского шару, товщина якого 2*D* значно менша за ширину (пластина). Розрахунок показує, що цей розподіл залежить від величини

$$d = \frac{1}{\sqrt{\pi \nu \mu \mu_0 \sigma}}, \qquad (7.25)$$

яка називається глибиною проникнення струму. Тут μ — магнітна проникність речовини провідника, μ_0— магнітна стала вакууму, σ − питома електропровідність, ν − частота змінного струму. Виражаючи ці величини в одиницях системи СІ, d дістанемо в м.

Якщо $d \gg D$, то густина струму практично стала по перерізу провідника. Якщо $d \ll D$ (значний скін-ефект), то закон розподілу струму має приблизно такий вигляд:

$$j(y) = j_{пов} e^{-\frac{y}{d}}, \qquad (7.26)$$

де $j(y)$ — густина струму на відстані y від поверхні провідника, $j_{пов}$— густина струму біля поверхні. З (7.26) видно, що глибина проникнення d є така відстань від поверхні провідника, на якій густина струму зменшується в e = 2,72... рази.

Завдяки скін-ефекту електричний струм при високих частотах протікає переважно через поверхневий шар провідника. Це призводить до зменшення діючого перерізу провідника і, як наслідок, до *збільшення опору* провідника. При високих частотах або товстих проводах це збільшення опору може бути значним.

Внаслідок нерівномірного розподілу струму в провіднику змінюється не тільки його опір, а й індуктивність. Справді, при постійному струмі магнітне поле виникає як у зовнішньому просторі, так і всередині провідника. За наявності значного скін-ефекту струм існує практично тільки в поверхневому шарі і магнітного поля всередині провідника немає. Це означає, що магнітна енергія стає меншою на величину енергії поля всередині провідника, а отже, індуктивність провідника зменшується.

Існування скін-ефекту завжди враховують у техніці швидкозмінних струмів. Оскільки такі струми практично не проходять у глибину провідника, то провідні лінії для них складають з порожнистих труб. У сучасній радіотехніці надвисоких частот багато деталей (хвилеводи, коаксіальні лінії) покривають тонким шаром срібла (який є добрим провідником), бо їхній опір практично зумовлений тільки поверхневим шаром.

Скін-ефект у поєднанні з нагріванням вихровими струмами використовують тепер з великим успіхом для *поверхневого гартування сталі.* При цьому методі гартування на деталі діють магнітним полем змінних струмів, які дістають від спеціальних машин, і потім швидко охолоджують. Внаслідок скін-ефекту розігрівання і наступне гартування мають місце тільки в поверхневому шарі, а вся основна маса залишається незагартованою. Це дає змогу виготовити вироби з високою стійкістю поверхні проти стирання, але не крихкі, як при звичайному гартуванні.

Контрольні питання

1.В чому суть явища електромагнітної індукції (ЕМІ)?

2.Сформулювати закон Фарадея для ЕМІ.

3.Сформулювати і пояснити правило Ленца.

4.В чому суть явища самоіндукції?

5.Записати вираз для е.р.с. самоіндукції?

6.Як визначається кількість електрики, що переноситься індукційним струмом? Самостійно вивести необхідну формулу із закону Фарадея для ЕМІ.

7.Пояснення явища ЕМІ за допомогою сили Лоренца.

8.Пояснення явища ЕМІ за допомогою вихрового електричного поля.

9.Самоіндукція. Індуктивність. Е.р.с. самоіндукції.

10.Взаємна індукція. Пояснити залежність сили струму від часу в колах, які містять індуктивність.

11.Енергія магнітного поля.

12.Вивести формулу для енергії магнітного поля соленоїда.

13.Вивести формулу для густини енергії магнітного поля.

14.Привести приклади використання явища ЕМІ в техніці зв'язку ?

15.Які струми називаються вихровими? Привести приклади використання вихрових струмів.

16.Чому швидкозмінні струми "витісняються" на поверхню провідника?

Розділ 8.Рівняння Максвелла

§8.1.Перше рівняння Максвелла

У 60-х роках XIX століття Д. Максвелл, взявши за основу ідеї Фарадея про електричне і магнітне поля, узагальнив закони, виведені експериментально, і розробив завершену теорію єдиного електромагнітного поля, яке утворюється довільною системою зарядів і струмів.

Теорія Максвелла є феноменологічною теорією електромагнітного поля. Це означає, що у цій теорії не розглядається внутрішній механізм явищ, які відбуваються в середовищі і викликають появу електричних і магнітних полів.

Електричні і магнітні властивості середовища характеризуються у теорії Максвелла трьома величинами: відносною діелектричною проникністю ε , відносною магнітною проникністю μ і питомою електропровідністю σ. В теорії Максвелла досліджується залежність ε, μ, σ від властивостей середовища, фізичний зміст цих явищ, які відбуваються у ньому при його поляризації і намагнічуванні.

Теорія Максвелла є макроскопічною теорією електромагнітного поля. У ній розглядаються електричні і магнітні поля, які утворюються в об'ємах набагато більших, ніж об'єми окремих атомів і молекул. Крім того, припускається, що відстані від джерел полів до розглядуваних точок у багато разів більші від розмірів молекул.

У теорії Максвелла розглядаються усереднені електричні і магнітні поля, причому усереднення відповідних мікрополів виконується для інтервалів часу, значно більших від періодів обертання або коливання

електричних зарядів і для ділянок поля, об'єми яких у багато разів більші від об'ємів атомів і молекул.

Теорія Максвелла є теорією близькодії, згідно з якою електричні і магнітні взаємодії здійснюються за допомогою електричних і магнітних полів і у яких вони поширюються зі скінченною швидкістю.

Максвелл узагальнив закон електромагнітної індукції для замкненого нерухомого провідного контуру, що знаходиться у змінному магнітному полі, і рівняння (7.10) є *перше рівняння Максвелла.*

$$\oint_{\Gamma} \vec{E}d\vec{l} = -\int_{S} \frac{\partial \vec{B}}{\partial t}d\vec{S}. \qquad (8.1)$$

Фактично рівняння (8.1) є математичним формулюванням закону електромагнітної індукції Фарадея.

§8.2.Струм зміщення. Друге рівняння Максвелла

У попередніх параграфах ми бачили, що всяке змінне магнітне поле спричинює вихрове електричне поле. Аналізуючи різні електромагнітні процеси, Максвелл дійшов висновку, що повинно існувати й обернене явище*: всяка зміна електричного поля спричинює виникнення вихрового магнітного поля.* Це твердження визначає найважливіші властивості електромагнітного поля і є другим основним положенням теорії Максвелла. Оскільки магнітне поле є основною, обов'язковою ознакою всякого струму, то Максвелл назвав змінне електричне поле *струмом зміщення,* на відміну від струму провідності, зумовленого рухом заряджених частинок (електронів і іонів). Слід зауважити, що цей термін (*струм зміщення*) не є вдалим. Він має деякий сенс у випадку діелектриків, бо в них насправді зміщуються заряди в атомах і молекулах. Однак поняття струму зміщення ми застосовуємо і для вакууму, де ніяких зарядів, а отже, і їх «зміщення» немає. Однак цей термін зберігся внаслідок історичних традицій.

Поняття струму «зміщення» можна пояснити за допомогою такого досліду зображеного на рис.8.1. Конденсатор з позитивно зарядженою правою пластиною розряджається через опір *R*. Навколо проводу створюється магнітне поле $\vec{H}$. Згідно закону повного струму (6.14)

$$\oint_{\Gamma} \vec{H}d\vec{l} = \sum I_i = \int_{S} \vec{j}d\vec{S}, \qquad (8.2)$$

де $\vec{j}$ – густина струму, який проходить через поверхню *S*, яка зв'язана з контуром Г. Через

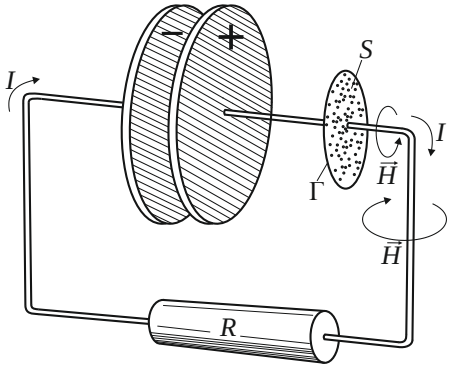


Рис.8.1.

поверхню *S*, яка показана на рис.8.1 і на рис.8.2, протікає струм *I,* і для цієї поверхні (8.2) може бути записана як $\oint_{\Gamma} \vec{H}d\vec{l} = I$. Поверхня S' на рис.8.2 теж пов'язана з контуром Г і має однакові права на участь в законі (теоремі) (8.2). Однак через цю поверхню не протікає ніякий струм. Тим не менше циркуляція $\vec{H}$ по контуру Г не може дорівнювати нулю, а це значить, що і $\int_{S'} \vec{j}d\vec{S} \neq 0$.

Тому рівняння (8.2) повинне бути замінене якимось іншим, яке підходить для

випадку, коли відбувається зміна розподілу зарядів (сила струму I змінюється з часом).

Перепишемо (8.2) в такому вигляді:

$$\oint_{\Gamma} \vec{H}d\vec{l} = \int_{S'} \vec{j}'d\vec{S} \qquad (8.3)$$

і спробуємо вияснити, що представляє собою $\vec{j}'$.

За теоремою Остроградського-Гауса (2.19)

$$\Phi_D = \oint_{S'+S} \vec{D}d\vec{S} = q, \qquad (8.4)$$

де q – заряд правої пластини конденсатора, яка знаходиться всередині замкнутої поверхні $S'+S$. Продиференціюємо вираз (8.4) за часом:

$$\frac{dq}{dt} = \frac{d\Phi_D}{dt} = \frac{d}{dt}\oint_{S+S'} \vec{D}d\vec{S}.$$

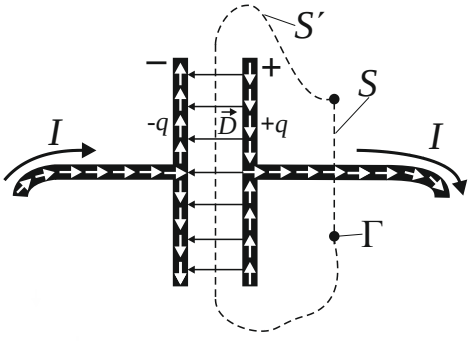


Рис.8.2.Білими стрілками показаний струм у провіднику.

Якщо поверхня $S'+S$ нерухома і не деформується, то зміни з часом потоку зміщення через поверхню $S'+S$ зумовлюються лише зміною електричного зміщення $\vec{D}$ з бігом часу. Тому повну похідну $\frac{d}{dt}$ можна замінити частинною похідною по часу $\frac{\partial}{\partial t}$ і диференціювання внести під знак інтеграла:

$$\frac{dq}{dt} = I' = \oint_{S'+S} \frac{\partial \vec{D}}{\partial t}d\vec{S} = \oint_{S'+S} \vec{j}'d\vec{S}.$$

Із цього виразу ми бачимо, що $\partial\vec{D}/\partial t$ має розмірність густини струму. Максвелл запропонував назвати вираз $\partial\vec{D}/\partial t$ густиною струму зміщення:

$$\vec{j}_{зм} = \frac{\partial \vec{D}}{\partial t}. \qquad (8.5)$$

Густина струму зміщення в даній точці простору дорівнює швидкості зміни вектора електричного зміщення в цій точці.

Струмом зміщення через довільну поверхню S називають фізичну величину, яка дорівнює потоку вектора густини струму зміщення крізь цю поверхню:

$$I_{зм} = \int_S \vec{j}_{зм}d\vec{S} = \int_S \frac{\partial \vec{D}}{\partial t}d\vec{S} = \frac{\partial \Phi_D}{\partial t}. \qquad (8.6)$$

Таким чином, закон (8.2) необхідно записати так:

$$\oint_{\Gamma} \vec{H}d\vec{l} = \int_S \vec{j}d\vec{S} + \int_S \vec{j}_{зм}d\vec{S} = \int_S \vec{j}_{повн}d\vec{S}, \qquad (8.7)$$

де $\vec{j}_{повн} = \vec{j} + \vec{j}_{зм}$ – повний струм, що дорівнює сумі струмів провідності і конвекційних, а також струму зміщення.

Тепер співвідношення (8.7) можна назвати *законом повного струму* в загальному випадку. Поняття струму зміщення усуває складність, що виникла в ситуації, яка показана на рис.8.2. Оскільки заряд конденсатора зменшується, то величина індукції електричного поля, конфігурація якого для будь-якого моменту часу показана на рис.8.3, зменшується. В цьому випадку вектор $\frac{\partial \vec{D}}{\partial t}$ направлений в сторону протилежну до напрямку вектора $\vec{D}$. Векторна функція $\partial\vec{D}/\partial t$ (струм зміщення) на рис.8.4 зображена чорними стрілками. Значить, у співвідношенні (8.3) густина струму $\vec{j}'$ є густина струму зміщення і

$$\int_{S'} \vec{j}' \, d\vec{s} = \int_{S'} \frac{\partial \vec{D}}{\partial t} d\vec{S},$$

якщо поверхня S є на значній відстані від конденсатора.

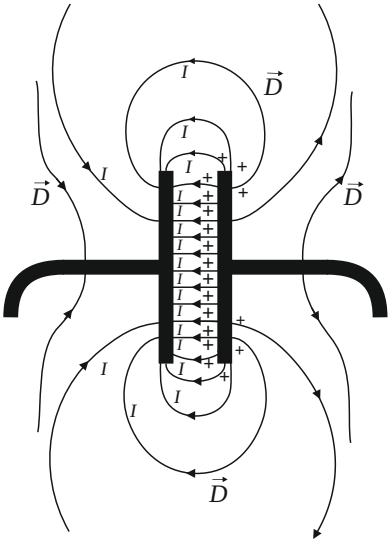


Рис.8.3.

Увівши уявлення про струм зміщення, Максвелл зробив висновок про те, що кола будь-яких непостійних струмів замкнені так само, як і кола постійного струму. Замкненість таких кіл забезпечується струмами зміщення, які «проходять» у тих ділянках, де немає провідників, наприклад, між обкладинками конденсатора під час його зарядки або розрядки.

За Максвеллом, струм зміщення, як і звичайні струми провідності, є джерелом виникнення вихрового магнітного поля, циркуляція напруженості $\vec{H}$ якого по замкненому контуру не дорівнює нулю.

В діелектрику вектор зміщення $\vec{D}$, згідно (2.18), складається з двох доданків: $\vec{D} = \varepsilon_0 \vec{E} + \vec{P}$, де $\vec{E}$ – напруженість електричного поля, $\vec{P}$ – вектор поляризації, який характеризує дійсне зміщення електричних зарядів у неполярних молекулах і поворот полярних молекул, що містяться в одиниці об'єму діелектрика. Густина струму зміщення в діелектрику

$$\vec{j}_{зм} = \varepsilon_0 \frac{\partial \vec{E}}{\partial t} + \frac{\partial \vec{P}}{\partial t}, \qquad (8.8)$$

де $\varepsilon_0 \frac{\partial \vec{E}}{\partial t}$ – густина струму зміщення у вакуумі, $\frac{\partial \vec{P}}{\partial t}$ – густина струму поляризації, що обумовлений зміщенням або поворотом зарядів в діелектрику.

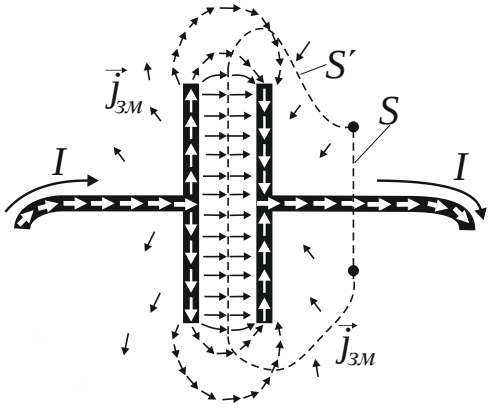


Рис.8.4.Струм провідності (білі стрілки), струм зміщення (чорні стрілки).

Струм зміщення у вакуумі не виділяє тепла. Густина цього струму тим більша, чим більша швидкість зміни напруженості електричного поля. Густину повного струму запишемо так:

$$j_{повн} = \vec{j} + \frac{\partial \vec{D}}{\partial t}. \qquad (8.9)$$

Повний струм в колах змінного струму завжди замкнений. Тепер співвідношення (8.7) можемо записати так:

$$\oint_\Gamma \vec{H} d\vec{l} = \int_S \left(\vec{j} + \frac{\partial \vec{D}}{\partial t}\right) d\vec{S}. \qquad (8.10)$$

Якщо струми провідності відсутні ($\vec{j}$=0), то (8.10) перепишеться як:

$$\oint_\Gamma \vec{H} d\vec{l} = \int_S \frac{\partial \vec{D}}{\partial t} d\vec{S}. \qquad (8.11)$$

Порівнявши (8.11) з (7.10), можемо сказати, що (8.11) визначає наявність *нового явища індукції*, в якому змінне електричне поле супроводжується магнітним полем. Якщо це явище індукції існує, то чому Фарадей його не відкрив? Перш за все, він його не шукав; але існує і більш фундаментальна причина, чому досліди, подібні до фарадеївських, не могли виявити яких-небудь нових явищ, що пов'язані із струмом зміщення (8.5).

Щоб побачити нове явище індукції (8.11), потрібні швидкозмінні електричні поля. Ось чому від теоретичного відкриття струмів зміщення Максвеллом до експериментального доказу їх існування в дослідах Герца пройшло багато років.

Введення струму зміщення, який визначається виразом (8.5), «урівняло у правах» електричне і магнітне поля. Із явища електромагнітної індукції (8.1) витікає, що змінне магнітне поле породжує електричне поле. Із рівняння (8.11) витікає, що змінне електричне поле породжує магнітне поле. Співвідношення (8.10) є другим рівнянням Максвелла.

§8.3.Рівняння Максвелла

Відкриття струму зміщення дозволило Максвеллу створити єдину теорію електричних і магнітних явищ. Ця теорія пояснила всі відомі того часу експериментальні факти і передбачила ряд нових явищ, існування яких підтвердилось пізніше.

Рівняння Максвелла пов'язують зміни основних характеристик електромагнітного поля – векторів $\vec{E}, \vec{D}, \vec{H}, \vec{B}$ в даному матеріальному середовищі з розподілом у ньому електричних зарядів і струмів. Середовище в рівняннях Максвелла враховується феноменологічно, тобто без розкриття внутрішнього механізму взаємодії речовин з полем.

В основі теорії Максвелла лежать чотири рівняння. Першим рівнянням Максвелла є співвідношення (8.1), яке зв'язує значення $\vec{E}$ із змінною $\vec{B}$ з часом і є виразом закону електромагнітної індукції.

Перше рівняння Максвелла вказує на те, що джерелами електричного поля можуть бути не тільки електричні заряди, але і змінні з часом магнітні поля.

Другим рівнянням Максвелла є закон повного струму (8.10), який показує, що магнітні поля можуть збуджуватись або рухомими зарядами (струмами), або змінними електричними полями.

Третім рівнянням Максвелла є теорема Остроградського-Гауса (2.19) для потоку вектора електричного зміщення $\vec{D}$ крізь довільну замкнену поверхню *S*, що охоплює сумарний заряд *q*:

$$\oint_S \vec{D} d\vec{S} = q.$$

Якщо заряд розподілений всередині замкненої поверхні з об'ємною густиною ρ, то

$$\oint_S \vec{D} d\vec{S} = \int_V \rho dV. \qquad (8.12)$$

Четвертим рівнянням Максвелла є теорема Остроградського-Гауса (5.24) для магнітного потоку крізь довільну замкнену поверхню :

$$\oint_S \vec{B} d\vec{S} = 0. \qquad (8.13)$$

Отже, повна система рівнянь Максвелла в інтегральній формі має такий вигляд:

$$1.\ \oint_\Gamma \vec{E}d\vec{l} = -\int_S \frac{\partial \vec{B}}{\partial t}d\vec{S},$$
$$2.\ \oint_\Gamma \vec{H}d\vec{l} = \int_S \left(\vec{j}+\frac{\partial \vec{D}}{\partial t}\right)d\vec{S}, \qquad (8.14)$$

$$3.\ \oint_S \vec{D}d\vec{S} = \int_V \rho dV,$$
$$4.\ \oint_S \vec{B}d\vec{S} = 0. \qquad (8.14)$$

Величини, що входять в рівняння Максвелла (8.14), не є незалежними. Між ними є зв'язок (2.21), (6.19), (4.14):

$$\vec{D} = \varepsilon\varepsilon_0\vec{E},\ \vec{B} = \mu\mu_0\vec{H},\ \vec{j} = \sigma\vec{E}.$$

Зазначимо, що до першого та четвертого рівняння Максвелла входить лише основні характеристики поля $\vec{E}$ і $\vec{B}$, а в друге і третє – лише допоміжні величини $\vec{D}$ і $\vec{H}$. Рівняння Максвелла несиметричні відносно полів. Це зв'язано з тим, що в природі існують лише електричні заряди, а магнітних зарядів немає.

Для стаціонарних полів ($\vec{E} = \text{const}$ і $\vec{B} = \text{const}$) рівняння Максвелла мають такий вигляд

$$\oint_\Gamma \vec{E}d\vec{l} = 0,\quad \oint_S \vec{D}d\vec{S} = q,$$
$$\oint_\Gamma \vec{H}d\vec{l} = I,\ \oint_S \vec{B}d\vec{S} = 0. \qquad (8.15)$$

В даному випадку електричні і магнітні поля існують незалежно одне від одного.

§8.4.Рівняння Максвелла в диференціальній формі

Рівняння Максвелла (8.14) застосовні до поверхні будь-якої величини; тому величини, які входять до них, стосуються різних точок поля. Так, наприклад, у лівій частині рівняння (8.10) $\vec{H}$ є напруженість магнітного поля в точках контура Г, який обмежує розглядувану поверхню *S*, тоді як потік у правій частині залежить від значень $\vec{D}$ в точках самої поверхні. Однак можна перетворити ці рівняння в таку форму, щоб усі величини стосувалися однієї точки поля. Для цього рівняння Максвелла необхідно застосувати до контурів і поверхонь, які зв'язані між собою, нескінченно малих величин.

Для цього розглянемо деякі математичні поняття. Оператор набла, або оператор Гамільтона позначають символом $\vec{\nabla}$ і під ним розуміють вектор з компонентами $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$. Тобто:

$$\vec{\nabla} = \vec{\imath}\frac{\partial}{\partial x} + \vec{\jmath}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}. \qquad (8.16)$$

Сам по собі цей вектор смислу не має. Він набуває змісту в поєднанні із скалярною або векторною функцією, на яку він символічно множиться. Так, якщо помножити вектор $\vec{\nabla}$ на скаляр φ, то отримається вектор:

$$\vec{\nabla}\varphi = \vec{\imath}\frac{\partial\varphi}{\partial x} + \vec{\jmath}\frac{\partial\varphi}{\partial y} + \vec{k}\frac{\partial\varphi}{\partial z} = \text{grad}\varphi, \qquad (8.17)$$

який є градієнтом функції φ (див. співвідношення (1.33)).

Якщо вектор $\vec{\nabla}$ помножити скалярно на вектор $\vec{E}$, то отримається скаляр

$$\vec{\nabla}\vec{E} = \nabla_x E_x + \nabla_y E_y + \nabla_z E_z = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \text{div}\vec{E}, \qquad (8.18)$$

який є не що інше, як дивергенція вектора $\vec{E}$ (див.1.36). Якщо помножити $\vec{\nabla}$ на вектор $\vec{H}$ векторно, то отримається вектор, компоненти якого знаходяться за допомогою визначника

$$\vec{\nabla}\times\vec{H} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ H_x & H_y & H_z \end{vmatrix} = \text{rot}\vec{H}. \qquad (8.19)$$

Такий векторний добуток називають ротором вектора (в нашому випадку вектора $\vec{H}$, як приклад). Із (8.19) знаходимо, що

$$\begin{cases} \text{rot}_x\vec{H} = \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z}, \\ \text{rot}_y\vec{H} = \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}, \\ \text{rot}_z\vec{H} = \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y}. \end{cases} \quad (8.20)$$

Теорема Остроградського-Гаусса в математиці записується так:

$$\oint_S \vec{a}d\vec{S} = \int_V \text{div}\vec{a}dV, \quad (8.21)$$

де V – об'єм, обмежений замкненою поверхнею S, $\vec{a}$ – деяка векторна функція.

Є теорема Стокса, яка математично записується так:

$$\oint_\Gamma \vec{a}d\vec{l} = \int_S \text{rot}\vec{a} \cdot d\vec{S}, \quad (8.22)$$

де контур Г і поверхня S взаємопов'язані.

Тоді перше і друге рівняння Максвелла із урахування (8.22) можуть бути записані так:

$$\oint_\Gamma \vec{E}d\vec{l} = \int_S \text{rot}\vec{E}d\vec{S} = -\int_S \frac{\partial\vec{B}}{\partial t}d\vec{S},$$

$$\oint_\Gamma \vec{H}d\vec{l} = \int_S \text{rot}\vec{E}d\vec{S} = \int_s \left(\vec{j} + \frac{\partial\vec{D}}{\partial t}\right)d\vec{S}.$$

Оскільки S є довільною поверхнею, яка пов'язана з контуром Г, то рівність інтегралів можна замінити на рівність підінтегральних виразів і отримаємо:

$$\text{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t}, \quad (8.23)$$

$$\text{rot}\vec{H} = \vec{j} + \frac{\partial\vec{D}}{\partial t}. \quad (8.24)$$

Скориставшись теоремою Остроградського-Гаусса (8.21), третє (8.12) і четверте (8.13) рівняння Максвелла перепишемо так:

$$\oint_S \vec{D}d\vec{S} = \int_V \text{div}\vec{D} = \int_V \rho dV,$$

$$\oint_S \vec{B}d\vec{S} = \oint_V \text{div}\vec{B} = 0.$$

Оскільки замкнена поверхня S і пов'язаний з нею об'єм V є довільними, то рівність інтегралів замінимо на рівність підінтегральних виразів. Тому:

$$\text{div}\vec{D} = \rho, \quad (8.25)$$

$$\text{div}\vec{B} = 0. \quad (8.26)$$

Співвідношення (8.23 – 8.26) є рівняннями Максвелла в диференціальній формі, в яких характеристики поля $\vec{E}$, $\vec{B}$, $\vec{D}, \vec{H}$ належать до однієї точки простору. Кожне з векторних рівнянь (8.23) і (8.24) еквівалентно трьом скалярним рівнянням, які зв'язують компоненти векторів, що стоять в лівій і правій частинах рівності. Скориставшись формулами (8.18), (8.19), (8.20), представимо рівняння Максвелла в скалярній формі:

$$\begin{matrix} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\frac{\partial B_x}{\partial t} \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\frac{\partial B_y}{\partial t} \end{matrix}, \quad (8.27)$$

$$\begin{matrix} \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -\frac{\partial B_z}{\partial t} \\ \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0 \end{matrix}. \quad (8.28)$$

(перша пара рівнянь),

$$\begin{matrix} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = j_x + \frac{\partial D_x}{\partial t} \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = j_y + \frac{\partial D_y}{\partial t} \end{matrix}, \quad (8.29)$$

$$\begin{matrix} \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = j_z + \frac{\partial D_z}{\partial t} \\ \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = \rho \end{matrix}. \quad (8.30)$$

(друга пара рівнянь).

Всього вийшло 8 рівнянь, в які входять 12 функцій (по три компоненти

векторів $\vec{E}, \vec{B}, \vec{D}, \vec{H}$). Оскільки кількість рівнянь менша кількості невідомих функцій, то рівнянь (8.27 – 8.30) недостатньо, щоб знайти характеристики полів за заданим розподілом зарядів і струмів. Щоб виконати розрахунок полів, необхідно доповнити рівняння Максвелла співвідношеннями (2.21), (6.19) і (4.14), а саме:

$$\vec{D} = \varepsilon\varepsilon_0\vec{E}, \qquad (8.31)$$

$$\vec{B} = \mu\mu_0\vec{H}, \qquad (8.32)$$

$$\vec{j} = \sigma\vec{E}. \qquad (8.33)$$

Сукупність рівнянь (8.23 – 8.26) і (8.31 – 8.33) становить основу електродинаміки середовищ, які знаходяться в стані спокою.

§8.5. Значення теорії Максвелла

Теорія Максвелла відіграла значну роль у розвитку наших знань про електрику. Щоб краще зрозуміти значення цієї теорії, потрібно згадати історичну послідовність основних відкриттів у галузі електрики до праць Максвелла.

Кількісне вивчення електричних явищ почалося з праць Кулона (1785 р.), який встановив спочатку закон взаємодії електричних зарядів і поширив його пізніше на взаємодію "магнітних зарядів". Однак аж до 1820 р. електричні і магнітні явища розглядали як різні явища, не зв'язані між собою.

Відкриття Ерстедом у 1820 р. дії струму на магнітну стрілку показало, що між магнітними й електричними явищами існує зв'язок і що магнітні дії можна дістати за допомогою електричних струмів. Магнітну дію струмів докладно вивчив Ампер, який дійшов висновку, що всі магнітні явища в природі, в тому числі і пов'язані з постійними магнітами, спричинені електричними струмами (теорія молекулярних струмів Ампера).

Особливе значення мало відкриття електромагнітної індукції Фарадеєм. Фарадей виходив з основної ідеї про взаємозв'язок явищ природи. Він вважав, що коли струм здатний викликати магнітні явища, то і, навпаки, за допомогою магнітів або інших струмів можна дістати електричні струми. Внаслідок наполегливих і численних спроб, Фарадей справді відкрив в 1831 р. це явище, яке ще більше зміцнювало уявлення про зв'язок між електрикою і магнетизмом.

Іншою найважливішою ідеєю в працях Фарадея було визнання основної визначальної ролі проміжного середовища в електричних явищах. Фарадей не допускав дії на відстані, яка, як тепер добре відомо, фізично не має сенсу, і вважав, що електричні і магнітні взаємодії передаються проміжним середовищем і що саме в цьому середовищі відбуваються основні електричні і магнітні процеси.

У працях Максвелла ідеї Фарадея було далі поглиблено й розвинуто і перетворено в точну математичну теорію. У теорії Максвелла думка про тісний зв'язок електричних і магнітних явищ дістала остаточне оформлення у вигляді двох основних положень теорії, розглянутих в §7.4 і §8.2, і була в строгій формі виражена у вигляді рівнянь Максвелла (§8.3, §8.4). Тому теорія Максвелла була завершенням важливого етапу в розвитку вчення про електрику і привела до класичного уявлення про *електромагнітне поле*, яке містить у загальному випадку і електричне, і магнітне поля, пов'язані між собою і здатні взаємно перетворюватись одне в одне.

Рівняння Максвелла містять у собі всі основні закони електричного і магнітного

полів, включаючи електромагнітну індукцію, і тому є *загальними* рівняннями електромагнітного поля в середовищах, які перебувають у спокої.

Теорія Максвелла не тільки пояснила вже відомі факти, а й завбачила нові важливі явища. Зовсім новим у цій теорії було припущення Максвелла про магнітне поле струмів зміщення (§8.2). На основі цього припущення Максвелл теоретично завбачив існування *електромагнітних хвиль*, тобто змінного електромагнітного поля, яке поширюється в просторі з скінченною швидкістю. Теоретичне дослідження властивостей електромагнітних хвиль привело потім Максвелла до створення *електромагнітної теорії світла*, за якою світло є теж електромагнітними хвилями. У подальшому електромагнітні хвилі дійсно були виявлені під час досліду, а ще пізніше електромагнітна теорія світла, а з нею і вся теорія Максвелла, дістала повне і беззаперечне підтвердження.

Відкладемо розгляд електромагнітних хвиль до вивчення електричних коливань і пружних хвиль.

Контрольні питання:

1.Чим відрізняється струм зміщення від струму провідності?

2.Поясніть, чому конденсатор проводить змінний струм і не проводить постійний?

3.Які фізичні закони відображені в рівняннях Максвелла?

Розділ 9. Коливання

§9.1.Види коливань. Загальні поняття

Коливання це рухи або процеси, які мають в тій чи іншій мірі повторюваність із часом. Коливання властиві всім явищам природи: пульсуюче випромінювання зірок, всередині яких відбуваються циклічні ядерні реакції; з високою ступінню періодичності обертаються планети Сонячної системи; рух Місяця викликає припливи і відпливи на Землі; в земній іоносфері й атмосфері циркулюють потоки заряджених і нейтральних частинок; вітри збуджують коливання і хвилі на поверхнях водоймищ і т. д. Всередині будь-якого живого організму безперервно відбуваються різноманітні процеси, які циклічно повторюються, наприклад, з дивною надійністю б’ється людське серце, навіть психіка людей піддається коливанням. У вигляді складної сукупності коливань частинок і полів (електронів, протонів, фотонів та інших) можна представити «будову» мікросвіту.

В техніці коливання або виконують певні функціональні обов’язки (маятник, коливальний контур, генератор коливань і т. п.), або виникають як неминучий прояв фізичних властивостей (вібрації машин і споруд, нестійкості вихрових потоків під час русху тіл в газах і т. ін.).

У фізиці виділяють коливання механічні, електромагнітні та їх комбінації. Це обумовлено тією винятковою роллю, яку відіграють гравітаційна й електромагнітна взаємодія в масштабах, характерних для життєдіяльності людини. За допомогою розповсюдження механічних коливань густини і тиску повітря, які сприймаються нами як звук, а також дуже швидких коливань електричного і магнітного полів, які сприймаються нами як

світло, ми отримуємо більшість прямої інформації про оточуючий світ.

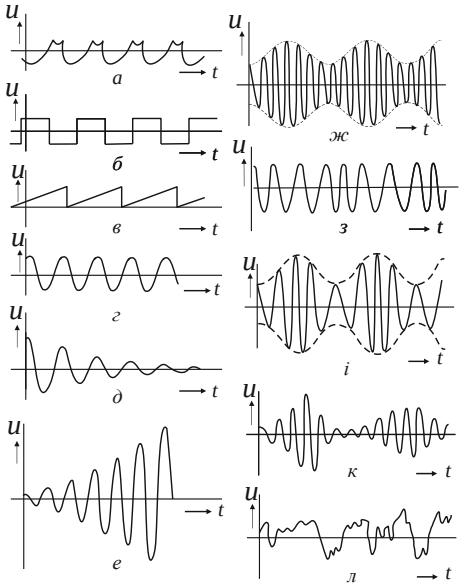


Рис.9.1. Різні види коливань:
а – періодичні коливання складної форми; *б* – прямокутні коливання; *в* – пилкоподібні; *г* – синусоїдальні; *д* – згасаючі; *е* – наростаючі; *ж* – амплітудно-модульвані; *з* – частотно-модульовані; *і* – коливання модульовані по амплітуді і по фазі; *к* – коливання, амплітуда і фаза яких випадкові функції; *л* – випадкові коливання; u - величина, яка коливається, t – час.

З вивченням коливань різної фізичної природи виникла впевненість про можливість загального, «позапредметного» підходу до них, який базується на властивостях і закономірностях коливальних процесів взагалі. В результаті з'явилась теорія коливань і хвиль, основним математичним апаратом якої спочатку були диференціальні рівняння із звичайними похідними.

Кінематика коливань дозволяє виділити декілька прикладів (рис.9.1). Будемо говорити про коливання, які описуються функцією часу $u(t)$. На рис.9.1, *а-г*, показані періодичні коливання різної форми, в яких будь-яке значення $u(t)$ повторюється через однакові проміжки часу T, які називаються **періодом** коливань, тобто $u(t) = u(t + T)$. Величина, яка обернена періоду T і дорівнює числу коливань за одиницю часу, називають частотою коливань $\left(\nu = \frac{1}{T}\right)$; користуються також круговою або циклічною частотою $\omega = 2\pi\nu$. У випадку просторових коливань вводяться аналогічні поняття просторового періоду (довжини хвилі λ) і хвильового числа $k = 2\pi/\lambda$.

Різновидами періодичних коливань є прямокутні (рис.9.1,*б*), пилкоподібні (рис.9.1, *в*) і найбільш важливі синусоїдальні, або **гармонічні** коливання (рис.9.1,*г*). Останні можуть бути записані у вигляді

$$\begin{aligned} u(t) &= A\cos\varphi = A\cos(\omega_0 t + \varphi_0), \\ u(t) &= A\sin\varphi = A\sin(\omega_0 t + \varphi_0'), \end{aligned} \quad (9.1)$$

де A – амплітуда, максимальне значення за модулем величини u, $\omega t + \varphi_0$ – фаза, φ_0 або φ_0' – початкове значення фази, ω_0 – циклічна частота.

Для показаного на рис.9.1,*д* згасаючого коливання

$$u(t) = Ae^{-\beta t}\cos(\omega t + \varphi_0), \quad (9.2)$$

де β – коефіцієнт згасання, ω – частота згасаючих коливань. На рис,9.1,*е* показано наростаюче коливання.

Для передачі інформації застосовуються модульовані коливання (рис.9.1,*ж-і*), амплітуда, фаза або частота яких змінюється за певним законом у відповідності з сигналом, що передається. Наприклад, у радіомовленні високочастотні (ВЧ) коливання модулюються коливаннями звукової частоти, які передають мовлення, музику.

На рис.9.1,*к* показаний приклад синусоїдального коливання, яке модульоване по амплітуді і фазі випадковими функціями, а

на рис. 9.1,*л* дана одна із реалізацій зовсім неупорядкованого процесу («білого шуму»), який тільки умовно можна віднести до коливань.

§9.2.Власні коливання

Найпростішою коливальною системою в механіці є тягарець, який підвішено на пружині і який рухається без тертя (рис.9.2). При цьому ми припускаємо, що масою пружини

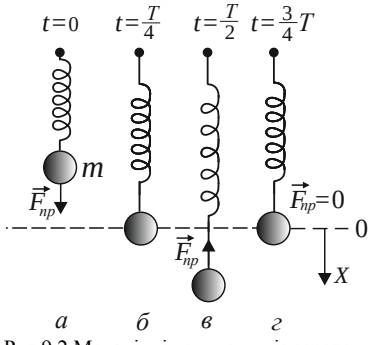


Рис.9.2.Механічні коливання із зосередженими масою і пружністю. $F_{пр}$ – сила пружності пружини.

можна знехтувати порівняно з масою тягарця і що вся пружність належить тільки пружині (система з зосередженими масою і пружністю). З механіки відомо, що тягарець, виведений із стану рівноваги, здійснює гармонічні коливання, при яких його відхилення від положення рівноваги змінюється з часом за законом синуса (косинуса). Коли тягарець перебуває в крайніх положеннях (*а* і *в*, рис.9.2), кінетична енергія тягарця дорівнює нулю, але потенціальна енергія пружини сягає максимуму. Навпаки, при проходженні тягарцем положення рівноваги (*б* і *г* рис.9.2), кінетична енергія тягарця має найбільше значення, а потенційна енергія пружини, яка в цьому положенні і не стискується, і не розтягується, дорівнює нулю. Тому розглядуване механічне коливання є **періодичне перетворення енергії** системи з кінетичної в потенціальну і навпаки.

Аналогічні процеси відбуваються при електричних коливаннях. Найпростіший електричний коливальний контур складається з конденсатора і індуктивності, сполучених між собою (рис.9.3). Ми вважатимемо, що ємність між витками котушки дуже мала порівняно з ємністю конденсатора, а індуктивність конденсатора і з'єднувальних проводів мала порівняно з індуктивністю котушки (контур із зосередженими ємністю і індуктивністю, або закритий коливальний контур). Припустимо, що ми, розімкнувши контур, зарядили конденсатор. Між пластинами конденсатора з'явиться електричне поле *E*, яке матиме певну

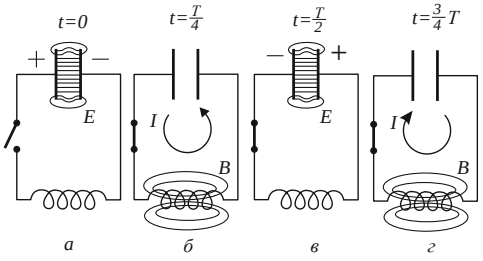


Рис.9.3.Електричні коливання в контурі із зосередженими індуктивністю і ємністю.

енергію (рис.9.3,*а*). Замкнемо тепер конденсатор на індуктивність. Конденсатор почне розряджатись і його електричне поле зменшуватиметься. При цьому в контурі виникне електричний струм розряджання конденсатора, внаслідок чого в котушці індуктивності виникне магнітне поле *B*. Через деякий час, що дорівнює чверті періоду коливання, конденсатор розрядиться повністю і електричне поле зникне зовсім. Але магнітне поле при цьому досягне максимуму, а отже, енергія електричного поля перетвориться в енергію магнітного поля.

В наступні моменти часу магнітне поле зникатиме, бо немає струмів, які б його підтримували. Це зникаюче поле приведе до появи екстраструму самоіндукції, який за законом Ленца прагнутиме підтримати струм заряджання конденсатора і матиме, отже, такий самий напрям, як і цей останній. Тому конденсатор знову перезаряджатиметься і між його пластинами виникне електричне поле протилежного напряму. Протягом часу, який дорівнює половині періоду коливання, магнітне поле зникне зовсім, а електричне поле досягне максимуму, і енергія магнітного поля знову перетвориться в енергію електричного поля. Далі конденсатор знову розряджатиметься, і в контурі виникне струм, направлений протилежно струму в попередній стадії процесу. Після $3T/4$ конденсатор знову буде розряджений, а енергія електричного поля знову перетвориться в енергію магнітного поля (рис.9.3,*г*) і т. д. Через проміжки часу, які дорівнюють повному періоду коливання T, електричний стан контура буде такий самий, як і на початку коливань (рис.9.3,*а*).

Якщо опір контура дорівнює нулю, то такий процес періодичного перетворення електричної енергії в магнітну і навпаки триватиме нескінченно довго і ми дістанемо незгасаючі електричні коливання.

Механічні коливання, що виникають під дією сил, які розвиваються в самій коливальній системі, називають **власними** коливанням. Вони виникають при всякому порушенні рівноваги коливальної системи. Подібно до цього, електричні коливання, що відбуваються під дією процесів у самому коливальному контурі, дістали назву **власних електричних коливань**. Розглянуті вище коливання є, очевидно, власними.

§9.3.Рівняння власних коливань. Коливання при відсутності згасання

Для тягарця, що підвішений на пружині і рухається вертикально (по осі x) без тертя, рівняння руху, згідно з другим законом Ньютона, можна записати так:

$$ma_x = F_{пр}\,, \qquad (9.3)$$

де $a_x = \frac{d^2x}{dt^2} = \ddot{x}$ – прискорення тягарця (друга похідна по часу від його координати зміщення x відносно положення рівноваги). $F_{пр} = -kx$ – сила пружності пружини, яка завжди направлена проти зміщення x, k – коефіцієнт пружності пружини, m – маса тягарця.

Тоді рівняння (9.3) можна переписати в такому вигляді:

$$\ddot{x} + \omega_0^2 x = 0\,, \qquad (9.4)$$

де

$$\omega_0 = \sqrt{\frac{k}{m}} \qquad (9.5)$$

циклічна частота коливань. Рівняння (9.4) називається диференційним рівнянням власних (гармонічних) коливань другого порядку, однорідним з постійними коефіцієнтами. Розв'язком цього рівняння є співвідношення (9.1), які називаються рівняннями власних коливань в інтегральному вигляді. Величини A і φ в (9.1) знаходяться із початкових умов.

Розглянемо тепер кількісно власні коливання в електричному контурі на рис.9.3. Для цього розглянемо коливальний контур (рис.9.4). Будемо вважати, що електричні процеси в контурі **квазістаціонарні**. Це означає, що миттєве значення величини струму I — одне й те саме в будь-кому місці контуру і що до миттєвих значень електричних величин можна застосовувати правила Кірхгофа. Нехай заряд конденсатора q_0, а струм в колі $I = \frac{dq}{dt} =$

$=\dot{q}$. Тоді, згідно другого правила Кірхгофа (4.40), можна записати, що

$$IR = \varphi_1 - \varphi_2 + \mathcal{E}_s, \qquad (9.6)$$

де $\varphi_1 - \varphi_2 = -(\varphi_2 - \varphi_1) = -U_c = -\frac{q}{C}$ — напруга на конденсаторі, $\mathcal{E}_s = -L\frac{dI}{dt} = -L\frac{d\dot{q}}{dt} = -L\ddot{q}$ — е.р.с. самоіндукції, яка виникає в котушці індуктивності L.

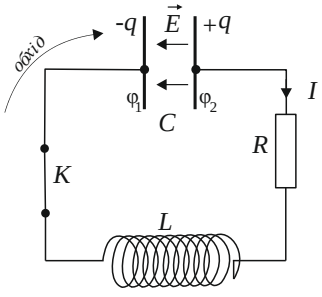


Рис.9.4.

Для випадку ідеалізованого контуру ($R = 0$) рівняння (9.6) запишеться так

$$\ddot{q} + \frac{1}{LC}q = 0,$$

або

$$\ddot{q} + \omega_0^2 q = 0\,, \qquad (9.7)$$

де

$$\omega_0 = \frac{1}{\sqrt{LC}}\cdot \qquad (9.8)$$

Рівняння (9.7) математично еквівалентне рівнянню (9.4), і воно є лінійним однорідним диференційним рівнянням другого порядку із звичайними похідними і зі сталими коефіцієнтами, що описує коливання заряду на конденсаторі в контурі на рис.9.4. Розв'яжемо рівняння (9.7). Для цього помножимо його на $I = \dot{q}$ і отримаємо:

$$\ddot{q}\dot{q} + \omega^2 q\dot{q} = 0\,,$$

або

$$\frac{d}{dt}(\dot{q}^2 + \omega_0^2 q^2) = 0\,.$$

Звідки отримаємо, що $\dot{q}^2 + \omega_0^2 q^2 = \text{const} = q_0^2\omega_0^2$. З цього співвідношення знаходимо, що

$$\frac{dq}{dt} = \dot{q} = \omega_0\sqrt{q_0^2 - q^2} \text{ і } \frac{dq}{\sqrt{q_0^2 - q^2}} = \omega_0 dt.$$

Інтегрування останнього виразу дає, що

$$\arccos\frac{q}{q_0} = \omega_0 t + \varphi_0\,,$$

або

$$q = q_0\cos(\omega_0 t + \varphi_0)\,. \qquad (9.9)$$

Вираз (9.9) є інтегральним рівнянням власних (гармонічних) коливань заряду на конденсаторі.

Таким чином, заряд на обкладках конденсатора змінюється за гармонічним законом з частотою, яка визначається виразом (9.8). Ця частота називається **власною частотою контуру**. Для періоду коливань отримується так звана **формула Томсона**:

$$T = 2\pi/\omega_0 = 2\pi\sqrt{LC}\,. \qquad (9.10)$$

Напруга на конденсаторі відрізняється від заряду на множник $1/C$:

$$U = \frac{q}{C} = \frac{q_0}{C}\cos(\omega_0 t + \varphi_0) = \\ = U_0\cos(\omega_0 t + \varphi_0). \qquad (9.11)$$

Продиференціювавши функцію (9.9) по часу, отримаємо вираз для сили струму:

$$I = -q_0\omega_0\sin(\omega_0 t + \varphi_0) = \\ = I_0\cos(\omega_0 t + \varphi_0 + \pi/2), \qquad (9.12)$$

де $I_0 = q_0\omega_0$ — амплітуда сили струму. Таким чином, сила струму випереджає по фазі напругу на конденсаторі на $\pi/2$. Врахувавши, що $U_0 = q_0/C$, а $I_0 = q_0\omega_0$, отримаємо, що

$$U_0 = \sqrt{\frac{L}{C}}I_0. \qquad (9.13)$$

Порівнявши це співвідношення із законом Ома (4.8), величину $\sqrt{\frac{L}{C}}$ називають **хвильовим опором** контуру.

Енергія електричного поля конденсатора дорівнює:

$$W_c = \frac{q^2}{2C} = \frac{q_0^2}{2C}\cos^2(\omega_0 t + \varphi_0). \quad (9.14)$$

Енергія магнітного поля котушки буде:

$$W_L = \frac{LI^2}{2} = \frac{Lq_0^2\omega_0^2}{2}\sin^2(\omega_0 t + \varphi_0)$$

або

$$W_L = \frac{q_0^2}{2C}\sin^2(\omega_0 t + \varphi_0) = \frac{q_0^2}{2C}\cos^2(\omega_0 t + \varphi_0 + \pi/2). \quad (9.15)$$

Сумарна енергія електричного і магнітного полів

$$W = W_C + W_L = q_o^2/(2C), \quad (9.16)$$

що відповідає енергії, яка була надана контуру через заряджання конденсатора зарядом q_0 (до напруги U_0). Співвідношення (9.16) відображає закон збереження енергії в ідеалізованому контурі. Зауважимо, що формули і співвідношення для механічних коливань можна отримати із формул і співвідношень для електричних коливань, і навпаки, зберігаючи таку заміну:

заряд конденсатора q – зміщення тягарця від положення рівноваги x;

сила струму $I = \frac{dq}{dt} = \dot{q}$ – швидкість коливань $\upsilon = \frac{dx}{dt} = \dot{x}$;

індуктивність контуру L – маса тягарця m;

величина $1/C$ –коефіцієнт пружності пружини k.

§9.4.Геометричне зображення гармонічних коливань. Додавання гармонічних коливань однакового напрямку і однакової частоти

Перш ніж розглядати додавання коливальних рухів, спинимось на способі зображення коливань за допомогою обертального вектора амплітуди. Розглянемо рівняння (9.1), вважаючи, що зміна величини U еквівалентна зміщенню від положення рівноваги коливальної системи x. Із довільної точки O, яка вибрана на осі X, під кутом φ_0, що дорівнює початковій фазі коливання, відкладемо вектор $\vec{A}$, модуль якого дорівнює амплітуді A коливання (рис.9.5).

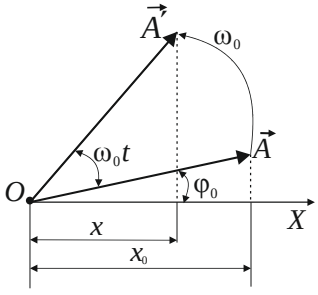


Рис.9.5.

Проекція вектора $\vec{A}$ на вісь OX дорівнює зміщенню x_0 у момент початку відліку часу ($t = 0$):

$$x_0 = A\cos\varphi_0.$$

Обертатимемо в площині рисунка вектор амплітуди навколо осі OO, яка перпендикулярна до площини рисунка, з постійною кутовою швидкістю ω_0. За проміжок часу t вектор амплітуди повертається на кут $\omega_0 t$. Проекція вектора $\vec{A}$ в цьому положенні на вісь OX дорівнює:

$$x = A\cos(\omega_0 t + \varphi_0).$$

За час T, що дорівнює періоду коливань, вектор амплітуди повертається на кут 2π, а проекція його кінця зробить одне повне коливання навколо положення рівноваги O. Отже, обертовий вектор амплітуди повністю характеризує гармонічне коливання.

Нехай точка бере участь у двох гармонічних коливаннях однакової частоти, які напрямлені вздовж однієї прямої:

$$x_1 = A_1\cos(\omega_0 t + \varphi_{01}), \quad (9.17)$$

$$x_2 = A_2\cos(\omega_0 t + \varphi_{02}). \quad (9.18)$$

Ці коливання зручно додати, користуючись методом обертального вектора амплітуди. Для цього відкладемо з точки O під кутом φ_{01} вектор амплітуди $\vec{A}_1$, а під кутом φ_{02} – вектор амплітуди $\vec{A}_2$ (рис.9.6). Оскільки вектори $\vec{A}_1$ і $\vec{A}_2$ обертаються з однаковою кутовою швидкістю, то різниця фаз $\varphi_{02} - \varphi_{01}$ між ними постійна. Оскільки сума проекцій двох векторів на одну вісь дорівнює проекції на ту саму вісь вектора, який є їх сумою, то результуюче коливання можна подати вектором амплітуди $\vec{A}$, що дорівнює сумі векторів $\vec{A}_1$ і $\vec{A}_2$:

$$\vec{A} = \vec{A}_1 + \vec{A}_2,$$

який обертається навколо точки O з тією самою кутовою швидкістю ω_0, що й вектори $\vec{A}_1$ і $\vec{A}_2$. Результуюче коливання описується рівнянням

$$x = x_1 + x_2 = A\cos(\omega_0 t + \varphi), \quad (9.19)$$

де A – амплітуда результуючого коливання, а φ – його початкова фаза.

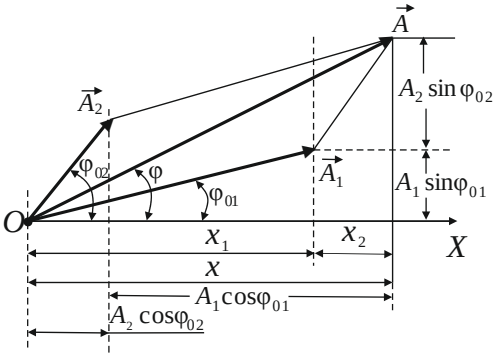


Рис.9.6.

Застосовуючи теорему косинусів до одного з трикутників, на які паралелограм розбивається діагоналлю, з рис. 9.6 видно, що

$$A^2 = A_1^2 + A_2^2 + 2A_1A_2\cos(\varphi_{02} - \varphi_{01}), \quad (9.20)$$

$$\operatorname{tg}\varphi = \frac{A_1\sin\varphi_{01} + A_2\sin\varphi_{02}}{A_1\cos\varphi_{01} + A_2\cos\varphi_{02}}. \quad (9.21)$$

Амплітуда A результуючого коливання залежить від різниці фаз початкових фаз $\varphi_{02} - \varphi_{01}$ коливань, що додаються. Можливі значення A лежать в межах

$$A_1 + A_2 \geq A \geq |A_2 - A_1|. \quad (9.22)$$

§9.5. Додавання гармонічних коливань однакового напрямку з різними частотами. Биття

Якщо частоти коливань (9.17) і (9.18) не однакові, то вектори $\vec{A}_1$ і $\vec{A}_2$ будуть обертатися з різною швидкістю. В цьому випадку результуючий вектор $\vec{A}$ пульсує за величиною і обертається зі змінною швидкістю. Результуючим рухом буде в цьому випадку не гармонічне коливання, а деякий складний коливальний процес.

Особливий інтерес становить випадок коли два гармонічні коливання однакового напрямку, що додаються, мало відрізняються за частотою.

Періодичні зміни амплітуди коливання, які виникають при додаванні двох гармонічних коливань одного напрямку з близькими частотами, називаються **биттями**.

Нехай амплітуди коливань $A_1 = A_2 = A$, а частоти дорівнюють $\omega_1 = \omega_0$, $\omega_2 = \omega_0 + \Delta\omega$ і $\Delta\omega \ll \omega_0$.

Тоді рівняння коливань матимуть вигляд:

$$x_1 = A\cos\omega_0 t; \quad x_2 = A\cos(\omega_0 t + \Delta\omega t). \quad (9.23)$$

Порівнюючи (9.23) з (9.17) і (9.18), можемо вважати, що $\varphi_{01} = 0$, а $\varphi_{02} = \Delta\omega t$.

Використаємо формулу (9.20) і знайдемо амплітуду результуючого коливання (биття).

$$A_б^2 = 2A^2(1 + \cos\Delta\omega t) = 4A^2\cos^2\frac{\Delta\omega}{2}t.$$

Звідки знайдемо, що

$$A_б = 2A\left|\cos\frac{\Delta\omega}{2}t\right|. \qquad (9.24)$$

Фазу результуючого коливання знайдемо за допомогою формули (9.21), а саме:

$$\mathrm{tg}\varphi = \frac{\sin\Delta\omega t}{1+\cos\Delta\omega t} = \mathrm{tg}\frac{\Delta\omega}{2}t.$$

Звідки

$$\varphi = \frac{\Delta\omega}{2}t.$$

Тоді результуюче коливання можна записати в такому вигляді:

$$x = 2A\left|\cos\frac{\Delta\omega}{2}t\right|\cos\left(\omega_0 t + \frac{\Delta\omega t}{2}\right). \quad (9.25)$$

Для випадку, коли $\omega_0 \gg \Delta\omega$, (9.25) записують так:

$$x = 2A\left|\cos\frac{\Delta\omega t}{2}\right|\cos\omega_0 t\,. \qquad (9.26)$$

Отриманий вираз (9.26) є добутком двох коливань. Оскільки $\Delta\omega \ll \omega_0$, то множник $2A\left|\cos\frac{\Delta\omega}{2}t\right|$ майже не зміниться, коли множник $\cos\omega_0 t$ здійснює кілька повних коливань. Тому результуюче коливання x можна розглядати як гармонічне з частотою ω_0 і амплітудою (9.24).

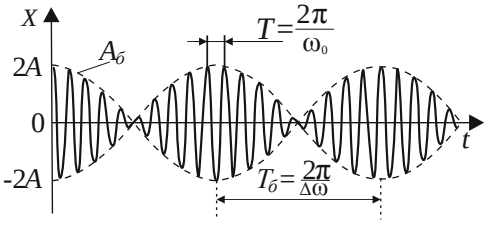


Рис.9.7.

Частота зміни $A_б$ удвоє більша від частоти зміни косинуса (оскільки береться за модулем). Частота биття дорівнює різниці частот коливань, що додаються, тобто $\omega_б = \Delta\omega$. Період биття

$$T_б = 2\pi/\Delta\omega. \qquad (9.27)$$

Суцільні лінії на рис.9.7 дають графік результуючого коливання (9.26) у випадку $\omega_0/\Delta\omega = 10$ і графік амплітуди $A_б$ (пунктирна лінія).

§9.6. Додавання взаємно-перпендикулярних коливань

Нехай матеріальна точка M одночасно бере участь у двох гармонічних коливаннях з однаковою частотою у двох взаємно-перпендикулярних напрямках, як вздовж осі X, так і вздовж осі Y (рис 9.8).

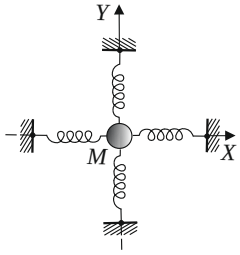


Рис 9.8.

Якщо збудити обидва коливання, матеріальна точка буде рухатись вздовж деякої криволінійної траєкторії, форма якої залежить від різниці фаз обох коливань.

Виберемо початок відліку часу так, щоб початкова фаза першого коливання дорівнювала нулю. Тоді рівняння коливань матимуть такий вигляд:

$$\begin{aligned} x &= A\cos\omega_0 t, \\ y &= B\cos(\omega_0 t + \varphi), \end{aligned} \qquad (9.28)$$

де φ – різниця фаз обох коливань, A і B – їх амплітуди коливань.

Ці вирази – параметрична форма рівняння траєкторії, вздовж якої рухається точка, що бере участь в обох коливаннях. Щоб отримати рівняння траєкторії у звичайному вигляді, треба виключити з цих рівнянь параметр t. Проведемо наступні перетворення:

$$\frac{x}{A} = \cos\omega_0 t,\ \sin\omega_0 t = \sqrt{1-\frac{x^2}{A^2}},$$

$$\frac{y}{B} = \cos(\omega_0 t + \varphi) = \cos\omega_0 t \cdot \cos\varphi -$$

$$-\sin\omega_0 t \cdot \sin\varphi =$$

$$= \frac{x}{A}\cos\varphi - \sqrt{1-\frac{x^2}{A^2}}\sin\varphi;$$

$$\left(\frac{y}{B} - \frac{x}{A}\cos\varphi\right)^2 = \left(-\sqrt{1-\frac{x^2}{A^2}}\sin\varphi\right)^2.$$

$$\frac{x^2}{A^2}\cos^2\varphi - 2\frac{x}{A}\frac{y}{B}\cos\varphi + \frac{y^2}{B^2} =$$

$$= -\frac{x^2}{A^2}\sin^2\varphi + \sin^2\varphi.$$

В результаті отримаємо:

$$\frac{x^2}{A^2} - 2\frac{x}{A}\frac{y}{B}\cos\varphi + \frac{y^2}{B^2} = \sin^2\varphi\ . \quad (9.29)$$

Це рівняння еліпса, осі якого повернуті відносно координатних осей OX і OY. Орієнтація еліпса і величини його півосей залежать від амплітуд A і B і різниці фаз φ.

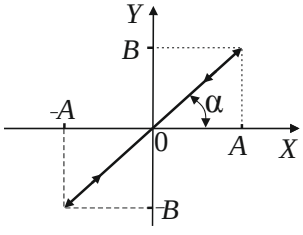


Рис.9.9.

Розглянемо частинні випадки.

1). $\varphi = \pm 2m\pi\ ,\quad m = 0,1,2,3,\ldots\ .$

Тоді

$$\left(\frac{x}{A} - \frac{y}{B}\right)^2 = 0\ .$$

Звідси $y = \frac{B}{A}x$. Результуюче коливання є гармонічним вздовж прямої з частотою ω і амплітудою $\sqrt{A^2+B^2}$ (рис.9.9). Пряма утворює з віссю OX кут $\alpha = \operatorname{arctg}\frac{B}{A}$.

2).$\varphi = \pm(2m+1)\pi\ , m = 0,1,2,3,\ldots.$

У цьому випадку $\left(\frac{x}{A} + \frac{y}{B}\right)^2 = 0$ і $y = -\frac{B}{A}x.$ Результуючий рух – це гармонічне коливання вздовж прямої $y = -\frac{B}{A}x$ (рис.9.10).

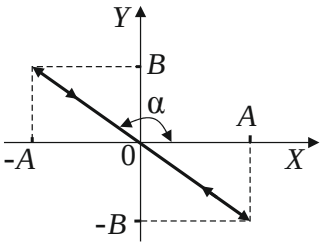


Рис. 9.10.

3).$\varphi = \pm(2m+1)\frac{\pi}{2},\ m = 0,1,2,3,\ldots.$

Тоді рівняння траєкторії буде:

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1.$$

Це рівняння еліпса, осі якого збігаються з осями координат, а його півосі дорівнюють відповідним амплітудам (рис.9.11).

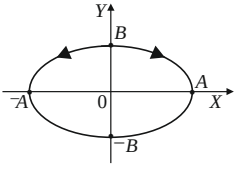


Рис.9.11.

Якщо $A = B$, то еліпс вироджується в коло. Випадки

$$\varphi = (2m+1)\frac{\pi}{2} \text{ і } \varphi = -(2m+1)\frac{\pi}{2}$$

відрізняються напрямком руху по еліпсу (колу) (за годинниковою стрілкою, чи проти неї).

Якщо частоти взаємно перпендикулярних коливань, що додаються, різні, то замкнена траєкторія результуючого коливання досить складна.

Замкнені траєкторії, що кресляться точкою, яка здійснює одночасно два взаємно перпендикулярні коливання з різними частотами, називаються фігурами **Ліссажу**. Форма цих кривих залежить від співвідношення амплітуд, частот і різниці фаз коливань, що додаються (рис.9.12).

Фігури Ліссажу можна спостерігати, наприклад, на екрані електронно-променевого

осцилографа, якщо до двох пар пластин, які відхиляють електронний промінь, підведені змінні напруги з рівними або кратними

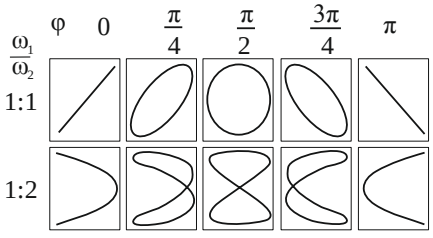


Рис. 9.12.

частотами. Спостереження фігур Ліссажу – зручний метод дослідження співвідношень між частотами і фазами коливань, а також і форми коливань. За допомогою фігур Ліссажу знаходять невідомі частоти коливань.

§9.7.Диференційне рівняння згасаючих коливань і його розв'язання

Розглянемо цей випадок на прикладі електричних коливань, що виникають в контурі, показаному на рис.9.4, що має активний опір R. Енергія, що запасається в контурі, в процесі коливань в контурі поступово витрачається **на нагрівання** опору R, внаслідок чого вільні коливання згасають.

Диференційне рівняння згасаючих коливань отримаємо із (9.6), розділивши його на L:

$$\ddot{q} + \frac{R}{C}\dot{q} + \frac{1}{LC}q = 0.$$

Прийнявши до уваги формулу (9.8) і ввівши позначення

$$\beta = \frac{R}{2L}, \qquad (9.30)$$

який носить назву *коефіцієнта згасання*, отримаємо рівняння

$$\ddot{q} + 2\beta\dot{q} + \omega_0^2 q = 0. \qquad (9.31)$$

Ми дістали лінійне диференційне рівняння другого порядку із звичайними похідними і з сталими коефіцієнтами. Таке саме рівняння можна отримати для напруги U і для величини струму I. Зауважимо, що коливання, які описуються лінійними диференційними рівняннями, називаються лінійними коливаннями, а відповідні коливальні системи – лінійними системами.

Розв'язок рівняння (9.31) будемо шукати у вигляді

$$q = u(t)e^{-\beta t}. \qquad (9.32)$$

Підставимо (9.32) в (9.31), врахувавши, що

$$\dot{q} = (\dot{u} - \beta u)\, e^{-\beta t}$$

і

$$\ddot{q} = (\ddot{u} - 2\beta\dot{u} + \beta^2 u)e^{-\beta t}.$$

Тоді отримаємо для змінної $u(t)$ рівняння:

$$\ddot{u} + (\omega_0^2 - \beta^2)u = 0. \qquad (9.33)$$

Розв'язок цього рівняння має різний вигляд залежно від співвідношення між коефіцієнтами.

1). Допустимо спочатку, що

$$\omega_0^2 > \beta^2, \qquad (9.34)$$

і тоді

$$\omega = \sqrt{\omega_0^2 - \beta^2}. \qquad (9.35)$$

При такій умові рівняння (9.33) перепишеться так:

$$\ddot{u} + \omega^2 u = 0,$$

розв'язок якого по аналогії з (9.7) має вигляд

$$u = q_0 \cos(\omega t + \varphi),$$

де q_0 і φ – сталі, які визначаються з початкових умов. Отже, розв'язком рівняння (9.31) буде

$$q = q_0 e^{-\beta t}\cos(\omega t + \varphi), \qquad (9.36)$$

де $q_0 e^{-\beta t}$ – «амплітуда» згасаючих коливань, q_0 – початкова амплітуда (максимальний заряд конденсатора).

Амплітуда згасаючих коливань зменшується з плином часу, і тим швидше, чим більший коефіцієнт β.

Величина (9.35) називається *власною циклічною частотою згасаючих коливань*. Графік залежності q від часу наведений на рис.9.13.

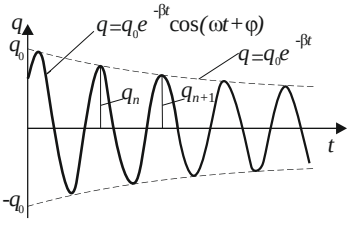


Рис. 9.13.

Щоб знайти силу струму в контурі, продиференціюємо (9.36) по часу:

$$I = \frac{dq}{dt} = \dot{q} = q_0 e^{-\beta t} \cdot (-\beta \cos(\omega t + \varphi) - \omega \sin(\omega t + \varphi)). \quad (9.37)$$

Помножимо і розділимо цей вираз на

$$\sqrt{\omega^2 + \beta^2} = \sqrt{\omega_0^2} = \omega_0. \text{ Тоді}$$

$$I = \omega_0 q_0 e^{-\beta t} \left(-\frac{\beta}{\sqrt{\omega^2 + \beta^2}} \cos(\omega t + \varphi) - \right.$$

$$\left. - \frac{\omega}{\sqrt{\omega^2 + \beta^2}} \sin(\omega t + \varphi) \right).$$

Якщо ввести кут ψ, який визначається умовами:

$$\cos\psi = -\frac{\beta}{\sqrt{\omega^2 + \beta^2}} = -\frac{\beta}{\omega_0},$$

$$\sin\psi = \frac{\omega}{\sqrt{\omega^2 + \beta^2}} = \frac{\omega}{\omega_0},$$

то можна написати, що

$$I = \omega_0 q_0 e^{-\beta t} \cos(\omega t + \varphi + \psi). \quad (9.38)$$

Так як $\cos\psi < 0$, а $\sin\psi > 0$, то значення ψ знаходиться в межах від $\frac{\pi}{2}$ до π ($\frac{\pi}{2} < \psi < \pi$). Таким чином, при наявності в контурі активного опору сила струму випереджає по фазі напругу (заряд) на конденсаторі більше, ніж на $\frac{\pi}{2}$ (при $R = 0$ випередження складає $\frac{\pi}{2}$).

Дослідимо докладніше розв'язок (9.36) і знайдемо насамперед ті моменти часу, в які заряд q досягає максимумів і мінімумів. Для цього за правилом знаходження екстремумів прирівняємо (9.37) нулю. Це дає:

$$\operatorname{tg}(\omega t + \varphi) = -\frac{\beta}{\omega}.$$

Нехай $t = t_1$ є який-небудь розв'язок цього рівняння. Тоді його розв'язками також будуть:

$$\dots, t_1 - 2T, t_1 - T, t_1 + T, t_1 + 2T, \dots;$$

$$\dots, t_1 - \frac{3}{2}T, t_1 - \frac{T}{2}, t_1 + \frac{T}{2}, t_1 + \frac{3}{2}T, \dots$$

де

$$T = 2\pi/\omega = 2\pi/\sqrt{\omega_0^2 - \beta^2}. \quad (9.39)$$

При цьому легко впевнитись, що коли при $t = t_1$ маємо максимум q, то всі значення t, що стоять у першому рядку, відповідають також максимуму ($\frac{d^2q}{dt^2} < 0$), а всі значення другого рядка – мінімуму q. Отже, хоч згасаючі коливання не є періодичним процесом у строгому розумінні цього слова, однак цей процес має певну **повторюваність** у тому розумінні, що максимальних і мінімальних значень заряд (а також струм і напруга) набувають через однакові проміжки часу T. Цей проміжок часу T (формула (9.39)) ми і називаємо **періодом** згасаючих коливань.

Нехай q_n і q_{n+1} – максимальні значення заряду конденсатора (рис.9.13) в двох послідовних максимумах з номерами n і (n+1) в моменти часу t_n і t_{n+1}, причому $t_{n+1} = t_n + T$. Згідно з (9.36) і (9.39), маємо:

$$q_n = q_0 e^{-\beta t_n} \cos(\omega t_n + \varphi),$$

$$q_{n+1} = q_0 e^{-\beta(t_n+T)} \cos\left[\omega\left(t_n + \frac{2\pi}{\omega}\right) + \varphi\right] =$$

$$= q_0 e^{-\beta(t_n+T)} \cos(\omega t_n + \varphi).$$

Поділивши почленно обидві ці рівності, знаходимо:

$$D = \frac{q_n}{q_{n+1}} = e^{\beta T}. \qquad (9.40)$$

Ми бачимо, що відношення двох послідовних максимальних значень заряду не залежить від номера максимумів, або, інакше кажучи, ці максимальні значення зменшуються в геометричній прогресії. Відношення (9.40) називається **декрементом згасання**, а його натуральний логарифм

$$\lambda = \ln D = \beta T \qquad (9.41)$$

логарифмічним декрементом згасання.

Логарифмічний декремент згасання λ можна визначити й інакше. Позначимо через τ час, протягом якого амплітуда коливань зменшується в $e = 2{,}71828\ldots$ раз. Тоді згідно (9.36)

$$\frac{q_0 e^{-\beta(t+\tau)}}{q_0 e^{-\beta t}} = e^{-\beta\tau} = \frac{1}{e},$$

а отже,

$$\beta\tau = 1.$$

Поділивши почленно (9.41) на останнє співвідношення, отримаємо:

$$\frac{T}{\tau} = \frac{1}{N_e} = \lambda.$$

Тут N_e – число повних коливань за час τ, який називається **часом релаксації**. Отже, логарифмічний декремент є величина обернена до числа коливань, після якого амплітуда зменшується в e раз.

Для характеристики згасання коливальних контурів часто користуються, особливо в радіотехніці, ще іншою величиною, яка називається **добротністю контура** і позначають буквою Q. З логарифмічним декрементом вона зв'язана співвідношенням:

$$Q = \pi/\lambda. \qquad (9.42)$$

Через те, що $\lambda = 1/N_e$, то

$$Q = \pi N_e. \qquad (9.43)$$

Добротність контура є помножене на π число повних коливань, після закінчення яких амплітуда зменшується в e раз. Отже, добротність контура тим вища, чим менше згасання коливань у ньому.

Формула (9.35) показує, що частота електричних коливань ω залежить від коефіцієнта згасання β і не дорівнює частоті коливань ω_0 того самого контура при опорі $R = 0$ ($\beta = 0$). Із збільшенням опору контура частота ω зменшується, а період коливань T збільшується.

Покажемо, що при слабкому згасанні (коли $\beta^2 \ll \omega_0^2$) добротність коливальної системи з точністю до множника 2π дорівнює відношенню енергії, яка запасена в системі в даний момент, до зменшення цієї енергії за один період коливань. Згідно (9.36), амплітуда заряду конденсатора зменшується за законом $e^{-\beta t}$. Енергія, яка запасається в конденсаторі, пропорційна квадрату амплітуди заряду; значить енергія конденсатора W зменшується за законом $e^{-2\beta t}$. Відносне зменшення енергії за період коливань дорівнює

$$\frac{\Delta W}{W} = \frac{W(t) - W(t+T)}{W(t)} = \\ = \frac{1 - e^{-2\beta T}}{1} = 1 - e^{-2\lambda}.$$

При слабому згасанні (тобто при умові, що $\lambda \ll 1$) можна за формулами приблизного підрахунку вважати, що $e^{-2\lambda} \approx 1 - 2\lambda$. Тоді

$$\frac{\Delta W}{W} = 1 - (1 - 2\lambda) = 2\lambda.$$

Скориставшись формулою (9.42) отримаємо, що

$$Q = 2\pi \frac{W}{\Delta W}. \qquad (9.44)$$

2). Розглянемо випадок, коли $\omega_0^2=\beta^2$. Це буде мати місце, коли опір коливального контуру

$$R = 2\sqrt{L/C} = R_{кр},$$

де $R_{кр}$ називається критичним опором. Тоді (9.33) прийме вигляд: $\ddot{u} = 0$. Розв'язком цього рівняння є: $u = a+bt$. Це означає, що розв'язок (9.32) буде мати вид:

$$q = (a + bt)e^{-\beta t},$$

де a і b визначаються із початкових умов, наприклад, t=0, q=0, I= I_0. Тоді a=0, а b=I_0 і

$$q = I_0 t \cdot e^{-\beta t}.$$

Цей розв'язок описує так званий аперіодичний процес, графік якого показаний на рис.9.14,*а*. Коливальна система, що виведена із положення рівноваги, повертається (експоненційно) в рівноважний стан. Такий процес використовується в балістичних гальванометрах, приладах для вимірювання кількості заряду, який протік через поперечний переріз провідника в електричному колі

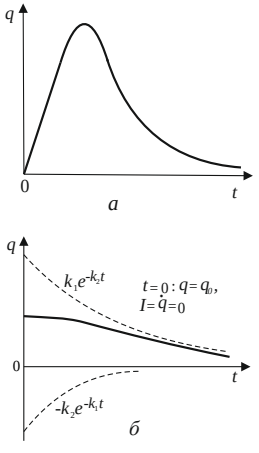


Рис.9.14. Аперіодичний процес в коливальному контурі (*а*) та аперіодичний розряд конденсатора (*б*).

3). Припустимо тепер, що опір контура великий наскільки, що

$$\omega_0^2 < \beta^2.$$

У цьому випадку частота ω, яка виражається формулою (9.35), буде уявною. Це означає, що розв'язок (9.36) вже не існує, а отже, електричних коливань у контурі не буде. В цьому випадку розв'язок основного рівняння (9.31) має вигляд:

$$q = A_1 e^{-k_1 t} + A_2 e^{-k_2 t}, \qquad (9.45)$$

де

$$k_1 = \beta + \sqrt{\beta^2 - \omega_0^2},$$

$$k_2 = \beta - \sqrt{\beta^2 - \omega_0^2},$$

а A_1 і A_2 – довільні сталі. Підставляючи (9.45) у (9.31), можна впевнитись, що рівняння при цьому задовільняється тотожно, а отже, (9.45) є справді шуканим розв'язком. Через те, що $\omega_0^2 < \beta^2$, то k_1 і k_2 обидва дійсні і додатні.

Значення сталих A_1 і A_2 визначаються з початкових умов задачі. Якщо такими є умови:

$$t = 0:\ q = q_0,\ \ I = \dot{q} = 0,\ \text{то}$$

$$q|_{t=0} = A_1 + A_2 = q_0,$$

$$\dot{q}|_{t=0} = -A_1 k_1 - A_2 k_2 = 0.$$

Це дає, що

$$A_1 = -q_0 \frac{k_2}{k_1 - k_2},$$

$$A_2 = q_0 \frac{k_1}{k_1 - k_2}.$$

Після чого розв'язок (9.45) набирає вигляду:

$$q = \frac{q_0}{k_1 - k_2}(k_1 e^{-k_2 t} - k_2 e^{-k_1 t}).$$

На рис.9.14,*б* зображено графічно обидва доданки цієї формули (пунктир) і їх сума (суцільна крива), яка відображає так званий аперіодичний розряд конденсатора.

Якщо опір контура дуже великий, так що $\beta^2 \gg \omega_0^2$, то $k_1 \gg k_2$, і в останньому виразі можна знехтувати другим доданком порівняно з першим, а в знаменнику $-k_2$ порівняно з k_1. Тоді

$$q = q_0 e^{-k_2 t}.$$

Цей випадок відповідає розряду конденсатора.

З вищезазначеного видно, що для можливості електричних коливань необхідно, щоб виконувалась умова (9.34). Підставляючи замість β і ω_0 їх значення (9.30) і (9.8) знаходимо умову можливості коливань у вигляді:

$$\frac{1}{LC} > \frac{R^2}{4L^2},$$

або

$$R < 2\sqrt{L/C} = R_{кр}. \qquad (9.46)$$

Формула (9.46) визначає *критичний опір,* при якому коливальний процес переходить в аперіодичний.

§9.8.Вимушені електричні коливання. Резонанс

1.Щоб викликати вимушені коливання, необхідно, щоб на коливальну систему впливала зовнішня періодично змінна дія. У випадку електричних коливань це можна здійснити, якщо в електричне коло (рис.9.4) включити послідовно з елементами контуру змінну е.р.с. або, розірвавши контур, подати на контакти, які утворились, періодичну напругу (рис. 9.15):

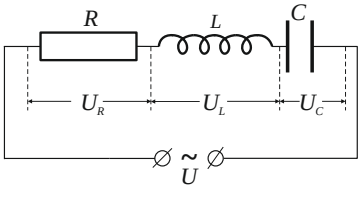


Рис.9.15.

$$U = U_m \cos \omega t, \qquad (9.47)$$

де U_m − амплітуда, а ω − частота вимушуючої сили. Цю напругу необхідно додати до е.р.с. самоіндукції в рівнянні (9.6). В результаті формула (9.6) приймає вид:

$$IR = -\frac{q}{c} - L\frac{dI}{dt} + U_m \cos \omega t. \quad (9.48)$$

Після перетворення отримаємо рівняння

$$\ddot{q} + 2p\dot{q} + \omega_0^2 q = \frac{U_m}{L} \cos \omega t, \quad (9.49)$$

де ω_0 і β визначаються формулами (9.8) і (9.30). Рівняння (9.49) називається диференційним неоднорідним другого порядку рівнянням вимушених коливань.

Загальним розв'язком рівняння (9.49) є сума загального розв'язку лінійного однорідного рівняння (9.31), тобто розв'язок (9.36), і частинного розв'язку рівняння (9.49). Розв'язок (9.36) відіграє помітну роль лише на початковій стадії виникнення вимушених коливань. З часом внаслідок експоненційного множника $e^{-\beta t}$ роль доданка (9.36) зменшується, амплітуда вимушених коливань зростає, доки не досягнення певного значення.

Серед частинних розв'язків рівняння (9.49) є такий, який змінюється з часом синусоїдально з частотою ω зовнішньої вимушуючої сили (9.47). Будемо шукати його в комплексній формі, що можна робити, так як всі математичні операції, з якими прийдеться мати справу, є лінійні і дійсні. Замінимо праву частину рівняння (9.49) на комплексну величину $U_m e^{i\omega t}$, тобто напишемо, що

$$\ddot{q} + 2\beta\dot{q} + \omega_0^2 q = \frac{U_m e^{i\omega t}}{L}\cdot \qquad (9.50)$$

Будемо шукати частинний розв'язок рівняння (9.50) у комплексному вигляді. Знайшовши його, відкинемо його уявну частину і отримаємо розв'язок рівняння (9.49). Частинний розв'язок рівняння (9.50) шукаємо у вигляді

$$q = q_1(\omega) e^{i\omega t}, \qquad (9.51)$$

звідки отримаємо, що

$$\dot{q} = \frac{dq}{dt} = i\omega q_1 e^{i\omega t} = i\omega q,$$

$$\ddot{q} = \frac{d^2q}{dt^2} = i^2\omega^2 q_1 e^{i\omega t} = -\omega^2 q.$$

Підставимо (9.51) в (9.50) і отримаємо

$$q_1 = \frac{U_m/L}{\omega_0^2 - \omega^2 + i2\beta\omega}. \qquad (9.52)$$

Тоді частинний розв'язок (9.51) запишеться так:

$$q = \frac{U_m/L}{\omega_0^2 - \omega^2 + i2\beta\omega} e^{i\omega t}. \qquad (9.53)$$

Розв'язок (9.53) тільки символічно представляє вимушене коливання. В ньому повинна бути залишена тільки **дійсна частина**. Для знаходження останньої введемо позначення

$$\omega_0^2 - \omega^2 + i2\beta\omega = \rho e^{i\psi}, \qquad (9.54)$$

де ρ і ψ – величини дійсні і визначаються згідно теорії комплексних чисел за формулами:

$$\rho = \sqrt{(\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2}, \qquad (9.55)$$

$$\operatorname{tg}\psi = \frac{2\beta\omega}{\omega_0^2 - \omega^2} = \frac{R}{\frac{1}{\omega C} - \omega L}. \qquad (9.56)$$

Таким чином, розв'язок (9.53) може бути записаний, як

$$q = q_m e^{i(\omega t - \psi)}, \qquad (9.57)$$

де

$$q_m = \frac{U_m/L}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\beta^2\omega^2}} \qquad (9.58)$$

є амплітуда вимушених коливань. Виділивши у (9.57) дійсну частину, отримаємо частинний розв'язок рівняння (9.49) у вигляді

$$q = q_m \cos(\omega t - \psi), \qquad (9.59)$$

який для часу $t \gg \tau = 1/\beta$ стане *загальним розв'язком рівняння* (9.49), оскільки вільні коливання, що описуються рівнянням (9.36), практично зовсім згаснуть.

Таким чином, вимушені коливання будуть гармонійними з частотою ω вимушуючої сили (9.47). Амплітуда q_m цих коливань залежить від ω і визначається формулою (9.58). Коливання (9.59) відстають по фазі від прикладеної напруги (9.47) на величину ψ, яка визначається формулою (9.56) і теж залежить від ω.

2. Дослідимо отриманий результат. Продиференцюємо розв'язок (9.59) по t і знайдемо силу струму в контурі для коливань, які встановились:

$$I = \dot{q} = -\omega q_m \sin(\omega t - \psi) = \\ = I_m \cos(\omega t - \psi + \pi/2),$$

де $I_m = \omega q_m$.

Запишемо цей вираз у вигляді

$$I = I_m \cos(\omega t - \varphi), \qquad (9.60)$$

де $\varphi = \psi - \pi/2$, є зсув по фазі між струмом і прикладеною напругою (9.47). У відповідності з (9.56)

$$\operatorname{tg}\varphi = \operatorname{tg}\left(\psi - \frac{\pi}{2}\right) = -\frac{1}{\operatorname{tg}\psi} = \\ = \frac{\omega L - \frac{1}{\omega C}}{R}. \qquad (9.61)$$

Із цієї формули витікає, що струм відстає за фазою від напруги (9.47) ($\varphi > 0$) в тому випадку, коли $\omega L > 1/(\omega C)$, і випереджає напругу (9.47) ($\varphi < 0$) при умові, що $\omega L < 1/(\omega C)$. Згідно (9.58)

$$I_m = \omega q_m = \frac{U_m}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}. \qquad (9.62)$$

Порівнюючи (9.62) із законом Ома (4.8), бачимо, що $\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}$ визначає опір кола. R–називається активним опором, ωL і $1/(\omega C)$ –реактивними опорами (розглядаються в подальшому).

Представимо співвідношення (9.48) у вигляді

$$IR + \frac{q}{C} + L\frac{dI}{dt} = U_m \cos\omega t. \qquad (9.63)$$

Добуток IR дорівнює напрузі U_R на активному опорі, q/C є напруга на конденсаторі U_c, вираз $L(dI/dt)$ визначає напругу на індуктивності U_L з урахуванням цього можна записати

$$U_R + U_C + U_L = U_m \cos \omega t. \quad (9.64)$$

Таким чином, сума напруг (це векторна сума коливань) на окремих елементах контуру дорівнює в кожний момент часу напрузі, яка прикладена зовні (рис.9.15).

Згідно (9.60)

$$U_R = RI_m \cos(\omega t - \varphi). \quad (9.65)$$

Розділивши вираз (9.59) на ємність, отримаємо напругу на конденсаторі:

$$U_C = \frac{q_m}{C}\cos(\omega t - \psi) = \\ = U_{cm}\cos(\omega t - \varphi - \pi/2). \quad (9.66)$$

Тут $U_{cm} = q_m/C = I_m/(\omega C)$, де враховано, що $I_m = \omega q_m$.

Помноживши похідну по часу від функції (9.59) на L, отримаємо напругу на індуктивності

$$U_L = L(dI/dt) = -\omega L I_m \sin(\omega t - \varphi) = \\ = U_{Lm}\cos(\omega t - \varphi + \pi/2). \quad (9.67)$$

Тут

$$U_{Lm} = \omega L I_m. \quad (9.68)$$

Співставлення формул (9.60), (9.65), (9.66), (9.67) показує, що напруга на ємності відстає по фазі від сили струму на $\pi/2$, а напруга на індуктивності випереджає струм на $\pi/2$. Напруга на активному опорі змінюється у фазі зі струмом. Згідно з геометричним представленням гармонічних коливань фазові співвідношення можна наглядно представити за допомогою векторної діаграми. Візьмемо пряму, від якої відраховується початкова фаза, за вісь струмів. Тоді вийде діаграма, яку наведено на рис.9.16. Згідно (9.63), три функції U_R, U_C і U_L в сумі повинні дорівнювати прикладеній напрузі U. У відповідності з цим, напруга U зображується на діаграмі вектором,

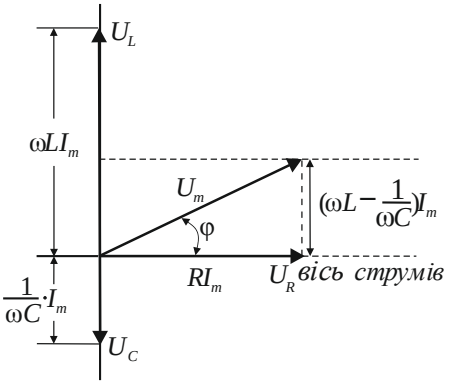


Рис.9.16.

який дорівнює сумі векторів U_R, U_C і U_L. Зауважимо, що із прямокутного трикутника, який утворений на діаграмі векторами U, U_R і різницею $U_L - U_C$, легко отримати формулу (9.62).

3. Резонанс (французьке resonance, від латинського resono – звучу у відповідь, відзиваюсь) відносно великий селективний (вибірковий) відклик коливальної системи на періодичну дію з частотою, близькою до частоти її власних коливань. При резонансі відбувається різке зростання амплітуди вимушених коливань. Резонанс як механічне і акустичне явище вперше описав італійський учений Г.Галілей, а в електромагнітних системах – на прикладі коливального контуру – англійський учений Дж.Максвелл (1868р.).

Резонансну частоту для заряду q і напруги на конденсаторі U_C знайдемо з умови мінімуму підкореневого виразу у формулі (9.58). Тобто похідна по ω від цього виразу повинна дорівнювати нулю:

$$\frac{d}{d\omega}((\omega_0^2 - \omega^2)^2 + 4\beta^2\omega^2) = 0.$$

Звідки знаходимо, що

$$\omega_{q_{рез}} = \omega_{U_{рез}} = \sqrt{\omega_0^2 - 2\beta^2} = \\ = \sqrt{\frac{1}{LC} - \frac{R^2}{2L^2}} \le \omega_0. \qquad (9.69)$$

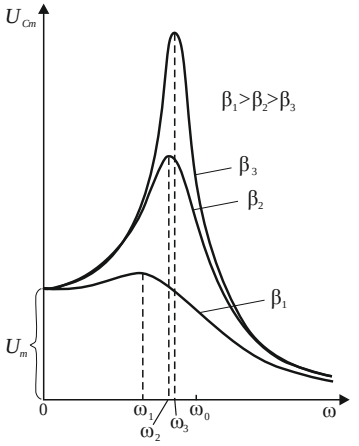


Рис.9.17

Резонансні криві для U_C зображені на рис.9.17 (резонансні криві для q мають такий же вид). При $\omega \to 0$ резонансні криві сходяться в одній точці з ординатою $U_{C_m} = U_m$. Це означає, що конденсатор підключений до джерела постійної напруги U_m. Максимум при резонансі тим вищий і гостріший, чим менший коефіцієнт згасання $\beta = R/(2L)$, тобто чим менший активний опір і більша індуктивність контуру.

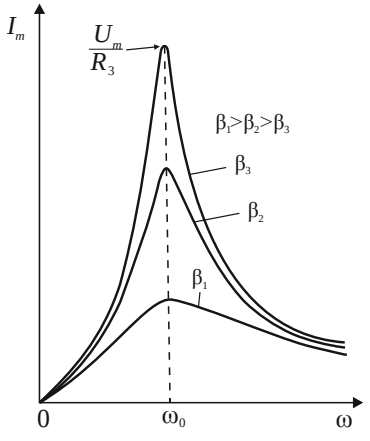


Рис.9.18.

Резонансні криві для сили струму зображено на рис.9.18, для різних коефіцієнтів згасання. Із формули (9.62) видно, що амплітуда сили струму має максимальне значення при $\omega L - 1/(\omega C) = 0$. Отже, резонансна частота для сили струму співпадає із власною частотою контуру ω_0:

$$\omega_{I_{рез}} = \omega_0 = 1/\sqrt{LC}. \qquad (9.70)$$

При $\omega \to 0$, $I_m = 0$, оскільки при постійній напрузі постійний струм в колі з конденсатором протікати не може. Згідно (9.62) при $\omega \to \omega_0$ $I_{m\,рез} = U_m/R$.

Резонансне значення напруги на конденсаторі знайдемо з формули (9.58), поділивши її на ємність C і підставивши замість частоти ω її резонансне значення з формули (9.69). Тоді отримаємо, що

$$U_{cm\,рез} = \frac{U_m \omega_0^2}{2\beta\sqrt{\omega_0^2 - \beta^2}}. \qquad (9.71)$$

Особливий інтерес представляє випадок слабкого згасання (коли $\beta^2 \ll \omega_0^2$). Тоді

$$\frac{U_{cm\,рез}}{U_m} = \frac{1}{\omega_0 CR} = \\ = \frac{\sqrt{LC}}{CR} = \frac{1}{R}\sqrt{\frac{L}{C}} = Q. \qquad (9.72)$$

При цьому враховано, що для випадку слабкого згасання $Q = \frac{\pi}{\lambda} = \frac{\pi}{\beta T} \cong \frac{\pi\omega_0}{2\pi\beta} =$

$$= \frac{2L}{2\sqrt{LC}\cdot R} = \frac{1}{R}\sqrt{\frac{L}{C}}.$$

Таким чином, добротність контуру (для слабого згасання) показує, в скільки разів напруга на конденсаторі може перевищувати прикладену напругу.

Добротність контуру визначає також гостроту резонансних кривих. На рис.9.19 показана одна із резонансних кривих для сили струму в контурі. По вертикальній осі відкладено не значення I_m (формула (9.62)), що

відповідає даній частоті, а відношення $I_m/I_{m\,рез}$. Взагалі кажучи, чим більша добротність контуру Q, тим вужчий і вищий максимум на кривій залежності амплітуди струму від частоти вимушуючої сили. Щоб з'ясуватити це питання з більшою точністю,

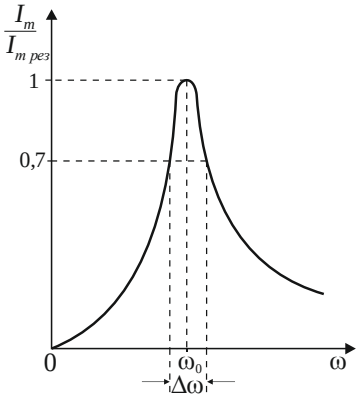


Рис.9.19.

розглянемо частоти, які близькі до ω_0, наприклад, $\omega = \omega_0 \pm \Delta\omega/2$. Потім вирахуємо величину $\omega L - 1/(\omega C)$, яка входить в знаменник рівняння (9.62), з точністю до першого степеня відношення $\Delta\omega/(2\omega_0)$. Тоді отримаємо:

$$\omega L - \frac{1}{\omega C} = \omega_0 L\left(1 \pm \frac{\Delta\omega}{2\omega_0}\right) - $$
$$- \frac{1}{\omega_0 C\left(1 \pm \frac{\Delta\omega}{2\omega_0}\right)} =$$
$$= \omega_0 L\left(1 \pm \frac{\Delta\omega}{2\omega_0} - \frac{1}{1 \pm \frac{\Delta\omega}{2\omega_0}}\right) =$$
$$= \pm\omega_0 L \cdot \frac{\Delta\omega}{\omega_0}.$$

При цьому враховано, що

$$\omega_0 L = \frac{1}{(\omega_0 C)} \text{ і}$$
$$\frac{1}{1 \pm \frac{\Delta\omega}{2\omega_0}} = \left(1 \pm \frac{\Delta\omega}{2\omega_0}\right)^{-1} \approx 1 \mp \frac{\Delta\omega}{2\omega_0}.$$

Розглянемо ширину $\Delta\omega$ кривої $\frac{I_m}{I_{m\,рез}}(\omega)$ на рис.9.19, яку взято на висоті 0,7 (відношення амплітуд струмів, яке дорівнює 0,7, відповідає відношенню потужностей, яке дорівнює $0{,}7^2 \approx 0{,}5$, оскільки потужність пропорційна I^2). Тоді

$$\left(\frac{I_m}{I_{m\,рез}}\right)^2 = \frac{R^2}{R^2 + \left(\omega_0 L\frac{\Delta\omega}{\omega_0}\right)^2} =$$
$$= \frac{1}{1 + \left(\frac{\omega_0 L}{R}\frac{\Delta\omega}{\omega_0}\right)^2} = \frac{1}{1 + \left(Q\frac{\Delta\omega}{\omega_0}\right)^2} = \frac{1}{2}.$$

Звідки отримаємо, що $Q \cdot \frac{\Delta\omega}{\omega_0} = 1$ і

$$Q = \omega_0/\Delta\omega. \qquad (9.73)$$

Точки на осі ω, яким відповідають значення $\omega_0 \pm \Delta\omega/2$, називають точками "половинної енергії". За ширину резонансної кривої часто приймають відстань $\Delta\omega$ між точками половинної енергії. Радіоприймач налаштовують на певну станцію і розрізняють її від інших за допомогою резонансного контуру з добротністю Q, яка дорівнює декільком сотням. Неважко виготовити мікрохвильовий резонансний контур з добротністю Q порядка 10^4 або навіть 10^5.

Звернемося тепер до різниці фаз φ між струмом (9.60) і прикладеною напругою (9.47). З (9.61) видно, що при дуже малих частотах, коли $\omega L \ll 1/(\omega C)$, $\mathrm{tg}\varphi$ дуже великий і від'ємний, а отже, $\varphi \cong -\pi/2$. У цьому разі струм випереджає напругу і коло має ємнісний характер. При збільшенні частоти ω реактивний опір $(\omega L - 1/(\omega C))$, залишаючись від'ємним, зменшується за абсолютним значенням і різниця фаз φ зменшується. Коли $\omega = \omega_0$ формула (9.61) дає $\mathrm{tg}\varphi = 0$, а отже $\varphi = 0$. При подальшому збільшенні ω реактивний опір $\left(\omega L - 1/(\omega C)\right) > 0$ і збільшується із зростанням ω. Тому $0 < \mathrm{tg}\varphi < +\infty$ і $0 < \varphi < +\pi/2$. Отже, при $\omega > \omega_0$ струм відстає від напруги і коло набуває індуктивного характеру, причому кут

φ при збільшені частоти ω асимптотично прямує до граничного значення $+\pi/2$. Залежність різниці фаз φ від частоти коливань зображено графічно на рис.9.20. Так само як і I_m φ залежить ще й від активного опору контура R. Чим менше R, тим швидше змінюється φ біля $\omega = \omega_0$ і в граничному випадку при R=0 зміна фази набуває стрибкоподібного характеру (рис.9.20).

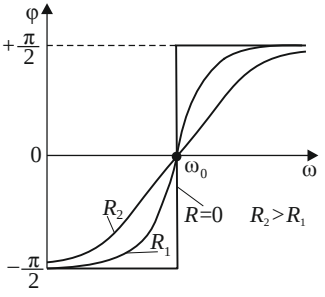


Рис.9.20

Зазначимо, що частота ω_0, при якій настає резонанс, не дорівнює частоті власних коливань контура $\sqrt{\omega_0^2 - \beta^2}$ (формула (9.35)). Проте в переважній більшості практичних випадків $\beta^2 \ll \omega_0^2$, і тому з великим наближенням цією різницею можна знехтувати. Вище ми припускали, що змінюється частота прикладеної напруги ω, а параметри контура залишаються незмінними. Однак зрозуміло, що для одержання резонансу можна робити й інакше: змінювати в контурі індуктивність або ємність (тобто змінювати ω_0), залишаючи частоту ω сталою.

§9.9. Параметричний резонанс

При вивченні електричного резонансу (напруг і струмів) ми розглядали дію періодичної е.р.с. генератора змінної напруги (9.47) на коливальний контур. Проте виявляється, що аналогічні явища спостерігаються і при зовнішньому впливі інших типів, причому коливання, які виникають, як і при дії е.р.с. , істотно залежать від частоти дій на контур. Тому поняття резонансу можна узагальнити і поширити на більшу кількість явищ.

Припустимо, що дія зовнішніх факторів змінює один з параметрів коливальної системи, і розглянемо спочатку простий механічний приклад. Нехай є маятник, довжину якого можна змінювати (рис.9.21),

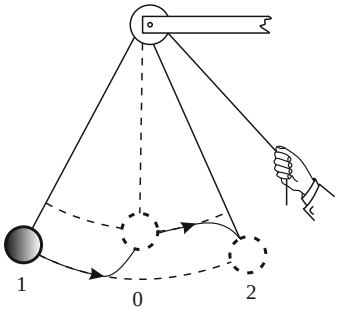


Рис.9.21.Параметричний резонанс у механіці.

підтягуючи кінець нитки, перекинутої через блок, або, навпаки, опускаючи його. Періодично змінюватимемо довжину маятника, підтягуючи нитку (зменшуючи довжину) кожного разу, коли маятник перебуватиме поблизу положення рівноваги (0), і відпускати нитку (збільшувати довжину) при крайніх положеннях маятника (1 і 2), тобто з частотою, яка дорівнює подвоєній частоті власних коливань маятника. Ми знайдемо, що маятник почне здійснювати коливання із *зростаючою* амплітудою, яка збільшуватиметься доти, поки нитка не зіскоче з блока. Це спостерігатиметься і тоді, коли частота зміни довжини не дорівнює подвоєній власній частоті, але близька до неї.

Отже, в цьому досліді ми маємо наростаючі коливання, як і у випадку резонансу під дією періодичної зовнішньої

сили, однак вони виникають внаслідок періодичної зміни одного з *параметрів* системи (довжини). Тому описане явище дістало назву *параметричного резонансу.*

Причину зростання коливань можна пояснити, виходячи з енергетичних міркувань. Коли ми вкорочуємо нитку в положенні 0 (рис.9.21), то зовнішня сила (сила руки) виконує роботу не тільки проти сили тяжіння, а й проти відцентрової сили, бо, проходячи через положення рівноваги 0, маятник має найбільшу швидкість. При видовженні нитки роботу виконує маятник. Проте ця робота виконується тільки за рахунок сили тяжіння, бо в положеннях 1 і 2 відцентрова сила дорівнює нулю (швидкість дорівнює нулю) і тому вона менша від роботи при вкороченні нитки. Отже, у коливальну систему (маятник) безперервно вводиться енергія за рахунок роботи зовнішньої сили, що й веде до зростання коливань. Такі явища параметричного резонансу спостерігаються і в електричних коливальних контурах, якщо параметри контура (ємність чи індуктивність) змінюються періодично. Розглянемо, наприклад, коливальний контур (рис.9.22), що має конденсатор з рухомою пластиною, яку можна

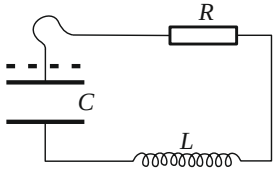


Рис.9.22.Коливальний контур із ємністю, яка змінюється.

періодично наближати до другої пластини або віддаляти від неї. Далі, припустимо, що в контурі внаслідок яких-небудь випадкових причин виникли коливання і що в момент часу, коли заряд конденсатора дорівнює нулю, ми зближуємо пластини. Це явище не супроводитиметься ніякою роботою, бо заряд конденсатора дорівнює нулю, а значить, і сила притягання між пластинами також дорівнює нулю. Через час, який дорівнює чверті періоду власних коливань $T/4$, заряд конденсатора буде найбільшим. Якщо в цей час розсунути пластини, то зовнішні сили виконають роботу, яка витрачається на перемагання взаємного притягання пластин. При цьому ємність конденсатора зменшиться, а напруга між його обкладками збільшиться, сприяючи коливанням у контурі. Якщо потім знову через час $T/4$ зблизити пластини, то енергія контура не зміниться, бо заряд конденсатора в цей момент знову дорівнює нулю. При подальшому розсуванні пластин у контур знову буде введена певна енергія і т.п. Тому, змінюючи значною мірою ємність конденсатора з частотою, яка дорівнює подвоєній частоті власних коливань (чи близька до неї), ми дістанемо в контурі електричні коливання із зростаючою амплітудою, яка збільшуватиметься доти, поки не буде пробитий конденсатор.

Зауважимо, що початкові малі коливання і в механічних, і в електричних системах завжди виникають під дією випадкових зовнішніх факторів, чи флуктуацій. Тому при досить значній періодичній зміні параметрів системи спостерігається *самозбудження* коливань. Правильне співвідношення між фазою коливань і фазою зміни параметра здійснюється при цьому автоматично, бо підсилюються тільки ті з коливань, які мають потрібну початкову фазу.

Описаний дослід з електричним параметричним резонансом вперше здійснили Л.І. Мандельштам і М.Д. Папалекси в 1933 р. Збудована ними ємнісна «параметрична

машина» мала обертовий конденсатор, який складався з двох систем пластин, що мали радіальні вирізи. Одна з систем була нерухома, а друга оберталася за допомогою електродвигуна. Така машина розвивала напруги до багатьох тисяч вольт.

Явище параметричного резонансу можна використати для технічного добування змінних струмів.

§9.10. Змінний струм

Описані в §9.8 вимушені коливання, які встановились в контурі, можна розглядати, як протікання в колі, яке має ємність, індуктивність і активний опір, змінного струму, обумовленого змінною напругою (9.47).

Сила струму визначається за формулами (9.60), (9.61) і (9.62). У випадку, коли $\varphi < 0$, струм фактично випереджає напругу.

Вираз

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}, \qquad (9.73)$$

який є знаменником у формулі (9.61), називається *повним електричним опором* або *імпедансом*.

Якщо коло складається тільки з активного опору *R*, то рівняння закону Ома має вид

$$IR = U_m \cos \omega t.$$

Звідси виходить, що струм в цьому випадку змінюється у фазі з напругою, а амплітуда сили струму дорівнює

$$I_m = U_m/R.$$

Порівняння цього виразу з (9.62) показує, що заміна конденсатора закороченою ділянкою кола означає перехід не до $C = 0$, а до $C = \infty$.

Всяке реальне коло має кінцеві R, L і C. В окремих випадках деякі із цих параметрів бувають такі, що їхнім впливом на струм можна знехтувати. Допустимо, що R кола дорівнює нулю, а C – дорівнює нескінченності. Тоді із формули (9.61) і (9.62) отримаємо, що

$$I_m = \frac{U_m}{\omega L}, \qquad (9.74)$$

а $\mathrm{tg}\varphi = \infty$ (відповідно $\varphi = \pi/2$). Величину

$$X_L = \omega L \qquad (9.75)$$

називають *реактивним індуктивним опором* або просто *індуктивним опором*. Якщо L виразити в генрі, а ω – в рад/с, то X_L буде виражено в Омах. Із (9.75) видно, що індуктивний опір зростає із збільшенням частоти ω. Постійному струму ($\omega = 0$) індуктивність не чинить опору. Напруга на індуктивності випереджає струм на $\pi/2$ (рис.9.16).

Тепер допустимо, що $R = 0$ і $L = 0$. Тоді згідно формул (9.61) і (9.62)

$$I_m = \frac{U_m}{1/(\omega C)}, \qquad (9.76)$$

$\mathrm{tg}\varphi = -\infty$ (тобто $\varphi = -\pi/2$). Величину

$$X_c = \frac{1}{\omega C} \qquad (9.77)$$

називають *реактивним ємнісним опором.* Якщо C виразити у фарадах, а ω – в рад/с, то X_c буде виражено в Омах. Із (9.77) видно, що ємнісний опір зменшується із ростом частоти ω. Для постійного струму ($\omega = 0, X_c = \infty$) – постійний струм через конденсатор протікати не може. Так як $\varphi = -\pi/2$, то струм, який протікає через конденсатор, випереджає напругу на $\pi/2$. Відповідно напруга на

конденсаторі відстає від струму на $\pi/2$ (рис.9.16).

Накінець, допустимо, що R = 0. В цьому випадку формула (9.62) приймає вид:

$$I_m = \frac{U_m}{\left|\omega L - \frac{1}{\omega C}\right|}. \qquad (9.78)$$

Величина

$$X = \omega L - \frac{1}{\omega C} = X_L - X_c \qquad (9.79)$$

називається *реактивним опором* або *реактансом*.

Формули (9.60) і (9.61) можна представити у вигляді

$$\operatorname{tg}\varphi = \frac{X}{R};\ \ I_m = \frac{U_m}{\sqrt{R^2+X^2}} = \frac{U_m}{Z}. \qquad (9.80)$$

Знайдемо потужність, яка виділяється в колі змінного струму. Миттєве значення потужності дорівнює добутку миттєвих значенню напруги і сили струму:

$$P(t)=U(t)\cdot I(t)= \\ =U_m I_m \cos\omega t \cdot \cos(\omega t-\varphi). \qquad (9.81)$$

Скориставшись формулою

$$\cos\alpha \cdot \cos\beta = \frac{1}{2}\cos(\alpha-\beta)+\frac{1}{2}\cos(\alpha+\beta),$$

вираз (9.81) можна представити у вигляді

$$P(t) = \frac{1}{2}U_m I_m \cos\varphi + \\ + \frac{1}{2}U_m I_m \cos(2\omega t-\varphi). \qquad (9.82)$$

Практичний інтерес має середнє за часом значення $P(t)$, яке позначимо просто P. Так як середнє значення $\cos(2\omega t-\varphi)$ дорівнює нулю, то

$$P = \frac{U_m I_m}{2}\cos\varphi. \qquad (9.83)$$

Із (9.82) видно, що миттєва потужність коливається біля середнього значення з частотою, яка у два рази перевищує частоту струму (рис.9.23). Так як $\cos\varphi = (1+\operatorname{tg}\varphi)^{-1/2}$, то, використавши формулу (9.61), отримаємо:

$$\cos\varphi = \frac{R}{\sqrt{R^2+\left(\omega L-\frac{1}{\omega C}\right)^2}} = \frac{R}{Z}. \qquad (9.84)$$

Підставивши це значення $\cos\varphi$ у формулу (9.83) і врахувавши (9.80), отримаємо

$$P = \frac{RI_m^2}{2}. \qquad (9.85)$$

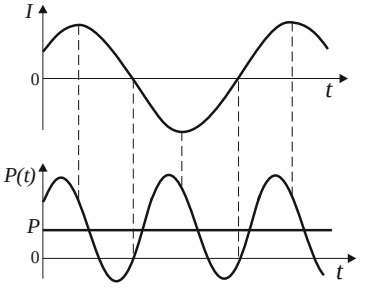


Рис.9.23

Таку ж потужність розвиває постійний струм, сила якого дорівнює

$$I = I_m/\sqrt{2}. \qquad (9.86)$$

Величина (9.86) називається *діючим (або ефективним) значенням сили струму*. Аналогічно величина

$$U = U_m/\sqrt{2}. \qquad (9.87)$$

називається *діючим значенням напруги*.

Формула (9.83) через діючі значення сили струму і напруги має вид

$$P = IU\cos\varphi. \qquad (9.88)$$

Множник $\cos\varphi$, який входить у формулу (9.88) називають *коефіцієнтом потужності*. В техніці стараються зробити $\cos\varphi$ як можна більшим. При малому $\cos\varphi$ для виділення в колі необхідної потужності необхідно пропускати струм великої сили, що призводить до зростання втрат у підвідних проводах.

§9.11.Автоколивання

Автоколивання – це незгасаючі коливання, які підтримуються зовнішніми джерелами енергії в нелінійних коливальних системах, де є розсіювання енергії, вид і властивості яких визначаються самою системою.

Автоколивання принципово відрізняються від інших коливальних процесів в коливальних системах, де є розсіювання енергії тим, що для їхньої підтримки, не потрібно періодичної дії зовні. Коливання струни скрипки при рівномірному русі смичка, струму в радіотехнічному генераторі, повітря в органній трубі, маятника в годиннику, спів соловейка – приклади автоколивань. В найпростіших автоколивальних системах можна виділити коливальну систему із згасанням (коливальний контур), підсилювач коливань, нелінійний обмежувач і ланку *оберненого зв'язку*. Одну з найпростіших класичних схем лампового генератора незгасаючих коливань подано на рис.9.24. Коливальний контур, що має ємність C і індуктивність L, увімкнений в коло сітки електронної лампи. В колі анода, крім. батареї живлення, є ще котушка L', розміщена в безпосередній близькості до котушки L, так, що між обома котушками існує індуктивний зв'язок.

Принцип дії генератора такий. Коли в коливальному контурі виникають коливання (при вмиканні батареї або під впливом яких-небудь випадкових причин), то між обкладками конденсатора виникає змінна напруга. Така сама напруга виникає між сіткою і катодом K, бо вони приєднані до обкладок конденсатора. Внаслідок цього в колі анода виникає змінний струм I_a. Але котушки L' і L індуктивно зв'язані між собою і тому змінний струм I_a спричинює в котушці L змінну е.р.с. взаємної індукції

$$\mathcal{E} = -L_{12}\frac{dI_a}{dt}, \qquad (9.89)$$

де L_{12} – коефіцієнт взаємної індукції. Ця е.р.с., залежно від взаємного напрямку витків котушок L' і L, може або перешкоджати коливанням струму в контурі, або сприяти їм.

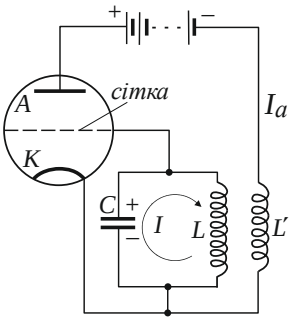


Рис.9.24.Ламповий генератор з коливальним контуром у колі сітки.

Очевидно, що, перемикаючи при потребі кінці котушки L, можна завжди домогтися, щоб е.р.с. $\mathcal{E}$, яка виникає, мала такий самий знак, як і струм I в контурі. Тоді робота цієї е.р.с. буде додатною і коливальний контур одержуватиме енергію за рахунок джерела струму, так що згасання контура зменшиться. Якщо зв'язок між котушками досить сильний, одержувана енергія може бути більша за ту, яка витрачається в контурі. У цьому випадку амплітуда коливань зростатиме доти, поки одержувана енергія не дорівнюватиме втратам, після чого встановляться незгасаючі коливання.

Подивимося тепер, яку кількісну умову повинні задовольняти параметри лампового генератора (опір контура R коефіцієнт взаємоіндукції L_{12} і т.д.), щоб автоколивання були можливі.

Знову звернемось до схеми на рис.9.24. Коливальний контур цього генератора можна подати еквівалентною схемою рис.9.25, в якій зворотний зв'язок здійснюється деяким генератором із змінною е.р.с. (9.89).

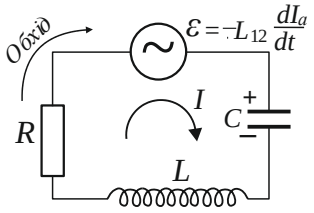


Рис.9.25.Еквівалентна схема генератора, зображеного на рис.9.24.

Застосовуючи до цього контуру друге правило Кірхгофа і міркуючи так само, як і в §9.6, отримаємо:

$$IR - U_C = -L\frac{dI}{dt} - L_{12}\frac{dI_a}{dt}.$$

Враховуючи, що $U_C = q/C$ і $I = -\frac{dq}{dt}$, це рівняння перепишемо так:

$$\frac{d^2q}{dt^2} + \frac{R}{L}\frac{dq}{dt} + \frac{q}{LC} - \frac{L_{12}}{L}\cdot\frac{dI_a}{dt} = 0. \quad (9.90)$$

У цих рівняннях анодний струм I_a є функція сіткової напруги u_c (що виражається сітковою характеристикою лампи $I_a(u_c)$), яка в свою чергу залежить від заряду конденстора. Але сіткова характеристика лампи нелінійна. Тому і написане рівняння є нелінійним рівнянням, а отже, автоколивання є *нелінійними коливаннями.* Однак, щоб розв'язати поставлене питання, задачу можна спростити. Припустимо, що ми маємо *малі* коливання і що робочу точку вибрано в середній частині сіткової характеристики. Тоді малий відрізок характеристики можна вважати відрізком прямої лінії і рівняння коливань буде *лінійним*. Крім того, для простоти вважатимемо, що проникність сітки лампи дуже мала, так, що I_a практично залежить тільки від потенціалу сітки u_c (але не від потенціалу анода U_a). Тоді потенціал сітки визначається напругою на конденсаторі U_C і

$$I_a = a + SU_C = a + \frac{S}{C}q,$$

де S – так звана крутість сіткової характеристики лампи, a – постійна . Тому

$$\frac{dI_a}{dt} = \frac{S}{C}\frac{dq}{dt}.$$

Підставивши цей вираз у рівняння коливань (9.90), знаходимо:

$$\ddot{q} + \left(\frac{R}{L} - \frac{SL_{12}}{LC}\right)\dot{q} + \frac{1}{LC}q = 0. \quad (9.91)$$

Це рівняння того самого виду, що й (9.31), і його розв'язком є (9.36):

$$q = q_0 e^{-\beta' t}\cos(\omega t + \varphi). \quad (9.92)$$

Але в цьому випадку коефіцієнт згасання β' дорівнює

$$\beta' = \frac{1}{2}\left(\frac{R}{L} - \frac{SL_{12}}{LC}\right). \quad (9.93)$$

Він менший за коефіцієнт $\beta = R/(2L)$ при відсутності зворотного зв'язку і тому можна сказати, що дія зворотного зв'язку еквівалентна введенню в контур *від'ємного опору.*

З (9.93) видно, що при певних умовах коефіцієнт згасання β' може дорівнювати нулю або навіть бути від'ємним, що відповідає виникненню автоколивань. Це буде тоді, коли

$$SL_{12}/C > R. \quad (9.94)$$

Записана умова є *умовою самозбудження* лампового генератора.

Якщо виконується умова (9.94), то $\beta' < 0$ і розв'язок (9.92) відображає *наростаючі* коливання, амплітуда яких збільшується з часом до нескінченності. Причина цього дивного результату полягає в

тому, що ми нелінійне рівняння коливань (9.90) замінили наближеним лінеаризованим рівнянням (9.91), яке справджується тільки для малих коливань (початкової стадії процесу), але не відповідає всім властивостям генератора. Насправді ж після самозбудження амплітуда наростаючих коливань прагнутиме до деякої *скінченної* границі, що визначається властивостями генератора і не залежить від початкових умов.

Контрольні питання

1.Дати визначення коливального руху.

2.Які коливання називаються гармонічними? Записати рівняння гармонічних коливань.

3.Характеристики гармонічних коливань: амплітуда, фаза, циклічна частота, лінійна частота, початкова фаза. Дати їх визначення.

4.Як проявляється закон збереження енергії при гармонічних коливаннях?

5.Геометричне зображення гармонічного коливання.

6.Додавання однаково спрямованих гармонічних коливань з однаковими та різними частотами. При якій умові утворюються коливання, що носять назву «биття»?

7.Додавання взаємно перпендикулярних гармонічних коливань з однаковими та різними частотами. Що таке фігури Ліссажу?

8.Яка відповідність між гармонічними механічними коливаннями і коливаннями, які виникають в ідеалізованому електричному контурі?

9.Записати рівняння згасаючих коливань.

10.Дати визначення коефіцієнта та логарифмічного декремента згасання.

11.Розкрити фізичний зміст добротності коливального контура .

12.Як відрізняється циклічна частота згасаючих коливань від “власної” циклічної частоти коливальної системи?

13.Дати визначення аперіодичного процесу. Записати формулу для визначення критичного опору коливального контуру.

14.Записати рівняння вимушених коливань.

15.З якою частотою відбуваються вимушені коливання?

16.Записати і пояснити формулу для амплітуди вимушеного коливання.

17.Що таке резонанс?

18.Чим пояснюється зсув по фазі між вимушуючою силою (напругою) і зміщенням (зарядом) при вимушених коливаннях?

19.Який зв'язок між вимушеними коливаннями та змінним струмом?

20.Використання параметричних коливань і автоколивань в техніці зв'язку.

Розділ 10.Пружні хвилі

Хвилі

Хвилі – це зміни стану середовища (збурення), які розповсюджуються в цьому середовищі і несуть з собою енергію. Найбільш важливі хвилі, які часто зустрічаються, це такі: пружні хвилі, хвилі на поверхні рідини і електромагнітні хвилі. Окремим випадком пружних хвиль є звук і сейсмічні хвилі, а з електромагнітних – радіохвилі, світло, рентгенівські промені та інші.

Основна властивість всіх хвиль, незалежно від їхньої природи, в тому, що у хвилях здійснюється перенос енергії без переносу речовини (останнє може мати місце

тільки як побічне явище). Хвильові процеси зустрічаються майже у всіх областях фізичних явищ, тому їх вивчення має велике значення.

§10.1.Утворення хвиль у пружному середовищі. Поздовжні і поперечні хвилі

Розглянемо пружне середовище між частинками якого існують сили взаємодії, що перешкоджають тому або іншому виду його деформації. Тіло, яке коливається в пружному середовищі, періодично діє на прилеглі до нього частинки середовища, виводячи їх з положення рівноваги і змушуючи здійснювати вимушені коливання. При цьому середовище поблизу тіла деформується і в ньому виникають пружні сили. Ці сили діють як на прилеглі до тіла частинки, намагаючись повернути їх у положення рівноваги, так і на віддаленіші від тіла частинки, виводячи їх з положення рівноваги. Віддаленіші від тіла області середовища поступово втягуються в коливальний рух. *Процес поширення коливань в суцільному середовищі, яке неперервно розподілене в просторі і має пружні властивості, називається механічним хвильовим процесом, або механічною хвилею*.

При поширенні хвилі частинки середовища не рухаються разом з хвилею, а коливаються біля своїх положень рівноваги.

Пружними (або механічними) хвилями називаються механічні збурення, що поширюються у пружному середовищі. Пружні хвилі бувають поперечні і поздовжні.

У поперечних хвилях частинки середовища коливаються в площинах, які перпендикулярні до напрямку поширення хвилі (рис.10.1.). Поперечні хвилі можуть поширюватись в середовищі, в якому виникають пружні сили при деформації зсуву, тобто лише у твердих тілах.

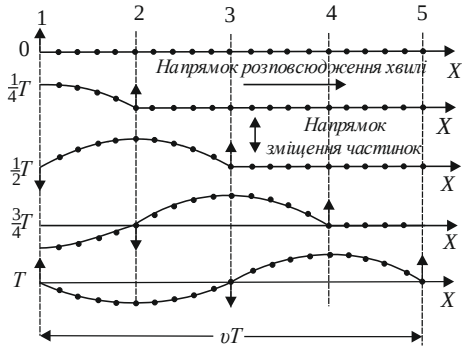


Рис.10.1.

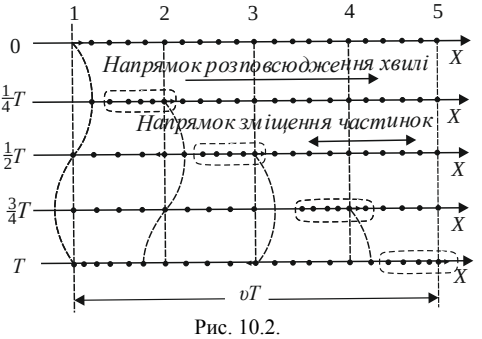


Рис. 10.2.

У поздовжніх хвилях частинки коливаються в напрямку поширення хвилі (рис.10.2). Ці хвилі можуть поширюватись в середовищах, в яких виникають пружні сили при деформації стиску і розтягу, тобто у твердих, рідких і газоподібних тілах. У твердих тілах обмежених розмірів (наприклад, в стрижнях ,пластинках) картина розповсюдження хвиль більш складна: тут виникають ще і інші типи хвиль, які є комбінацією перших двох основних типів.

Загальні характеристики і властивості хвиль

Хвилі можуть мати різну форму. Поодинокою хвилею, або імпульсом, називають порівняно коротке збурення, яке не має регулярного характеру (рис.10.3,*а*).

Обмежений ряд збурень, які

повторюються, називається цугом хвиль. Як правило, поняття цуга відноситься до відрізка синусоїди (рис.10.3,б). Особливе значення в теорії хвиль має уявлення про гармонійні хвилі, тобто нескінченні синусоїдальні хвилі, в яких всі зміни стану середовища відбуваються за законом синуса або косинуса (рис.10.3,*в*), так як такі хвилі могли би розповсюджуватись в однорідному середовищі без спотворення форми (якщо амплітуда їх невелика).

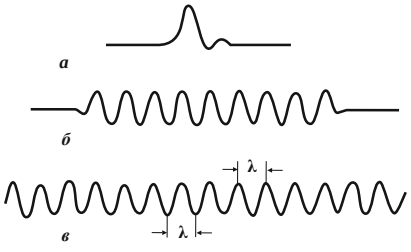


Рис.10.3.*а* – поодинока хвиля; *б* – цуг хвиль; *в* – нескінченна синусоїдальна хвиля.

Пружна хвиля називається *гармонічною*, якщо відповідні їй коливання частинок середовища є гармонійними.

Нехай поперечна гармонійна хвиля поширюється вздовж осі *OX*. На рис.10.4 наведена гармонійна хвиля (поперечна або поздовжня), яка поширюється вздовж осі *OX* зі швидкістю v. Тобто наведена залежність

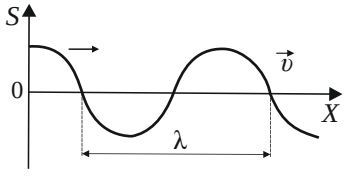


Рис.10.4.

зміщень *S* всіх частинок середовища, що беруть участь у хвильовому процесі, від відстані x цих частинок від джерела коливань *O* для якогось фіксованого моменту часу. Основними характеристиками гармонійної хвилі є *довжина хвилі* λ – відстань між двома максимумами або мінімумами (рис.10.3,*в*, рис.10.4) і *період T* – час, за який здійснюється один повний цикл коливань. Довжина хвилі λ зв'язана з періодом коливань *T* співвідношенням

$$\lambda = vT, \qquad (10.1)$$

де v – швидкість розповсюдження хвилі (рис.10.1; 10.2; 10.4). Співвідношення (10.1) справедливе для гармонійної хвилі любої природи. Довжину хвилі можна визначити ще й як відстань, на яку розповсюджується хвиля за час, що дорівнює періоду коливань частинок середовища, або як відстань між найближчими точками середовища, які коливаються з різницею фаз, що дорівнює 2π (рис.103,*в*). Замість періоду *T* нерідко користуються *частотою* ν, яка дорівнює числу періодів за одиницю часу:

$$\nu = 1/T, \qquad (10.2)$$

при цьому

$$\lambda\nu = v. \qquad (10.3)$$

В теорії хвиль користуються також поняттям *хвильового вектора* $\vec{k}$, за абсолютною величиною

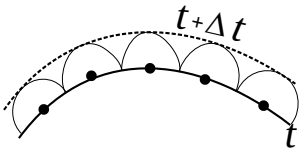
$$k = \frac{2\pi}{\lambda} = \frac{2\pi\nu}{v} = \frac{\omega}{v}, \qquad (10.4)$$

тобто він дорівнює числу довжин хвилі на відрізку 2π і орієнтований в напрямку розповсюдження хвилі.

Нехай точка, від якої йдуть коливання, коливається в суцільному середовищі. Коливання поширюються від центра у всі боки.

Поверхня, до якої доходять коливання в деякий момент часу, називається *фронтом хвилі.*

Фронт хвилі – це поверхня, яка відокремлює частину простору, уже залучену у хвильовий процес, від області, в якій коливання ще не виникли. *Принцип Гюйгенса*



встановлює посіб побудови фронту хвилі в момент часу $t+\Delta t$, якщо відоме положення фронту в момент часу t. Згідно принципу Гюцгенса кожна точка, до якої доходить хвильовий рух, стає центром вторинних хвиль; огинаюча цих хвиль дає положення фронту хвилі в наступний момент часу (див.рис.).

Поверхня, всі частинки якої коливаються з однаковими фазами, називається *хвильовою поверхнею*.

Хвильову поверхню можна провести через довільну точку простору, який охоплений хвильовим процесом. Отже, хвильових поверхонь існує нескінченна множина, а хвильовий фронт в кожний момент часу лише один. Хвильові поверхні залишаються нерухомими, а хвильовий фронт весь час переміщується.

Хвильові поверхні можуть бути довільної форми.

Хвиля називається *плоскою*, якщо її хвильові поверхні мають вигляд площин, які паралельні до площини, що проходить через джерело хвиль.

Хвиля називається *сферичною*, якщо її хвильові поверхні мають вигляд концентричних сфер. Центр цих сфер називається центром хвилі.

Напрямки, в яких поширюються коливання, називаються *променями*. В ізотропному середовищі промені перпендикулярні до фронту хвилі.

Поширення в пружному середовищі механічних збурень, збуджених джерелом хвиль, пов'язане з перенесенням хвилями енергії. Тому такі хвилі називаються *біжучими хвилями*.

§10.2.Рівняння плоскої і сферичної хвиль

Рівнянням хвилі називається вираз, який описує зміщення частинки, яка коливається, як функцію її рівноважних координат x, y, z і часу t.

Розглянемо плоску хвилю, яка поширюється вздовж осі *OX* і збуджується в площині $x = 0$ (рис.10.5). Нехай коливання в площині $x = 0$ мають вигляд:

$$S(0,t) = A\cos(\omega t + \varphi_0).$$

Знайдемо вигляд рівняння коливань частинок у площині, що відповідає довільному значенню x. Для того, щоб пройти шлях від площини $x = 0$ до площини на відстані x, хвилі потрібен час $\tau = \frac{x}{v}$, де v – швидкість поширення хвилі. Отже, коливання частинок,

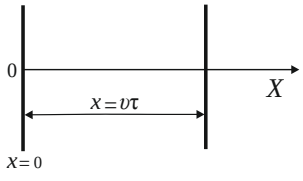


Рис.10.5.

що лежать у площині x, будуть запізнюватись на час τ від коливань частинок в площині $x=0$, тобто матимуть вигляд:

$$S(x,t) = A\cos[\omega(t-\tau)+\varphi_0] =$$
$$= A\cos\left(\omega t - \frac{\omega x}{v} + \varphi_0\right) =$$
$$= A\cos\left(\omega t - \frac{2\pi}{vT}x + \varphi_0\right) =$$
$$= A\cos\left(\omega t - \frac{2\pi}{\lambda}x + \varphi_0\right).$$

Тоді рівняння біжучої плоскої хвилі, що поширюється вздовж осі *OX*, має такий вигляд із врахуванням (10.4):

$$S(x,t) = A\cos(\omega t - kx + \varphi_0), \quad (10.5)$$

де A – амплітуда коливань, яка називається амплітудою хвилі; $\omega = \frac{2\pi}{T}$ – циклічна частота хвилі; k–хвильове число хвилі; φ_0 – початкова фаза коливань в площині x=0. Величина $\omega t - kx + \varphi_0$ визначає фазу коливань в довільній площині з координатою x і називається фазою плоскої хвилі. Зафіксуємо певне значення фази:

$$\omega t - kx + \varphi_0 = \text{const}. \quad (10.6)$$

Цей вираз визначає зв'язок між часом t і тим місцем x, в якому фаза має зафіксоване значення. Величина $\frac{dx}{dt}$ дає швидкість, з якою переміщується дане значення фази. Продиференціюємо вираз (10.6):

$$\omega dt - k dx = 0.$$

Звідси

$$\frac{dx}{dt} = \frac{\omega}{k} = \frac{\omega\lambda}{2\pi} = \frac{\lambda}{T} = v. \quad (10.7)$$

Отже, швидкість v поширення хвилі є ніщо інше, як швидкість переміщення фази хвилі і її називають ***фазовою швидкістю.***

Якщо плоска хвиля поширюється в довільному напрямку, то рівняння хвилі запишеться так:

$$S(\vec{r},t) = A\cos(\omega t - \vec{k}\vec{r} + \varphi_0), \quad (10.8)$$

де $\vec{k} = k\vec{n}$ – вектор, який дорівнює за модулем хвильовому числу і має напрям нормалі $\vec{n}$ до хвильової поверхні.

У випадку сферичної хвилі рівняння хвилі буде:

$$S(r,t) = \frac{A}{r}A\cos(\omega t - kr + \varphi_0), \quad (10.9)$$

де r – відстань від центра хвилі до точки середовища, яка розглядається. У випадку сферичної хвилі навіть у середовищі, яке не поглинає енергію, амплітуда коливань не залишається постійною, а зменшується з відстанню за законом A/r. Це співвідношення справедливе лише для r, що значно більші за розміри джерела.

Згідно (10.7) $dx/dt > 0$. Це означає, що рівняння (10.5) описує хвилю, яка розповсюджується в сторону зростання x. Хвиля, яка розповсюджується в протилежному напрямку, описується рівнянням

$$S(x,t) = A\cos(\omega t - kx + \varphi_0). \quad (10.10)$$

Дійсно, прийнявши фазу хвилі за константу і продиференціювавши по часу, отримаємо, що $\frac{dx}{dt} = -v$. Це означає, що хвиля, яка описується рівнянням (10.10) розповсюджується в сторону зменшення x.

§ 10.3. Хвильове рівняння

Рівняння довільної хвилі є розв'язком ***рівняння,*** яке називається ***хвильовим.***

Для встановлення виду цього рівняння використаємо рівняння плоскої хвилі (10.8), що поширюється в довільному напрямку:

$$S(\vec{r},t) = A\cos\left(\omega t - \vec{k}\vec{r} + \varphi_0\right) =$$
$$= A\cos(\omega t - k_x x - k_y y - k_z z + \varphi_0).$$

Продиференціюємо цю функцію двічі за кожною змінною:

$$\frac{\partial^2 S}{\partial t^2} = -\omega^2 A\cos\left(\omega t - \vec{k}\vec{r} + \varphi_0\right) = -\omega^2 S,$$

$$\frac{\partial^2 S}{\partial x^2} = -k_x^2 A\cos\left(\omega t - \vec{k}\vec{r} + \varphi_0\right) = -k_x^2 S,$$

$$\frac{\partial^2 S}{\partial y^2} = -k_y^2 A\cos\left(\omega t - \vec{k}\vec{r} + \varphi_0\right) = -k_y^2 S,$$

$$\frac{\partial^2 S}{\partial z^2} = -k_z^2 A\cos\left(\omega t - \vec{k}\vec{r} + \varphi_0\right) = -k_z^2 S.$$

Додамо ліві і праві частини похідних за координатами:

$$\frac{\partial^2 S}{\partial x^2}+\frac{\partial^2 S}{\partial y^2}+\frac{\partial^2 S}{\partial z^2}= =-\left(k_x^2+k_y^2+k_z^2\right)S=-k^2 S.$$

Зіставимо цю суму з похідною за часом і врахуємо згідно (10.7), що $\omega^2 = v^2 k^2$.

Тоді поширення хвиль в однорідному ізотропному середовищі описується ***хвильовим рівнянням*** – диференційним рівнянням в частинних похідних:

$$\frac{\partial^2 S}{\partial x^2}+\frac{\partial^2 S}{\partial y^2}+\frac{\partial^2 S}{\partial z^2}=\frac{1}{v^2}\frac{\partial^2 S}{\partial t^2}, \quad (10.11)$$

або

$$\Delta S = \frac{1}{v^2}\frac{\partial^2 S}{\partial t^2}, \quad (10.12)$$

де v – фазова швидкість, а $\Delta = \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$ – оператор Лапласа.

Легко переконатись в тому, що хвильовому рівнянню (10.11) задовільняє не тільки функція (10.8), але і будь-яка функція виду

$$S(x,y,z,t)= =f(\omega t-k_x x-k_y y-k_z z+\varphi_0). \quad (10.13)$$

Будь-яка функція, яка задовільняє рівнянню виду (10.11), описує деяку хвилю, причому корінь квадратний із величини, оберненої коефіцієнту при $\frac{\partial^2 S}{\partial t^2}$, дає фазову швидкість цієї хвилі.

В подальшому будемо розглядати плоскі хвилі, які розповсюджуються вздовж осі *X*, і рівняння (10.11) буде мати вид:

$$\frac{\partial^2 S}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 S}{\partial t^2}. \quad (10.13)$$

ОСТРОГРАДСЬКИЙ МИХАЙЛО ВАСИЛЬОВИЧ
(1801 – 1862)
розв'язав у 1826 р. задачу поширення хвиль на поверхні води.

§10.4.Енергія пружної хвилі

Нехай в деякому середовищі поширюється в додатному напрямку осі *OX* плоска хвиля (10.5).

Визначимо зміну енергії малого об'єму *dV* пружного середовища, пов'язану з поширенням у середовищі плоскої хвилі. Оскільки об'єм *dV* дуже малий, то можна вважати, що всі частинки середовища, які містяться в цьому об'ємі, коливаються з однаковою фазою так, що їх швидкості однакові і $v = \frac{dS}{dt} = -A\omega\sin(\omega t - kx + +\varphi_0)$. Тому кінетичну енергію об'єму *dV*, яка пов'язана з коливальним рухом, запишемо так:

$$dW_k = \frac{v^2}{2}dm = \frac{v^2}{2}\rho dV = = \frac{1}{2}\rho A^2\omega^2 dV\sin^2\left(\omega t - kx + \varphi_0\right),$$

де ρ – густина середовища.

Визначаючи роботу деформації об'єму *dV* середовища під час хвильового руху, можна показати, що потенціальна енергія dW_p об'єму *dV* середовища дорівнює його кінетичній енергії:

$$dW_p = dW_k.$$

Повна механічна енергія коливального руху об'єму *dV* дорівнює

$$dW = dW_k + dW_p = = \rho A^2\omega^2 dV\sin^2\left(\omega t - kx + \varphi_0\right).$$

Об'ємна густина енергії хвиль у пружному середовищі

$$w = \frac{dW}{dV} = =\rho A^2\omega^2\sin^2(\omega t-kx+\varphi_0). \quad (10.14)$$

Густина енергії в кожний момент часу в різних точках простору різна. В одній і тій же точці густина енергії змінюється з часом за законом квадрату синуса. Середнє значення квадрата синуса дорівнює $1/2$. Відповідно, середнє за часом значення об'ємної густини енергії в кожній точці середовища дорівнює:

$$\langle w\rangle = (1/2)\rho A^2\omega^2. \quad (10.15)$$

Поширення хвиль у пружному середовищі нерозривно пов'язане з процесом

передавання енергії від одних ділянок середовища до інших. Саме тому при хвильовому русі об'ємна густина енергії w коливань у кожній точці середовища змінюється з часом.

Об'ємну густину енергії (10.14) можна записати так:

$$w=(1/2)\rho A^2\omega^2\cdot \cdot\left(1-\cos 2\left(\omega t-kx+\varphi_0\right)\right). \quad (10.16)$$

Швидкість u поширення енергії хвилі дорівнює швидкості переміщення в просторі поверхні, яка відповідає максимальному значенню об'ємної густини енергії w. Рівняння поверхні $w = w_{max}$ має вигляд:

$$2(\omega t - kx + \varphi_0) = \pi.$$

Продиференціюємо цей вираз і прирівняємо до нуля:

$$\omega dt - kdx = 0.$$

Звідси швидкість переміщення поверхні, що відповідає максимальному значенню об'ємної густини енергії w буде

$$u = \frac{dx}{dt} = \frac{\omega}{k} = v.$$

Отже, швидкість поширення енергії хвилі збігається з фазовою швидкістю хвилі.

Для характеристики процесу перенесення енергії хвилями введемо поняття про потік енергії.

Потоком енергії Φ крізь яку-небудь поверхню площею S називається фізична величина, яка чисельно дорівнює кількості енергії dW, яка переноситься через цю поверхню за одиницю часу:

$$\Phi = \frac{dW}{dt}. \quad (10.17)$$

Потік енергії – скалярна величина, розмірність якої дорівнює розмірності енергії поділеній на розмірність часу, тобто співпадає з розмірністю потужності. Тому Φ вимірюється у ватах.

Знайдемо потік енергії хвилі, що рухається з фазовою швидкістю $\vec{v}$, через площину dS (рис.10.6). За час dt хвиля перенесе енергію, що міститься всередині косого циліндра, об'єм якого

$$dV = vdt\, dS\cos\alpha = vdt\, dS_n.$$

Тоді $dW = wdV = wv\, dt\, dS_n$ і потік енергії, $d\Phi = wvdS_n = w\vec{v}\cdot d\vec{S}$, де w – об'ємна густина енергії хвилі, $d\vec{S} = \vec{n}dS$ – вектор площини dS, $\vec{n}$ – одинична нормаль до dS.

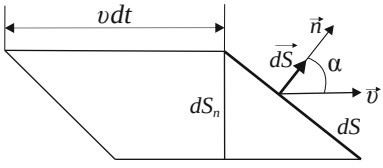


Рис.10.6.

В різних точках простору вводиться векторна величина $\vec{j}$, яка називається густиною потоку енергії.

Густина потоку енергії – векторна величина, яка напрямлена у бік поширення хвилі і числово дорівнює потоку енергії $d\Phi$ через одиницю площі dS поверхні, яка розташована перпендикулярно до напрямку поширення хвилі:

$$j = \frac{d\Phi}{dS} = \frac{wvdS_n}{dS_n} = wv.$$

Оскільки швидкість $\vec{v}$ – це вектор, модуль якого дорівнює фазовій швидкості хвилі, а напрямок збігається з напрямком поширення хвилі (і перенесення енергії), то

$$\vec{j} = w\vec{v}. \quad (10.18)$$

Вектор густини потоку енергії хвилі, який називається вектором Умова, дорівнює добутку вектора швидкості поширення енергії хвилі на величину її об'ємної густини.

Вектор $\vec{j}$ в різних точках простору має

неоднакові значення, а в даній точці простору змінюється з часом за законом квадрата синуса. Середнє значення вектора Умова:

$$\langle \vec{J} \rangle = \langle w \rangle \vec{v} = (1/2)\,\rho A^2 \omega^2 \vec{v}. \quad (10.19)$$

Знаючи $\vec{j}$ у всіх точках довільної поверхні S, можна обчислити потік енергії через цю поверхню:

$$\Phi = \int_S \vec{j} d\vec{S}.$$

Скалярна величина *I,* яка дорівнює модулю середнього значення вектора Умова називається ***інтенсивністю хвилі*:**

$$I = |\langle \vec{j} \rangle|.$$

Інтенсивність хвилі чисельно дорівнює енергії, яка переноситься хвилею за одиницю часу через одиницю площі поверхні, яка перпендикулярна до напрямку поширення хвилі

$$I = \langle w \rangle v = (1/2)\rho\omega^2 A^2 v. \quad (10.20)$$

Інтенсивність синусоїдальної хвилі пропорційна квадрату її амплітуди.

§10.5.Групова швидкість поширення хвилі

*Швидкість поширення гармонійної хвилі***,** що описується рівняннями (10.5) і (10.9), *дорівнює фазовій швидкості* v (10.7).

Постає запитання, чи для всіх видів хвиль справедлива така залежність? Усі реальні хвилі не є гармонійними. Проте можна довести, що будь-яку несинусоїдальну хвилю можна розкласти на еквівалентну їй систему синусоїдальних хвиль. Таке розкладання називають *гармонійним,* або *спектральним аналізом* несинусоїдальної хвилі. Сукупність значень амплітуд, початкових фаз і частот еквівалентної системи синусоїдальних хвиль називають *спектром* відповідно *амплітуди, початкових фаз* і *частот* несинусоїдальної хвилі, що розглядається. Спектр може бути як *дискретним,* або *лінійчастим,* тобто множиною окремих різних значень, так і *неперервним (суцільним).*

Встановлено, що всі синусоїдальні (гармонічні) хвилі поширюються в середовищі незалежно одна від одної, а результуюче зміщення будь-якої частинки середовища дорівнює векторній сумі її зміщень, зумовлених кожною з хвиль. Цей результат справедливий для хвиль будь-якої природи, його називають принципом суперпозиції хвиль (див.**§**10.7**.).**

Суперпозицію гармонійних хвиль, які мало відрізняються одна від одної за частотою (чи довжиною хвиль), називають групою хвиль, або *хвильовим пакетом.*

Якщо швидкість поширення синусоїдальних хвиль у середовищі не залежить від їх частоти (довжини хвилі), то швидкість поширення пакета дорівнюватиме фазовій швидкості. Проте існують і такі гармонічні хвилі, фазова швидкість яких залежить від їх частоти (довжини хвилі). Залежність фазової швидкості від частоти (довжини хвилі) називають *дисперсією хвиль,* а середовище, де спостерігається це явище, – *дисперсійним.*

Якщо хвильовий пакет поширюється в дисперсійному середовищі, то швидкість окремих його синусоїдальних хвиль різна. Внаслідок цього безперервно змінюватимуться різниці фаз між хвилями, імпульс складного коливання дещо змінюватиме свою форму, а амплітудне значення хвилі зміщуватиметься відносно складових хвиль. При цьому швидкість поширення хвильового пакета не збігається зі швидкістю жодної з них.

Для прикладу розглянемо простий хвильовий пакет, утворений двома плоскими синусоїдальними хвилями, що поширюються

вздовж осі *X*. Не порушуючи загальності результату, вважатимемо, що амплітуди цих хвиль однакові, початкові фази $\varphi_{01} = \varphi_{02} = 0$, а циклічні частоти і хвильові числа їх різні, але мало відмінні між собою:

$$\begin{cases} S_1(x,t) = A_0 \cos(\omega_1 t - k_1 x) \\ S_2(x,t) = A_0 \cos(\omega_2 t - k_2 x). \end{cases} \quad (10.21)$$

Згідно з принципом суперпозиції хвиль рівняння результуючої хвилі буде із урахуванням, що

$$\cos\alpha + \cos\beta = 2\cos\frac{1}{2}(\alpha+\beta)\cdot\cos\frac{1}{2}(\alpha-\beta).$$

$$\begin{aligned} S(x,t) &= S_1(x,t) + S_2(x,t) = \\ &= 2A_0 \cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right)\cdot \\ &\cdot \cos(\omega_{сер}\, t - k_{сер} x), \end{aligned} \quad (10.22)$$

де

$$\begin{aligned} \Delta\omega &= \omega_1 - \omega_2,\ \omega_{сер} = (\omega_1+\omega_2)/2, \\ \Delta k &= k_1 - k_2,\ \ k_{сер} = (k_1+k_2)/2. \end{aligned} \quad (10.23)$$

З виразу (10.22) випливає, що результуюча хвиля є плоскою, її циклічна частота і хвильове число дорівнюють півсумі циклічних частот і хвильових чисел синусоїдальних хвиль, що утворюють хвильовий пакет. Однак амплітуда цієї хвилі не стала, а залежить від координати *x* і часу *t*:

$$A(x,t) = 2A_0 \cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right). \quad (10.24)$$

Цей вираз для амплітуди хвильового пакета сам по собі є виразом рівняння плоскої синусоїдальної хвилі – хвилі амплітуди коливань, фаза якої

$$\varphi_A = \frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x. \quad (10.25)$$

Таке коливання, як зазначалось у §9.5, називають *биттям*. Графік його для коливань, що поширюються в не дисперсійному середовищі, наведено на рис.9.7.

Знайдемо, чому дорівнює швидкість результуючої хвилі у дисперсійному середовищі. Хвиля, поширюючись у середовищі, переносить енергію.

Швидкість v поширення енергії хвилі дорівнює швидкості переміщення в просторі поверхні, через яку переноситься максимальна кількість енергії за одиницю часу. Оскільки енергія в разі коливального процесу пропорційна квадрату амплітуди, то швидкість поширення енергії хвильового пакета збігається з фазовою швидкістю хвилі амплітуди. Вочевидь, що для того щоб амплітуда *A(x,t)* (див.10.24) мала стале значення, наприклад максимальне, її фаза φ_A мусить залишатися сталою, тобто

$$\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x = \text{const}. \quad (10.26)$$

Взявши повний диференціал від цього виразу, одержимо

$$\frac{\Delta\omega}{2}dt - \frac{\Delta k}{2}dx = 0,$$

звідки

$$v_{гр} = \frac{dx}{dt} = \frac{\Delta\omega}{\Delta k} = \frac{\omega_1-\omega_2}{k_1-k_2}\cdot \quad (10.27)$$

За умови, що $\Delta\omega$, а отже, і Δk, прямують до нуля,

$$v_{гр} = \frac{d\omega}{dk}\cdot \quad (10.28)$$

Цю швидкість називають *груповою швидкістю пакета хвиль.*

Швидкість гармонійних хвиль (див.(10.7)) дорівнює $v = \omega/k$. Отже, згідно зі співвідношеннями (10.7) і (10.28), групова швидкість пакета хвиль не дорівнює жодній із швидкостей синусоїдальних складових цього пакета.

Вираз для групової швидкості (10.28) можна подати ще й у такій формі. Згідно з (10.7) $\omega = vk$.

Підставивши це значення у (10.28), отримаємо

$$v_{гр} = \frac{d(vk)}{dk} = v + k\frac{dv}{dk}\cdot \quad (10.29)$$

Якщо врахувати, що $k = 2\pi/\lambda$ і $dk = -\frac{2\pi}{\lambda^2}d\lambda$, то отримаємо такий вираз для групової швидкості:

$$v_{гр} = v - \lambda \frac{dv}{d\lambda}. \qquad (10.30)$$

За відсутності дисперсії хвиль у середовищі ($\frac{dv}{d\lambda}=0$) фазові швидкості хвиль однакові і не залежать від довжини хвилі. Тому в таких середовищах групова швидкість хвиль збігається з фазовою швидкістю.

§10.6.Інтерференція хвиль

Інтерференція хвиль (від латинського inter – взаємно, між собою і ferio – вдаряю, уражаю), *складання в просторі двох (або декількох) хвиль, при якому в різних його точках виникає підсилення або послаблення амплітуди результуючої хвилі*. Інтерференція характерна для хвиль любої природи; хвиль на поверхні рідини, пружних (наприклад, звукових), електромагнітних (наприклад, радіохвиль або світла).

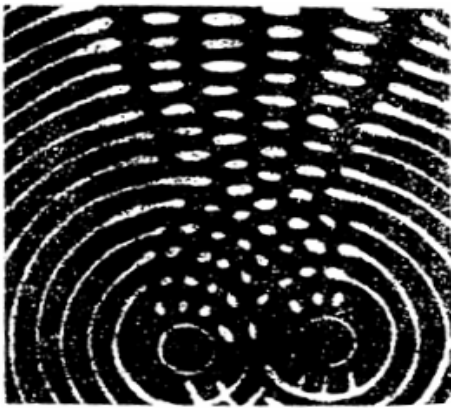


Рис.10.7.

Інтерференційна картина– регулярне чергування областей підвищеної і пониженої інтенсивності хвиль. Як приклад наведена інтерференційна картина хвиль на поверхні води, які збуджуються у двох різних точках (рис.10.7). В основі явища інтерференції лежить *принцип суперпозиції хвиль*, справедливість якого підтверджується на дослідах. Суть його в наступному.

Якщо в середовищі розповсюджується одночасно декілька хвиль, то коливання частинок середовища є геометрична сума коливань, які здійснювали б частинки при розповсюдженні кожної хвилі окремо. Значить, хвилі просто накладаються одна на другу, не спотворюючи одна одну.

В місцях зустрічі хвиль коливання середовища, які викликані кожною з хвиль, складаються одне з одним. Результат додавання (результуюча хвиля) залежить від співвідношення фаз, періодів, напрямів і амплітуд хвиль, що накладаються.

Розглянемо накладання двох косинусоїдальних сферичних хвиль, які збуджуються точковими джерелами S_1 і S_2 (рис.10.8).

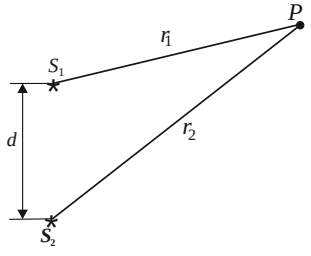


Рис.10.8.

Ці хвилі в точці накладання P викликають коливання частинок середовища. Ці коливання будуть описуватись згідно (10.9) рівняннями:

$$S_1 = \frac{A_1}{r_1}\cos(\omega_1 t - k_1 r_1 + \varphi_{01}) = \frac{A_1}{r_1}\cos\varphi_1,$$
$$S_2 = \frac{A_2}{r_2}\cos(\omega_2 t - k_2 r_2 + \varphi_{02}) = \frac{A_2}{r_2}\cos\varphi_2.$$

Результат додавання цих коливань залежить від спрямованості. Будемо вважати, що відстань між джерелами коливань S_1 і S_2, $d << r_1$ і r_2. Тоді можна вважати, що в точці P

додаються коливання однаково спрямовані і, згідно формули (9.20), амплітуда результуючого коливання (хвилі) в точці *P* дорівнює:

$$A_P^2=(\frac{A_1}{r_1})^2+(\frac{A_2}{r_2})^2+2\frac{A_1}{r_1}\frac{A_2}{r_2}\cos\varphi, \quad (10.31)$$

де

$$\varphi=\varphi_1-\varphi_2=(\omega_1-\omega_2)t+ \\ +(k_2r_2-k_1r_1)+\varphi_{01}-\varphi_{02} \quad (10.32)$$

різниця фаз хвиль в точці *P*.

Якщо частоти хвиль різні, тобто $\omega_1 \neq \omega_2$, $k_1 \neq k_2$ та ще й ω_1, ω_2, k_1, k_2 φ_{01}, φ_{02} хаотично залежать від часу, то φ є функція часу і $\cos\varphi$ може змінюватись хаотично від -1 до $+1$. Тоді середнє значення $\langle\cos\varphi\rangle = 0$ і середнє значення результуючої амплітуди хвилі в точці *P* буде

$$\langle A_p^2\rangle = \langle\left(\frac{A_1}{r_1}\right)^2\rangle + \langle\left(\frac{A_2}{r_2}\right)^2\rangle. \quad (10.33)$$

Оскільки, згідно (10.20) квадрат амплітуди хвилі пропорційний її інтенсивності, то в цьому випадку в точці *P* будемо мати додавання інтенсивності двох хвиль. Такі хвилі називаються *некогерентними.*

Дамо визначення, що ж таке когерентність.

Когерентність (від латинського cohaerens *– що знаходиться в зв'язку), узгоджене протікання з часом і в просторі декількох коливальних або хвильових процесів, яке проявляється при їхньому додаванні.* Коливання (хвилі) називаються когерентними, якщо різниця фаз залишається сталою або закономірно змінюється з часом і при додаванні коливань (хвиль) визначає амплітуду сумарного коливання.

Таким чином, хвилі, а значить і коливання в точці *P* будуть когерентними, якщо різниця фаз φ (формула 10.32) не буде залежати від часу. Це буде у випадку, якщо $\omega_1=\omega_2$ (монохроматичні хвилі) і $\varphi_{01}-\varphi_{02}=\text{const}$. За умови, що швидкості розповсюдження хвиль однакові $(v_1=v_2)$ $k_1=k_2=\frac{\omega_1}{v_1}=\frac{\omega_2}{v_2}=k$. Хвильове число не залежатиме від часу, якщо частота ω і швидкість v не залежатимуть від часу. При цих умовах

$$\varphi = k(r_2-r_1)+\varphi_{01}-\varphi_{02} \quad (10.34)$$

не буде залежати від часу і коливання в точці *P*, а значить і хвилі, які збудили ці коливання, будуть когерентними.

Тоді результат інтерференції двох хвиль, що визначається формулою (10.31), в різних точках залежить від величини $\Delta= r_2-r_1$, яка називається *геометричною різницею ходу хвиль.*

В точках, для яких $\cos\varphi=+1$, буде максимум інтерференції. Амплітуда результуючого коливання в таких точках буде згідно (10.31)

$$A_P=\frac{A_1}{r_1}+\frac{A_2}{r_2}\cdot \quad (10.35)$$

Умова максимуму згідно (10.34) може бути записана так:

$$k(r_2-r_1)+\varphi_{01}-\varphi_{02}=\pm 2m\pi, \quad (10.36)$$

де m=0,1,2,3,... .

В точках, для яких $\cos\varphi=-1$, буде мінімум інтерференції: амплітуда результуючого коливання в таких точках згідно (10.31) буде:

$$A_P=\left|\frac{A_1}{r_1}-\frac{A_2}{r_2}\right|\cdot \quad (10.37)$$

Умову мінімуму отримаємо із (10.34):

$$k(r_2-r_1)+\varphi_{01}-\varphi_{02}=\pm(2m+1)\,\pi, \quad (10.38)$$

де m=0,1,2,3,... .

Дуже часто реалізується випадок коли $\varphi_{01}-\varphi_{02}=0$. Тоді умова (10.36) набирає

такого вигляду із врахування, що $k = 2\pi/\lambda$, де λ – довжина хвилі в даному середовищі:

$$\Delta = r_2 - r_1 = \pm 2m\frac{\lambda}{2}. \qquad (10.39)$$

Умови мінімуму (10.38) тоді набуде вигляду:

$$\Delta = (r_2 - r_1) = \pm(2m+1)\frac{\lambda}{2}. \qquad (10.40)$$

При інтерференції хвиль їхня інтенсивність (енергія) механічно не підсумовується. Інтерференція хвиль призводить до перерозподілу інтенсивності (енергії) коливань між сусідніми областями середовища.

На рис.10.7 показані дві хвилі, які інтерферують; гребені хвиль – світлі лінії, западини – темні місця.

У місцях перетину двох гребенів або двох западин розміщені максимуми коливань (світлі плями), в місцях перетину гребенів і западин розміщені мінімуми (темні місця).

Особливим випадком інтерференції є стоячі хвилі.

Стоячі хвилі *– це хвилі, які утворюються при накладанні двох біжучих хвиль, що поширюються назустріч одна одній з однаковими частотами і амплітудами.*

Нехай дві плоскі хвилі поширюються назустріч одна одній вздовж осі *OX* в середовищі без згасання. Рівняння цих хвиль

$$S_1 = A_0\cos(\omega t - kx),$$
$$S_2 = A_0\cos(\omega t + kx + \varphi),$$

де φ – різниця фаз хвиль у точці $x = 0$, $k = 2\pi/\lambda$ (рис.10.9).

Додавши ці рівняння і перетворивши результат по формулі для суми косинусів з урахуванням, що $k = 2\pi/\lambda$, отримаємо:

$$S = S_1 + S_2 = 2A_0\cos\left(\frac{2\pi x}{\lambda} + \frac{\varphi}{2}\right)\cos\left(\omega t + \frac{\varphi}{2}\right). \qquad (10.41)$$

Множник $\cos(\omega t + \frac{\varphi}{2})$ показує, що в точках середовища виникає коливання з тією ж самою частотою ω, що і коливання зустрічних хвиль.

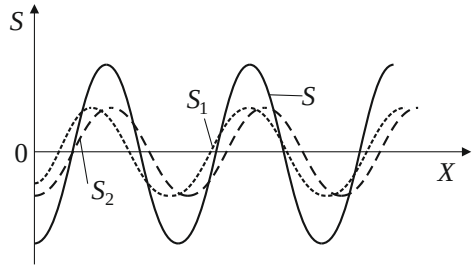


Рис.10.9.

Множник $2A_0\cos\left(\frac{2\pi x}{\lambda} + \frac{\varphi}{2}\right)$, який не залежить від часу, виражає амплітуду $A_{ст}$ результуючих хвиль, точніше – амплітуда, як величина позитивна, дорівнює абсолютному значенню цього множника:

$$A_{ст} = \left|2A_0\cos\left(\frac{2\pi x}{\lambda} + \frac{\varphi}{2}\right)\right|.$$

Амплітуда результуючого коливання залежить від координати x, що визначає положення точок середовища.

У точках середовища, де

$$\frac{2\pi x}{\lambda} + \frac{\varphi}{2} = \pm m\pi \quad (m=0,1,2,3,...),$$

амплітуда $A_{ст}$ досягає максимального значення $2A_0$. Точки, в яких $A_{ст}$ максимальна, називаються пучностями стоячої хвилі.

У точках середовища, де

$$\frac{2\pi x}{\lambda} + \frac{\varphi}{2} = \pm(2m+1)\frac{\pi}{2} \quad (m=0,1,2,3,...)$$

$A_{ст} = 0$. Ці точки називаються вузлами стоячої хвилі. Точки середовища, що знаходяться у вузлах, не коливаються. Виберемо початок відліку x так, щоб φ дорівнювало нулю. Тоді координати пучностей

$$x_n = \pm m\frac{\lambda}{2}, \qquad (10.42)$$

а вузлів

$$x_в = \pm(2m+1)\frac{\lambda}{4}\cdot \qquad (10.43)$$

Відстань між двома сусідніми пучностями отримаємо, якщо знайдемо різницю двох значень x_n для двох послідовних значень *m:*

$$x_n(m{+}1) - x_n(m) = (m{+}1)\frac{\lambda}{2} - m\frac{\lambda}{2} = \frac{\lambda}{2},$$

тобто відстань між сусідніми пучностями дорівнює половині довжини хвилі тих хвиль, в результаті інтерференції яких утворюється дана стояча хвиля.

Відстань від вузла до найближчої пучності дорівнює:

$$x_в - x_n = (2m{+}1)\frac{\lambda}{4} - m\frac{\lambda}{2} = \frac{\lambda}{4}\cdot$$

Множник $(2A_0 \cos 2\pi\frac{x}{\lambda})$ у виразі для амплітуди при переході через нульове значення змінює знак. У відповідності з цим фаза коливань по різні сторони від вузла відрізняється на π. Це означає, що точки, які лежать по різні сторони від вузла коливаються

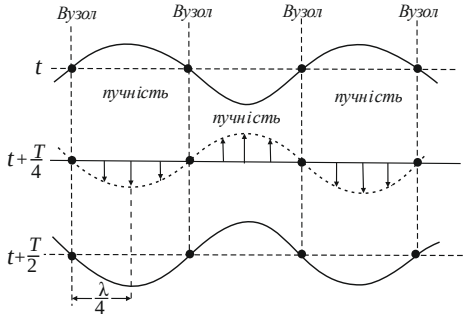


Рис.10.10.

в протифазі. Всі точки, які знаходяться між сусідніми вузлами, коливаються синфазно (тобто в однаковій фазі). На рис.10.10 приведений ряд «моментальних фотографій» відхилення точок від положення рівноваги. Перша «фотографія» відповідає моменту, коли відхилення досягає найбільшого абсолютного значення. Наступні «фотографії» зроблені з інтервалом в четверть періоду. Стрілками показані напрямки швидкостей частинок.

Отже, в стоячій хвилі є ряд нерухомих вузлових точок, які розміщені на відстані півхвилі одна від одної. Частинки між вузлами коливаються з різними амплітудами, від нуля у вузлі до подвійної амплітуди у пучності. Всі частинки одночасно проходять через положення рівноваги і одночасно досягають максимальних відхилень, отже, коливаються в однакових фазах. В суміжному інтервалі між вузлами характер коливань такий самий, але фази протилежні.

У стоячій хвилі, енергія не переноситься – повна енергія коливань кожного елемента об'єму середовища, обмеженого сусіднім вузлом і пучністю, не залежить від часу. Вона лише переходить з кінетичної енергії в потенціальну енергію пружно деформованого середовища і навпаки. Відсутність перенесення енергії стоячою хвилею є результатом того, що падаюча і відбита хвилі, які утворюють цю стоячу хвилю, переносять енергію в рівних кількостях в протилежних напрямках.

Стоячі хвилі утворюються, наприклад, на струні, повітряному стовпі в духових музичних інструментах, в камертоні. Вони є результатом інтерференції біжучої хвилі в одному напрямку і відбитої – в протилежному. Як засвідчують досвід і теорія, у разі відбивання хвиль від густішого середовища на межі поділу двох середовищ стояча хвиля матиме вузол, у разі її відбивання від середовища з меншою густиною – пучність.

Утворення вузла стоячої хвилі на межі поділу менш густого і більш густішого середовища можна пояснити так. На межі двох середовищ коливання, спричинені біжучою хвилею, додаються. Оскільки на межі таких

середовищ утворюється вузол (тобто сума зміщень точок дорівнює нулю), звідси можна дійти висновку, що в разі відбивання хвилі від густішого середовища її фаза змінюється на протилежну. Як відомо, фаза змінюється на протилежну на відстані, що дорівнює половині довжини хвилі. Тому можна стверджувати, що під час відбивання хвилі від густішого середовища «втрачається півхвилі».

Відбившись від поверхні менш густого середовища, хвиля не змінює фазу у місці відбивання. Тому фази падаючої і відбитої хвиль на межі поділу таких двох середовищ однакові. В цьому разі зміщення точок буде вдвічі більшим, ніж зміщення, спричинене падаючою хвилею, тобто виникне пучність.

У закріпленої з обох кінців натягнутої струни при збудженні поперечних коливань встановлюються стоячі хвилі, причому в місцях закріплення струни повинні утворюватись вузли. Тому в струні збуджуються із значною інтенсивністю тільки такі коливання, половина довжини хвилі яких вкладається на довжині струни ціле число раз (рис.10.11).

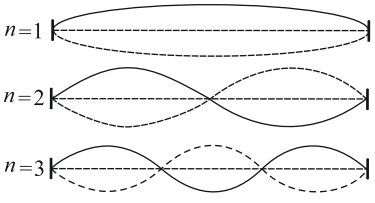


Рис.10.11.

Звідси витікає умова:

$$l = n\lambda/2 \text{ або } \lambda_n = 2l/n \quad (10.44)$$

(l – довжина струни). Довжинам хвиль (10.44) відповідають частоти

$$\nu_n = \frac{v}{\lambda_n} = \frac{v}{2l}n \quad (n = 1, 2, 3, \dots) \quad (10.45)$$

(v – фазова швидкість хвилі, яка визначається силою натягу струни і масою одиниці довжини, тобто лінійною густиною струни).

Частоти ν_n називаються власними частотами струни. Власні частоти є кратними частоті $\nu_1 = \frac{v}{2l}$, яка називається основною частотою.

Гармонічні коливання з частотами (10.45) називаються власними або нормальними коливаннями. Їх називають також гармоніками. В загальному випадку коливання струни є накладання різних гармонік.

Коливання струни дивні в тому сенсі, що для них, за класичними уявленнями, отримуються дискретні значення однієї із величин, яка характеризує коливання (частоти). Для класичної фізики така дискретність є винятком. Для квантових процесів дискретність є швидше правилом, ніж винятком.

§10.7.Звук

Звук в широкому сенсі – коливальний рух частинок пружного середовища, що розповсюджується у вигляді хвилі в газоподібному, рідкому чи твердому середовищі – це теж саме, що й пружні хвилі; у вузькому сенсі – явище, що суб'єктивно сприймається органом слуху людини і тварин.

Людина чує звук в діапазоні частот від 16 Гц до 20 кГц. Людина не чує звуки з частотою нижче 16 Гц (інфразвук) і більшою 20 кГц (ультразвук). Високочастотні пружні хвилі в діапазоні від 10^9 до 10^{12}–10^{13} Гц називаються гіперзвуком.

Важливою характеристикою звуку є його спектр, який отримується в результаті розкладання звуку на прості гармонійні

коливання (так званий частотний аналіз звуку). Основна частота, яка сприймається на слух, визначає при цьому висоту звуку, а набір гармонійних складових – тембр звуку. Набір частот коливань, які присутні в даному звуці, називається його акустичним спектром. Якщо у звуці є коливання всіх частот у деякому інтервалі від ν' до ν", то спектр називається *суцільним*. Якщо звук складається із коливань дискретних частот ν_1, ν_2, ν_3 і т.д., то спектр називається *лінійчастим*. Суцільний акустичний спектр мають шуми. Коливання з лінійчастим спектром викликають відчуття звуку з більш − менш визначеною частотою. Такий звук називається *тональним.*

Висота тонального звуку визначається основною (найменшою) частотою ν_1. Відносна інтенсивність *обертонів* (тобто коливань з частотами ν_2, ν_3 і т.д.) визначає забарвлення, або тембр звуку. Різний спектральний склад звуків, які збуджуються різними музикальними інструментами, дозволяє відрізнити на слух, наприклад, флейту від скрипки.

Енергетичною характеристикою звукових коливань є *інтенсивність звуку I*, тобто середня за часом енергія, яка переноситься хвилею за одиницю часу через одиничну площадку, що перпендикулярна до напрямку розповсюдження хвилі і виражається через амплітуду звукового тиску:

$$I = \frac{\Delta p_0^2}{2v_{\parallel}\rho}, \qquad (10.46)$$

де Δp_0 – амплітуда звукового тиску, $v_{\parallel}$ – швидкість звуку в середовищі, ρ – густина середовища.

Звуковий тиск – це змінна частина тиску, яка виникає при проходженні звукової хвилі в середовищі. Розповсюджуючись в середовищі, звукова хвиля утворює його згущення і розрідження, які створюють додаткові зміни тиску по відношенню до його середнього значення в середовищі. Звуковий тиск здійснюється з частотою, яка дорівнює частоті звукової хвилі. Звуковий тиск – основна кількісна характеристика звуку. Амплітуда звукового тиску Δp_0 і амплітуда швидкості коливань v_m частинок середовища у звуковій хвилі зв'язані співвідношенням

$$\Delta p_0 = v_{\parallel}\rho v_m. \qquad (10.47)$$

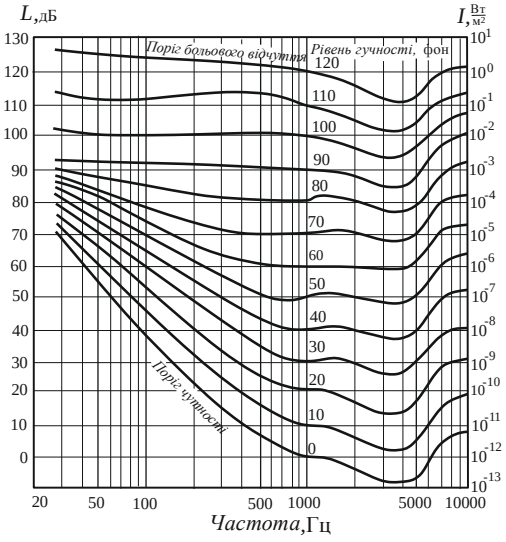


Рис.10.12.Криві однакової гучності, які виражають залежність рівня звукового тиску (в децибелах) від частоти при заданій гучності (у фонах). Нижня крива – поріг чутності, верхня крива – поріг больового відчуття.

Для того, щоб викликати звукове відчуття, хвиля повинна мати деяку мінімальну інтенсивність, яка називається *порогом чутності* (рис.10.12, нижня крива). Поріг чутності дещо різний для різних осіб і сильно залежить від частоти звуку. Найбільш чутливе людське вухо до частот 1000-5000 Гц (рис.10.12). В цій області частот поріг чутності дорівнює середній інтенсивності звуку біля $5\cdot10^{-13}$ Вт/м^2. При інших частотах поріг

чутності має більшу інтенсивність звуку (рис.10.12, нижня крива).

При інтенсивностях звуку порядку 0,5−5 Вт/м^2 хвиля перестає сприйматись як звук і викликає тільки відчуття болі і тиску. Значення інтенсивності, при якому це має місце, називається *порогом больового відчуття.* Поріг больового відчуття, так само як і поріг чутності, залежний від частоти (верхня крива на рис.10.12; дані, приведені на цьому рисунку, відносяться до середнього нормального слуху).

Гучність звуку, величина, яка характеризує слухове відчуття для даного звуку. Гучність звуку складним чином залежить від інтенсивності (тиску) звуку, частоти і форми коливань (рис.10.12). При незмінній частоті і формі коливань гучність звуку зростає із збільшенням інтенсивності (тиску) звуку (рис.10.12). При однаковій інтенсивності (тиску) звуку гучність звуку чистих тонів (гармонійних коливань) різних частот різна, тобто на різних частотах однакову гучність можуть мати звуки різної інтенсивності (тиску).

Гучність звуку, яка оцінюється суб'єктивно, зростає значно повільніше, ніж інтенсивність (тиск) звукових хвиль. При зростанні інтенсивності (тиску) в геометричній прогресії гучність зростає приблизно в арифметичній прогресії, тобто лінійно.

Виходячи з цього *рівень гучності L* визначають як логарифм відношення інтенсивності (тиску) даного звуку I (або Δp) до інтенсивності (тиску) I_0 (або Δp_0) прийнятих за вихідні:

$$L = \lg\frac{I}{I_0}\cdot \qquad (10.48)$$

Вихідна інтенсивність I_0 приймається за величину 10^{-12}Вт/м^2, що відповідає пороговому звуковому тиску $\Delta p_n = 2\cdot 10^{-5}$Па. Як видно із рис.10.12 поріг чутності при частоті чистого тону 1000 Гц лежить на нульовому рівні *($L = 0$).*

Одиниця рівня гучності L, яка визначається формулою (10.48) називається белом (Б). Як правило, користуються в 10 раз меншими одиницями – децибелами (дБ).

Значення L в децибелах визначається за формулою

$$L = 10\lg\frac{I}{I_0}\cdot \qquad (10.49)$$

Децибел (дБ) – рівень звукового тиску Δp, для якого виконується співвідношення $201\text{g}(\Delta p/\Delta p_{\text{п}}) = 1$, де $\Delta p_{\text{п}} = 2\cdot 10^{-5}$Па – пороговий звуковий тиск (враховано, що $I \sim \Delta p^2$). Одному децибелу відповідає рівень звукового тиску Δp=2,244·10^{-5}Па. За формулою (10.49) один децибел відповідає зміні інтенсивності звуку в $10^{0,1}$ = 1,26 рази. Різні рівні гучності, які зустрічаються в повсякденному житті, приведені в таблиці 10.1.

Таблиця 10.1.

Звук	Рівень гучності L,дБ
Найбільш тихий звук, який здатне вловити людське вухо	10-15
Тікання годинника	20
Шепіт на відстані 1м	30
Негучна розмова	35-40
Крик	60
Шум працюючого поряд двигуна грузового автомобіля	80
Допустима гучність для людського вуха	Не більше 85
Шум в метро і шум від грузового залізнодорожного потяга	90
Навушники *mp*3 плеєра (*max*)	115-120
Шум літака, що злітає, на відстані 30м від нього	130
Можливість контузії	130-155
Шок	155-160
Розрив барабанних перетинок	160-175
Розрив легенів	175-200
Вибух атомної бомби (причина смерті)	>200

Рівень гучності звуку L_N (у фонах) залежить від інтенсивності звуку, його частоти і розраховується за формулою

$$L_N = 10 \lg\left(\frac{I_N}{I_0}\right), \qquad (10.50)$$

де I_N – інтенсивність звуку з частотою ν = 1000 Гц, який має однакову гучність із звуком, який досліджується. Залежність L_N від інтенсивності звуку і його частоти складна, і простим аналітичним співвідношенням не виражається. Визначається L_N за кривими рівня гучності (рис.10.12). Фон – одиниця рівня гучності даного звуку. Рівень гучності даного звуку у фонах дорівнює рівню звукового тиску в децибелах для чистого тону з частотою 1000 Гц, гучність якого при порівняні на слух дорівнює гучності даного звуку. Для звуку еталонної частоти (ν = 1000 Гц) рівень гучності звуку, який виражений у фонах, чисельно дорівнює рівню інтенсивності, що виражається у децибелах, $L = L_N$ (числова рівність справедлива тільки для частоти ν = 1000 Гц).

Весь діапазон інтенсивностей, при яких хвиля викликає в людському вусі звукове відчуття від $2 \cdot 10^{-13}$ Вт/м^2 при 3000−4000 Гц до майже 10 Вт/м^2 біля частоти 30 Гц. Відповідні значення рівня гучності від −8 до +130 дБ приведені на рис.10.12.

Джерелами звуку можуть бути будь-які явища, які викликають місцеву зміну тиску або механічних напружень. Широко розповсюджені джерела звуку у вигляді твердих тіл, які коливаються (наприклад, дифузори гучномовців і мембрани телефонів, струни і деки музичних інструментів); в ультразвуковому діапазоні частот це пластинки і стрижні із п'єзодіелектричних матеріалів або магнітострикційних матеріалів. Широкий клас джерел звуку – електроакустичні перетворювачі.

До приймачів звуку відносяться, зокрема, слухові апарати людини і тварин. В техніці для прийому звуку застосовують, головним чином, електроакустичні перетворювачі: у повітрі – мікрофони, у воді – гідрофони, в земній корі – геофони.

Розповсюдження звукової хвилі характеризується в першу чергу *швидкістю звуку*. В ряді випадків спостерігається дисперсія швидкості звуку, тобто залежність швидкості його розповсюдження від частоти. При розповсюдженні звукової хвилі відбувається поступове згасання звуку, тобто зменшення його інтенсивності і амплітуди, яке обумовлюється в значній степені поглинанням звуку, що зв'язане з необоротним переходом звукової енергії в інші форми (головним чином, в теплоту).

Швидкість звуку:

– в газах

$$v = \sqrt{\gamma \frac{p}{\rho}}, \qquad (10.51)$$

де γ – відношення теплоємкості газу при постійному тиску до теплоємності при постійному об'ємі ($\gamma \approx 1{,}5$), p і ρ – тиск і густина незбуреного хвилею газу:

– в ізотропних твердих тілах для поздовжньої хвилі

$$v_{\parallel} = \sqrt{\frac{E}{\rho}}, \qquad (10.52)$$

а для поперечної

$$v_{\perp} = \sqrt{\frac{G}{\rho}}, \qquad (10.53)$$

де E – модуль Юнга, G – модуль зсуву, ρ – густина речовини твердого тіла.

§10.8.Ефект Доплера для звукової хвилі

В 1842 році в журналі Празького наукового товариства з'явилась стаття під назвою «Про кольорове світло подвійних зірок

і деяких інших сузір'їв неба». Автор статті, австрійський професор математики Християн Доплер, запропонував нове і неочікуване пояснення забарвлення небесних об'єктів. Свою теорію Доплер будував на *твердженні про те, що частота хвильових коливань повинна залежати від швидкості руху джерела відносно спостерігача* (тобто приладу, який сприймає ці хвилі). Іншими словами, один і той самий звук буде мати різну висоту, а світло – різне забарвлення (колір) для джерела, яке рухається, або знаходиться в стані спокою.

Для того, щоб краще зрозуміти ефект Доплера для звуку, розглянемо такий наочний випадок. В озері плавають два човни. На одному з них у воду занурений прилад, який може випромінювати звукові коливання з частотою ν_0. З другого човна у воду занурений приймач, який може реєструвати частоту сигналу, який доходить.

Яку частоту звуку буде реєструвати приймач в наступних випадках:

1.обидва човни знаходяться в стані спокою на воді;

2.човен з джерелом знаходиться в стані спокою, човен із приймачем рухається;

3.човен із джерелом рухається, човен із приймачем знаходиться в стані спокою.

Розглянемо кожен випадок окремо.

1.Відповідь на це питання очевидне: приймач реєструє звук, частота якого ν_0. Дійсно, якщо швидкість звуку у воді v_o, то довжина звукової хвилі, яку випромінює джерело, дорівнює $\lambda_0 = v_o/\nu_0$. Спостерігач у другому човні, який вимірює швидкість звуку і довжину хвилі, записує саме ці величини – $v_1 = v_o$ і $\lambda_1 = \lambda_0$. Так що частота звуку, яку зареєструє спостерігач, буде дорівнювати

$$\nu_1 = \frac{v_1}{\lambda_1} = \frac{v_o}{\lambda_0} = \nu_0. \qquad (10.54)$$

2.Нехай тепер човен із спостерігачем почав рухатись вздовж прямої, яка з'єднує обидва човни. Позначимо швидкість човна із спостерігачем через $v_{сп}$ і будемо вважати, що вона додатня, якщо відстань між човнами скорочується. Швидкість звуку, яку тепер буде вимірювати спостерігач, буде дорівнювати не v_o, а

$$v_2 = v_o + v_{сп} = v_o(1 + \frac{v_{сп}}{v_0}).$$

Довжина хвилі звуку у воді не залежить від того, рухається спостерігач чи ні. Тому $\lambda_2 = \lambda_0 = v_o/\nu_0$. З другої сторони, $\lambda_2 = v_2/\nu_2$; значить, $v_o/\nu_0 = v_2/\nu_2$, тобто

$$\nu_2 = \frac{v_2}{v_o}\nu_0 = \frac{v_o + v_{сп}}{v_o}\nu_0 = \\ = \left(1 + \frac{v_{сп}}{v_o}\right)\nu_0. \qquad (10.55)$$

Це значить, що якщо відстань між човнами скорочується, то приймач зареєструє більш високу частоту, ніж частота джерела звуку. Навпаки, якщо човен із спостерігачем віддаляється від човна з джерелом, то частота звуку, який приймається приймачем, зменшується:

$$\nu_2' = \left(1 - \frac{v_{сп}}{v_o}\right)\nu_0. \qquad (10.55')$$

3.Розглянемо випадок, коли човен із джерелом рухається по напрямку до спостерігача із швидкістю $v_{дж}$. Швидкість звуку у воді як і раніше дорівнює v_o. Довжину хвилі, яку буде вимірювати спостерігач, тепер буде менша, ніж в тому випадку, коли човни знаходились в стані спокою. Дійсно, довжина хвилі – це відстань між двома сусідніми точками, які коливаються у фазі, наприклад, між сусідніми гребнями хвилі. За час $T = \lambda_0/ v_o$, поки один гребінь хвилі віддалиться від точки, в якій він був «висланий», на відстань λ_0, джерело зміститься в тому ж напрямку на

відстань $v_{дж} \cdot T = v_{дж}\lambda_0 / v_o$. Тому відстань між «першим» гребнем і наступним буде дорівнювати не λ_0, а

$$\lambda_3 = \lambda_0 - \frac{v_{дж}\lambda_0}{v_o} = \lambda_0(1 - v_{дж}/v_o).$$

Значить, частота звуку, яку зареєструє спостерігач, буде дорівнювати

$$\nu_3 = \frac{v_o}{\lambda_3} = \frac{v_o}{\lambda_0(1 - v_{дж}/v_o)},$$

або

$$\nu_3 = \frac{\nu_0}{(1 - v_{дж}/v_o)} \cdot \qquad (10.56)$$

В тому випадку, коли човен із джерелом віддаляється від човна з приймачем, частота звуку, яку буде реєструвати приймач дорівнює

$$\nu_3' = \frac{\nu_0}{1 + v_{дж}/v_o} \cdot \qquad (10.56')$$

Ми бачимо, що формули (10.56) відрізняються від формул (10.55), хоч основні закономірності зберігаються: частота звуку підвищується, коли човни зближаються, і частота знижується, коли човни розходяться.

Однак, якщо швидкість човна мала порівняно із швидкістю звуку у воді, то ми можемо скористатись наближеними рівностями:

$$\frac{1}{1-\alpha} \approx 1 + \alpha, \quad \frac{1}{1+\alpha} \approx 1 - \alpha \text{, де } \alpha \ll 1.$$

В цьому випадку формули (10.56) приблизно співпадають з формулами (10.55):

$$\nu_3 \approx \nu_2 = \left(1 \pm \frac{v_{дж}}{v_o}\right) \nu_0.$$

Можна написати і загальну формулу для випадку, коли рухаються обидва човни – і джерело, і приймач. Така формула отримається послідовним застосуванням двох формул (10.55) і (10.56), одержаних вище:

$$\nu = \frac{1 \pm v_{сп}/v_o}{1 \mp v_{дж}/v_o} \nu_0, \qquad (10.57)$$

або наближено, враховуючи що $(1 + \alpha)(1 + \beta) \approx 1 + \alpha + \beta$ при $\alpha, \beta \ll 1$, отримаємо, що

$$\nu \approx \left(1 \pm \frac{v_{сп} \pm v_{дж}}{v_o}\right) \nu_0. \qquad (10.58)$$

Неважко бачити, що $v_{сп} \pm v_{дж}$ є відносна швидкість джерела і приймача. Знаки «+» відповідають їх руху в протилежних напрямках, знаки «−» – в одному напрямку. Тому формулу для ефекту Доплера в наближенні малих швидкостей можна записати так:

$$\nu \approx \left(1 \pm \frac{v}{v_o}\right) \nu_0 , \qquad (10.59)$$

де v – відносна швидкість джерела і приймача звуку.

Отриману формулу (10.59) можна ще раз узагальнити – на той випадок, коли і середовище рухається. (Це узагальнення означає, що в розглянутому нами прикладі човни знаходяться не в озері, а в річці). В цьому випадку загальний вид формули (10.59) не змінюється, але під величинами $v_{дж}$ і $v_{сп}$ тепер слід розуміти швидкості джерела і спостерігача відносно річки; у відповідності з цим необхідно підставляти і значення v у формулу (10.59).

Ми розглядали випадок, коли джерело і приймач рухаються вздовж однієї прямої і вздовж цієї ж прямої розповсюджується фронт звукової хвилі. Якщо ж напрямок розповсюдження звуку складає кут θ з напрямком відносної швидкості v, то частота, яку реєструє приймач, дорівнює

$$\nu \approx \left(1 + \frac{v}{v_0}\cos\theta\right) \nu_0. \qquad (10.60)$$

Зміна частоти ν_0 визначається проекцією $v\cos\theta$ відносної швидкості на напрямок розповсюдження звуку. Якщо $\theta = 90^\circ$, то $\nu = \nu_0$ – так званий поперечний ефект

Доплера у звичайній акустиці відсутній.

Ефект Допплера для звукових хвиль вперше експериментально перевірили у 1845 р.

Музиканти на слух оцінювали звучання музичного інструменту, встановленого на платформі потяга, що проходив повз них.

Контрольні питання

1.Дати визначення рівняння хвилі.

2.Записати рівняння біжучої плоскої та сферичної хвиль. В чому різниця між цими рівняннями?

3.Хвильове рівняння для одномірного випадку.

4.Характеристики хвилі: амплітуда, фаза, частота, довжина хвилі (хвильове число), фазова швидкість хвиль у пружному середовищі.

5.Густина потоку енергії хвиль – вектор Умова.

6.Групова швидкість хвиль. Чим вона відрізняється від фазової швидкості?

7.Звукові хвилі. Характеристики звукових хвиль: амплітуда звукового тиску, інтенсивність звуку, гучність звуку.

8.Когерентні хвилі, інтерференція хвиль. Умови максимуму та мінімуму інтерференції.

9.Стоячі хвилі. Рівняння стоячої хвилі. Вузли та пучності стоячої хвилі та відстань між ними.

10.Основні характеристики звуку.

11.Як проявляється ефект Доплера для звукових хвиль?

Розділ 11.Електромагнітні хвилі

§11.1.Загальні відомості про електромагнітні хвилі

Електромагнітні хвилі – це електромагнітні коливання, які розповсюджуються в просторі з кінцевою швидкістю. Існування електромагнітних хвиль було передбачено англійським фізиком М.Фарадеєм в 1832 році. Англійський фізик Дж.Максвелл в 1865 році теоретично показав, що електромагнітні коливання розповсюджуються в вакуумі із швидкістю світла. В 1888 році максвелловська теорія електромагнітних хвиль отримала підтвердження в дослідах німецького фізика Г.Герца, що відіграло вирішальну роль в її утвердженні.

Теорія Максвелла дозволила встановити, що радіохвилі, світло, рентгенівське і гамма-випромінювання є електромагнітні хвилі з різною довжиною хвилі (див.табл.11.1), причому між сусідніми

Спектр електромагнітних хвиль

Таблиця 11.1.

Частота ν,Гц	Довжина хвилі λ,м	Назва діапазону	Джерела. Основні методи збудження
$0 - 10^3$	$\infty - 3\cdot10^5$	Низькочастотні хвилі	Змінні струми. Генератори
$10^3 - 10^{12}$	$3\cdot10^5 - 3\cdot10^{-4}$	Радіохвилі	Генератори радіочастот. Генератори НВЧ
$8{,}6\cdot10^{11} - 4\cdot10^{14}$	$3{,}5\cdot10^{-4} - 7{,}5\cdot10^{-7}$	Інфрачервоні промені	Теплове випромінювання атомів і молекул
$4\cdot10^{14} - 7{,}5\cdot10^{14}$	$7{,}5\cdot10^{-7} - 4\cdot10^{-7}$	Видиме світло	Випромінювання атомів і молекул при нагріванні і при електричних розрядах
$7{,}5\cdot10^{14} - 6\cdot10^{16}$	$0{,}4\cdot10^{-6} - 5\cdot10^{-9}$	Ультрафіолетові промені	Випромінювання атомів під час дії на них швидких електронів
$3\cdot10^{16} - 3\cdot10^{20}$	$10^{-8} - 10^{-12}$	Рентгенівські промені	Випромінювання атомі при збудженні зовнішніх і внутрішніх електронів
$3\cdot10^{19}$ і більше	10^{-11} і менше	γ– промені	Ядерні процеси, радіоактивний розпад, космічні процеси

діапазонами шкали електромагнітних хвиль не має різких границь (рис.11.1). Назви різним діапазонам спектрів дані за способом їх збудження і за методами спостереження. Властивості різних діапазонів хвиль також відрізняються за їх дією, за проникною здатністю. Кожний з цих діапазонів має широке практичне застосування. В електродинаміці розглядають переважно

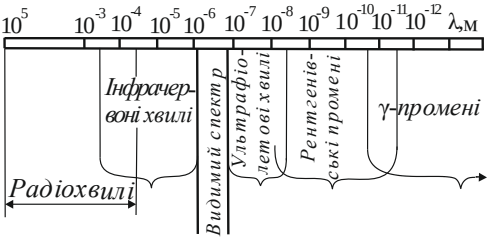


Рис.11.1. Шкала електромагнітних хвиль.

низькочастотні радіохвилі. Інші частотні діапазони вивчають в оптиці та в атомній і ядерній фізиці. Особливості електромагнітних хвиль, закони їх збудження і розповсюдження описуються рівняннями Максвелла.

§11.2.Доведення необхідності існування електромагнітних хвиль

В розділі 8 ми з'ясували, що змінне електричне поле породжує магнітне, яке, взагалі кажучи, теж є змінним. Це змінне магнітне поле породжує змінне електричне і т.д. Таким чином, якщо в якому-небудь місці простору виникає змінне електричне або змінне магнітне поле, то можливе збудження в послідовних точках простору змінного електричного поля змінним магнітним полем і навпаки, тобто, розповсюдження цих полів у вигляді єдиного електромагнітного поля в просторі, що і є утворенням електромагнітної хвилі. Покажемо, що існування електромагнітних хвиль витікає із рівнянь Максвелла. Розглянемо випадок однорідного, нейтрального ($\rho=0$), непровідного ($\vec{j}=0$) середовища з постійними проникностями ε і μ. Будемо розглядати одномірну задачу: вектори $\vec{E}, \vec{D}, \vec{B}$ і $\vec{H}$ залежать тільки від однієї координати (від x) і від часу. Тоді, оскільки всі похідні по y і z у формулах (8.29)−(8.27) дорівнюють нулю, то їх можна переписати так:

$$\frac{\partial B_x}{\partial t} = 0, \qquad (11.1)$$

$$\frac{\partial E_z}{\partial x} = \frac{\partial B_y}{\partial t}, \qquad (11.2)$$

$$\frac{\partial E_y}{\partial x} = -\frac{\partial B_z}{\partial t}, \qquad (11.3)$$

$$\frac{\partial B_x}{\partial x} = 0, \qquad (11.4)$$

$$\frac{\partial D_x}{\partial t} = 0, \qquad (11.5)$$

$$\frac{\partial H_z}{\partial x} = -\frac{\partial D_y}{\partial t}, \qquad (11.6)$$

$$\frac{\partial H_y}{\partial x} = \frac{\partial D_z}{\partial t}, \qquad (11.7)$$

$$\frac{\partial D_x}{\partial x} = 0. \qquad (11.8)$$

Із рівнянь (11.1) і (11.5) витікає, що складові векторів $\vec{B}$ і $\vec{D}$ по осі OX не залежить від часу, а із (11.4) і (11.8) витікає, що B_x і D_x не залежать також і від x. Однак, вище припускалось, що вектори $\vec{E}, \vec{D}, \vec{B}$ і $\vec{H}$ не залежать від y і z. Таким чином , B_x і D_x не залежать ні від часу, ні від однієї із координат, тобто постійні з часом і в просторі:

$$B_x = \text{const}, \quad D_x = \text{const}.$$

B_x і D_x – постійні однорідні статичні поля, які накладаються на змінне електромагнітне поле. Статичні поля не впливають на переміщення електромагнітного поля, тобто на утворення електромагнітної хвилі, і можуть не враховуватись. Будемо вважати, що $B_x = 0$ і $D_x = 0$.

Для випадку, який ми розглядаємо, диференціальні рівняння Максвелла можуть бути записані так :

$$\frac{\partial D_y}{\partial t} = -\frac{\partial H_z}{\partial x}, \qquad (11.9)$$

$$\frac{\partial B_z}{\partial t} = -\frac{\partial E_y}{\partial x}, \qquad (11.10)$$

$$\frac{\partial D_z}{\partial t} = \frac{\partial H_y}{\partial x}, \qquad (11.11)$$

$$\frac{\partial B_y}{\partial t} = \frac{\partial E_z}{\partial x}. \qquad (11.12)$$

Із (11.9) виходить, що змінне з часом електричне поле, що направлене вздовж осі *OY* викликає появу магнітного поля, яке направлене вздовж осі *OZ*. Зміна останнього з часом, як випливає із (11.10) призводить до появи електричного поля, яке направлене вздовж осі *OY*. Із (11.11) витікає, що змінне з часом електричне поле, яке направлене вздовж осі *OZ*, викликає появу магнітного поля, направленого вздовж осі *OY*, а змінне з часом магнітне поле, яке направлене вздовж осі *OY* (співвідношення (11.12)), призводить до появи направленого вздовж осі *OZ* електричного поля.

Із сказаного можна зробити висновок: змінні з часом електрична і магнітна складові електромагнітного поля взаємно перпендикулярні. Так як $B_x = 0$ і $D_x = 0$, то складові електромагнітного поля обидві перпендикулярні до осі *OX*. Поля E_y і H_z не викликають появу полів E_z і H_y. Тому для опису плоскої електромагнітної хвилі достатньо взяти одну пару рівнянь (11.9) і (11.10), а компоненти, що є у рівняннях (11.11) і (11.12), прирівняємо до нуля, тобто $D_z = 0$ і $H_y = 0$. Врахуємо співвідношення (8.28) і (8.29) і (11.9) і (11.10) перепишемо так:

$$\varepsilon\varepsilon_0 \frac{\partial E_y}{\partial t} = -\frac{\partial H_z}{\partial x}, \qquad (11.13)$$

$$\mu\mu_0 \frac{\partial H_z}{\partial t} = -\frac{\partial E_y}{\partial x}. \qquad (11.14)$$

Продиференціюємо перше рівняння по t і виконаємо перестановку : $\frac{\partial}{\partial t}\frac{\partial H_z}{\partial x} = \frac{\partial}{\partial x}\frac{\partial H_z}{\partial t}$. У це співвідношення замість $\frac{\partial H_z}{\partial t}$ із рівняння (11.14) підставимо $\frac{\partial H_z}{\partial t} = -\frac{1}{\mu\mu_0}\frac{\partial E_y}{\partial x}$. Тоді із (11.13) отримаємо рівняння:

$$\frac{\partial^2 E_y}{\partial x^2} = \varepsilon\mu\varepsilon_0\mu_0 \frac{\partial^2 E_y}{\partial t^2}. \qquad (11.15)$$

Якщо продиференціювати рівняння (11.14) по t і виконати аналогічні перетворення отримаємо таке рівняння для H_z:

$$\frac{\partial^2 H_z}{\partial x^2} = \varepsilon\mu\varepsilon_0\mu_0 \frac{\partial^2 H_z}{\partial t^2}. \qquad (11.16)$$

Порівняння (11.15), (11.16) з (10.13) дає підстави стверджувати, що (11.15) і (11.16) є хвильові рівняння, які представляють хвилі E_y і H_z, тобто процеси розповсюдження в просторі одночасно існуючих і взаємно обумовлених змінних з часом електричного і магнітних полів, тобто електромагнітного поля, що є електромагнітною хвилею. В даному випадку хвильові рівняння (11.15) і (11.16) описують розповсюдження електромагнітної хвилі вздовж осі *OX*. Нагадаємо, що $E_x = E_z = 0$ і $H_x = H_y = 0$, так що $E_y = E$ і $H_z = H$. В рівняннях (11.15) і (11.16) збережені індекси y і z при E і H , щоб підкреслити ту обставину, що вектори $\vec{E}$ і $\vec{H}$ направленні вздовж взаємно перпендикулярних осей *OY* і *OZ*. Таким чином, $\vec{E} \perp \vec{H} \perp OX$ і це значить, що *електромагнітна хвиля – поперечна хвиля*. В електродинаміці розглядається більш загальна задача, без викладених нами обмежень, і досягається аналогічний результат: із рівнянь Максвелла випливає необхідність існування електромагнітних хвиль.

Розглянемо якісну картину утворення електромагнітних хвиль. Уявімо собі, що в якійсь точці *O* (рис.11.2) всередині

нескінченного непровідного середовища створено яким-небудь способом електричне поле $\vec{E}$. Якщо електричних зарядів, які підтримували б це поле немає, то воно зникатиме. Але спадне поле $\vec{E}$ за Максвеллом спричинює магнітне поле $\vec{H}$.

Через те що поле $\vec{E}$ спадає, то густина струму зміщення $\vec{j}_{зм} = \varepsilon\varepsilon_0 \frac{\partial \vec{E}}{\partial t}$ напрямлена протилежно $\vec{E}$ і лінії сил магнітного поля напрямлені за рухом стрілки годинника (коли дивитись зверху по напрямку струму, рис.11.2). Через те, що в середовищі немає постійних струмів, які підтримували б поле $\vec{H}$, то це останнє в свою чергу зникатиме і спричинюватиме вихрове електричне поле $\vec{E}_1$. Силові лінії цього поля будуть напрямлені проти руху стрілки годинника, як показано на рис.11.2. Поле $\vec{E}_1$ знищить початкове поле $\vec{E}$ в точці 0, але зате виникне в сусідній точці 1.

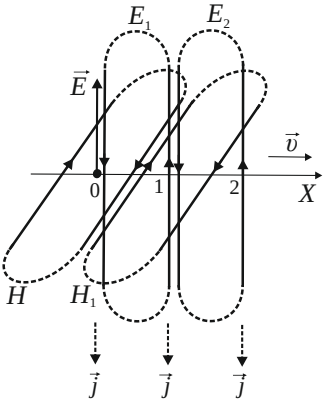


Рис.11.2.Вільні електромагнітні хвилі.

Зникаючи в точці 1, електричне поле $\vec{E}_1$ спричинить виникнення магнітного поля $\vec{H}_1$, яке буде напрямлене, як і поле $\vec{H}$, за рухом стрілки годинника. Поле $\vec{H}_1$ знищить поле $\vec{H}$ і виникне в більш віддаленій точці. Зникаючи, воно спричинить вихрове електричне поле $\vec{E}_2$, яке знищить поле $\vec{E}_1$ у точці 1, але виникне в точці 2, і т. д. Таким чином, замість початкового поля $\vec{E}$ ми дістанемо і електричне поле, і магнітне поле, які взаємно зв'язані одне з одним і які поширюються в просторі, тобто електромагнітну хвилю.

З рис.11.2 також видно, що $\vec{E}$ перпендикулярне до $\vec{H}$, причому обидва ці вектори перпендикулярні до швидкості поширення хвилі $\vec{v}$. Всі три вектори зв'язані між собою правилом правого свердлика : якщо обертати свердлик з правою нарізкою так, щоб його рукоятка переміщувалась від вектора $\vec{E}$ до вектора $\vec{H}$, то напрям поступального руху свердлика збігатиметься з напрямом $\vec{v}$.

Вище ми розглядали електромагнітні хвилі якісно. Проте теорія Максвелла не тільки завбачила існування електромагнітних хвиль, а й дала змогу точно, в кількісній формі встановити всі основні їх властивості.

§11.3. Властивості електромагнітних хвиль

1.Порівняння (11.15), (11.16) з (10.13) дає можливість стверджувати, що фазова швидкість електромагнітної хвилі

$$v = \frac{1}{\sqrt{\varepsilon\mu\varepsilon_0\mu_0}}. \qquad (11.17)$$

Для вакууму ($\varepsilon = 1, \mu = 1, \varepsilon_0 = 8{,}85 \cdot \cdot 10^{-12}$ Ф/м, $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м і тому фазова швидкість (позначимо її буквою c) дорівнює:

$$c = \frac{1}{\sqrt{\varepsilon_0\mu_0}} = 3 \cdot 10^8 \text{ м/с}, \qquad (11.18)$$

що дорівнює швидкості світла у вакуумі. Цей важливий результат був отриманий Максвеллом до експериментального відкриття електромагнітних хвиль. Це привело Максвелла до думки про електромагнітну природу світла. Для середовища із ε і μ

фазова швидкість електромагнітних хвиль з (11.17) буде дорівнювати:

$$v = \frac{c}{\sqrt{\varepsilon\mu}} = \frac{c}{n}, \qquad (11.19)$$

де $n = \sqrt{\varepsilon\mu}$ – показник заломлення середовища. Це співвідношення дає зв'язок між оптичними (n), електричними (ε) і магнітними (μ) властивостями середовища. Необхідно мати на увазі, що ε і μ залежать від частоти коливань $\vec{E}$ і $\vec{H}$.

2.Розв'язками рівнянь (11.15) і (11.16) є функції:

$$E = E_m \cos(\omega t - kx + \alpha_1) \qquad (11.20)$$

і

$$H = H_m \cos(\omega t - kx + \alpha_2). \qquad (11.21)$$

В цих формулах ω – циклічна частота хвилі, k – хвильове число, яке дорівнює ω/v, α_1 і α_2 – початкові фази коливань в точках з координатами $x = 0$, E_m і H_m – амплітуди хвилі. Підставимо (11.20) і (11.21) в рівняння (11.13) і (11.14):

$$\varepsilon\varepsilon_0\omega E_m \sin(\omega t - kx + \alpha_1) = kH_m \sin(\omega t - kx + \alpha_2),$$

$$\mu\mu_0\omega H_m \sin(\omega t - kx + \alpha_2) = kE_m \sin(\omega t - kx + \alpha_1).$$

Щоб рівняння (11.13) і (11.14) задовільнялись розв'язками (11.20) і (11.21) необхідна рівність початкових фаз α_1 і α_2. Крім того, повинні виконуватись співвідношення:

$$kE_m = \mu\mu_0\omega H_m,$$
$$\varepsilon\varepsilon_0\omega E_m = kH_m.$$

Перемноживши ці дві рівності, знайдемо що:

$$\varepsilon\varepsilon_0 E_m^2 = \mu\mu_0 H_m^2. \qquad (11.22)$$

Таким чином, коливання електричного і магнітного векторів в електромагнітній хвилі відбувається з однаковою фазою ($\alpha_1 = \alpha_2$), а амплітуди цих векторів зв'язані співвідношенням :

$$E_m\sqrt{\varepsilon\varepsilon_0} = H_m\sqrt{\mu\mu_0}. \qquad (11.23)$$

Оскільки вектори $\vec{E}$ і $\vec{H}$ коливаються у фазі, то співвідношення (11.23) справедливе не тільки для амплітудних значень E_m і H_m, а і для миттєвих значень E і H в якийсь момент часу, тобто :

$$E\sqrt{\varepsilon\varepsilon_0} = H\sqrt{\mu\mu_0}. \qquad (11.24)$$

Помноживши рівняння (11.20) на орт осі OY ($E\vec{j} = \vec{E}$), а рівняння (11.21) на орт осі OZ ($H\vec{k} = \vec{H}$), отримаємо рівняння плоскої електромагнітної хвилі у векторному вигляді:

$$\vec{E} = \vec{E}_m \cos(\omega t - kx),$$
$$\vec{H} = \vec{H}_m \cos(\omega t - kx). \qquad (11.25)$$

(ми прирівняли $\alpha_1 = \alpha_2 = 0$).

Косинусоїдальна або синусоїдальна електромагнітна хвиля називається монохроматичною хвилею.

В кожній точці електромагнітного поля монохроматичної хвилі проекції векторів $\vec{E}$ та $\vec{H}$ на осі координат інерціальної системи відліку здійснюють гармонійні коливання однакової лінійної частоти $\nu = \omega/(2\pi)$, яка називається частотою хвилі.

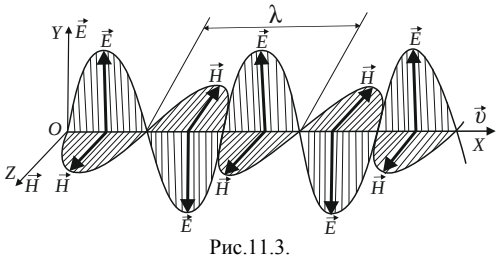


Рис.11.3.

На рис.11.3 наведені вектори $\vec{E}$ і $\vec{H}$ поля плоскої (вектор $\vec{E}$ коливається в площині YOX) монохроматичної хвилі в один і той самий момент часу. У фіксованій точці простору вектори $\vec{E}$ і $\vec{H}$ змінюються з часом за гармонічним законом. Вони одночасно збільшуються від нуля, потім через ¼T досягають найбільшого значення. Ще через ¼T обидва вектори одночасно дорівнюють нулю. Потім знову обидва вектори досягають

максимального значення, але протилежного напрямку, ніж півперіод тому. І через час, який дорівнює періоду коливання векторів, знову стають нульовими. Така зміна векторів $\vec{E}$ і $\vec{H}$ відбувається у всіх точках простору, але зі зсувом за фазою, що визначається відстанню між точками, яка відраховується вздовж осі *OX*.

Площина, яка проходить через вектор $\vec{E}$ і вектор швидкості $\vec{v}$, називається *площиною поляризації хвилі.* Довжина хвилі λ, період T, частота ν і швидкість поширення електромагнітної хвилі зв'язані між собою співвідношеннями

$$\lambda = vT = v/\nu \cdot \qquad (11.26)$$

Якщо плоска електромагнітна біжуча хвиля розповсюджується в додатному напрямі осі *OX*, то її можна, представити парою рівнянь (11.25). Тоді відбитій біжучій хвилі, яка розповсюджується в оберненому напрямку, від'ємному напрямку осі *OX*, відповідає одна пара із сукупності рівнянь:

$$\begin{aligned}\vec{E}&=\pm\vec{E}_m\cos(\omega t+kx),\\ \vec{H}&=\mp\vec{H}_m\cos(\omega t+kx).\end{aligned} \qquad (11.27)$$

В результаті попарного додавання (11.25) і (11.27) і відповідного математичного перетворення, отримаємо рівняння стоячої електромагнітної хвилі:

$$E = 2E_m \cos kx \cos \omega t, \qquad (11.28)$$

$$H = 2H_m \sin kx \sin \omega t. \qquad (11.29)$$

(11.28) – рівняння електричної, (11.29) – рівняння магнітної стоячої хвилі. Тут $|2E_m \cos kx|$ і $|2H_m \sin kx|$ – відповідно амплітуда електричних і амплітуда магнітних коливань. Оскільки максимальному значенню синуса відповідає нульове значення косинуса, то максимальному (нульовому) значенню амплітуди електричних коливань відповідає нульове (максимальне) значення амплітуди магнітних коливань, другими словами, пучності (вузлу) електричної стоячої хвилі відповідає вузол (пучність) магнітної стоячої хвилі. Із рівнянь (11.28) і (11.29) витікає, що коливання електричної компоненти хвилі зсунуті на $\pi/2$ по фазі відносно коливань магнітної компоненти хвилі ($\cos \omega t$ відносно $\sin \omega t$).

§11.4.Енергія електромагнітних хвиль. Потік енергії. Вектор Пойнтинга

Електромагнітне поле має енергію. Тому поширення електромагнітних хвиль пов'язане з перенесенням енергії в полі, подібно до того, як поширення пружних хвиль у речовині пов'язане з перенесенням механічної енергії.

Об'ємна густина енергії електромагнітної хвилі складається з об'ємних густин і електричного, і магнітного полів:

$$w = w_E + w_H = \frac{\varepsilon\varepsilon_0 E^2}{2} + \frac{\mu\mu_0 H^2}{2}.$$

Врахувавши (11.24), отримаємо, що густини енергії електричного і магнітного полів в кожен момент часу однакові, тобто $w_E = w_H$. Тому

$$\begin{aligned}w&=2w_E=\varepsilon\varepsilon_0 E^2=\\ &=\sqrt{\varepsilon\varepsilon_0\mu\mu_0}EH=\frac{1}{v}EH.\end{aligned} \qquad (11.30)$$

У випадку плоско поляризованої монохроматичної хвилі, що поширюється (рис.11.3) вздовж додатного напрямку осі *OX*, напруженість поля

$$E = E_m \cos(\omega t - kx).$$

Відповідно, об'ємна густина енергії цієї хвилі $w = \varepsilon\varepsilon_0 E_m^2 \cos^2(\omega t - kx)$. Значення w в кожній точці поля періодично змінюється з частотою 2ω в границях від 0 до $w_{max} =$

$\varepsilon\varepsilon_0 E_m^2$. Середнє значення w за період π/ω пропорційне квадрату амплітуди напруженості поля:

$$\langle w\rangle=\frac{\omega}{\pi}\int_0^{\pi/\omega} w\,dt= \\ =\varepsilon\varepsilon_0 E_m^2\langle\cos^2(\omega t-kx)\rangle=\frac{1}{2}\varepsilon\varepsilon_0 E_m^2. \quad (11.31)$$

Помноживши густину енергії w на швидкість v поширення хвилі в середовищі, отримуємо модуль густини потоку енергії.

Модуль густини потоку енергії чисельно дорівнює енергії, яку переносить хвиля за одиницю часу через одиницю площі поверхні, що розміщена перпендикулярно до напрямку поширення хвилі:

$$\Pi = wv = EH.$$

Оскільки вектори $\vec{E}$ і $\vec{H}$ взаємно перпендикулярні і утворюють з напрямком поширення хвилі правогвинтову систему, то напрямок вектора $\vec{E}\times\vec{H}$ збігаються з напрямком переносу енергії, а модуль цього вектора дорівнює EH. Отже, вектор густини потоку енергії електромагнітної хвилі, який називаються вектором **Пойнтинга** дорівнює:

$$\vec{\Pi} = \vec{E}\times\vec{H}\,. \quad (11.32)$$

Потік Ф електромагнітної енергії через деяку поверхню S можна знайти за допомогою інтегрування:

$$\Phi = \int_S \vec{\Pi}d\vec{S}. \quad (11.33)$$

Інтенсивність електромагнітної хвилі I дорівнює модулю середнього значення вектора Пойнтинга за проміжок часу, який дорівнює періоду T повного коливання:

$$I = \left|\langle\vec{\Pi}\rangle\right| = \left|\langle\vec{E}\times\vec{H}\rangle\right| = \frac{1}{T}\left|\int_0^T \vec{E}\times\vec{H}dt\right|,$$

або

$$I = \langle w\rangle v. \quad (11.34)$$

Таким чином інтенсивність плоско поляризованої монохроматичної біжучої хвилі прямо пропорційна квадрату амплітуди E_m коливань вектора $\vec{E}$ поля хвилі:

$$I = \frac{1}{2}\sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}}E_m^2. \quad (11.35)$$

§11.5. Генерування і експериментальне дослідження електромагнітних хвиль (радіохвиль)

Із попередніх параграфів відомо, що для отримання електромагнітної хвилі необхідно створити в просторі досить швидко змінне електричне поле. Очевидно, що для цієї мети електричні коливальні контури з зосередженими ємністю і індуктивністю (закриті контури), які розглянуті в розділі 9, не придатні. У таких випадках усе електричне поле зосереджене у вузькому зазорі конденсатора, а все магнітне поле – всередині індуктивності, а в навколишньому просторі електричне поле практично дорівнює нулю (рис.11.4,*а*).

Делокалізацію змінних електричного і магнітного полів, перенесення їх із частин простору, які обмежені обкладками конденсатора і котушкою, можна здійснити, розсуваючи обкладки конденсатора і зменшуючи їх площу, а також розсуваючи витки котушки і зменшуючи їх кількість (рис.11.4,*б,в*). При цьому одночасно зменшується ємність і індуктивність контуру, а значить, збільшується власна частота коливального контуру. Зрештою, приходимо до прямолінійного відрізка проводу (металічного стержня), як показано на рис.11.4,*г*. Це – відкритий коливальний контур (відкритий вібратор). Як видно із рисунка, змінним електричному і магнітному полям (електромагнітному полю) відкритого вібратора повністю забезпечений вихід в

оточуючий простір. Такий вібратор винайшов і вперше застосував для генерування електромагнітних хвиль (для перевірки теорії Максвелла) в 1887−1888 рр. німецький фізик Г.Герц.

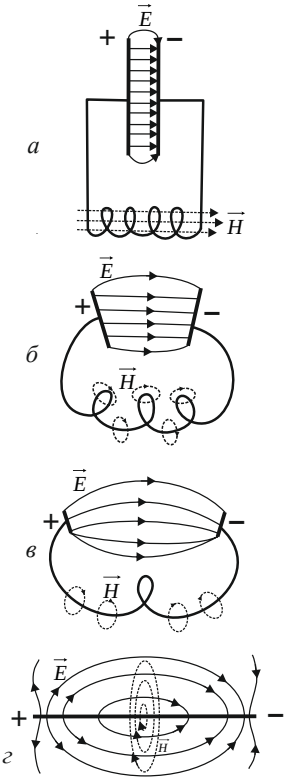


Рис.11.4.

Для вивчення електромагнітних хвиль Герц використав власні електричні коливання відкритого вібратора (диполя), який складався з двох однакових металевих стержнів *вв* (рис.11.5), розділених іскровим проміжком.

Обидві половини вібратора заряджались від джерела високої напруги індуктора. Коли різниця потенціалів досягала пробійного значення, в розряднику проскакувала іскра, яка замикала обидві половини вібратора, і у вібраторі виникали згасаючі електричні коливання високої частоти. Щоб швидкозмінні струми не переходили з вібратора в котушку, між вібратором і котушкою ввімкнено дроселі *Д*, реактивний опір яких дорівнює ωL. Так як зміна контура від виду, зображеного на рис.11.4,*а*, до виду, зображеного на рис.11.4,*г*, призводить до дуже сильного зменшення ємності відкритого контура, то з метою деякого збільшення ємності кінці половинок стержнів іскрового проміжку мали сферичні стовщення (рис.11.5).

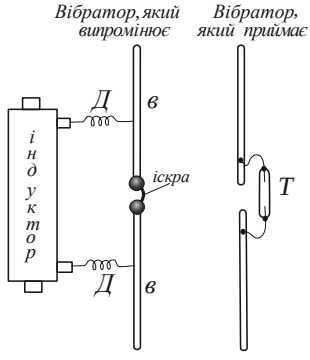


Рис.11.5.

Індуктор – котушка із залізним осердям і двома обмотками. За допомогою механічного переривача в первинній обмотці створювався пульсуючий струм з частотою його розриву $(10^{-3} \div 10^{-4})\text{с}^{-1}$. У вторинній обмотці індукувалась змінна е.р.с. (до 10^4В) тієї ж частоти, яка прикладалась до двох половинок вібратора. Коли напруга між половинками вібратора досягала пробивної величини, між ними проскакувала іскра, яка закорочувала обидві половинки вібратора і в ньому виникали високочастотні власні згасаючі електромагнітні коливання, частота яких визначалась величинами його ємності і індуктивності. Герц отримував з різними вібраторами коливання, частота яких була порядку $10^7 \div 10^8$Гц і електромагнітні хвилі з $\lambda = (0{,}6 \div 10)$м. Половинки вібратора

зазнавали багатократну перезарядку, після якої електромагнітні коливання повністю згасали.

Згасання коливань обумовлено тим, що енергія, яку отримав вібратор при його зарядці, витрачалась на виділення ленц-джоулевого тепла і на випромінювання хвиль. Внаслідок згасання коливань і погасання іскри контур розмикається, вібратор заряджається від індуктора до наступної іскри, процес багаторазової перезарядки половинок вібратора повторюється і т.д. Такі періодичні порції багатократних перезарядок половин вібратора (з періодом роботи переривача індуктора) супроводжується випромінюванням порції (***цуга***), електромагнітних хвиль в оточуючий простір.

Період роботи переривача індуктора набагато більший, ніж період електромагнітних коливань вібратора, тому то і використовують дроселі *Д* (рис.11.5) для захисту індуктора від високочастотних струмів. Для виявлення електромагнітних хвиль Герц застосовував вібратори різної форми. Найпростішим і найзручнішим є прямий відкритий вібратор, який за формою і розмірами тотожний випромінюючому вібратору (рис.11.5). Під дією змінного електричного поля прохідної електромагнітної хвилі електрони всередині вібратора починають здійснювати вимушені коливання, внаслідок чого у вібраторі виникає швидкозмінний струм, а між обома половинами вібратора – змінна напруга. Якщо довжина приймального вібратора дорівнює довжині випромінюючого вібратора і дорівнює $\lambda/2$, то власні частоти обох вібраторів збігаються і електричні коливання у приймальному вібраторі внаслідок резонансу підсилюються. Появу змінної напруги Герц виявляв за виникненням електричної іскорки у мікроскопічному зазорі всередині вібратора або за свіченням мініатюрної газорозрядної трубки *Т*, ввімкненої між обома половинками вібратора (рис.11.5).

Тепер для випромінювання електромагнітних хвиль користуються майже завжди генераторами, які дають змогу діставати електричні коливання практично будь-якої потужності і до того ж правильної синусоїдальної форми. Щоб збудити коливання у вібраторі, між обома його половинками ***вв*** можна ввімкнути один або кілька витків зв'язку (рис.11.6) і помістити їх поблизу котушки індуктивності *К* генератора *Г* (магнітний зв'язок). Є й інші способи зв'язку

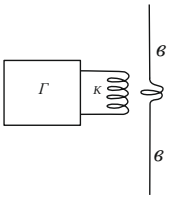


Рис.11.6.

між вібраторами і генераторами. Щоб підсилити коливання у вібраторі, користуються явищем резонансу, для чого частоту генератора роблять такою, щоб вона дорівнювала одній з власних частот вібратора, звичайно частоті основного його коливання.

Герц продовжив також дослідження властивостей електромагнітних хвиль. За допомогою великих металічних дзеркал і асфальтової призми (розміром більша 1м і масою 1200 кг) він здійснював відбивання і заломлення електромагнітних хвиль і виявив, що обидва ці явища підпорядковуються законам, які встановлені в оптиці для світлових хвиль. Відбиваючи біжучу плоску хвилю за допомогою металічного дзеркала в оберненому напрямку, Герц отримав стоячу хвилю. Відстань між вузлами і пучностями хвилі

дозволяла визначити довжину хвилі λ. Помноживши λ на частоту коливань вібратора ν, можна було знайти швидкість електромагнітних хвиль, яка виявилась близькою до швидкості світла c. Розташовуючи на шляху хвиль решітку із паралельних одна одній мідних дротин, Герц виявив, що при обертанні решітки навколо променя інтенсивність хвиль, які пройшли через решітку, сильно змінюється. Якщо дротини, що утворюють решітку, були перпендикулярні до вектора $\vec{E}$, хвиля проходила через решітку без завад. При розташуванні дротин паралельно $\vec{E}$ хвиля через решітку не проходила. Так було доведено поперечність електромагнітних хвиль.

Досліди Герца відіграли вирішальну роль для підтвердження і визнання теорії Максвелла. Експериментальне відкриття Герцом електромагнітних хвиль лягло в основу великого винаходу бездротової телефонії і телеграфії (радіо) в 1895-1896 рр. російським вченим А.С. Поповим і незалежно від нього італійським інженером Г.Марконі.

§11.6.Випромінювання елементарного диполя

Серед різних електричних систем, які випромінюють електромагнітні хвилі особливо важливе значення має електричний диполь. У §11.5 ми вже зустрічалися із застосуванням диполя (вібратора) для випромінювання електромагнітних хвиль. Проте там ми говорили про так званий *півхвильовий* диполь, довжина якого дорівнює половині довжини хвилі. А тепер ми розглянемо диполь, довжина якого мала порівняно з довжиною хвилі (елементарний диполь). Найпростішим прикладом елементарного диполя є дві металеві кулі, які заряджаються від якого-небудь генератора електричних коливань, якщо відстань між кулями $l \ll \lambda$ (рис.11.7). Момент такого диполя $p = ql$. Якщо генератор видає гармонічні коливання, то $q = q_0 \cos \omega t$ і момент диполя змінюється з часом також за гармонічним законом:

$$p = p_0 \cos \omega t. \qquad (11.36)$$

Тут $p_0 = q_0 l$ є амплітуда електричного момента диполя. ЇЇ можна подати також через амплітуду величини струму I_0. Справді, величина струму в диполі дорівнює $I = \frac{dq}{dt} = q_0 \omega \sin \omega t$, а амплітуда струму $I_0 = q_0 \omega$. Тому $q_0 = I_0/\omega$, а отже,

$$p_0 = \frac{l}{\omega} I_0. \qquad (11.37)$$

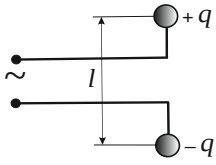


Рис.11.7.Елементарний диполь.

З елементарними диполями нам доводиться зустрічатися досить часто. Антени радіостанцій разом з поверхнею землі, на якій індукуються заряди протилежного знаку, в багатьох випадках можна розглядати як елементарні диполі. Проте найважливішим прикладом елементарних диполів є електрони всередині атомів. Коловий і еліптичний рухи електронів, які обертаються навколо позитивного ядра, можна розкласти на два прямолінійні гармонічні коливання. Але електрон, який здійснює прямолінійні гармонічні коливання, разом з позитивним ядром (яке не бере участі у випромінюванні) можна розглядати як диполь, момент якого змінюється згідно з формулою (11.36). Через те, що довжини хвиль, які випромінює атом

(для видимого світла округлено $5 \cdot 10^{-7}$м), набагато більші за розміри атомів ($\sim 10^{-10}$м), то диполі, які розглядаються, можна вважати з великою точністю елементарними.

Розглянемо тепер, який характер мають електромагнітні хвилі, що випромінюються диполем. При цьому ми не подаватимемо строгого виведення виразів для електромагнітного поля з рівнянь Максвелла, а обмежимося якісними результатами.

Характер електромагнітного поля диполя істотно залежить від того, близько чи далеко розміщена точка, яка розглядається. Якщо відстань r від центра диполя до цієї точки мала порівняно з довжиною хвилі ($r \ll \lambda$), то справедливі ті самі формули, що й для постійних електричного і магнітного полів. Електричне поле диполя виражається формулами і спадає з відстанню пропорційно $1/r^3$ (§2.2). Магнітне поле диполя виражається тією самою формулою (5.8) (чи скалярною формулою (5.9)), що й поле елемента струму, і пропорціональне $1/r^2$. Такі поля можна вважати квазістаціонарними. З ними не може бути пов'язане випромінювання енергії. Потік енергії через одиничну площину за одиницю часу визначається складовою вектора Пойтинга $\vec{\Pi} = \vec{E} \times \vec{H}$ (формула (11.32)), яка перпендикулярна до цієї площини. У квазістаціонарних полях $\vec{E}$ і $\vec{H}$ зсунуті по фазі на $\pi/2$ (як в стоячих хвилях), тому вектор Пойтинга, що змінюється з частотою 2ω, в середньому за період коливань точно дорівнює нулю. Відмінність $\vec{\Pi}$ від нуля може бути обумовлена тільки полями $\vec{E}$ і $\vec{H}$, які коливаються з однаковою фазою (як в біжучих хвилях, формула (11.25)) і зменшуються пропорціонально $1/r$ ($\Pi \sim 1/r^2$). Це безпосередньо витікає із закону збереження енергії, оскільки при відсутності втрат в середовищі повний потік енергії в просторі не повинен змінюватись із відстанню. Оскільки площа, яка охоплює диполь (площа сфери) збільшується як r^2, то необхідно, щоб П був пропорціональним $1/r^2$.

Таким чином, поле в ближній зоні ($r \ll \lambda$) диполя (зоні індукції) служить для формування біжучих складових полів, які відповідають за випромінювання. На великих відстанях від диполя ($r \gg \lambda$) закон зміни полів зовсім інший. Ця, так звана *хвильова* область, становить основний інтерес і тому

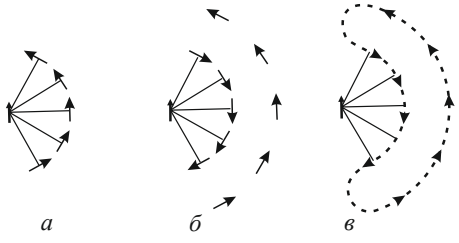


Рис.11.8.Форма електричних силових ліній поля диполя, який випромінює електромагнітні хвилі.

ми спинимося на ній докладніше. З'ясуємо, який вигляд має хвильовий фронт електромагнітної хвилі диполя. Через те що електромагнітне збурення поширюється на всі боки від диполя з однаковою швидкістю c (ми припускаємо, що диполь міститься у вакуумі), то час проходження хвилі до всіх точок, віддалених від диполя на одну й ту саму відстань r, однаковий. Тому в усіх точках сфери, центр якої збігається з диполем, фаза коливань однакова, тобто ми маємо *сферичний* хвильовий фронт, а отже хвиля, випромінювана диполем, є сферична чи кульова хвиля. Через те, що електричне поле E в хвилі перпендикулярне до напрямку поширення, то вектор $\vec{E}$ в різних точках перпендикулярний до радіусів-векторів (рис.11.8,*а*). Це поле періодично змінюється, і тому, переміщуючись

вздовж радіуса-вектора, ми знаходитимемо поля взаємно протилежного напряму (рис.11.8,*б*). З'єднавши на рис.11.8,*б* стрілки суцільною лінією, ми дістанемо одну з електричних силових ліній (рис.11.8,*в*).

На рис.11.9 наведена картина послідовного відокремлення силових ліній електричного поля $\vec{E}$, які створюються. У першій чверті періоду T коливань ($t = T/4$) виникає квазістатична частина поля (рис.11.9,*а*), яка при $t = T/2$ перетворюється в нуль, але від поля «відриваються» замкнуті самі на себе силові лінії поля $\vec{E}$ і «зчеплені» з ними кільцеві ортогональні магнітні лінії (на рис.11.9,*б* не показані). Разом вони утворюють автономну напівхвильову тороїдальну (в силу аксіальної симетрії) комірку сферично розбіжної хвилі, яка несе електромагнітну енергію (рис.11.9,*в,г*).

Напрям магнітного поля $\vec{H}$ у кожній точці перпендикулярний до $\vec{E}$ і до напряму поширення. Тому магнітними силовими лініями будуть концентричні кола, які лежать у площинах, перпендикулярних до диполя і мають центр на осі диполя.

Розглянемо докладніше, який вигляд має електромагнітне поле диполя в хвильовій області.

Відомо, що заряд, який рухається зі сталою швидкістю, спричиняє тільки постійне магнітне поле *H*. З іншого боку, згідно з основним положенням теорії Максвелла, для виникнення вихрового електричного поля необхідно, щоб $\frac{dH}{dt}$ було відмінне від нуля. Тому для випромінювання потрібно, щоб заряд рухався з певним *прискоренням* $\dot{v} = \ddot{l}$ (крапка означає диференціювання по часу) або, інакше, щоб друга похідна по часу від момента диполя $\ddot{p} = (\ddot{q}\vec{l})$ не дорівнювала нулю. Відповідно до

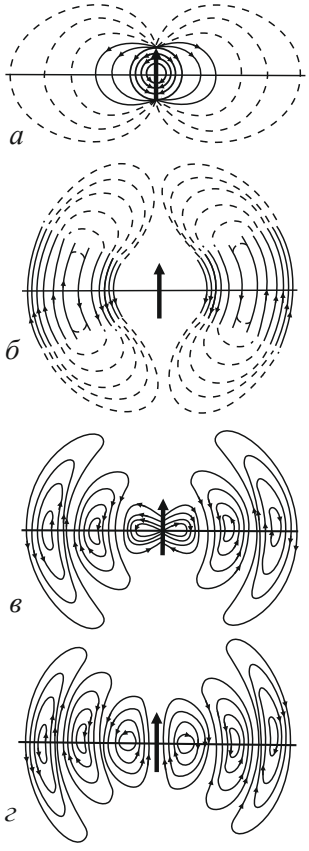


Рис.11.9.*а*-електричні силові лінії біля електричного диполя (при умові постійного заряду); *б-г*-силові лінії, які відділились від диполя: *б*-через ½ періода ($T/2$) після підключеннння генератора (заряд на диполі відсутній); *в*-через $(3/2)T$ (масштаб змінений); *г*-через $(7/4)T$ (масштаб змінений).

цього точний розрахунок показує, що електричне поле *E* хвилі (а отже, і магнітне поле *H*) пропорціональне $\ddot{p}(t - r/c)$. Аргумент $(t - -r/c)$ показує, що коливання поля в точці, яка лежить на відстані r від диполя, запізнюється відносно коливань $\ddot{p}$ в часі на r/c, так що $E(t)$ і $H(t)$ в розглядуваній точці в момент t визначаються значенням $\ddot{p}$ у більш ранній момент часу $(t - r/c)$. Раніше вже було показано, що виходячи із закону збереження енергії, *E* і *H* пропорціональні 1/*r*.

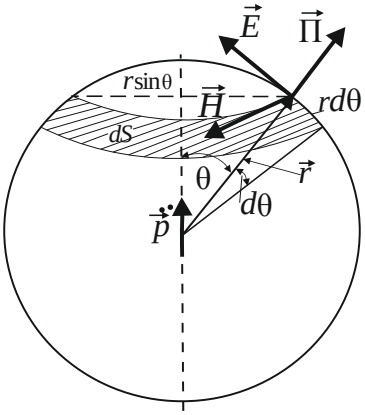


Рис.11.10.Електромагнітне поле у сферичній хвилі, яку випромінює диполь.

Величини полів E і H залежить ще від напряму випромінювання, тобто від кута θ, утвореного радіусом-вектором $\vec{r}$ з віссю диполя (рис.11.10). Вже говорилось, що поблизу диполя його магнітне поле виражається тією самою формулою (5.9), що й для елемента струму. Це магнітне поле пропорціональне $\sin\theta$ і дорівнює нулю в будь-якій точці, яка лежить на продовжені осі диполя. Так само і напруженості полів E і H у хвильовій зоні пропорціональні $\sin\theta$.

Точний розрахунок дає такий результат:

$$E = \frac{1}{4\pi\varepsilon_0}\frac{\ddot{p}(t-r/c)\sin\theta}{c^2 r},\ H = \sqrt{\frac{\varepsilon_0}{\mu_0}}E. \quad (11.38)$$

Тут $c = 3\cdot 10^8$ м/с – швидкість поширення електромагнітних хвиль у вакуумі. Всі інші величини в (11.38) мають розмірність в системі СІ.

Обидва поля $\vec{E}$ і $\vec{H}$ напрямлені перпендикулярно до напряму поширення, тобто до напряму $\vec{r}$, і зв'язані з ним правилом правого свердлика (рис.11.10). Зауважимо, що формула (11.38) справедлива тільки для таких рухів зарядів, швидкість яких v мала порівняно із швидкістю світла c (нерелятивістський випадок, $\beta = v/c \ll 1$). В окремому випадку, коли p змінюється за гармонічним законом (11.36)

$$\ddot{p}(t-r/c) = -p_0\,\omega^2\cos\omega(t-r/c) = \\ = -p_0\,\omega^2\cos(\omega t - kr),$$

і вираз (11.38) набирає вигляду:

$$E = -\frac{1}{4\pi\varepsilon_0}\frac{p_0\,\omega^2\sin\theta}{c^2 r}\cos(\omega t - kr). \quad (11.39)$$

Енергія, яка випромінюється диполем. Значення вектора потоку енергії в якій-небудь точці з координатами r, θ є

$$\Pi = EH = \frac{{p_0}^2\omega^4\sin^2\theta}{16\pi^2\varepsilon_0 c^3 r^2}\cos^2(\omega t - kr).$$

Інтерес має не миттєве значення Π, а його середнє значення $\langle\Pi\rangle$ за досить великий проміжокчасу, тобто інтенсивність випромінювання. Через те, що $\langle\cos^2(\omega t - kr)\rangle = 1/2$, то

$$\langle\Pi\rangle = I = \frac{1}{32\pi^2\varepsilon_0 c^3}\frac{{p_0}^2\omega^4\sin^2\theta}{r^2}. \quad (11.40)$$

Інтенсивність випромінювання пропорціональна частоті коливань ω у четвертій степені. Вона залежить також від напряму випромінювання (пропорціональна

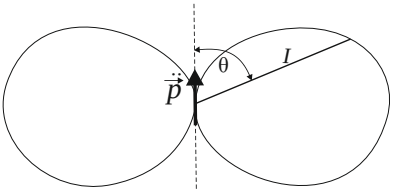


Рис.11.11.Діаграма направленості випромінювання елементарного диполя.

$\sin^2\theta$). Залежність інтенсивності випромінювання диполя від напряму подано на рис.11.11 («діаграма направленості» випромінювання). Таким чином, у напрямі своєї осі диполь нічого не випромінює. Навпаки, в напрямах, перпендикулярних до осі диполя, інтенсивність випромінювання найбільша. Ця діаграма будується так, щоб довжина відрізка, який відсікається нею на промені, що

проведений із центра диполя, давала інтенсивність випромінювання під кутом θ.

Нарешті, обчислимо повну енергію *P*, яку випромінює диполь в усіх напрямах за одиницю часу. Через те, що $\langle \Pi \rangle$ є енергія, яка проходить через одиницю поверхні за 1 сек, то

$$P = \int_S \langle \vec{\Pi} \rangle d\vec{S} = \int_S \langle \vec{\Pi}\vec{n} \rangle dS = \int_S \langle \Pi_n \rangle dS,$$

де інтегрування виконується по довільній поверхні *S*, яка охоплює диполь. Вибираючи сферу, радіус якої r, з центром у точці розміщення диполя і користуючись сферичними координатами r, θ (рис.11.10), отримаємо:

$$dS = 2\pi r^2 \sin\theta d\theta,$$

$$\langle \Pi_n \rangle dS = \langle \Pi \rangle dS = \frac{{p_0}^2 \omega^4}{16\pi^2 \varepsilon_0 c^3} \sin^3\theta d\theta,$$

$$P = \frac{{p_0}^2 \omega^4}{16\pi^2 \varepsilon_0 c^3} \int_0^{\pi} \sin^3\theta d\theta = \frac{1}{12\pi} \frac{{p_0}^2 \omega^4}{\varepsilon_0 c^3}. \quad (11.41)$$

Написаний вираз можна подати в іншому вигляді. Для цього виразимо амплітуду коливань момента p_0 через амплітуду струму в диполі I_0 згідно з (11.37). Тоді отримаємо :

$$P = \frac{1}{12\pi} \frac{l^2 \omega^2}{\varepsilon_0 c^3} {I_0}^2, \quad (11.42)$$

де *l* – довжина диполя. Цей вираз подібний до виразу для потужності, яку виділяє змінний струм з опором *R* :

$$P = \frac{1}{2} R {I_0}^2.$$

Тому величину

$$R_{випром} = \frac{1}{6\pi} \frac{l^2 \omega^2}{\varepsilon_0 c^3} \quad (11.43)$$

називають *опором випромінювання диполя*. Це поняття застосовують в радіотехніці до випромінювання антен передавальних станцій.

Контрольні питання

1.Рівнянні плоскої біжучої електромагнітної хвилі (ЕМХ).

2.Фазова швидкість ЕМХ у вакуумі та середовищі.

3.Співвідношення між напруженістю електричного та напруженістю магнітного поля в ЕМХ.

4.Густина потоку електромагнітної енергії – вектор Пойтинга. Чому вектор $\vec{E}$ ЕМХ називають світловим?

5.Якою характеристикою ЕМХ визначається її інтенсивність?

6.Пояснити, чому стоячі хвилі не переносять енергію.

7.Нарисувати і пояснити полярну діаграму випромінювання електричного диполя у хвильовій зоні. Які можна зробити висновки?

8.Яка необхідна умова, щоб електричний диполь і електричний заряд випромінювали електромагнітні хвилі?

Розділ 12. Інтерференція електромагнітних хвиль

§12.1.Основні відомості про хвилі оптичного діапазону. Інтенсивність електромагнітних хвиль

Світло, у *вузькому розумінні* – теж, що і видиме випромінювання, тобто електромагнітні хвилі в інтервалі частот, які сприймаються людським оком ($7{,}5\cdot10^{14} \div 4{,}0\cdot10^{14}$ Гц), що відповідає довжинам хвиль у вакуумі від ~0,4 до ~0,76 мкм. Світло високої інтенсивності око сприймає в дещо більш широкому діапазоні. Світлові хвилі різних частот сприймаються людиною як різні кольори.

Світло в *широкому розумінні* – синонім оптичного випромінювання, що включає, крім видимого випромінювання,

ультрафіолетову (УФ) та інфрачервону (ІЧ) області спектру.

Світло – це складне явище: в одних випадках воно поводить себе як електромагнітна хвиля, в інших – як потік особливих частинок (фотонів). В даному розділі розглядається хвильова оптика, тобто коло явищ, в основі яких лежить хвильова природа світла.

В електромагнітній хвилі коливаються вектори $\vec{E}$ і $\vec{H}$. Як показують дослідження фізіологічну, фотохімічну, фотоелектричну та інші дії світла обумовлені коливаннями електричного вектора $\vec{E}$. У відповідності до цього будемо в подальшому говорити про світловий вектор, розуміючи під ним вектор напруженості електричного поля $\vec{E}$. Про магнітний вектор світлової хвилі ми згадувати майже не будемо. Модуль амплітуди світлового вектора будемо позначати, як правило, буквою E_m. Відповідно зміни з часом і в просторі проекції світлового вектора на напрямок, вздовж якого він коливається, описується рівнянням (11.20):

$$\vec{E} = \vec{E}_m \cos(\omega t - kr + \alpha). \quad (12.1)$$

Тут $k = 2\pi/\lambda$ – *хвильове* число, r – відстань, яка відраховується вздовж напрямку розповсюдження світлової хвилі. Для плоскої хвилі, яка розповсюджується в середовищі, яке не поглинає світло, $E_m = \text{const}$, для сферичної хвилі E_m зменшується з відстанню пропорційно $1/r$.

Згідно (11.19) відношення швидкості світлової хвилі у вакуумі до фазової швидкості v в деякому середовищі називається абсолютним показником заломлення цього середовища. Таким чином

$$n = c/v. \quad (12.2)$$

Із формули (11.19) видно, що $n = \sqrt{\varepsilon\mu}$. Для більшості прозорих речовин μ практично не відрізняється від одиниці. Тому можна вважати, що

$$n = \sqrt{\varepsilon}. \quad (12.3)$$

Формула (12.3) зв'язує оптичні властивості речовин з його електричними властивостями. На перший погляд може здатись, що ця формула неправильна. Наприклад, для води $\varepsilon = 81$, а $n = 1{,}33$. Однак треба мати на увазі, що значення $\varepsilon = 81$ отримано із електростатичних вимірювань. У швидкозмінних електричних полях значення ε отримується іншим, причому воно залежить від частоти коливань поля. Цим пояснюється дисперсія світла, тобто залежність показника заломлення (або швидкості світла) від частоти (або довжини хвилі). Підстановка у формулу (12.3) значення ε, яке отримане для відповідної частоти, дає правильне значення n.

Значення показника заломлення характеризує *оптичну густину* середовища. Середовище з більшим n називається оптично більш густим, ніж середовище з меншим n. Відповідно середовище з меншим n називається оптично менш густим, ніж середовище з більшим n.

Довжини хвиль видимого світла $\lambda_0 = 0{,}40 \div 0{,}76$ мкм відносяться до світлових хвиль у вакуумі. В речовині довжини світлових хвиль будуть іншими. У випадку коливань з частотою ν довжина хвилі у вакуумі буде $\lambda_0 = \frac{c}{\nu}$. В середовищі, в якому фазова швидкість світлової хвилі $v = \frac{c}{n}$, довжина хвилі має значення $\lambda = \frac{v}{\nu} = \frac{c}{\nu n} = \lambda_0/n$. Таким чином, довжина світлової хвилі в середовищі з показником заломлення n зв'язана з довжиною хвилі у вакуумі співвідношенням

$$\lambda = \lambda_0/n \ . \qquad (12.4)$$

Інтенсивність світла I згідно (11.35) в данній точці простору буде:

$$I = \frac{1}{2}\sqrt{\frac{\varepsilon_0}{\mu_0}}\, nE_m^2 \ , \qquad (12.5)$$

тобто інтенсивність світла пропорційна показнику заломлення середовища і квадрату амплітуди світлової хвилі.

Зауважимо, що при проходженні світла в однорідному середовищі можна вважати, що інтенсивність пропорційна квадрату амплітуди світлової хвилі

$$I \sim E_m^2 \ . \qquad (12.6)$$

Однак у випадку проходження світла через границю розділу середовищ вираз для інтенсивності, який не враховує множник n, приводить до незбереження світлового потоку.

Лінії, вздовж яких розповсюджується світлова енергія, називаються променями. Усереднений вектор Пойтинга $\langle\vec{\Pi}\rangle$ направлений в кожній точці по дотичній до променя. В ізотропних середовищах (властивості в усіх напрямках однакові) напрямок $\langle\vec{\Pi}\rangle$ співпадає з нормаллю до хвильової поверхні, тобто з напрямом хвильового вектора $\vec{k}$. Значить промені перпендикулярні до хвильових поверхонь. В анізотропних середовищах (властивості в різних напрямках є різними) нормаль до хвильової поверхні в загальному випадку не співпадає з напрямком вектора Пойтинга, тобто промені не перпендикулярні хвильовим поверхням.

§12.2. Інтерференція світла

Хвильові властивості світла найбільш виразно проявляють себе в інтерференції. Ці явища характерні для хвиль будь-яої природи і порівняно просто спостерігаються в досліді для хвиль на поверхні води (§10.7) або для звукових хвиль. Спостерігати ж інтерференцію світлових хвиль можна тільки при певних умовах.

Визначення інтерференції хвиль дано в §10.7. *Під інтерференцією світла, як правило, розуміють широке коло явищ, в яких при накладанні пучків світла результуюча інтенсивність не дорівнює сумі інтенсивностей окремих пучків: в одних місцях вона більша, в інших менша, тобто виникають світлі і темні ділянки – інтерференційні смуги, які чергуються.*

Умови, при яких має місце інтерференція, розглядались у §10.7. Перш за все хвилі повинні бути монохроматичними ($\omega_1 = \omega_2$). Світло, яке випромінюється звичайними (не лазерними) джерелами, не буває строго монохроматичним. Тому для спостереження інтерференції світло від одного джерела необхідно розділити на два пучки, а потім накласти їх один на одного. Існуючі експериментальні методи отримання когерентних пучків із одного світлового пучка можна поділити на два класи. В методі ділення хвильового фронту пучок пропускається, наприклад, через два близько розташованих отвори в непрозорому екрані. Такий метод годиться тільки при досить малих розмірах джерела. В іншому методі пучок ділиться однією або декількома поверхнями, які частково відбивають і пропускають світло. Цей метод поділу амплітуди може застосовуватись і для протяжних джерел. Він забезпечує більшу інтенсивність і лежить в основі дії різних інтерферометрів. В залежності від числа пучків, які інтерферують, розрізняють двопроменеві і багатопроменеві інтерферометри, які мають

важливе практичне застосування в техніці, метрології і спектроскопії.

Розглянемо інтерференцію монохроматичного світла. Світлові коливання в деякій точці, через яку проходить строго монохроматична хвиля, повинні відбуватись нескінченно довго і мати незмінну частоту і амплітуду. Просте світло, яке випромінюється любим реальним джерелом, не має цієї властивості. Тим не менше монохроматична ідеалізація є достатньою для розв'язання багатьох задач. Зокрема при вивченні явищ інтерференції вона годиться для визначення положення максимумів і мінімумів інтерференційної картини.

Нехай розділення на дві когерентні хвилі відбувається в точці O (рис.12.1). До точки P, в якій спостерігається інтерференційна картина, одна хвиля в середовищі з показником n_1, проходить шлях r_1, а друга – в середовищі з показником заломлення n_2 – шлях r_2. Згідно (11.20) і (12.1) перша хвиля викличе в точці P коливання

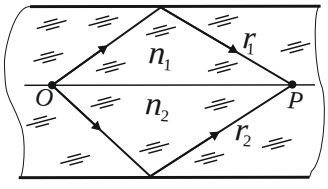


Рис.12.1.

$$\vec{E}_1=\vec{E}_{m1}\cos\left(\omega t-k_1r_1+\varphi_{01}\right), \quad (12.7)$$

а друга

$$\vec{E}_2=\vec{E}_{m2}\cos\left(\omega t-k_2r_2+\varphi_{02}\right), \quad (12.8)$$

де

$$k_1=\omega/v_1=\omega n_1/c=2\pi n_1/\lambda_0\,,\quad k_2=\omega/v_2=\omega n_2/c=2\pi n_2/\lambda_0 \quad (12.9)$$

хвильові числа першої і другої хвиль, c – швидкість світла у вакуумі, $\lambda_0=c/\nu$ – довжина хвилі у вакуумі, φ_{01} і φ_{02} – початкові фази.

Згідно принципу суперпозиції, напруженість результуючого поля в точці P дорівнює їх векторній сумі:

$$\vec{E}_P=\vec{E}_1+\vec{E}_2. \quad (12.10)$$

В результаті додавання двох гармонічних коливань однакової частоти отримується коливання тієї ж частоти, а незмінна з часом його амплітуда залежить від співвідношення фаз коливань, які додаються (§9.4), і тому в різних точках спостереження амплітуда має, взагалі кажучи, різні значення.

Зважаючи на дуже велику частоту оптичних коливань ($\nu\sim10^{14}$Гц) напруженість $\vec{E}_P$ неможливо виміряти безпосередньо. Всі приймачі випромінювання вимірюють енергетичні величини (інтенсивність світла або освітленість поверхні) усереднені за проміжок часу, який набагато більший за період оптичних коливань. Тому величини, які експериментально вимірюють, пропорційні середньому значенню квадрату напруженості електричного поля $I_p\sim\langle E_P^2\rangle$ за час, який визначається інерційністю приймача випромінювання:

$$\langle E_P^2\rangle=\langle\left(\vec{E}_1+\vec{E}_2\right)^2\rangle= =\langle E_1^2\rangle+\langle E_2^2\rangle+2\langle\vec{E}_1\cdot\vec{E}_2\rangle. \quad (12.11)$$

Вираз для результуючої інтенсивності крім суми інтенсивностей $I_1\sim\langle E_1^2\rangle$ і $I_2\sim\langle E_2^2\rangle$ кожної із хвиль містить ще один доданок, який пропорційний $2\langle\vec{E}_1\cdot\vec{E}_2\rangle$, що носить назву інтерференційного члена. В тих випадках, коли він дорівнює нулю, результуюча інтенсивність $I_P=I_1+I_2$ і інтерференція відсутня.

Скалярний добуток $\vec{E}_1\cdot\vec{E}_2$ дорівнює нулю, коли $\vec{E}_1\perp\vec{E}_2$, тобто хвилі, що додаються, поляризовані у взаємно перпендикулярних напрямках (поляризація електромагнітних хвиль розглядається в

подальшому). *Відсутність інтерференції променів, які поляризовані у взаємно перпендикулярних напрямках, була виявлена Френелем і Араго в* 1816р. *і інтерпретована в* 1817р. *Юнгом як доказ поперечності світлових хвиль. Електромагнітна теорія світла це підтвердила.*

Будемо вважати, що вектори $\vec{E}_1$ і $\vec{E}_2$ в точці спостереження P здійснюють коливання вздовж однієї прямої. Тоді інтерференційний член можна записати так: $2\langle\vec{E}_1\cdot\vec{E}_2\rangle=2\langle E_1E_2\rangle$.

Додавання однаково спрямованих з однаковими частотами коливань розглянуто у §9.4**.** Згідно формули (9.20) амплітуда результуючого коливання E_{mP} в точці P із врахуванням (12.7) і (12.8) буде:

$$E_{mP}^2=E_{m1}^2+E_{m2}^2+ +2E_{m1}E_{m2}\cos\left(k_2r_2-k_1r_1+\varphi_{01}-\varphi_{02}\right). \quad (12.12)$$

Цей же результат можна отримати із співвідношення (12.11), якщо врахувати, що

$$\begin{cases}\langle E_P^2\rangle=\frac{1}{2}E_{mP}^2;\ \langle E_1^2\rangle=\frac{1}{2}E_{m1}^2,;\langle E_2^2\rangle=\frac{1}{2}E_{m2}^2;\\ \langle\vec{E}_1\vec{E}_2\rangle=\frac{1}{2}E_{m1}E_{m2}\cdot\\ \cdot\cos\left(k_2\,r_2-k_1r_1+\varphi_{01}-\varphi_{02}\right).\end{cases}$$

Враховуючи, що інтенсивності хвиль пропорційні квадратам їхніх амплітуд, для інтенсивності результуючого коливання в точці P отримаємо:

$$I_P=I_1+I_2+2\sqrt{I_1I_2}\cos\varphi, \quad (12.13)$$

де

$$\varphi=k_2\,r_2-k_1r_1+\varphi_{01}-\varphi_{02} \quad (12.14)$$

різниця фаз двох когерентних хвиль від одного джерела. Із урахуванням (12.9) запишемо (12.14) у наступному вигляді:

$$\varphi=\frac{2\pi}{\lambda_0}(r_2n_2-r_1n_{1)}+\varphi_{01}-\varphi_{02}). \quad (12.15)$$

Добуток геометричної довжини шляху світлової хвилі на показник заломлення цього середовища називається *оптичною довжиною шляху*, а різниця оптичних довжин шляхів, що пройшли хвилі, $\Delta=r_2n_2-r_1n_1$ – *оптичною різницею ходу*.

У випадку когерентних хвиль $\varphi_{01}-\varphi_{02}$=const і для зручності будемо вважати, що $\varphi_{01}=\varphi_{02}$. В результаті (12.13) перепишеться так:

$$I_P=I_1+I_2+2\sqrt{I_1I_2}\cos\frac{2\pi}{\lambda_0}\Delta. \quad (12.16)$$

Якщо оптична різниця ходу дорівнює парному числу півхвиль у вакуумі:

$$\Delta=2m\frac{\lambda_0}{2}\ \ (m=0,\pm1,\pm2,\pm3,\dots), \quad (12.17)$$

то $\varphi=\frac{2\pi}{\lambda_0}\Delta=2m\pi$, і коливання, що збуджуються в точці P обома хвилями, відбуваються в однакових фазах. *Число* m *називається порядком інтерференції.* Отже $\cos\frac{2\pi}{\lambda_0}\Delta=1$ і результуюча інтенсивність – максимальна:

$$\begin{matrix}I_{Pmax}=I_1+I_2+2\sqrt{I_1I_2}\ \ (E_{m1}\neq E_{m2});\\ I_{Pmax}=4I_1\qquad(E_{m1}=E_{m2}).\end{matrix} \quad (12.18)$$

Таким чином, максимум інтенсивності при накладанні двох когерентних хвиль буде у точках, для яких виконується умова (12.17), тобто на оптичній різниці ходу укладається парне число півхвиль ($\lambda_0/2$). Сукупність таких точок утворює інтерференційні лінії (смуги), порядок яких визначається числом m.

Якщо оптична різниця ходу:

$$\Delta=(2m+1)\frac{\lambda_0}{2}\ (m=0,\pm1,\pm2,\pm3,\dots), \quad (12.19)$$

то $\varphi=(2m+1)\pi$ і коливання, що збуджуються в точці P обома хвилями, знаходяться у протифазі. Це означає,що $\cos\frac{2\pi}{\lambda_0}\Delta=-1$ і результуюча інтенсивність в цій точці згідно (2.16) буде мінімальна:

$$\begin{matrix}I_{Pmin}=I_1+I_2-2\sqrt{I_1I_2}\ \ (E_{m1}\neq E_{m2});\\ I_{Pmin}=0\qquad(E_{m1}=E_{m2}).\end{matrix} \quad (12.20)$$

Мінімум інтенсивності отримаємо в точках, для яких оптична різниця ходу

променів вміщує непарне число півхвиль ($\lambda_0/2$). Таким чином (12.17) – *умова інтерференційного максимуму,* а (12.19) – *умова інтерференційного мінімуму*.

§12.3. Розрахунок інтерференційної картини від двох когерентних джерел

Розглянемо випадок інтерференції хвиль від двох однакових синфазних ($\alpha_{01} = \alpha_{02}$) монохроматичних точкових джерел S_1 і S_2 (рис.12.2), які знаходяться на відстані d одне від одного в середовищі з показником заломлення n=1. Якщо відстань l до екрана, де спостерігаються інтерференційні смуги,

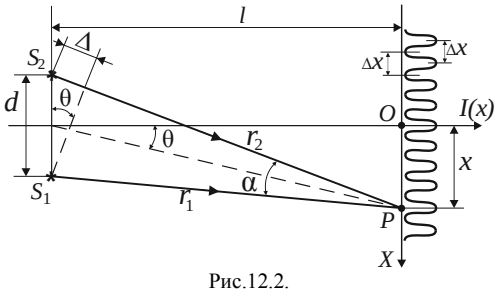


Рис.12.2.

набагато більша відстані між джерелами ($l \gg d$), то коливання в точці P, що збуджуються хвилями від S_1 і S_2 мають однаковий напрямок і однакові амплітуди ($I_1 = I_2 = I_0$). Інтенсивність результуючого коливання в точці P згідно (12.16) буде:

$$I_P = 2I_0\left(1+\cos\frac{2\pi}{\lambda_0}\Delta\right) = 4I_0\cos^2\left(\frac{\pi}{\lambda_0}\Delta\right), \quad (12.21)$$

де I_0 – інтенсивність коливань від одного джерела, $\Delta = (r_2 - r_1)$ – різниця ходу хвиль, що інтерферують. Освітленість екрана в мінімумах дорівнює нулю, а в максимумах – $4I_0$. Положення максимумів визначається умовою (12.17).

В тому випадку, якщо хвилі від джерел розповсюджуються не у вакуумі, а в середовищі з показником заломлення n, формула (12.21) залишається справедливою, але в ній під Δ необхідно розуміти не геометричну, а оптичну різницю ходу хвиль, що інтерферують: $\Delta = n(r_2 - r_1)$.

Щоб знайти залежність освітленості екрана від координати x (рис.12.2), необхідно різницю ходу Δ виразити через координату x точки спостереження P. Для зручності введемо кут θ, що утворюється напрямком на точку P з перпендикуляром до лінії, яка з'єднує джерела (тобто з оптичною віссю схеми, яка розглядається). Для практично важливого випадку малих значень θ ($\theta \ll 1$), для різниці ходу можна записати $\Delta \approx d\sin\theta \approx d\cdot\theta$. Так як $\mathrm{tg}\theta \approx \theta \approx x/l$, то $\Delta \approx xd/l$. Підставляючи цей вираз для Δ в (12.21), отримаємо

$$I_P(x) = 2I_0\left(1+\cos\frac{2\pi dx}{\lambda_0 l}\right). \quad (12.22)$$

В точці, для якої x=0, розташований максимум, який відповідає нульовій різниці ходу. Для цього максимуму порядок інтерференції m=0. Це центр інтерференційної картини. Відстань між *сусідніми максимумами* або *мінімумами* (просторовий період інтерференційної картини) визначається із умови:

$$\frac{2\pi d\cdot\Delta x}{\lambda_0 l} = 2\pi.$$

Звідки

$$\Delta x = \frac{\lambda_0 l}{d}. \quad (12.23)$$

Формула (12.23) визначає так звану *ширину інтерференційної смуги.*

Якщо ввести кут сходження променів $\alpha \approx d/l$, тобто кут, під яким видні джерела S_1 і S_2 із точки спостереження, то вираз для Δx можна записати так:

$$\Delta x = \lambda_0/\alpha. \qquad (12.24)$$

Ця формула використовується для випадку інтерференції плоских хвиль (паралельних променів), які розповсюджуються під кутом α. Дійсно, для великих відстаней від джерел, сферичні хвилі на невеликих ділянках наближено можна вважати як плоскі, кут між напрямками яких при $\theta \ll 1$ приблизно дорівнює d/l.

Відзначимо, що середнє значення освітленості по екрану в інтерференційній картині, згідно формули (12.22) буде:

$$\langle I(x)\rangle = \langle 2I_0\rangle + \langle 2I_0 \cos\frac{2\pi dx}{\lambda_0 l}\rangle = 2I_0,$$

що дорівнює подвоєній освітленості від одного джерела. Це значить, що при інтерференції відбуваються тільки перерозподіл енергії в просторі, а повний потік енергії залишається незмінним.

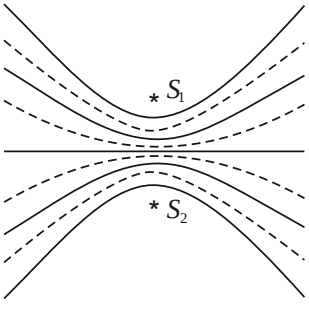


Рис.12.3.

До цього часу розглядались тільки точки екрана, які лежать в площині рис.12.2. В просторі поверхнями максимальної і мінімальної інтенсивності є гіперболоїди обертання з фокусами в точках S_1 і S_2, так як відповідають множині точок, для яких різниця відстаней від двох заданих точок S_1 і S_2 має одне і теж значення, тобто $r_2 - r_1 = \text{const}$. Для максимальної інтенсивності $r_2 - r_1 = 2m\frac{\lambda_0}{2}$, для мінімальної – $r_2 - r_1 = (2m+1)\frac{\lambda_0}{2}$.

Переріз таких гіперболічних поверхонь наведено на рис.12.3, де суцільними лініями показано розміщення максимумів інтенсивності, штрихованими – мінімуми інтенсивності.

Контрастність інтерференційної картини характеризують за допомогою параметра

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \qquad (12.25)$$

де I_{max} і I_{min} – інтенсивність світлих і темних смуг. Параметр V вперше ввів Майкельсон і його називають *функцією видності* (або просто *видністю*). Якщо інтенсивність темної смуги дорівнює нулю ($I_{min} = 0$), то $V = 1$, тобто когерентність променів максимальна. Для рівномірно освітленого екрана ($I_{max} = I_{min}$) видність $V = 0$ (промені не когеренті).

§12.4. Способи спостереження інтерференції світла

1. Метод поділу хвильового фронту

Вперше експериментальна установка для демонстрації інтерференції світла була виготовлена Томасом Юнгом на початку XIX століття. Яскравий пучок сонячних променів освітлював екран *A* з малим отвором *S* (рис.12.4). Світло, яке пройшло через отвір *S*, має вид розбіжного пучка, який падає на другий екран *B* з двома малими отворами S_1 і S_2, які розташовані близько один від одного на однакових відстанях від *S*. Ці отвори діють як вторинні точкові когерентні (синфазні) джерела і хвилі, які виходять із них, перекриваються і створюють інтерференційну картину, яку спостерігають на віддаленому екрані **E**.

Розташування світлих і темних смуг на екрані знаходять, користуючись вище розглянутою монохроматичною ідеалізацією.

Відстань Δx між сусідніми смугами визначається за формулою (12.23). Вимірюючи ширину інтерференційних смуг, Юнг у 1802 році вперше визначив довжини світлових хвиль для різних кольорів, хоч ці виміри і не були точними.

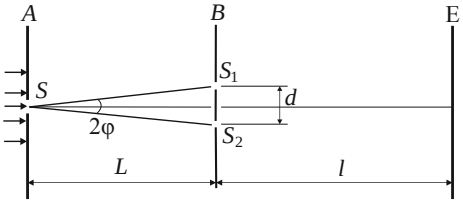


Рис.12.4.Схема досліду Юнга.

Інший інтерференційний дослід був здійснений Френелем у 1816 році. Дві когерентні світлові хвилі отримувались в результаті відбивання від двох дзеркал, площини яких нахилені під невеликим кутом α одна до одної (рис.12.5). Джерелом світла є вузька яскраво освітлена щілина *S*, яка паралельна ребру між дзеркалами. Відбиті від дзеркал пучки падають на екран і в тій області, де вони перекриваються, виникає інтерференційна картина. Від прямого попадання променів від джерела *S* екран захищений ширмою. Для розрахунку

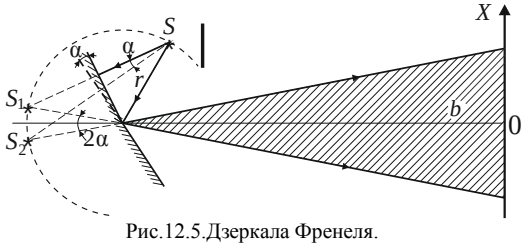


Рис.12.5.Дзеркала Френеля.

освітленості екрана можна вважати, що хвилі, які інтерферують, випромінюється вторинними джерелами S_1 і S_2, які є уявними зображеннями щілини *S* в дзеркалах. Тому $I(x)$ визначається формулою (12.22), в якій відстань *l* від джерел до екрана необхідно замінити на $r+b$, де $r-$ відстань від *S* до ребра дзеркал, $b-$ відстань від ребра до екрана (рис.12.5). Відстань між джерелами, як видно із рис.12.5, дорівнює $d \approx 2r\alpha$. Тому ширина інтерференційної смуги на екрані дорівнює

$$\Delta x \approx \frac{\lambda l}{d} = \frac{\lambda(r+b)}{2r\alpha}. \qquad (12.26)$$

В іншому інтерференційному досліді Френеля для розділення початкової світлової хвилі на дві використовують біпризму з кутом при вершині α, який дуже малий (біпризма Френеля). Джерелом світла служить яскраво освітлена вузька щілина *S*, яка паралельна заломному ребру біпризми (рис.12.6). Можна вважати, що тут утворюються два уявних зображення S_1 і S_2 джерела *S*, які близько розташовуються, так як кожна половинка біпризми відхиляє промені на невеликий кут $(n-1)\alpha$.

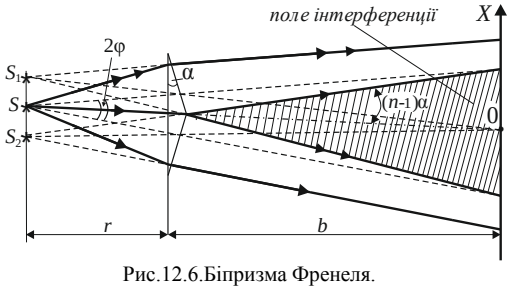


Рис.12.6.Біпризма Френеля.

Аналогічний пристрій, в якому роль когерентних джерел відіграють дійсні зображення яскраво освітленої щілини, отримуються, якщо збиральну лінзу розрізати по діаметру і половинки трохи розсунути (білінза Бі'є, рис.12.7). Проріз закривають непрозорим екраном, а промені, які падають на лінзу, проходять через дійсні зображення щілини *S* S_1 і S_2 і в подальшому перекриваються, утворюючи інтерференційне поле. У всіх наведених пристроях інтерференційні смуги можна спостерігати в будь-якому місці в області перекриття

розбіжних променів від джерел S_1 і S_2 (ці області заштриховані на рис.12.5−12.7).

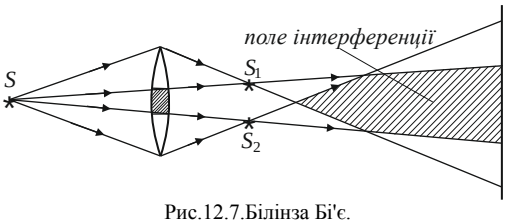


Рис.12.7.Білінза Бі'є.

2. Метод поділу амплітуд

Інтерференцію світла за методом поділу амплітуди в багатьох відношеннях спостерігати простіше, ніж в дослідах з поділом хвильового фронту. В дослідах Поля світло від джерела *S* відбивається двома поверхнями тонкої прозорої плоско-паралельної пластинки (рис.12.8). В будь-яку точку *P*, яка знаходиться з тієї ж сторони від пластинки, що і джерело, приходять два промені з різницею ходу $\Delta = 2h\cos\theta$ (рис. 12.8). При малому коефіцієнті відбивання повторне відбивання від внутрішніх поверхонь пластинки можна не враховувати із-за малої енергії променів, які зазнають два відбивання і більше. Перші два промені утворюють інтерференційну картину. Для визначення виду смуг можна уявити собі, що промені йдуть із уявних зображень S_1 і S_2

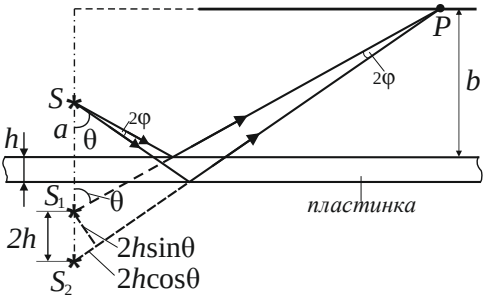


Рис.12.8.

джерела *S*, які утворюються поверхнями пластинки. На віддаленому екрані, який розташований паралельно до пластинки, інтерференційні смуги мають вид концентричних кілець з центрами на перпендикулярі до пластинки, що проходить через джерело *S*. За допомогою листочка слюди (товщиною 0,03-0,05 мм) можна отримати яскраву інтерференційну картину прямо на стелі і стінах аудиторії. Чим тоньша пластинка, тим більша відстань між смугами.

Особливо важливий частинний випадок інтерференції світла, яке відбивається двома поверхнями плоскопаралельної пластинки, коли точка спостереження *P* знаходиться на нескінченності, тобто спостереження ведеться або оком, яке акомодоване на нескінченність, або на екрані, що розташований у фокальній площині збиральної лінзи (рис.12.9). В цьому випадку обидва промені, які йдуть від *S* до *P*, породжені одним падаючим променем і після відбивання від передньої і задньої поверхонь пластинки паралельні між собою (рис.12.9). Оптична

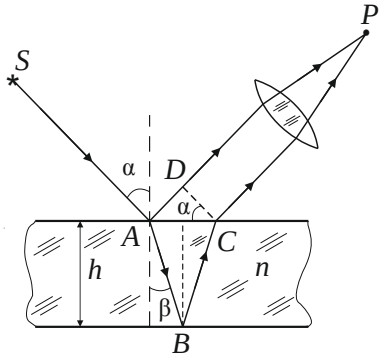


Рис.12.9.

різниця ходу між ними в точці *P* така ж, як на лінії *DC*:

$$\Delta = n(|AB| + |BC|) - |AD| \pm \frac{\lambda}{2}.$$

Тут n − показник заломлення матеріалу пластинки. Вважається, що над пластинкою знаходиться повітря, тобто $n_1 \approx 1$.

Так як $|AB| = |BC| = \frac{h}{\cos\beta}$, $|AD| = 2h \cdot \mathrm{tg}\beta \cdot$ $\cdot \sin\alpha$ (h–товщина пластинки, α і β – кути падіння і заломлення на верхній грані; $\sin\alpha = n\sin\beta$), то для різниці ходу отримаємо

$$\Delta = 2nh\cos\beta \pm \frac{\lambda_0}{2} = \\ = 2h\sqrt{n^2 - \sin^2\alpha} \pm \frac{\lambda_0}{2}. \quad (12.27)$$

Тут враховано, що при відбиванні хвилі в точці A, вона "втрачає півхвилі", тобто фаза хвилі змінюється на π. Тому різниця фаз φ хвиль, які складаються в точці P дорівнює

$$\varphi = \frac{4\pi nh\cos\beta}{\lambda_0} \pm \pi = \\ = \frac{4\pi h\sqrt{n^2 - \sin^2\alpha}}{\lambda_0} \pm \pi, \quad (12.28)$$

де λ_0 – довжина хвилі у вакуумі. Різниця фаз визначається кутом α, який однозначно зв'язаний з положенням точки P у фокальній площині лінзи (рис.12.9).

Згідно формули (12.17) світлі смуги розташовані в місцях, для яких

$$2h\sqrt{n^2 - \sin^2\alpha} \pm \frac{\lambda_0}{2} = 2m\frac{\lambda_0}{2}, \quad (12.29)$$

де m – ціле число, яке називається порядком інтерференції. Смуга, яка відповідає даному порядку інтерференції, обумовлена світлом, яке падає на пластинку під певним кутом α. Тому такі смуги називають *інтерференційними смугами однакового нахилу*. Смуги однакового нахилу можна отримати не тільки у відбитому світлі, але і у світлі, яке проходить через пластинку. В цьому випадку один із променів проходить прямо, а другий – після двох відбивань на внутрішній стороні пластинки.

3. Інтерференція при протяжних джерелах

Ми розглянули інтерференційні досліди, в яких поділ амплітуди світлової хвилі від джерела відбувалось в результаті часткового відбивання на поверхнях плоскопаралельної пластинки. У випадку точкового джерела смуги спостерігаються скрізь, тобто вони не локалізовані. Але на нескінченності або у фокальній площині збиральної лінзи смуги спостерігаються і при протяжному джерелі. Локалізовані смуги при протяжному джерелі можна спостерігати і в інших умовах. Виявляється, що для достатньо тонкої пластинки або плівки (поверхні якої не обов'язково повинні бути паралельними і взагалі плоскими) можна спостерігати інтерференційну картину, яка локалізована біля поверхні, що відбиває світло. В білому світлі інтерференційні смуги забарвлені. Тому таке явище називають *кольори тонких плівок*. Його легко спостерігати на мильних бульбашках, на тонких плівках мастила або бензину, що плавають на поверхні води, на плівках окислів на поверхні металів і плівках лазерних дисків і т.п.

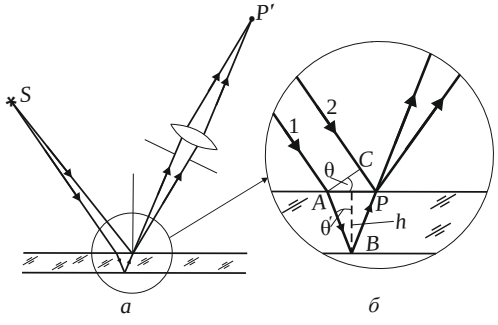


Рис.12.10.

Локалізовані біля поверхні плівки або тонкої пластинки інтерференційні смуги можна спостерігати неозброєним оком або за допомогою лупи або мікроскопа. За допомогою збиральної лінзи інтерференційну картину з поверхні плівки можна відобразити на екрані (рис.12.10,*а*).

Дійсно, промені, які виходять із точки P (рис.12.10,*б*), знову збираються в точці P', яка є зображенням точка P, що створюється лінзою. Інтерферуючі промені прийдуть в точку P' з тією ж різницею фаз, яку вони мали в точці P.

Щоб зрозуміти причину виникнення інтерференційних смуг знайдемо різницю ходу двох променів, які приходять в точку *P* від джерела *S*.

Із рис.12.10,*б* видно, що

$$\Delta = n(|AB| + |BP|) - |CP| \pm \frac{\lambda_0}{2}.$$

Якщо плівка мало вирізняється від плоскопаралельної, трикутники можна вважати прямокутними і вважати, що $|AB| \approx |BP| \approx h/\cos\theta'$, $|CP| \approx 2h\mathrm{tg}\theta' \cdot \sin\theta$. Тут $\theta -$ кут падіння променів на плівку, θ' – кут заломлення променів в точці *A*. За законом заломлення $\sin\theta = n\sin\theta'$. Тоді

$$\Delta = \frac{2nh}{\cos\theta'} - 2h\mathrm{tg}\theta'\sin\theta \pm \frac{\lambda_0}{2} \approx$$
$$\approx 2hn\cos\theta' \pm \frac{\lambda_0}{2},$$

де враховано "втрату півхвилі" при відбиванні променя 2 в точці *P*. Різниця фаз φ двох променів в точці *P* буде:

$$\varphi = \frac{4\pi nh\cos\theta'}{\lambda_0} \pm \pi. \qquad (12.30)$$

Формула (12.30) відповідає іншим умовам інтерференції, ніж формула (12.28). Для даної точки *P* товщина h має певне значення і, якщо використовувати протяжне джерело світла, то різні величини φ (формула (12.30)) для променів, що приходять від різних точок джерела, зв'язані з різним для них значенням $\cos\theta'$. В точці *P* (і, відповідно, в *P*') буде максимум інтенсивності, якщо різниця ходу Δ згідно (12.17) буде:

$$2nh\langle\cos\theta'\rangle \pm \frac{\lambda_0}{2} = m\lambda_0 \quad (m=0,1,2,3,\dots), \qquad (12.31)$$

де $\langle\cos\theta'\rangle$ означає усереднення по тих точках джерела, світло від яких попадає в *P'*. Співвідношення (12.31) є справедливим і при неплоских поверхнях плівки при умові, що кут між ними залишається малим. Тому інтерференційні смуги відповідають сукупності місць плівки, де її оптична товщина має одне і теж значення (при умові, звичайно, що $\langle\cos\theta'\rangle$ є достатньо однаковий для всієї області спостереження). По цій причині такі смуги називають *смугами однакової товщини.*

Кільцеві смуги однакової товщини, які спостерігаються у повітряному зазорі між випуклою сферичною поверхнею лінзи малої кривизни і плоскою поверхнею скла, які дотикаються (рис.12.11), називаються *кільцями Ньютона*. Їх спільний центр розташований в точці дотику. При нормальному падінні світла на лінзу для відбитих променів 1 і 2 (рис.12.11) різниця ходу буде дорівнювати

$$\Delta = 2h + \frac{\lambda_0}{2},$$

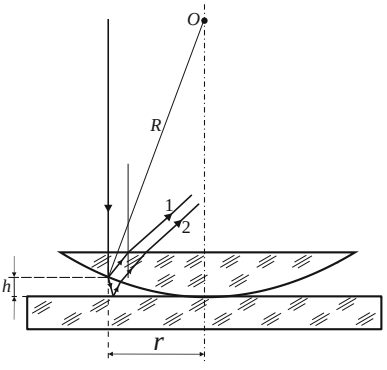


Рис.12.11.

де враховано "втрату півхвилі" променя 2 при відбиванні від поверхні скла. Товщина h повітряного зазору зв'язана із відстанню r до точки дотику (рис.12.11):

$$h = R - \sqrt{R^2 - r^2} =$$
$$= R\left(1 - \sqrt{1 - \frac{r^2}{R^2}}\right) \approx r^2/(2R).$$

Тут використано умову $(r^2/R^2) \ll 1$. Тоді різниця ходу

$$\Delta = r^2/R + \frac{\lambda_0}{2}.$$

Згідно умови (12.17) світлі кільця будуть спостерігатись при умові, що

$$\Delta = \frac{r_{св}^2}{R} + \frac{\lambda_0}{2} = 2m\frac{\lambda_0}{2} \quad (m = 1, 2, 3, \dots),$$

а радіуси світлих кілець будуть:

$$r_{св} = \sqrt{\frac{(2m-1)\lambda_0 R}{2}}. \qquad (12.32)$$

Темні кільця, згідно умови (12.19), будуть при

$$\Delta = \frac{r_{тем}^2}{R} + \frac{\lambda_0}{2} = (2m+1)\frac{\lambda_0}{2}$$

$(m = 0,1,2,3,\dots)$,

а радіуси темних кілець визначаються за формулою

$$r_{тем} = \sqrt{mR\lambda_0}\,. \qquad (12.33)$$

У відбитому світлі центр темний ($r_{тем} = 0$), так як при товщині повітряного зазору $h \ll \lambda_0$, різниця ходу $\Delta \approx \lambda_0/2$, а різниця фаз хвиль, які інтерферують, близька до π. Якщо лінзу поступово відсувати від поверхні скла, то інтерференційні кільця будуть стягуватись до центру. Для світла, яке проходить через лінзу і пластинку, розташування кілець Ньютона обернене до розташування кілець Ньютона у відбитому світлі. Тобто, за формулою (12.32) визначають радіуси темних кілець, а за формулою (12.33) – радіуси світлих кілець.

За допомогою кілець Ньютона, як і в досліді Юнга, можна порівняно простими засобами приблизно визначити довжину хвилі світла.

§12.5. Інтерференція квазімонохроматичного світла. Часова когерентність

У всіх описаних вище інтерференційних дослідах при їхній інтерпретації вважалось, що джерело випромінює монохроматичне світло (світло однієї частоти $\nu(\Delta\nu = 0)$, або однієї довжини хвилі $\lambda\,(\Delta\lambda = 0)$). Природно, що результати отримані для монохроматичного світла, мають обмежене застосування. В цьому параграфі ми з'ясуємо, до яких змін в інтерференційних явищах призводить врахування спектрального складу ($\Delta\lambda \neq 0$) реальних джерел світла.

Розглянемо найпростіший випадок, коли джерело випромінює дві дуже вузькі, близькі одна до одної спектральні лінії з частотами ω_1 і ω_2. Якби випромінювання на кожній із частот представляло собою нескінченну синусоїдальну хвилю, то результуюче випромінювання було б хвилею середньої частоти з амплітудою, що періодично змінювалась би. Але в дійсності випромінювання кожної із спектральних компонентів є хаотична послідовність більш–менш довгих хвильових цугів. Як правило, за час спостереження проходить багато цугів, коливання в яких ніяк не зв'язані за фазою. Тому можна вважати, *що замість одного є два розташованих в одному місці джерела і незалежно одне від одного випромінюють хвилі з частотами* ω_1 і ω_2. При виконанні інтерференційних дослідів з такими джерелами кожна із хвиль створює свою інтерференційну картину, і ці картини просто накладаються одна на одну.

Якщо частоти ω_1 і ω_2 (λ_1 і λ_2) мало відрізняються (*таке світло називають квазімонохроматичним, тобто майже монохроматичним*), то інтерференційні смуги в кожній картині мають майже однакову ширину (формула (12.23)). В тих випадках, де світлі смуги однієї картини накладаються на світлі смуги іншої, чіткість смуг в сумарній картині найбільша. Навпаки, там, де світлі смуги однієї картини співпадають з темними смугами іншої картини, чіткість смуг зменшується майже до повного їх зникнення.

Дослідимо це питання кількісно.

Інтенсивність результуючого коливання в точці P (рис.12.1) екрану визначається формулою (12.13), де різниця фаз двох коливань (див.§10.7 формула (10.32)) буде:

$$\varphi=(\omega_1-\omega_2)t+ +(k_2r_2-k_1r_1)+\varphi_{01}-\varphi_{02}, \quad (12.34)$$

або

$$\varphi = \Delta\omega t + \varphi_{\text{р}} , \quad (12.35)$$

де $\varphi_{\text{р}} = \left(k_2r_2 - k_1r_1\right) + \varphi_{01} - \varphi_{02}$ – різниця фаз двох коливань в точці P , що визначається формулою (12.14). Якщо інтенсивності хвиль з частотами ω_1 і ω_2 однакові ($I_1 = I_2 = I_0$), то формула (12.13) перепишеться так:

$$I_P = 2I_0\left(1 + \cos\left(\Delta\omega t + \varphi_{\text{р}}\right)\right),$$

або

$$I_P = 4I_0\cos^2\left(\frac{\Delta\omega}{2}t + \frac{\varphi_{\text{р}}}{2}\right)\cdot \quad (12.36)$$

Приблизний вид залежності інтенсивності світла в точці від часу показаний на рис.12.12 згідно формули (12.36). Якщо час спостереження в точці значний, то ні максимуму, ні мінімуму незафіксуємо (рис.12.12). Для часу $t_{спостереж} > t_{когер} = \frac{2\pi}{\Delta\omega}$ в точці P будуть пульсації інтенсивності. Якщо час спостереження дорівнює $t_{когер} = 2\pi/\Delta\omega$, то в точці буде або максимум, або мінімум інтерференційної картини. Час $t_{когер}$ називається *часом когерентності*. Це такий проміжок часу, на протязі якого випадкова зміна фази хвилі досягає значення π. Тоді

$$\frac{\Delta\omega}{2}(t+t_{когер}) + \frac{\varphi_{\text{р}}}{2} - \left(\frac{\Delta\omega}{2}t + \frac{\varphi_{\text{р}}}{2}\right) = \pi.$$

Звідси знаходимо, що

$$t_{когер} = 2\pi/\Delta\omega = 1/\Delta\nu . \quad (12.37)$$

Якщо врахувати, що $\nu = c/\lambda$, то $|\Delta\nu| = c\Delta\lambda/\lambda^2$ і

$$t_{когер} = \lambda^2/(c\Delta\lambda). \quad (12.38)$$

Відстань $l_{когер} = ct_{когер}$, на яку переміщується хвиля за час $t_{когер}$ називається *довжиною когерентності* (або довжиною цуга). Довжина когерентності є та відстань, на якій випадкова зміна фази хвилі досягає значення приблизно π. Для отримання

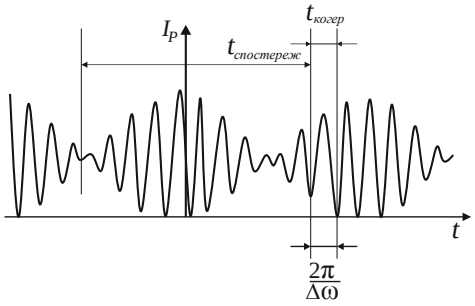


Рис.12.12.

інтерференційної картини методом поділу природної хвилі на дві частини необхідно, щоб оптична різниця ходу Δ була менша, ніж довжина когерентності $l_{когер}$. Ця вимога обмежує число видимих інтерференційних смуг, що спостерігаються за схемою, яка показана на рис.12.2. Із збільшенням номера смуги m різниця ходу зростає, внаслідок чого чіткість смуг (їх видність) стає меншою і меншою.

Довжині когерентності відповідає максимально можливий порядок інтерференції:

$$m_{max}\lambda \approx l_{когер} = \lambda^2/\Delta\lambda .$$

Звідки

$$m_{max} = \lambda/\Delta\lambda . \quad (12.38)$$

Для білого світла (сонце, лампа розжарювання, дуга з вугільними електродами) ефективним діапазоном довжин хвиль для візуального спостереження є приблизно довжини хвиль від 0,4 до 0,7 мкм, тобто $\Delta\lambda \sim \lambda$. В цьому випадку $m_{max} \sim 1$ і інтерференційні смуги, здавалось би, спостерігатись не повинні. Дійсно, приймач випромінювання, який має

однакову чутливість на різних ділянках спектру, наприклад термоелемент, покаже при переміщенні в полі зору поперек смуг майже рівномірний розподіл освітленості. Але око є селективний приймач із сильно змінною чутливістю в залежності від довжини хвилі, що дає деяким хвилям перевагу перед іншими. Візуальне спостереження смуг в білому світлі полегшується, дякуючи здатності нашого зору розрізняти кольори, а не тільки інтенсивність світла. Тому в білому світлі око розрізняє біля десятка забарвлень інтерференційних смуг. Коли різниця ходу дорівнює нулю для деяких місць, куди обидві інтерферуючі хвилі приходять в однаковій фазі, умова максимуму виконується для всіх довжин хвиль. В цьому місці виникає ахроматична (тобто незабарвлена) світла смуга. По обидва боки від неї знаходяться забарвлені максимуми і мінімуми, а за ними поле зору здається оку рівномірно освітленим білим світлом. Таке походження чудових інтерфереційних кольорів в тонких плівках мастила або бензину на поверхні води. Легко оцінити максимальну товщину плівки, при якій можливе візуальне спостереження інтерференції в білому світлі. Приймаючи $m_{max} \approx 10$ для максимальної можливої різниці ходу отримуємо $\Delta_{max} \sim 10\langle\lambda\rangle$, (де $\langle\lambda\rangle = 0{,}5$ мкм). При цьому товщина плівки вдвічі менша: $h_{max} \sim 5\langle\lambda\rangle \sim$ 2,5мкм.

§12.6. Роль кінцевих розмірів джерела світла. Просторова когерентність

При аналізі інтерференційних дослідів первинне джерело хвиль вважалось точковим. Однак всі реальні джерела світла мають кінцеві розміри. Збільшення розмірів джерела, як і розширення спектру світла, яке випромінюється, приводить до погіршення контрастності (видності) інтерференційних смуг і навіть до їх повного зникнення. Щоб вияснити роль розмірів джерела будемо вважати тут випромінювання монохроматичним ($\Delta\lambda = 0$).

Протяжне джерело, яке світиться, складається із великого числа точкових взаємно некогерентних елементів. Тому інтенсивність в любому місці дорівнює сумі інтенсивностей в інтерференційних картинах, що створюються окремими точковими елементами джерела.

В інтерференційних дослідах за методом поділу хвильового фронту (§12.4) смуги на екрані перпендикулярні площині, в якій знаходяться первинне точкове джерело S і вторинні джерела S_1 і S_2. Використання замість S лінійного джерела (достатньо вузької щілини, яка витягнута перпендикулярно до цієї площини), збільшує інтенсивність і не приводить до погіршення чіткості інтерференційних смуг. Однак збільшення ширини щілини приводить до того, що смуги стають менш чіткими або пропадають зовсім.

Якщо розміри джерела (тобто ширина щілини S) набагато менша довжини світлової хвилі, то інтерференційна картина буде різкою, оскільки різниця ходу променів, що інтерферують, від любої точки джерела (по його ширині) до деякої точки спостереження P буде практично одна і таж. Але, як правило, розміри джерела значно більші довжини хвилі, тому однакові інтерференційні картини від різних його елементів зсунуті одна відносно одної. В результаті накладання цих картин інтерференційні смуги стають більш менш розмитими. Їх можна спостерігати тільки при використанні певних умов, які накладаються на геометрію експерименту. Ці умови розглянемо далі.

Розглянемо протяжне джерело світла, розміри якого D (це ширина щілини). Кожна точка такого джерела є точковим джерелом світла. Розглянемо дві такі точки S_1 і S_2 (рис.12.13,*а*), які є некогерентними світними точками. В інтерференційних дослідах світло від кожного джерела попадає в деяку точку спостереження на екрані двома різними шляхами. Розглянемо випадок, коли промені,

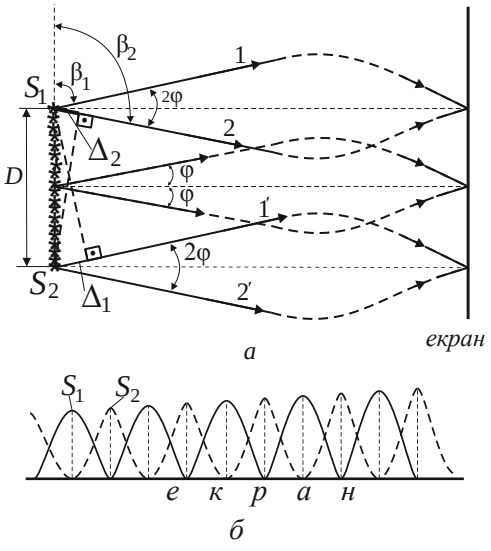


Рис.12.13.

що інтерферують, виходять із кожної точки джерела симетрично відносно нормалі до лінії S_1S_2. Нехай промінь, який іде від S_1 по одному із цих шляхів, утворює кут β_1 (промінь 1) із лінією $S_1\,S_2$, а іншим шляхом – кут β_2 (промінь 2). Для джерела S_2 відповідні промені $1'$і $2'$. Зсув двох інтерференційних картин на екрані, що створюються джерелами S_1 і S_2 визначається сумою різниць ходу Δ_1 променів 1 і $1'$ і Δ_2 променів 2 і $2'$. $\Delta_1 = D\sin\left(\frac{\pi}{2} - \beta_1\right) = D\cos\beta_1$, $\Delta_2 = D\sin\left(\beta_2 - \frac{\pi}{2}\right) = -D\cos\beta_2$. Тоді $\Delta_1 + \Delta_2 = D(\cos\beta_1 - \cos\beta_2)$. Для симетричного випадку (рис.12.13,*а*) кут $\beta_1 = \frac{\pi}{2} - \varphi$, кут $\beta_2 = \frac{\pi}{2} + \varphi$. Тоді $\Delta_1 + \Delta_2 = 2D\sin\varphi$. Якщо $\Delta_1 + \Delta_2 = 0$, то максимуми однієї картини співпадають з максимумами іншої. При $\Delta_1 + \Delta_2 = \frac{\lambda}{2}$ світлі смуги інтерференційної картини від джерела S_1 суміщаються із темними смугами інтерференційної картини від джерела S_2 (рис.12.13,*б*). Тому умовою отримання хорошої інтерференційної картини від протяжного джерела можна прийняти нерівність

$$\Delta_1 + \Delta_2 = 2D\sin\varphi \le \frac{\lambda}{2}\cdot \quad (12.39)$$

Для випадку малого кута $\sin\varphi \approx \varphi$ і (12.39) можна переписати так:

$$D \le \lambda/(2\cdot 2\varphi)\,, \quad (12.40)$$

де 2φ – кут між інтерферуючими променями, що виходять із джерела, називається *апертурою інтерференції*.

При великих апертурах спостерігати інтерференцію можна тільки від джерел, розміри яких менші довжини світлової хвилі. Якщо $\varphi \cong \pi/2$, тобто промені, що інтерферують, виходять із джерела майже в протилежних напрямках, то із (12.40) отримуємо, що його протяжність D повинна бути менша $\lambda/4$. Для спостереження інтерференції з використанням джерела, розміри якого набагато більші довжини хвилі світла, геометрія експерименту повинна бути такою, щоб промені, які інтерферують, виходили із джерела під малим кутом один до одного.

Застосуємо критерій (12.40) до інтерференційних дослідів, які розглядались вище. В досліді Юнга (рис.12.4) $2\varphi \approx \frac{d}{L}$, тому ширина додаткової щілини S повинна задовільняти умові $D \le \frac{\lambda L}{2d}$. При $\lambda = 5\cdot 10^{-7}$м, $L = 1$м, d = 0,5мм знаходимо, що ширина щілини повинна бути менша 0,5 мм.

Розглянемо дослід із дзеркалами Френеля (рис.12.5). Для точки спостереження

P, яка лежить в центрі інтерференційного поля кут 2φ між інтерферуючими променями, які виходять із джерела S, легко знайти із побудови, що приведена на рис.12.14. Кут α

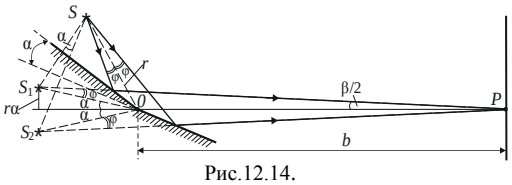


Рис.12.14.

(кут між дзеркалами) є зовнішнім для трикутника S_1OP і тому $\alpha = \varphi + \beta/2$. Для половини кута сходження променів можемо записати $\beta/2 = r\alpha/(r+b)$. Виключаючи із цих рівнянь $\beta/2$, знаходимо, що $\varphi = \alpha b/(r+b)$.

Підставимо φ в (12.40) і отримаємо наступне обмеження на ширину D щілини джерела S:

$$D \le \lambda(r+b)/(4b\alpha).$$

Якщо $b \gg r$, то ця умова приймає вид $D \le \lambda/(4\alpha)$. Щоб можна було спостерігати смуги інтерференції із джерелом, для якого $D \gg \lambda$, кут між дзеркалами повинен бути дуже малим ($\alpha \ll 1$). Можна показати, що в дослідах із дзеркалами Френеля апертура інтерференції 2φ має практично одне і те саме значення при будь-якому положенні точки спостереження P на екрані в області, де перекриваються інтерферуючі промені. Тому чіткість смуг однакова по всьому інтерференційному полю.

В досліді з біпризмою Френеля (рис.12.6.) апертура інтерференції також практично однакова по всьому полю і дорівнює $2\varphi = \alpha(n-1)b/(r+b) \approx 2\alpha(n-1)$ (останнє справедливе при $r \ll b$). Підставляючі 2φ у (12.40), знаходимо, що $D \le \lambda/(4\alpha(n-1))$.

При аналізі інтерференційного досліду Поля (рис.12.8) для спрощення будемо нехтувати заломленням в слюді, тобто замінимо пластинку двома паралельними площинами, які відбивають світло, відстань між якими дорівнює товщині h пластинки. Тоді відстань між вторинними джерелами S_1 і S_2 (уявними зображеннями джерела S) дорівнює $2h$. Кут сходження променів в точку спостереження в цьому досліді дорівнює апертурі інтерференції 2φ. Відстань $|S_1P|$ дорівнює $\frac{a+b}{\cos\theta}$, тому $2\varphi = \frac{2h\sin\theta}{|S_1P|} = \frac{h\sin 2\theta}{a+b}$. Товщина h листочка слюди дуже мала ($\approx 0{,}05$ мм) порівняно з $a+b \approx 5$м, тому мала і апертура інтерференції 2φ (при будь-якому положенні точки спостереження P, включаючи $\theta = 45°$). Значить, розмір джерела S може бути достатньо великим: із (12.40) знаходимо $D \le \frac{\lambda(a+b)}{2h\sin 2\theta}$. При $\lambda = 5\cdot 10^{-7}$м, $h = 0{,}05$мм, $a+b = 5$м і $\theta = 45°$ отримаємо, що $D \ll 25$мм. Для демонстрації досліду Поля можна використати велику лампу без всяких щілин. За допомогою листочка слюди, площа якого декілька квадратних сантиметрів, можна отримати яскраву інтерференційну картину великих розмірів, що покриває стелю і стіни аудиторії. Розмір джерела (≈ 10мм) набагато більший відстані $|S_1S_2|(2h \approx 0{,}1$мм), так що уявні зображення джерела майже повністю перекриваються.

При спостереженні локалізованих на нескінченності смуг однакового нахилу (рис.12.9) обидва промені, які інтерферують, виходять із джерела в одному напрямку, тобто апертура інтерференції дорівнює нулю. Тому тут немає ніяких обмежень на розміри джерела.

Залежність видності (12.25) смуг від апертури інтерференції можна наглядно продемонструвати на досліді із дзеркалом Ллойда (рис.12.15). Тут прямий пучок світла від джерела S інтерферує з пучком, який відбитий від дзеркала при майже ковзаючому

падінні. На відміну від дослідів Юнга або Френеля, в досліді Ллойда апертура інтерференції сильно залежить від положення точки спостереження P на екрані, який встановлений перпендикулярно до площини дзеркала. Із (рис.12.15) видно, що $2\varphi \approx 2x/l$. Апертура тим менша, чим ближче точка P до площини дзеркала. При використанні протяжного джерела S видність смуг замітно спадає при віддаленні від точки $x = 0$.

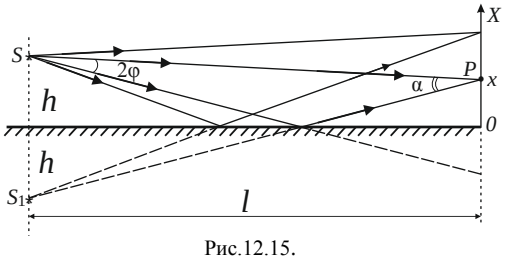


Рис.12.15.

Критерій (12.40) дозволяє визначити розмір області в межах якої видність смуг (12.25) більша 2/3 при данній ширині D щілини джерела S: $x \le \lambda l/(4D)$. На вдвічі більшій відстані $\lambda l/(2D)$ смуги пропадають. Кут сходження променів, як видно із рис.12.15, дорівнює $\alpha \approx 2h/l$. Ширина інтерференційної смуги $\Delta x \approx \lambda/\alpha \approx \lambda l/(2h)$. Тому повне число смуг на цій відстані буде $N = 2x/\Delta x = h/D$. Це дає оцінку найвищого порядку інтерференції в монохроматичному світлі від протяжного джерела в досліді Ллойда.

§12.7. Практичне застосування інтерференції світла

Явище інтерференції світла використовують для створення різних оптичних приладів, за допомогою яких можна виконувати такі операції:

−вимірювати довжини світлової хвилі з великою точністю;

−вимірювати лінійні розміри тіл і коефіцієнти лінійного розширення тіл;

−визначати якість і точність шліфування різних поверхонь;

−вимірювати показники заломлення газоподібних, рідких і твердих тіл;

−вимірювати кутові розміри небесних тіл;

−вивчати і контролювати однорідність речовин;

−вивчати структури спектрів різних речовин;

−досліджувати ударні хвилі у газах.

За допомогою явища інтерференції можна виготовляти інтерференційні світлофільтри і вдається значно зменшити відбивання світла від оптичних систем.

Розглянемо більш детально деякі практичні застосування явища інтерференції.

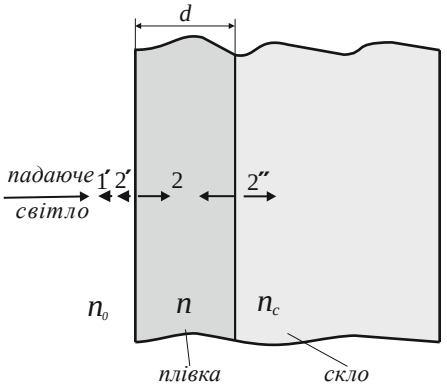


Рис.12.16.

О.Смакула розробив спосіб просвітлення оптики для зменшення втрат світла, зумовлених його відбиванням на межі поділу двох середовищ, оскільки проходження світла через кожну заломну поверхню лінзи супроводжується відбиттям $\approx 4\%$ падаючого потоку світла. У складних об'єктивах число відбивань велике і сумарна втрата світлового потоку досить значна, що приводить до зменшення світлосили оптичних приладів. Крім

того, відбивання від поверхні лінз приводить до виникнення бліків.

Щоб елементи оптичних систем зробити просвітленими, їх поверхні покривають прозорими плівками речовин, показник заломлення яких менший, ніж скла. При відбиванні світла від межі поділу повітря – плівка і плівка – скло виникає інтерференція когерентних хвиль $1'$ і $2'$ (рис.12.16.). Товщину плівки d і показники заломлення скла n_c і плівки n можна підібрати так, щоб хвилі, які інтерферують, гасила одна одну. Для цього їх амплітуди повинні бути рівні, а оптична різниця ходу $\Delta = (2m+1)\frac{\lambda_0}{2}\cdot$

Розрахунки показують, що амплітуди відбитих хвиль рівні, якщо $n = \sqrt{n_c}$. Речовину плівки підбирають так, щоб виконувалась умова $n_c > n > n_0$, де n_0 – показник заломлення повітря. Втрата півхвилі в цьому випадку відбувається на обох поверхнях. Умова мінімуму при куті падіння $\alpha = 0$ має такий вигляд

$$2dn = (2m+1)\frac{\lambda_0}{2}\cdot$$

Для мінімальної товщини плівки $m = 0$ і тоді

$$dn = \frac{\lambda_0}{4}\cdot$$

Ця формула показує, що неможливо добитися одночасного гасіння для всіх довжин хвиль видимого спектра. Тому ця умова повинна виконуватись для довжини хвиль λ_0=0,555мкм, яка найбільше сприймається людським оком.

Отже якщо $n = \sqrt{n_c}$ і оптична товщина плівки дорівнює $\lambda_0/4$, то в результаті інтерференції спостерігається гасіння відбитих хвиль.

Явище інтерференції використовують у ряді дуже точних вимірювальних приладів, які називають **інтерферометрами.**

Розглянемо схему **інтерферометра Жамена**, який використовують для точних вимірювання показників заломлення газів і їх залежності від температури, тиску і вологості (рис.12.17). Дві зовсім однакові скляні товсті плоскопаралельні пластини P_1 та P_2 установлені майже паралельно одна до одної. Поверхні *AB* і *CD* – дзеркальні. Промені світла довжиною λ_0 від монохроматичного джерела S падають на поверхню пластини P_1 під різними кутами α близькими до $45°$. На рисунку показано лише один падаючий промінь. Внаслідок його відбивання від

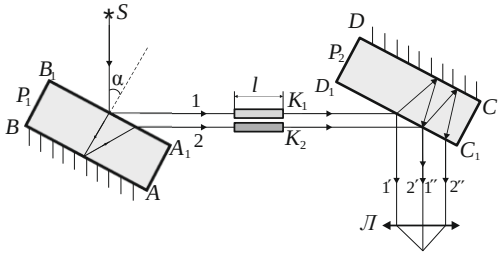


Рис.12.17.

поверхонь AB і A_1B_1 пластини P_1 з неї виходять дві когерентні паралельні хвилі 1 і 2. Пройшовши крізь дві цілком однакові закриті скляні кювети K_1 і K_2 довжиною l, ці хвилі відбиваються поверхнями CD і C_1D_1 пластини P_2. В результаті виникають чотири промені $1'$ і $1''$, $2'$ і $2''$, з яких промені $1''$ і $2'$ накладаються один на одного і інтерферують.

Якщо одну з кювет K_1 заповнено газом з відомим показником заломлення n, а другу – газом з показником заломлення n_x, який треба знайти, то оптична різниця ходу між інтерферуючими хвилями $\Delta = (n_x - n)l$. Якщо ця різниця ходу $\Delta = 2m\frac{\lambda_0}{2}$, то вся інтерференційна картина зміститься на m смуг.

Визначивши m, можна знайти показник заломлення n_x:

$$n_x = n + \frac{m\lambda_0}{l}.$$

Внаслідок великої чутливості інтерферометра, його використовують для вимірювання показників заломлення газів, що мало відрізняються від одиниці.

Розглянемо спрощену схему інтерферометра Майкельсона (рис.12.18). Монохроматичне світло від джерела S падає під кутом $45°$ на плоскопаралельну пластину P_1. Сторона пластинки AB, яка посріблена і напівпрозора, розділяє промінь на дві частини: промінь 1, який відбивається від посрібленого шару і промінь 2, що проходить крізь нього. Промінь 1 відбивається від дзеркала M_1 і повертається назад, знову проходить через пластинку $P_1(1')$. Промінь 2 поширюється до дзеркала M_2, відбивається від нього, повертається назад і відбивається від пластинки

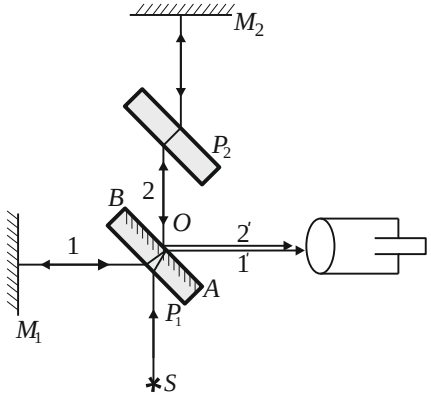


Рис.12.18.

$P_1(2')$. Оскільки промінь 1 проходить пластину P_1 двічі, то для компенсації різниці ходу, яка виникає на шляху другого променя, ставиться пластинка P_2, яку називають компенсатором.

Дзеркала M_1 і M_2 розміщують перпендикулярно одне до одного. Одне з них може переміщатися за допомогою мікрометричних гвинтів паралельно само до себе.

Хвилі $1'$ і $2'$ – когерентні: отже, за допомогою зорової труби буде спостерігатися інтерференційна картина. Оптична різниця хвиль $1'$ і $2'$ $\Delta = 2n_0(l_1 - l_2)$, де n_0- абсолютний показник заломлення повітря, а l_1 і l_2 - відстані від точки O до дзеркал M_1 і M_2.

Якщо $l_1 = l_2$, то буде інтерференційний максимум. Переміщення одного з дзеркал на відстані $\lambda_0/4$ призводить до виникнення інтерференційного мінімуму. Отже, за зміною інтерференційної картини можна зробити висновки про малі переміщення одного з дзеркал і тим самим використати інтерферометр Майкельсона для точних вимірювань довжини світлової хвилі, довжини тіла при зміні температури і т.п. Похибки при вимірюванні довжини складають порядка 10^{-8}мм.

Інтерферометр Майкельсона можна використати для знаходження незначних змін показника заломлення прозорих тіл залежно від тиску, температури, домішок. Такий інтерферометр називається *інтерференційним рефрактометром.*

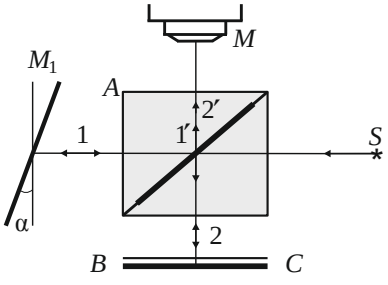


Рис.12.19.

В.П.Лінник використав принцип дії інтерферометра Майкельсона для створення *мікроінтерферометра* – високочутливого приладу, який використовується для контролю

чистоти обробки поверхонь металевих виробів. Світло від джерела S падає на скляний кубик A (рис.12.19), який складається з двох половин склеєних по діагональній площині. Одну із діагональних поверхонь напівпосріблено так, що вона частково пропускає, а частково відбиває промені, які падають на неї. Промінь 1, який пройшов через кубик, попадає на дзеркало M_1, відбивається і, пройшовши знову через кубик, відбивається від діагональної площини і попадає в мікроскоп M. Промінь 2 відбивається від напівпрозорої діагональної пощини, падає на поверхню BC, яка досліджується, відбивається від неї і проходить через кубик в мікроскоп.

Дзеркало M_1 нахилено під невеликим кутом α. Завдяки цьому при ідеальній гладкій поверхні BC в полі зору мікроскопа видно інтерференційні смуги однакової товщини повітряного проміжку між кубиком і дзеркалом. У тих місцях поверхні BC, де є виступи або заглибини, інтерференційні смуги викривлятимуться. За допомогою цього приладу можна виявити штрихи на поверхні деталі, глибина яких дорівнює $(2\div3)\cdot10^{-7}$мм.

Контрольні питання

1.Дати визначення оптичної довжини шляху та різниці ходу променів.

2.В чому суть явища інтерференції?

3.При яких умовах можливе явище інтерференції?

4.Які хвилі є когерентними?

5.Записати та пояснити умову максимального підсилення та послаблення електромагнітних хвиль внаслідок інтерференції.

6.Дати визначення ширини інтерференційної смуги. Від яких величин залежить ширина інтерференційної смуги?

7.Які відомі способи отримання когерентних хвиль оптичного діапазону? Пояснити суть цих способів.

8.Пояснити явище «втрати півхвилі» при відбиванні хвиль від перешкоди.

9.Які є характеристики когерентних властивостей хвиль? Дати їх визначення.

10.Як впливають немонохроматичність світла та розміри джерела світла на інтерференційну картину?

11.Де зустрічається явище інтерференції світла в побуті та яке її практичне застосування?

Розділ 13. Дифракція електромагнітних хвиль

§13.1. Дифракція хвиль

Дифракція хвиль *(від латинського* diffractus *– розламаний, заломлений), в початковому вузькому смислі – огинання хвилями перешкод, в сучасному більш широкому – будь-яке відхилення від законів геометричної оптики.* При такому загальному тлумаченні дифракція хвиль переплітається з явищами розповсюдження і розсіяння хвиль в неоднорідних середовищах. Внаслідок дифракції хвилі можуть попадати в область геометричної тіні: огинати перешкоди, стелитись вздовж поверхонь, проникати через невеликі отвори в екранах і т.п. Наприклад, звук можна почути за кутом будинку або радіохвиля може проникати за горизонт навіть без відбивання від іоносфери. В явищах дифракції світла, як і в інтерференції, на перший план виступають хвильові властивості світла. Для спостереження дифракції світлових

хвиль необхідно створити спеціальні умови. Це зумовлено малою довжиною світлових хвиль.

Між інтерференцією і дифракцією немає суттєвої фізичної різниці. Суть обох явищ в перерозподілі світлового потоку в результаті суперпозиції хвиль. Через історичні причини перерозподіл інтенсивності, що виникає в результаті суперпозиції хвиль, що збуджується кінцевим числом дискретних когерентних хвиль, прийнято називати інтерференцією хвиль. ***Перерозподіл інтенсивності, що виникає внаслідок суперпозиції хвиль, які збуджуються когерентними джерелами, що розташовані неперервно, прийнято називати дифракцією хвиль.*** Тому і говорять, наприклад, про інтерференційну картину від двох вузьких щілин і про дифракційну картину від однієї щілини.

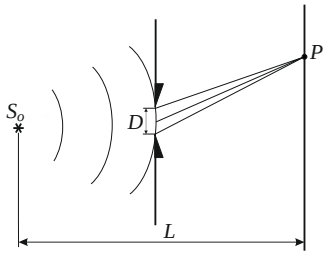


Рис.13.1.

Спостереження дифракції здійснюється, як правило, за такою схемою (рис.13.1). На шляху світлової хвилі, що розповсюджується від деякого джерела S_0, розміщують непрозорий екран з малим отвором, розміри якого D (наприклад, порядку або менші довжини хвилі λ). За перешкодою розташований екран, на якому виникає дифракційна картина. Структура дифракційного поля на екрані суттєво залежить від відстані L між джерелом S_0 і точкою спостереження P. Розрізняють ***дифракцію Френеля,*** якщо $L \sim D^2/\lambda$ і ***дифракцію Фраунгофера*** при $L \gg D^2/\lambda$. Дифракцію Фраунгофера можна спостерігати, якщо між S_0 і отвором, а також між отвором і екраном розмістити по збиральній лінзі так, щоб точка S_0 і P опинились у фокальній площині відповідної лінзи.

§13.2.Принцип Гюйгенса–Френеля

Перше пояснення дифракції світла було дане Френелем в 1818 р. У своєму мемуарі він показав, що кількісне описання дифракційних явищ можливе на основі принципу Гюйгенса, якщо його доповнити принципом інтерференції вторинних хвиль. Кірхгоф в 1882 р. дав строге математичне обґрунтування принципу Гюйгенса – Френеля.

В рамках електромагнітної теорії світла для описання дифракційних явищ не потрібно вводити які-небудь нові принципи. Але точний розв'язок задачі про розповсюдження світла на основі рівнянь Максвелла із відповідними граничними умовами містить великі математичні складнощі. В більшості випадків, які представляють практичний інтерес, наближений метод розв'язання задачі про розподіл світла біля границі між світлом і тінню, який оснований на принципі Гюйгенса – Френеля, є достатнім.

Згідно ***принципу Гюйгенса***, кожну точку, до якої прийшла хвиля від джерела, можна прийняти за центр вторинних хвиль, що розповсюджуються у всі сторони. Результуюча хвиля розглядається як накладання вторинних хвиль. Гюйгенс вважав, що окремі вторинні хвилі не мають періодичності, що вони дуже слабкі і помітну дію виконує тільки їх обвідна. При такому довільному припущенні принцип Гюйгенса дає тільки деякий рецепт побудови

хвильових фронтів, тобто поверхонь, до яких дійшло світлове збурення. Побудови Гюйгенса наглядно пояснюють закони прямолінійного розповсюдження, відбивання і заломлення світла. Але в цих побудовах не використовується поняття довжини хвилі, тому вони не дозволяють визначити умови застосування згаданих законів.

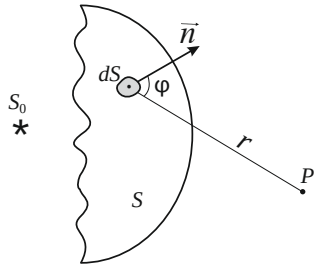


Рис.13.2.

Френель вклав в принцип Гюйгенса ясний фізичний зміст, відказавшись від штучного припущення про обвідну вторинних хвиль, і розглядав повне світлове поле як результат ***інтерференції вторинних хвиль.*** При цьому не тільки отримав фізичне пояснення рецепт Гюйгенса (до точок на обвідній всі вторинні хвилі приходять в однаковій фазі), але і з'явилась можливість розрахунку розподілу світлового поля в просторі. Вивчаючи розподіл світла біля межі між світлом і тінню на основі принципу Гюйгенса – Френеля, можна отримати кількісний опис дифракційних явищ.

Принцип Гюйгенса – Френеля можна виразити такими положеннями:

1) під час розрахунку амплітуди світлових коливань, що збуджуються джерелом S_0 в довільній точці P, джерело S_0 можна замінити еквівалентною йому системою вторинних джерел – малих ділянок dS будь-якої замкненої допоміжної поверхні S проведеної так, щоб вона охоплювало джерело S_0 і не охоплювала розглядувану точку P (рис.13.2);

2) вторинні джерела, які еквівалентні джерелу S_0, когерентні між собою, тому вторинні хвилі, збуджені ними, інтерферують і розрахунок інтерференції найбільш простий, якщо S – хвильова поверхня для світла джерела S_0, оскільки при цьому фази коливань всіх вторинних джерел однакові;

3) амплітуда dE_p коливань, що збуджуються в точці P вторинним джерелом, пропорційна до площі відповідної ділянки хвильової поверхні, обернено пропорційна до відстані r від неї до точки P і залежить від кута φ між зовнішньою нормаллю $\vec{n}$ до хвильової поверхні і напрямком від елемента dS до точки P (рис.13.2):

$$dE_p = K(\varphi)\frac{a_0}{r}dS\cos(\omega t - kr + \alpha_0)\,, \quad (13.1)$$

де $(\omega t + \alpha_0)$ – фаза коливань вторинних джерел dS хвильової поверхні, a_0 – величина, яка пропорційна амплітуді первинних хвиль одиниці хвильової поверхні в точках елемента поверхні *dS; K(φ)* коефіцієнт, який монотонно спадає від 1 при φ = 0 до 0 при $\varphi \geq \pi/2$ (вторинні джерела не випромінюють назад); кут φ називається кутом дифракції. Багато практично важливих задач дифракції можна розв'язати при цих досить загальних припущеннях відносно *K(φ),* не уточнюючи конкретної залежності *K* від φ.

4) Якщо частина поверхні S закрита непрозорими екранами, то вона не випромінює енергію, а інші випромінюють так само, як і за відсутності екранів. Врахування амплітуд і фаз вторинних хвиль дозволяє в кожному конкретному випадку знайти амплітуду результуючої хвилі в довільній точці простору, тобто визначити закономірності

поширення світла. В загальному випадку розрахунок інтерференції вторинних хвиль досить складний і громіздкий, однак для ряду випадків знаходження амплітуди результуючого коливання здійснюється алгебраїчним або геометричним додаванням. Результуюче коливання в точці P є суперпозицією коливань dE_p, взятих для всієї хвильової поверхні S:

$$E_p = \int_S K(\varphi)\frac{a_0}{r}\cos(\omega t - kr + \alpha_0)dS. \quad (13.2)$$

Ця формула є аналітичним виразом принципу Гюйгенса-Френеля.

§13.3.Метод зон Френеля. Прямолінійне поширення світла

За допомогою принципу Гюйгенса-Френеля можна обгрунтувати з хвильових властивостей світла *закон прямолінійного поширення світла в однорідному середовищі.* Френель розв'язав цю задачу, розглянувши взаємну інтерференцію вторинних хвиль і застосував прийом, який отримав назву ***методу зон Френеля.*** Щоб зрозуміти суть методу Френеля, визначимо амплітуду світлового коливання, яке збуджується в точці P сферичною хвилею, що розповсюджується в однорідному середовищі від точкового джерела S_0 (рис.13.3).

Згідно з принципом Гюйгенса – Френеля замінимо дію джерело S_0 дією уявних джерел, які розміщені на допоміжній поверхні S, що є однією з хвильових поверхонь хвилі, які поширюються від джерела S_0.

Ця допоміжна поверхня є поверхнею сфери з центром в S_0. Френель розбив хвильову поверхню S на кільцеві зони такого розміру, щоб відстані від країв зони до точки P відрізнялись на $\lambda/2$, тобто

$$PM_1 - PO = PM_2 - PM_1 = PM_3 - PM_2 = \cdots = \lambda/2.$$

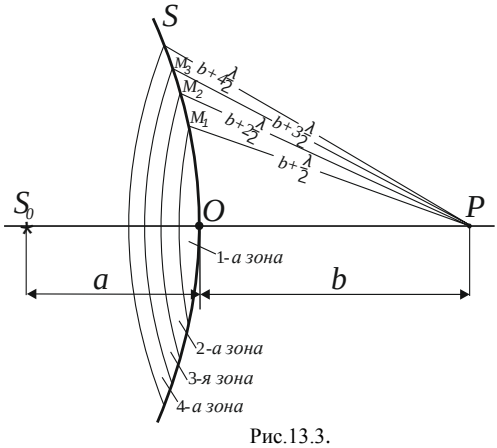


Рис.13.3.

Подібне розбивання хвильової поверхні S на зони можна виконати, провівши з точки P концентричні сфери, радіуси яких $b + \lambda/2$; $b + 2\lambda/2$; $b + 3\lambda/2$; … $b + m\lambda/2$, де $b=PO$. Точки сфери S, що лежать від точки P на відстанях $b + \lambda/2$; $b + 2\lambda/2$; $b + 3\lambda/2$ і т.д. утворюють межі 1-ї, 2-ї, 3-ї і т.д. зон Френеля. Оскільки коливання від сусідніх зон приходять до точки P відстані, які відрізняються на $\lambda/2$, то в точку P вони приходять з протилежними фазами і при накладанні ці коливання будуть взаємно ослаблюватися. Тому амплітуда результуючого коливання в точці P буде

$$E_P = E_1 - E_2 + E_3 - E_4 + \cdots \pm E_m, \quad (13.3)$$

де $E_1, E_2, \ldots, E_m$ – амплітуди коливань в точці P, що збуджуються 1–ю, 2–ю, …, m–ю зонами. В цей вираз всі амплітуди коливань від непарних зон входять зі знаком «+», а від парних зон – зі знаком «–».

Величина E_m залежить від площі σ_m $m-$ї зони, відстані $b + \frac{m\lambda}{2}$ і кута φ_m між зовнішньою нормаллю до поверхні зони в якій-небудь її точці і прямою, яка

напрямлена з цієї точки в точку P. На рис.13.4 точки B і B_1 відповідають зовнішній межі

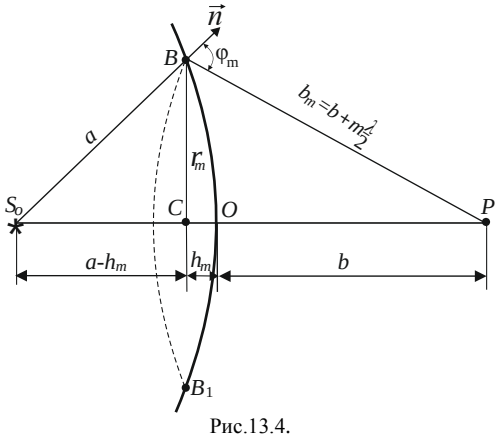


Рис.13.4.

m−ї зони: $BC = r_m$ − радіус зовнішньої границі m−ї зони. $CO = h_m$ − висота кульового сегмента BOB_1. З трикутників S_0BC і PBC видно, що

$$r_m^2 = a^2 - (a-h_m)^2 = \\ = (b+m\,\lambda/2)^2 - (b+h_m)^2. \qquad (13.4)$$

Звідки

$$a^2 - a^2 + 2ah_m - h_m{}^2 = b^2 + \\ +2bm\,\lambda/2 + (m\,\lambda/2)^2 - b^2 - 2bh_m - h_m^2.$$

Тоді

$$2(a+b)h_m = m\lambda b + (m\,\lambda/2)^2\,.$$

Оскільки $\lambda \ll b$, то при не дуже великих m доданком $(m\,\lambda/2)^2$ можна знехтувати і

$$h_m = \frac{m\lambda b}{2(a+b)}\,. \qquad (13.5)$$

Бічна поверхня кульового сегмента BOB_1, яка є сумою площ усіх m зон, починаючи з першої, дорівнює $S_m = \sigma_1+\sigma_2+\ldots+\sigma_m = 2\pi a h_m = \frac{\pi ab\lambda m}{a+b}$, а площа m−ої зони Френеля

$$\sigma_m = S_m - S_{m-1} = \\ = \frac{\pi ab\lambda}{a+b}m - \frac{\pi ab\lambda}{a+b}(m-1).$$

Звідки отримаємо, що

$$\sigma_m = \frac{\pi ab\lambda}{a+b}\,. \qquad (13.6)$$

Цей вираз не залежить від m, отже, при не дуже великих m площі зон Френеля однакові. У такий спосіб побудова зон Френеля розбиває поверхню сферичної хвилі на рівні за площею зони. Відстань $b_m = b + m\lambda/2$ від зони до точки P повільно зростає, а це значить, що згідно (13.1) амплітуда коливань, збуджених в точці P зонами Френеля, зменшується із зростанням номера зони m. Зростання m супроводжується зростанням кута φ_m (рис.13.4) і зменшенням коефіцієнта $K(\varphi)$ у формулі (13.1). Таким чином, амплітуди коливань, які збуджуються в точці P зонами Френеля, утворюють монотонну спадну послідовність :

$$E_1 > E_2 > E_3 > \cdots > E_{m-1} > E_m > E_{m+1} > \cdots.$$

Загальне число N зон Френеля, які вміщуються на частині сфери, яка повернена до точки P (рис.13.5), дуже велике.

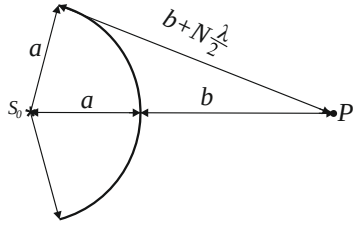


Рис. 13.5.

З рис.13.5 видно, що

$$b + N\lambda/2 = \sqrt{(b+a)^2 - a^2}.$$

Звідси:

$$N = \frac{2}{\lambda}\left(\sqrt{b^2 + 2ba} - b\right).$$

Якщо $a = b = 1$ м і $\lambda = 5 \cdot 10^{-7}$м, то $N \sim 3 \cdot 10^6$. Тому можна вважати, що в межах не дуже великих змін m, залежність величини E_m від m є лінійною і амплітуда коливань, яка викликана якою-небудь m−ю зоною, дорівнює півсумі амплітуд коливань, що викликані m-1−ю і m+1−ю зонами. Тобто

$$E_m = (E_{m-1} + E_{m+1})/2. \qquad (13.7)$$

Тоді амплітуда результуючого коливання в точці *P* буде:

$$E_p = E_1/2 + (E_1/2 - E_2 + E_3/2) + \\ +(E_3/2 - E_4 + E_5/2\ ) + \cdots = E_1/2,$$

оскільки усі вирази, що стоять у дужках, дорівнюють нулю. Отже, амплітуда коливань, що створюється в точці *P* сферичною хвильовою поверхнею, дорівнює половині амплітуди коливань, що створюється однією центральною зоною. Дія всієї хвильової поверхні в точці *P* зводиться до дії її малої ділянки, меншої, ніж центральна зона (1-ша зона на рис.13.3).

Якщо у виразі (13.4) покладемо, що висота сегмента $h_m \ll a$ (при не дуже великих *m*), тоді $r^2{}_m = 2ah_m$ і з урахуванням (13.5) радіус зовнішньої границі *m*-ї зони Френеля

$$r_m = \sqrt{2ah_m} = \sqrt{\frac{m\lambda ab}{a+b}}\ . \qquad (13.8)$$

При $a = b = 1$ м і $\lambda = 0{,}5$ мкм, $r_1 = 0{,}5$ мм, а згідно (13.6) $\sigma_1 = 0{,}8$ мм^2. Отже, поширення світла від S_0 до *P* відбувається так, немов би світловий потік поширюється всередині дуже вузького каналу вздовж прямої S_0P, тобто ***прямолінійно***. У такий спосіб хвильовий принцип Гюйгенса-Френеля дозволяє пояснити прямолінійне поширення світла в однорідному середовищі.

Виразимо кількість зон Френеля *m* через радіус зовнішньої границі (13.8):

$$m = \frac{r^2{}_m}{\lambda}\frac{(a+b)}{ab} = \frac{r^2{}_m}{\lambda}\left(\frac{1}{b} + \frac{1}{a}\right).$$

Кількість зон *m* симетрична відносно a і b. Це означає, що точкове джерело викликає в точці *P* таку дію, яку викликало б воно в точці S_0, якщо його розмістити у точці *P*.

Наведемо наглядну геометричну інтерпретацію розрахунку напруженості поля в точці *P* на основі принципу Гюйгенса-Френеля. Розіб'ємо хвильову поверхню на кільцеві зони, аналогічні зонам Френеля, але значно менші за шириною (різниця ходу від країв зони до точки *P,* складає однакову для всіх зон малу долю λ). Коливання, що створюються в точці *P* кожною із цих зон, зобразимо у вигляді вектора $d\vec{E}$, довжина якого дорівнює амплітуді коливань, а кут який утворює вектор $d\vec{E}$ з напрямком, що приймається за початок відліку, дає початкову фазу коливань. Амплітуди коливань, що створюються такими зонами у точці *P*, повільно зменшуються при переході від зони до зони. Кожне наступне коливання відстає по фазі від попереднього на одну і ту ж величину . Наприклад, коливання в точці *P*, що збуджується хвилею, що приходить від першої зони біля точки *O* (рис.13.3), відповідає вектор

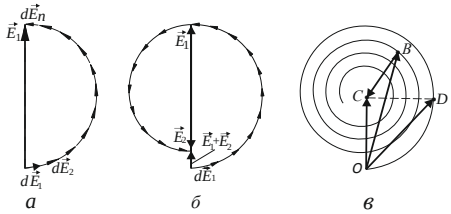


Рис.13.6. Вектори для результуючого коливання в точці *P*.

$d\vec{E}_1$ на рис.13.6(*а*). Коливання в точці *P*, яке збуджується хвилею від наступної (такої ж за площею) елементарної зони, зображується таким же за модулем вектором $d\vec{E}_2$, але повернутим відносно $d\vec{E}_1$ на невеликий кут, оскільки друге коливання дещо відстає по фазі від попереднього коливання (рис.13.6,*а*).

Коливанню в точці *P*, що збуджується хвилею, яка приходить від ділянки, яка прилягає до границі першої зони Френеля (точка M_1 на рис.13.3), буде відповідати вектор $d\vec{E}_n$, що повернутий відносно $d\vec{E}_1$ на π, так як згідно визначення зон Френеля різниця ходу $M_1P - OP = \lambda/2$.

Результуюче коливання в точці P, що збуджується хвилями від всієї 1-ої зони Френеля, зобразиться на діаграмі (рис.13.6,*а*) вектором $\vec{E}_1$, який замикає ломану лінію утворену векторами $d\vec{E}_1, d\vec{E}_2, \ldots, d\vec{E}_n$. Йому відповідає перший доданок в (13.3).

В граничному випадку, коли площі кільцевих зон дуже малі ($d\sigma \to 0$), вектор $\vec{E}_1$ проходить по діаметру напівкола. Продовжимо побудову дальше. Векторна діаграмі результуючого коливання в точці P від двох перших зон Френеля показана на рисунку 13.6,*б* і йому відповідає амплітуда $\left|\vec{E}_1 + \vec{E}_2\right| = E_1 - E_2$. Якщо б всі елементарні амплітуди dE_i, що складаються в точці P, були строго однаковими за величиною, то результуюча амплітуда коливань в точці P, що збуджуються хвилями від перших двох зон Френеля, дорівнювала б нулю (хвиля від першої і другої зон Френеля в точці P внаслідок інтерференції повністю гасила б одна одну). Але коефіцієнт $K(\varphi)$, що зменшується із збільшенням кута φ, характеризує поступове зменшення амплітуди коливань dE_i в точці P. Тому амплітуда $E_1 - E_2$ коливань від двох зон має кінцеве, хоч і мале значення. Цьому відповідає два перших доданки в (13.3).

Амплітуда коливань в точці P, що збуджується всією хвильовою поверхнею, зобразиться на рис.13,6,*в* вектором $\overrightarrow{OC}$. Із рисунка видно, що амплітуда в цьому випадку дорівнює половині амплітуди коливання, що збуджується хвилею від першої зони Френеля, **тобто $\left|\overrightarrow{OC}\right| = E_1/2$. Цей результат ми отримали** раніше алгебраїчно. Зауважимо, що коливання, яке збуджується внутрішньою половиною першої зони Френеля, зображується вектором $\overrightarrow{OD}$ (рис. 13.6,*в*). Таким чином, дія внутрішньої половини першої зони Френеля не еквівалентна половині дії першої зони. Вектор $\overrightarrow{OD}$ в $\sqrt{2}$ рази більший вектора $\overrightarrow{OC}$. Тобто інтенсивність світла, яка створюється внутрішньою половиною першої зони Френеля

$$|\overrightarrow{OD}|^2 = 2|\overrightarrow{OC}|^2 = 2(E_1/2)^2$$

в два рази перевищує інтенсивність $I_0 = (E_1/2)^2$, що створюється всією хвильовою поверхнею.

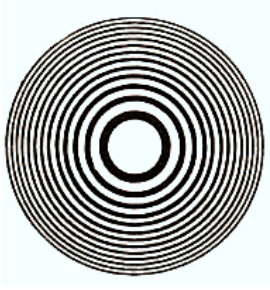


Рис.13.7.

Інтенсивність в точці P можна набагато підсилити, якщо виготовити екран, який буде відкривати, наприклад, першу, третю, п'яту і т. д., тобто всі непарні зони Френеля. Вторинні хвилі від всіх відкритих зон Френеля будуть приходити в точку спостереження P з однаковою фазою і в результаті інтерференції підсилять одна одну. Амплітуда результуючого коливання в точці P буде визначатись сумою

$$E_p = E_1 + E_3 + E_5 + \cdots .$$

Такий екран називають ***зонною пластинкою*** (рис.13.7). Можна виготовити аналогічний екран, який відкриває всі парні зони Френеля. Зонна пластинка, яка містить n відкритих зон Френеля створює в точці P освітленість приблизно в $4n^2$ раз більшу від освітленості, коли пластинка відсутня. Підсилення інтенсивності світла зонною пластинкою аналогічне дії збиральної лінзи. Відстані від пластинки до джерела S_0 і від джерела S_0 до його "зображення" (точка P) зв'язані таким же співвідношенням як і

відповідні величини для лінзи. В цьому легко переконатись, переписавши формулу (13.8) так:

$$\frac{1}{a}+\frac{1}{b}=\frac{1}{F}, \qquad (13.9)$$

де "фокусна відстань" $F = r_m^2/(m\lambda) \approx r_1^2/\lambda$ для даної пластинки величина постійна (так як $r_m \sim \sqrt{m}$, то F не залежить від m). На відміну від лінзи, зонна пластинка має декілька фокусів. Дійсно, якщо наблизити точку P до пластинки, можна знайти таке її положення, коли в центральному світлому крузі пластинки вміститься три перші зони Френеля (а не одна). Наступні три зони будуть перекриті темним кільцем. Ще наступні три зони будуть суміщатися зі світлим кільцем і т. д. Дії сусідніх зон Френеля практично знищують одна одну, тому результат буде такий, як у випадку, коли відкриті 1-ша, 7-ма, 13-та і т. д. зони. В цьому випадку в точці P буде максимум освітленості хоч і слабший, ніж в основному фокусі, Відповідна йому фокусна відстань $F_1 = F/3$. Аналогічно для фокусів вищих порядків можна записати $F_m = F/(2m+1)$, де m – цілі числа.

Інтенсивність світла в головному фокусі можна збільшити ще в чотири рази (порівняно з зонною пластинкою), якщо змінити на π фази вторинних хвиль, що виходять від всіх зон з парними (або, навпаки, з непарними) номерами. Така пластинка була виготовлена Вудом. Хімічним травленням в певних місцях товщина скляної пластини зменшувалась на $(m-1)\lambda/2$. В цьому випадку вторинні хвилі від усіх точок хвильової поверхні приходять в точку P в однаковій фазі.

§13.4.Дифракція Френеля на круглому отворі та диску

Дифракція Френеля – це дифракція сферичних світлових хвиль, а дифракційна картина спостерігається на скінченній віддалі від перешкоди, яка викликає дифракцію.

Дифракція на круглому отворі

Нехай сферична хвиля, що поширюється з точкового джерела S_0, зустрічає на своєму шляху екран з круглим отвором, радіус якого r_0 (рис.13.8,*а*). Дифракційну картину спостерігаємо на екрані **Е** в точці P, що лежить на лінії, яка з'єднує S_0 з центром отвору O. Екран **Е** паралельний до площини отвору і знаходиться від нього на відстані b. Побудуємо на відкритій частині фронту хвилі зони Френеля, що відповідають точці P. Вигляд дифракційної картини залежить від кількості зон Френеля, які укладаються в отворі. Якщо відстані a і b задовільняють співвідношення (13.8) так, що $r_o = r_m$, де m – ціле число, то в отворі укладається m зон Френеля. Тоді, згідно (13.3),

$$E_p = E_1 - E_2 + E_3 - E_4 + \cdots \pm E_m.$$

Із врахуванням (13.7) амплітуда E_p результуючих коливань у точці P залежатиме від парності або непарності m.

Якщо кількість зон Френеля m непарна, то

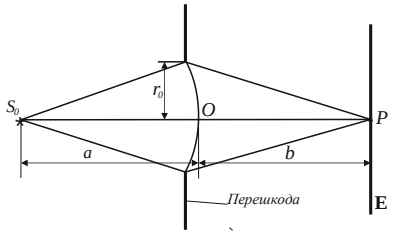
$$E_p = \frac{E_1}{2}+\left(\frac{E_1}{2}-E_2+\frac{E_3}{2}\right)+\left(\frac{E_3}{2}-E_4+\frac{E_5}{2}\right)+ \\ +\cdots+\left(\frac{E_{m-2}}{2}-E_{m-1}+\frac{E_m}{2}\right)+\frac{E_m}{2}=\frac{E_1}{2}+\frac{E_m}{2}. \qquad (13.10)$$

Якщо в отворі укладається парна кількість m зон, то

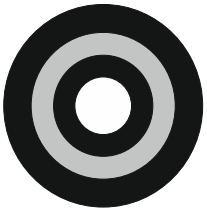
$$E_p = \frac{E_1}{2}+\left(\frac{E_1}{2}-E_2+\frac{E_3}{2}\right)+ \\ +\left(\frac{E_3}{2}-E_4+\frac{E_5}{2}\right)+\cdots+ \\ +\left(\frac{E_{m-3}}{2}-E_{m-2}+\frac{E_{m-1}}{2}\right)+ \\ +\frac{E_{m-1}}{2}-E_m=\frac{E_1}{2}+\frac{E_{m-1}}{2}-E_m. \qquad (13.11)$$

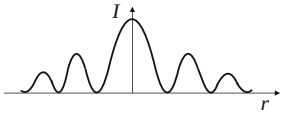
Отже, коли отвір відкриває непарну кількість зон Френеля, то амплітуда в точці P буде більша, ніж при вільному поширенні хвилі і у точці P буде інтерференційний максимум. Якщо m–парне, то у точці P буде

інтерференційний мінімум. Амплітуди коливань від двох сусідніх зон практично однакові. Тому

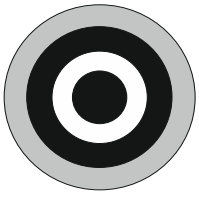


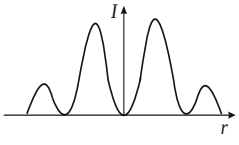
а





б





в

Рис.13.8. *а*−схема дифракції Френеля на круглому отворі; *б,в*− теоретичний розподіл інтенсивності світла на екрані **E,** якщо отвір відкриває непарне (*б*) і парне (*в*) число зон Френеля.

$$\frac{E_{m-1}}{2} - E_m = -\frac{E_m}{2}.$$

В результаті

$$E_p = \frac{E_1}{2} \pm \frac{E_m}{2}, \qquad (13.12)$$

де знак „+” стосується непарної, а знак „−” парної кількості зон Френеля. Якщо в отворі вкладається одна зона Френеля, то в точці P амплітуда $E_p = E_1$, тобто вдвоє більша, ніж за відсутності непрозорого екрану з отвором. Якщо в отворі вкладаються дві зони Френеля, то їх дія в точці P практично взаємно компенсується через інтерференцію. Таким чином, дифракційна картина від круглого отвору поблизу точки Pматиме вигляд темних і світлих кілець з центрами в точці P(якщо m − непарне, то в центрі буде світлий круг (рис.13.8,*б*), коли m − парне − темний (рис.13.8,*в*), причому інтенсивність $(I = E^2)$ максимумів зменшується з відстанню від центра картини.

При незмінному положенні джерела світла кількість зон m залежить від радіуса отвору r_0 і відстані b. При зміні радіуса отвору в отворі змінюватиметься і кількість зон Френеля, що вміщуються на відкритій ділянці хвильового фронту. Їх кількість буде змінюватись з парної на непарну і навпаки. Внаслідок цього інтенсивність світла в точці P періодично то зменшуватиметься, то збільшуватиметься. Така зміна інтенсивності також спостерігається, якщо екран **E** переміщувати вздовж лінії S_0P.

У випадку, коли діаметр отвору великий, так що $E_m \ll E_1$, то ніякої дифракційної картини на екрані не буде спостерігатись − світло в цьому випадку поширюється прямолінійно. Якщо отвір освітлюється не монохроматичним, а білим світлом, то кільця будуть мати кольорове забарвлення, бо кількість зон Френеля, які вкладаються в отворі, залежить від довжини хвилі світла.

Дифракція на диску

Нехай сферична хвиля, яка поширюється від точкового джерела S_0, зустрічає на своєму шляху непрозорий диск. У такому разі закриту диском ділянку фронту хвилі треба виключити з розгляду і будувати зони Френеля потрібно, починаючи з країв диска. На рис.13.9 показано побудову цих зон для точки P екрана **E,** яка лежить навпроти центра диска. Амплітуда E_m в точці P визначається спільною дією всіх відкритих зон, починаючи з першої:

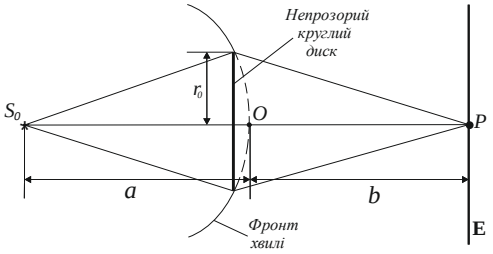


Рис.13.9.

$$\begin{cases} E_P = E_1 - E_2 + E_3 - E_4 + \cdots = \\ \quad = \dfrac{E_1}{2} + \left(\dfrac{E_1}{2} - E_2 + \dfrac{E_3}{2}\right) + \\ + \left(\dfrac{E_3}{2} - E_4 + \dfrac{E_5}{2}\right) + \cdots = \dfrac{E_1}{2}. \end{cases}$$

Отже, в точці P завжди буде інтерференційний максимум, який відповідає половині дії першої відкритої зони Френеля.

Центральний максимум оточений концентричними темними і світлими інтерференційними кільцями, що чергуються. Зміна інтенсивності світла I із відстанню від точки P на екрані приведена на рис.13.10. Із збільшенням радіуса диска перша відкрита зона віддаляється від точки P з одночасним збільшенням кута φ між нормаллю до поверхні цієї зони в якій-небудь її точці і напрямком випромінювання в бік точки P (§13.2). Тому інтенсивність центрального максимуму послаблюється при збільшенні розмірів диска. Якщо радіус диска набагато більший за радіус закритої ним центральної зони Френеля, то за диском буде звичайна тінь з дуже слабкою інтерференційною картиною в її межах.

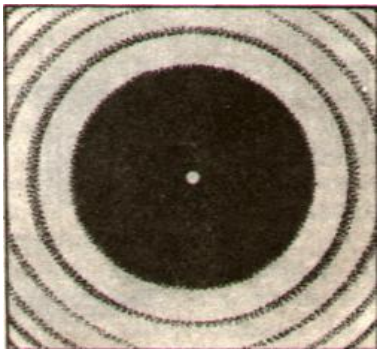


Рис.13.10.

Якщо диск закриває лише невелику частину центральної зони Френеля, то він зовсім не дає тіні – освітленість екрана скрізь залишається такою ж, як і за відсутності перешкоди. Вище отримані результати можна легко одержати за допомогою векторної діаграми (рис.13.6,*в*), якщо вважати, що вектор $\overrightarrow{OC}$, який представляє коливання напруженості поля в точці P при повністю відкритій хвильовій поверхні, як суму двох векторів, один з яких $\overrightarrow{OB}$ відображає коливання від тих зон Френеля, які перекриті диском, а інший $\overrightarrow{BC}$– коливання від відкритої ділянки хвильової поверхні, яка лежить за межами диску. Вектор $\overrightarrow{BC}$ тільки трохи менший за модулем від модуля вектора $\overrightarrow{OC}$.

Ці висновки теорії Френеля викликали сильні враження у його сучасників. В 1818 р. член конкурсного комітету французької академії наук Пуассон, який розглядав представлений на премію мемуар Френеля, прийшов до висновку про те , що в центрі тіні маленького диска повинна знаходитись світла пляма, і вважав цей висновок настільки

абсурдним, що висунув його як заперечення проти хвильової природи світла, яку розвивав Френель. Однак інший член того ж комітету Араго виконав експеримент, який показав, що це дивне передбачення Френеля правильне. Довга суперечка між прихильниками корпускулярної і хвильової теорії світла завершилась на користь хвильової теорії. Світла точка в центрі дифракційної картини від круглого диска отримала назву пляма Пуассона.

Вивчення дифракції Френеля в цьому параграфі було проведено з припущенням, що джерело світла точкове, а світло, що ним випромінюється, монохроматичне. У випадку протяжного джерела світло від кожного його елемента дає свою дифракційну картину. Внаслідок повної незалежності (некогерентності) світла окремих елементів відбувається просто додавання інтенсивностей в кожній точці, і результат дифракції визначається накладанням таких дещо зміщених одна відносно одної дифракційних картин. Для спостереження дифракції на досліді розміри джерела повинні бути достатньо малі, щоб темні і світлі смуги картин від його окремих елементів не перекривались. Аналогічно, у випадку немонохроматичного джерела різні спектральні компоненти його випромінювання створюють дифракційні картини, які не співпадають, так як розміри зон Френеля залежать від довжини хвилі. Розподіл інтенсивності світла, який спостерігається, відповідає накладанню цих дифракційних картин.

§13.5.Дифракція Френеля на прямолінійному краю напівплощини

Принцип Гюйгенса-Френеля можна застосовувати для знаходження розподілу інтенсивності біля межі тіні, яка відкидається краєм великого екрану. Обмежимось випадком плоскої хвильової поверхні падаючої хвилі, яка співпадає із напівплощиною (рис.13.11). На відстані b за напівплощиною поставимо паралельний їй екран, на якому візьмемо точку P (рис.13.11). Розіб'ємо відкриту частину хвильової поверхні на зони Френеля, які мають вид дуже вузьких прямолінійних смуг,

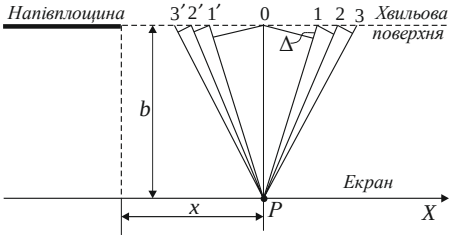


Рис.13.11.

паралельних краю напівплощини.Ширину зон виберемо так, щоб виміряні в площині рисунка відстані від точки P до країв любої зони відрізнялись на однакову величину Δ (Δ може дорівнювати і $\lambda/2$). При такій умові, коливання, які створюються в точці P сусідніми зонами, будуть відрізнятися за фазою на постійну величину ($\Delta\varphi = 2\pi\Delta/\lambda$).

Зонам, що розташовані справа від точки P, припишемо номера 1,2,3 і т. д., а зонам розташованим зліва – номера $1', 2', 3'$ і т. д. Зони з номерами m і m' мають однакову ширину і розташовані відносно точки P симетрично. Тому коливання в точці P, які вони створюють, співпадають за фазою і амплітудою. Щоб встановити залежність амплітуди від номера зони m, оцінимо площі зон. Із рис. 13.12 видно, що сумарна ширина перших m зон дорівнює

$$d_1 + d_2 + \cdots + d_m = \sqrt{(b + m\Delta)^2 - b^2} = \sqrt{2bm\Delta + m^2\Delta^2},$$

де $d_1, d_2, \ldots, d_m$ – ширини 1-ої, 2-ої і т. д. зон. Внаслідок того, що зони вузькі, $\Delta \ll b$. Тому

при не дуже великих m: $d_1 + d_2 + \cdots + d_m = \sqrt{2bm\Delta}$. При m = 1 за цією формулою отримаємо, що $d_1 = \sqrt{2b\Delta}$. Тоді сумарна ширина перших m зон буде

$$d_1 + d_2 + \cdots + d_m = d_1\sqrt{m}.$$

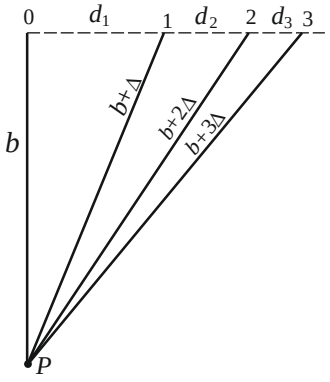


Рис.13.12.

Звідси

$$d_m = d_1\left(\sqrt{m} - \sqrt{m-1}\right). \quad (13.13)$$

Розрахунок за формулою (13.13) дає, що

$$d_1{:}d_2{:}d_3{:}d_4{:}\ldots = 1{:}0{,}41{:}0{,}32{:}0{,}27{:}\ \ldots . \quad (13.14)$$

В таких же співвідношеннях знаходяться і площі зон.

Із (13.14) витікає, що амплітуди коливань, які створюються в точці P окремими зонами, спочатку (для перших зон) спадають дуже швидко, потім це спадання стає повільнішим. Через це ламана лінія, яка отримується при графічному додаванні коливань, які збуджуються прямолінійними зонами, спочатку йде більш полого, ніж у випадку кільцевих зон (площі яких при аналогічній побудові приблизно однакові) (див.§13.3, рис.13.6). На рис.13.13 співставленні обидві векторні діаграми. В обох випадках відставання за фазою кожного наступного коливання є одним і тим же. Значення амплітуди для кільцевих зон (рис.13.13,*а*) прийнято постійним, а для прямолінійних зон (рис.13.13,*б*) – спадним згідно пропорції (13.14). Графіки на рис.13.13 є наближеними. При точній побудові графіків необхідно враховувати залежність амплітуди від r і φ (§13.3). Однак на загальному характері діаграм це не відбивається.

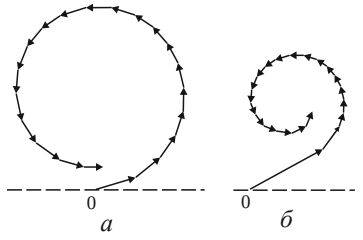


Рис.13.13.

На рис.13.13,*б* показані тільки коливання в точці P, обумовлені зонами, що лежать справа від неї. Зони з номерами m і m' розташовані симетрично відносно P. Тому природно при побудові діаграми вектори, які відображають відповідні цим зонам коливання, розташовані симетрично відносно початку координат O (рис.13.14). Якщо ширину зон спрямувати до нуля, то ломана лінія, яка зображена на рис.13.14, перетворюється в плавну криву (рис.13.15), яка називається **спіраллю Корню.**

Рівняння спіралі Корню в параметричній формі має вид:

$$\xi = \int_0^v \cos\frac{\pi u^2}{2}\,du, \qquad \eta = \int_0^v \sin\frac{\pi u^2}{2}\,du.$$

Ці інтеграли називаються інтегралами Френеля, значення яких приводяться в таблицях в залежності від параметра v. Зміст параметра v в тому, що $|v|$ дає довжину дуги кривої Корню, яка вимірюється від початку координат. Числа, які відмічені вздовж кривої на рис.13.15, дають значення параметра v. Точки F_1 і F_2, до яких асимптотично наближається крива при $v \to$

$\pm\infty$, називаються фокусами спіралі Корню. Їхні координати дорівнюють

$$\xi = +1/2, \eta = +1/2 \text{ для точки } F_1,$$

$$\xi = -1/2, \eta = -1/2 \text{ для точки } F_2.$$

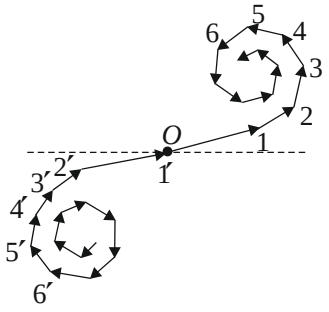


Рис.13.14.

Правий завиток спіралі (ділянка OF_1) відповідає зонам, які розташовані справа від точки P, лівий завиток (ділянка OF_2) – зонам, які розташовані зліва від точки P. Спіраль Корню дає можливість знайти амплітуду

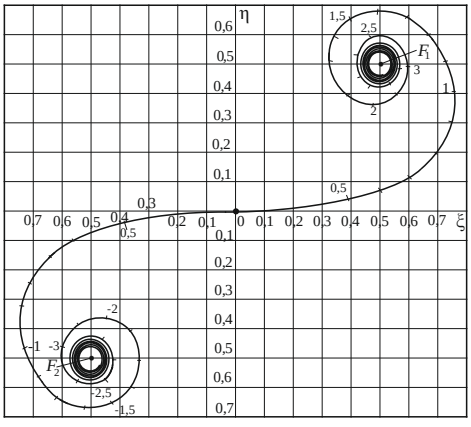


Рис.13.15.

світлового коливання в будь-якій точці екрану. Положення точки будем характеризувати координатою x, яка відраховується від границі геометричної тіні (рис.13.11). Для точки P, що лежить на границі геометричної тіні (x=0), всі штриховані зони будуть закриті. Коливання від нештрихованих зон відкриває правий завиток спіралі. Отже, результуюче коливання зобразиться вектором, початок якого знаходиться в точці O, а кінець – в точці F_1 (рис.13.16,*а*). При зміщенні точки P в область геометричної тіні напівплощина закриває все більше число нештрихованих зон.

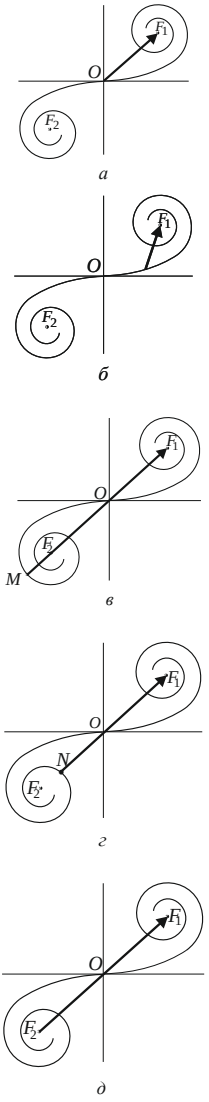


Рис.13.16.

Тому початок результуючого вектора переміщується по правому завитку по напрямку полюса F_1 (рис.13.16,*б*). В результаті амплітуда коливання монотонно прямує до нуля.

Якщо точка P зміщується від границі геометричної тіні вправо, в доповнення до

нештрихованих зон відкривається зростаюче число штрихованих зон. Тому початок результуючого вектора ковзає по лівому завитку спіралі в напрямку до полюса F_2. При цьому амплітуди проходять через ряд максимумів (перший із них дорівнює довжині відрізка MF_1 на рис.13.16,*в*) і мінімумів (перший із них дорівнює довжині відрізка NF_1 на рис. 13.16,*г*). При повністю відкритій хвильовій поверхні амплітуда дорівнює довжині відрізка F_2F_1 (рис.13.16,*д*), тобто в два рази перевищую амплітуду на границі геометричної тіні (див.рис.13.16,*а*). Відповідно, інтенсивність на границі геометричної тіні складає 1/4 інтенсивності I_0, яка отримується на екрані при відсутності перешкод.

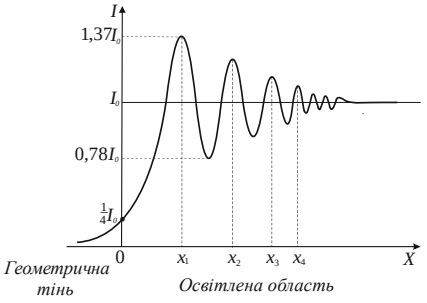


Рис.13.17.

Залежність інтенсивності світла I від координати x точки P приведена на рис.13.17. Із рисунка видно, що при переході в область геометричної тіні інтенсивність змінюється не скачком, а поступово прямує до нуля. За допомогою спіралі Корню можна також знайти відносну величину інтенсивності в максимумах і мінімумах. Для першого максимуму отримується значення $1{,}37I_0$, для першого мінімуму $0{,}78I_0$.

Зробимо декілька зауважень з приводу використаного вище наближеного методу розв'язку дифракційних задач. Детальний аналіз показує, що в основі методу Френеля лежать припущення, які справедливі, коли розміри перешкод (або отворів в екранах) великі порівняно з довжиною світлової хвилі.

В цьому випадку відхилення від геометричної оптики незначні, тобто помітна інтенсивність спостерігається тільки при малих кутах дифракції. Різниця в істинному і розрахованому напрямках $\vec{E}$ при цьому не суттєва. Вплив краю перешкоди (екран з отвором) на поле в отворі поширюється тільки на відстані порядку довжини хвилі і при великих розмірах отвору заміна істинних значень dE в підінтегральному виразі формули (13.2) на напруженість E хвилі в сумі (13.3) не приводить до суттєвих помилок, так як на більшій частині поверхні S ці значення співпадають. При великих порівняно з довжиною хвилі розмірах отворів розрахунки дифракційної картини методом Френеля добре підтверджується на досліді і узгоджується з результатами точних розрахунків. Підтверджується і припущення про незалежність дифракційної картини при цих умовах від матеріалу перешкоди. Для більшості задач метод Френеля є єдиний шлях розв'язку, який приводить до практично задовільних результатів.

§13.6.Дифракція Фраунгофера

Найбільший практичний інтерес мають дифракційні явища, які спостерігаються при падінні на екран (або на отвір в екрані) паралельного пучка світла. В результаті дифракції пучок втрачає паралельність, тобто появляється світло, яке розповсюджується в напрямках , що відрізняються від початкового. Розподіл його інтенсивності на дуже великій (в граничному випадку нескінченно великій) відстані від екрана відповідає *дифракції Фраунгофера.* Хвилі, які виникають в

результаті обмеження фронту падаючої «плоскої» хвилі при проходженні через отвір в екрані називаються *дифракційними,* а нормалі до їхніх хвильових поверхонь – дифракційними променями. Вони не існують в рамках геометричної оптики.

Дифракція на одній щілині

Нехай паралельний пучок монохроматичного світла падає нормально на непрозору плоску поверхню, в якій прорізано вузьку щілину, що має сталу ширину b і довжину $l >> b$ (рис.13.18).

Розмістимо за щілиною збиральну лінзу, а у фокальній площині лінзи – екран. Хвильова поверхня падаючої хвилі, площина щілини і екран паралельні один одному. Оскільки щілина вузька, картина, яка спостерігається в будь-якій площині, що перпендикулярна щілині, буде однаковою. Тому достатньо дослідити характер картини в одній такій площині, наприклад, в площині рис.13.18. Всі величини, що будуть вводитись в подальшому, наприклад, кут φ між променем і оптичною віссю лінзи, відносяться до цієї площини. Розіб'ємо відкриту частину хвильової поверхні на паралельні краю щілини елементарні зони, ширина яких dx.

Вторинні хвилі, що посилаються цими зонами в напрямку, який визначається кутом φ, зберуться в точці екрану P (рис.13.18). Лінза збирає у фокальній площині плоскі (а не сферичні) хвилі. Тому множник $1/r$ у виразі (13.1) для dE_P для випадку дифракції Фраунгофера буде відсутній. Якщо обмежитись розглядом невеликих кутів φ, коефіцієнт $K(\varphi)$ у формулі (13.1) можна вважати постійним і прийняти за одиницю. Тоді амплітуди коливань, які збуджуються зоною в любій точці екрану, будуть залежати від площі зон. Площа пропорційна ширині зони dx. Отже, амплітуда коливання dE_P, яке збуджується зоною, ширина якої dx, в любій точці екрану має вид $a_0 dx/b$. Тут a_0 амплітуда хвилі, яка падає на щілину.

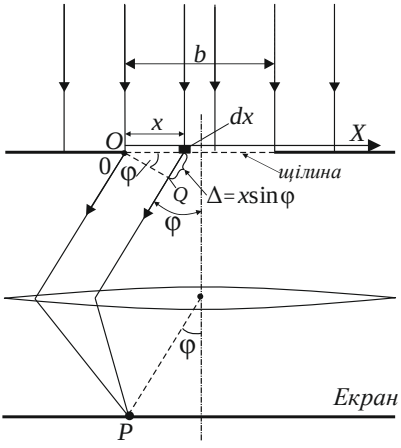


Рис.13.18.

Тепер визначимо фазові співвідношення між коливанням dE_P. Співставимо фази коливань, які збуджуються в точці P елементарними зонами з координатами 0 і x (рис.13.18). Оптичні шляхи OP і QP однакові. Тому різниця фаз між коливаннями, які розглядаються, утворюється на шляху Δ, який дорівнює $x\sin\varphi$ (рис.13.18). Якщо фазу коливань, що збуджується елементарною зоною, яка примикає до лівого краю щілини ($x = 0$), прийняти ωt, то фаза коливань, яка збуджується зоною з координатою x, буде дорівнювати

$$\omega t - 2\pi\frac{\Delta}{\lambda} = \omega t - \frac{2\pi}{\lambda}x\sin\varphi,$$

де λ – довжина хвилі в даному середовищі.

Таким чином, коливання, що збуджуються елементарною зоною з координатою x, в точці P, положення якої на екрані визначається кутом φ, може бути представлене у вигляді

$$dE_P = \frac{a_0}{b}\cos\left(\omega t - \frac{2\pi}{\lambda}x\sin\varphi\right)dx.$$

Згідно принципа Гюйгенса – Френеля (13.2) результуюче коливання, яке збуджується в точці P відкритою ділянкою хвильової поверхні, знайдемо, проінтегрувавши dE_P по всій ширині щілини:

$$\begin{cases} E_P = \int_0^b \frac{a_0}{b}\cos\left(\omega t - \frac{2\pi}{\lambda}x\sin\varphi\right)dx = \\ = \frac{a_0}{b}\left(-\frac{\lambda}{2\pi\sin\varphi}\right)\left(\sin\left(\omega t - \frac{2\pi}{\lambda}b\sin\varphi\right) - \sin\omega t\right) = \\ = a_0\frac{\sin\left((\pi b\sin\varphi)/\lambda\right)}{(\pi b\sin\varphi)/\lambda}\cos\left(\omega t - \frac{\pi b\sin\varphi}{\lambda}\right). \end{cases}$$

При цьому була використана формула:

$$\sin\alpha - \sin\beta = 2\sin\frac{\alpha-\beta}{2}\cdot\cos\frac{\alpha+\beta}{2}.$$

Амплітуда A_φ результуючого коливання в точці P, положення якої визначається кутом φ, буде:

$$A_\varphi = \left|a_0\frac{\sin\left((\pi b\sin\varphi)/\lambda\right)}{(\pi b\sin\varphi)/\lambda}\right|. \qquad (13.15)$$

Для точки, яка лежить навпроти центра лінзи, φ=0. Підстановка цього значення у формулу (13.15) дає значення амплітуди a_0. Цей результат можна отримати більш простим шляхом. При $\varphi = 0$ коливання від всіх елементарних зон приходять в точку P в однаковій фазі. Тому амплітуда результуючого коливання дорівнює алгебраїчній сумі амплітуд коливань, які додаються.

При значеннях φ, які задовільняють умові $(\pi b\sin\varphi)/\lambda = \pm k\pi$, тобто у випадку, якщо

$$b\sin\varphi = k\lambda\ (k = \pm1, \pm2, \pm3\ \ldots), \quad (13.16)$$

амплітуда A_φ дорівнює нулю. Таким чином, умова (13.16) визначає положення мінімумів амплітуди (інтенсивності). Відзначимо, що $b\sin\varphi$ є різниця ходу Δ променів, які йдуть в точку P від країв щілини (див. рис.13.18).

Умову (13.16) легко отримати із наступних міркувань. Якщо різниця ходу Δ від країв щілини дорівнює $k\lambda$, відкриту частину хвильової поверхні можна розбити на $2k$ однакових за шириною зон, причому різниця ходу від країв кожної зони буде дорівнювати $\lambda/2$ ((рис.13.19), який виконаний для k=2).

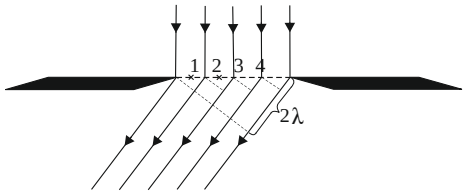


Рис. 13.19.

Коливання від кожної пари сусідніх зон взаємно компенсують одне одного, так що результуюча амплітуда дорівнює нулю. Якщо для точки P різниця ходу Δ дорівнює $(k + 1/2)\lambda$, число зон буде непарним, дія однієї із них буде нескомпенсованою і буде спостерігатись максимум інтенсивності. Інтенсивність світла пропорційна квадрату амплітуди. Отже, у відповідності з (13.15)

$$I_\varphi = I_0\frac{\sin^2((\pi b\sin\varphi)/\lambda)}{((\pi b\sin\varphi)/\lambda)^2}, \qquad (13.17)$$

де $I_0 = a_o^2$ – інтенсивність в центрі дифракційної картини (навпроти центра лінзи), I_φ – інтенсивність в точці, положення якої визначається даним значенням φ.

Із формули (13.17) отримується, що $I_{-\varphi} = I_\varphi$. Це означає, що дифракційна картина симетрична відносно центра лінзи.

Графік функції (13.17) зображений на рис.13.20. По осі абсцис відкладені значення $\sin\varphi$, по осі ординат – інтенсивність I_φ. Між мінімумами інтенсивності, які визначаються умовою (13.16), знаходяться максимуми різних порядків. Їх положення знаходяться із умови:

$\frac{dI_\varphi}{du} = 0$, де $u = \frac{\pi b\sin\varphi}{\lambda}$, а

$$I_\varphi = I_0((\sin u)/u)^2.$$

Диференціювання I_φ по u і прирівнювання похідної до нуля дає рівняння $\mathrm{tg}u = u$, яке має корені $u_0 = 0;\ u_1 = 1{,}43\pi; u_2 = 2{,}46\pi; u_3 =$

$3{,}47\pi$ і т. д. Це дає, що для максимумів $\sin\varphi_0 = 0, \sin\varphi_1 = \pm 1{,}43\lambda;\ \sin\varphi_2 = \pm 2{,}46\lambda,\ \sin\varphi_3 = \pm 3{,}47\lambda$ і т.д.

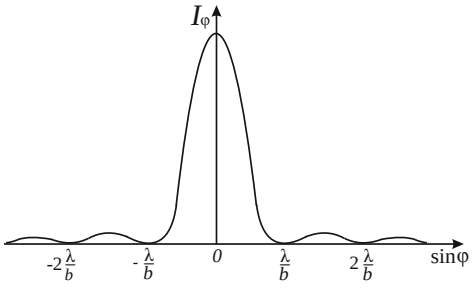


Рис.13.20.

Практично можна вважати, що максимуми знаходяться посередині між сусідніми мінімумами. Значення інтенсивності в максимумах швидко спадає із збільшенням порядку. Їх відношення приблизно можна виразити у вигляді

$$1: 0{,}0472: 0{,}0165: \ldots = \\ = 1: (2/(3\pi))^2 : (2/(5\pi))^2 : \ldots , \qquad (3.18)$$

що отримується із (13.17).

Таким чином, основна частина світлового потоку зосереджена в центральній смузі між мінімумами порядків k=±1, тобто в межах кутів $-\varphi_1 < \varphi < \varphi_1$, де $\sin\varphi_1 = \pm\lambda/b$. Кутова ширина максимума зменшується при збільшенні ширини b щілини: якщо $\varphi \ll 1$, то $\varphi_1 \approx \lambda/b$. Центральний максимум стає різким, перші мінімуми присуваються ближче до центра картини. Висота максимума інтенсивності пропорційна квадрату ширини щілини, так як зростаючий пропорційно b світловий потік розповсюджується в межах спадного кута ($\varphi \sim 1/b$). Відносна інтенсивність залишається незмінною: розподіл світла по максимумах різних порядків (13.18) не залежить від ширини щілини. При звуженні щілини картина розширюється, а її яскравість зменшується. Кількість мінімумів інтенсивності визначається відношенням ширини щілини b до довжини хвилі λ. Із умови (13.16) виходить, що $\sin\varphi_1 = \pm\lambda/b$.

Модуль $\sin\varphi$ не може перевищувати одиницю. Тому $k\lambda/b \le 1$, звідки

$$k \le b/\lambda. \qquad (3.19)$$

При ширині щілини, яка менша довжини хвилі, мінімуми взагалі не виникають. В цьому випадку інтенсивність світла монотонно спадає від середини картини до її країв. Якщо b=2λ, то k може приймати значення 1 і 2. Це значить, що будуть проявлятись чотири мінімуми і за щілиною буде розповсюджуватись *три когерентних пучки світла* (центральний і два бічних максимуми).

Отримані результати можна використати для оцінки дифракційної розбіжності пучків світла, ширина яких обмежена, наприклад, в результаті проходження через діафрагму. Основна частина світлого потоку приходиться на центральний дифракційний максимум, тому його ширину можна прийняти як оцінку кутової розбіжності $\Delta\varphi$ пучка з поперечним перерізом a.

$$\Delta\varphi \sim \lambda/a. \qquad (13.20)$$

Це кутове розширення пучка обумовлене хвильовою природою світла і його в принципі не можна усунути при заданій ширині перерізу пучка. Строго паралельних світлових пучків не існує. На шляху, довжина якої l, пучок зазнає дифракційного розширення порядку $l \cdot \Delta\varphi = \lambda l/a$. Цим розширенням можна знехтувати тільки тоді, коли воно мале порівняно з початковою шириною пучка, тобто при $l \cdot \Delta\varphi \ll a$, або $\lambda l \ll a^2$. В таких умовах пучок приблизно можна вважати паралельним і використовувати для його описання геометричну оптику.

Пучок світла з мінімально можливою при даному діаметрі a кутовою розбіжністю $\Delta\varphi \sim \lambda/a$ формується в результаті інтерференції вторинних хвиль від всього поперечного перерізу. Така інтерференція можлива тільки тоді, коли світлові коливання когерентні по всьому поперечному перерізу пучка. *Висока просторова когерентність лазерного пучка* обумовлена самою природою процесу випромінювання світла (вимушене випромінювання) і пояснює його малу розбіжність.

Якщо плоска хвиля падає на щілину похило під кутом φ' до нормалі, то різниця ходу між вторинними хвилями, які розповсюджуються від країв щілини в напрямку φ, буде $b(\sin\varphi - \sin\varphi')$. Умова дифракційних мінімумів замість (13.16) приймає вид $b(\sin\varphi - \sin\varphi') = k\lambda$. Кутовий розподіл інтенсивності як і раніше визначається формулою (13.17), в якій тепер $\sin\varphi$ слід замінити на $\sin\varphi - \sin\varphi'$. Центральний максимум дифракційної картини розташований при $\varphi = \varphi'$, тобто в напрямку падаючої хвилі.

§13.7.Дифракційна решітка

Важливе практичне застосування має явище дифракції Фраунгофера на системі із великої кількості паралельних щілин, які знаходяться на однакових відстанях одна від одної. Так побудована найпростіша *дифракційна решітка*, яку виготовив в 1786 р. астроном Ріттенгауз, була у вигляді натягнутих на рамку паралельних тонких дротин. Починаючи з Фраунгофера, який виконав в 1821р. перші дослідження за допомогою дифракційного спектроскопа, решітки виготовляють методом нанесення штрихів на поверхню скляної або металічної пластинки. Дифракційною решіткою можна вважати будь-який пристрій, який забезпечує просторову періодичну модуляцію падаючої світлової хвилі за амплітудою або фазою. Відстань d між серединами сусідніх щілин називається *періодом* решітки (рис.13.21).

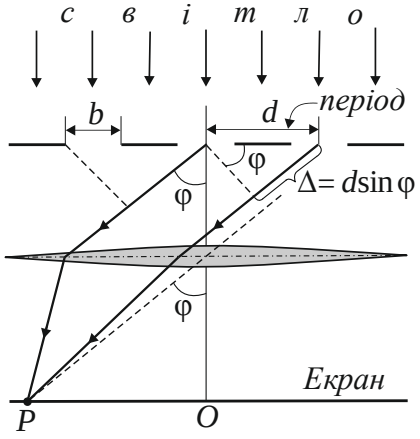


Рис.13.21.

Розташуємо паралельно до решітки збиральну лінзу, в фокальній площині якої поставимо екран. З'ясуємо характер дифракційної картини, яка отримується на екрані при падіння на решітку плоскої світлової хвилі (для простоти будемо вважати, що хвиля падає на решітку нормально). Кожна із щілин дасть на екрані картину, яка описується кривою, зображеною на рис.13.20. Картини від всіх щілин прийдуться на одне і те саме місце екрану (незалежно від положення щілини, центральний максимум лежить навпроти центра лінзи).

Якщо коливання, які приходять в точку P від різних щілин, були б некогерентні, то результуюча картина від N щілин відрізнялась би від картини, яка створюється однією щілиною, тільки тим, що всі інтенсивності зросли би в N раз. Однак коливання від різних щілин є в більшій чи меншій мірі когерентні; тому результуюча

інтенсивність буде відмінна від NI_φ (I_φ – інтенсивність, що створюється однією щілиною, формула (13.17)).

В подальшому ми будемо вважати, що світлові коливання падаючої хвилі по всьому її поперечному перерізу є когерентні і переріз набагато перевищує довжину решітки, так що коливання від всіх щілин можна вважати когерентними одне відносно одного. В цьому випадку результуюче коливання в точці *P*, положення якої визначається кутом φ, буде сума *N* коливань з однаковими амплітудами A_φ (13.15), зсунутими одне відносно одного по фазі на одну і ту ж величину δ.

Результатом такої багатопроменевої інтерференції є такий розподіл інтенсивності на екрані

$$I_{грат} = I_\varphi \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)}. \qquad (13.21)$$

Із рис.13.21 видно, що різниця ходу від сусідніх щілин дорівнює $\Delta = d\sin\varphi$. Отже, різниця фаз

$$\delta = \frac{2\pi}{\lambda}\Delta = \frac{2\pi}{\lambda} d\sin\varphi, \qquad (13.22)$$

де λ – довжина хвилі в даному середовищі.

Підставивши у формулу (13.21) вираз (13.17) для I_φ і (13.22) для δ, отримаємо, що

$$I_{грат} = I_0 \frac{\sin^2((\pi b\sin\varphi)/\lambda)}{((\pi b\sin\varphi)/\lambda)^2} \cdot \cdot \frac{\sin^2((N\pi d\sin\varphi)/\lambda)}{\sin^2((\pi d\sin\varphi)/\lambda)}, \qquad (13.23)$$

де I_0 – інтенсивність, що створюється однією щілиною навпроти центра лінзи. Перший множник в (13.23) перетворюється в нуль в точках, для яких виконується умова (13.16). В цих точках інтенсивність, яка створюється кожною із щілин окремо, дорівнює нулю.

Другий множник в (13.23) приймає значення N^2 в точках, які задовільняють умові

$$\boldsymbol{d\sin\varphi = m\lambda} \ (m = 0, \pm1, \pm2, \dots). \qquad (13.24)$$

Умова (13.24) визначає положення максимумів інтенсивності, які називаються ***головними***. Число m визначає ***порядок головного максимума.*** Максимум нульового порядку тільки один, максимумів 1-го, 2-го і т. д. порядків є по два. Із (13.23) витікає, що інтенсивність головних максимумів в N^2 раз більша інтенсивності I_φ, яку дає одна щілина в напрямку, що визначається кутом φ.

Крім мінімумів, які визначаються умовою (13.16), в проміжку між сусідніми головними максимумами є по $(N-1)$-му додатковому мінімуму. Ці мінімуми виникають в тих напрямках, для яких коливання від окремих щілин взаємно гасять одне одного. Умова додаткових мінімумів наступна:

$$d\sin\varphi = \pm\frac{k'}{N}\lambda, \qquad (13.25)$$

де $k' = 1,2,\dots,N-1,N+1,\dots,2N-1,\ 2N+1,\dots$.

У формулі (13.25) k' приймає всі цілочисленні значення, крім $0, N, 2N, \dots$, тобто тих, при яких умова (13.25) переходить в (13.24). Отже, між двома сусідніми головними максимумами знаходиться $N-1$ додаткових мінімумів і $N-2$ вторинних максимумів. На них накладатимуться мінімуми, що виникають при дифракції від однієї щілини. Із формул (13.16) і (13.24) видно, що головний максимум m-го порядку збігається з k-им мінімумом від однієї щілини, якщо виконується рівність:

$$\frac{k}{b} = \frac{m}{d}, \text{ або } \frac{m}{k} = \frac{d}{b}.$$

На рис.13.22 наведений розподіл інтенсивності I_φ від sinφ для $N = 4$ і $\frac{d}{b} = 3$. Пунктирна крива, що проходить через вершини головних максимумів, зображує інтенсивність, яка зумовлена дифракцією на одній щілині. Як видно з рис.13.22 при відношенні $d/b = 3$ головні максимуми 3-го,

6-го, тощо порядків збігаються з мінімумами інтенсивності від однієї щілини, тому ці максимуми зникають. Чим більше щілин N,

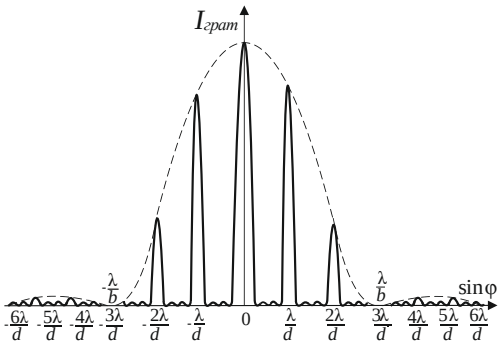


Рис.13.22.

тим більша кількість світлової енергії пройде через решітку, тим більше мінімумів утворюється між сусідніми головними максимумами, тим інтенсивнішими і гострішими будуть максимуми. Оскільки sinφ не може бути більше від одиниці, то кількість головних максимумів $m \le d/\lambda$.

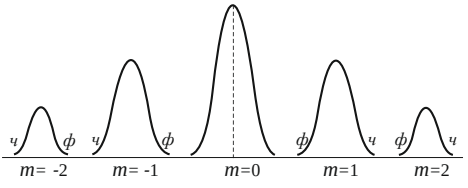


Рис.13.23.

Якщо дифракційну решітку освітлюють білим світлом, то для різних значень λ положення всіх головних максимумів, крім центрального, не збігаються один з одним. Тому центральний максимум має вигляд білої смужки, а всі інші – райдужних смужок, які називають дифракційними спектрами першого, другого і тощо порядків. У межах кожної смужки забарвлення змінюється від фіолетового біля внутрішнього краю, який найближчий до максимуму нульового порядку, до червоного – біля зовнішнього краю. На рис.13.23 зображено центральний та два головні максимуми і зазначено їхнє забарвлення ($ф$–фіолетовий, $ч$–червоний).

Кутовою дисперсією називається величина

$$D_\varphi = \frac{\delta\varphi}{\delta\lambda},$$

де δφ – кутова відстань між двома спектральними лініями, яким відповідають довжини хвиль λ і λ + δλ (рис.13.24).

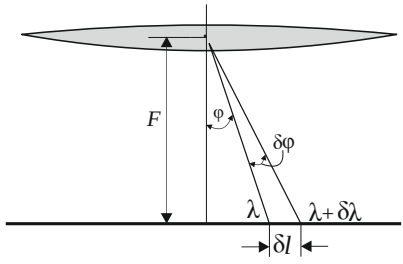


Рис.13.24.

Продиференціювавши формулу (13.24) по λ при сталому m отримаємо $d\cos\varphi \cdot \delta\varphi = m\delta\lambda$. Звідси

$$D_\varphi = \frac{\delta\varphi}{\delta\lambda} = \frac{m}{d\cos\varphi}. \qquad (3.26)$$

Якщо кути дифракції φ малі, то $\cos\varphi \approx 1$ і

$$D_\varphi = \frac{m}{d} = m\frac{N}{l}, \qquad (3.27)$$

де l – довжина робочої ділянки гратки, яка містить N – штрихів.

Із виразу (13.27) видно, що кутова дисперсія обернено пропорційна періоду гратки d і чим вищий порядок спектра m, тим більша дисперсія.

Лінійною дисперсією називають величину

$$D_l = \frac{\delta l}{\delta\lambda}, \qquad (13.28)$$

де δl – лінійна відстань на екрані між двома максимумами одного й того самого порядку m для хвиль λ і λ + δλ (рис.13.24). Якщо фокусна відстань лінзи, у фокальній площині якої спостерігається дифракційна картина дорівнює F, то $\delta l \approx F\delta\varphi$ і

$$D_l = F\frac{\delta\varphi}{\delta\lambda} = FD = F\frac{m}{d}\cdot \qquad (13.29)$$

Дисперсійною областю спектрального приладу, зокрема дифракційної решітки, називається ширина спектрального інтервалу $\Delta\lambda$, в якому спектри не перекриваються. Нехай довжини світлових хвиль, що падають на гратку знаходиться в інтервалі від λ до $\lambda_1 = \lambda + \Delta\lambda$. Інтервал $\Delta\lambda$ буде дисперсійною областю решітки тоді, коли лівий край спектра $(m+1)$–го порядку для довжини хвилі λ збігатиметься з правим краєм спектра m–го порядку для довжини хвилі λ_1.

Цю умову можна згідно (13.24) записати так:

$$d\sin\varphi = m\lambda_1, \quad d\sin\varphi = (m+1)\lambda.$$

Звідси

$$m\lambda_1 = (m+1)\lambda, \ \lambda_1 - \lambda = \Delta\lambda = \lambda/m.$$

Найменша різниця довжин хвиль двох спектральних ліній $\delta\lambda$, при якій спектральний прилад розділяє їх окремо, називається *спектральною розділяючою відстанню*, а величина

$$R = \frac{\lambda}{\delta\lambda}$$

– *роздільною здатністю* приладу.

Для дифракційної решітки Релей запропонував такий критерій спектрального розділення: спектральні лінії з довжинами хвиль λ і $\lambda_1 = \lambda + \delta\lambda$ вважаються розділеними, якщо головний максимум дифракційної картини для хвилі з довжиною λ збігається за своїм розміщенням з першим дифракційним мінімумом того самого порядку для хвилі з довжиною λ_1, а інтенсивність в проміжку між максимумами становить не більше, ніж 80% від інтенсивності максимуму (рис.13.25).

Нехай головний максимум m–го порядку для хвилі з $\lambda_1 = \lambda + \delta\lambda$ (формула (13.24)) знаходиться на місці першого мінімуму (формула (13.25) ($k' = 1$)) спектра того самого порядку для хвилі з λ:

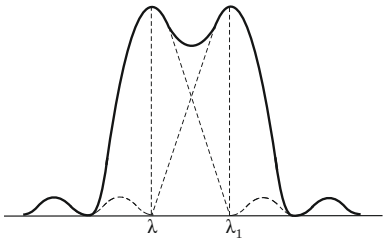


Рис.13.25.

$$d\sin\varphi = m\lambda_1, \quad d\sin\varphi = \left(m + \frac{1}{N}\right)\lambda.$$

Звідси

$$m(\lambda + \delta\lambda) = \left(m + \tfrac{1}{N}\right)\lambda.$$

Тобто

$$\delta\lambda = \frac{\lambda}{mN}\cdot$$

Тоді роздільна здатність гратки

$$R = \frac{\lambda}{\delta\lambda} = mN. \qquad (13.30)$$

Отже, роздільна здатність решітки пропорційна порядку спектра m і числу N щілин (штрихів). Оскільки $N = l/d$, а $m = (d\sin\varphi)/\lambda$, то

$$R = mN = \frac{l}{d}\cdot\frac{d\sin\varphi}{\lambda} \quad \text{і}$$

$$R_{max} = \frac{l\sin\varphi_{\max}}{\lambda},$$

де $\varphi_{\max}$ – максимальний кут дифракції ($\varphi_{\max} \le 90^\circ$). Тому, максимальна роздільна здатність решітки буде $R_{max} = (\lambda/\delta\lambda)_{\max} \le l/\lambda$. Таким чином, роздільна здатність дифракційної решітки, як видно із (13.30) пропорційна порядку спектра m і числу щілин решітки N.

На рис.13.26 співставленні дифракційні картини, які отримуються для двох спектральних ліній. За допомогою двох

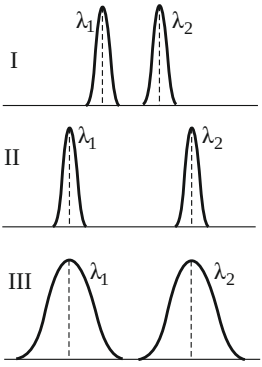


Рис.13.26.

решіток, які відрізняються значенням дисперсії D_φ і роздільною здатністю R. Решітки I і II мають однакову R (у них однакова кількість щілин N), але різну дисперсію (у решітки I період d у два рази більший, відповідно дисперсія D_φ в два рази менша, ніж у решітки II). Решітки II та III мають однакову дисперсію (у них однакові d), але різну R (у решітки II число щілин N і роздільна здатність R у два рази більша, ніж у решітки III).

Дифракційна решітка використовується як для пропускання світла так і для його відбивання. Прозорі решітки виготовляють із скляних або кварцових пластин, на поверхню яких за допомогою спеціальної машини наноситься алмазним різцем ряд паралельних штрихів. Щілинами є проміжки між штрихами.

Решітки, що відбивають світло, отримуються нанесенням алмазним різцем штрихів на поверхню металічного дзеркала. Світло падає на таку решітку похило. При цьому решітка з періодом d діє так, як при паралельному падінні діяла б прозора решітка з періодом $d\cos\theta$, де θ – кут падіння. Кращі решітки мають 1200 штрихів на 1 мм ($d \approx 0{,}8$ мкм). Оскільки кількість головних максимумів $m \le d/\lambda$, то спектри другого порядку у видимому світлі при такому періоді не спостерігаються. Загальна кількість штрихів в таких решітках досягає 200 тисяч (довжина біля 200 мм). Для фокусної відстані приладу F=2м довжина видимого спектру 1-го порядку складає в цьому випадку більше 0,7 м.

§13.8.Дифракція рентгенівського випромінювання

Рентгенівським випромінюванням є електромагнітні хвилі, довжина яких лежить в інтервалі $10^{-8} \div 10^{-12}$ м (§11.1**)**. Залежно від умов отримання рентгенівське випромінювання утворює суцільний або лінійчастий (характеристичний) спектр. Для спостереження дифракції на решітці необхідно, щоб її період був того самого порядку, що і довжина хвилі падаючого випромінювання.

Щоб зрозуміти як відбувається дифракція рентгенівських променів, поставимо дві дифракційні решітки одна за одною так, щоб штрихи були взаємно перпендикулярні. Перша решітка, штрихи якої будем вважати вертикальними, дасть в горизонтальному напрямку ряд максимумів, положення яких визначається умовою (13.24) (тут прийнято кут дифракції φ позначати буквою θ).

$$d_1 \sin\theta_1 = m_1\lambda \ (m_1 = 0, \pm1, \pm2, \dots). \quad (13.31)$$

Друга решітка (з горизонтальними штрихами) розіб'є кожний із пучків, що йдуть від першої решітки, на розташовані по вертикалі максимуми, положення яких визначається теж умовою (13.24).

$$d_2 \sin\theta_2 = m_2\lambda \ (m_2 = 0, \pm1, \pm2, \dots). \quad (13.32)$$

В результаті дифракційна картина буде мати вид правильно розташованих плям, кожній із яких відповідають два цілочисленних індекси m_1 і m_2 (рис.13.27).

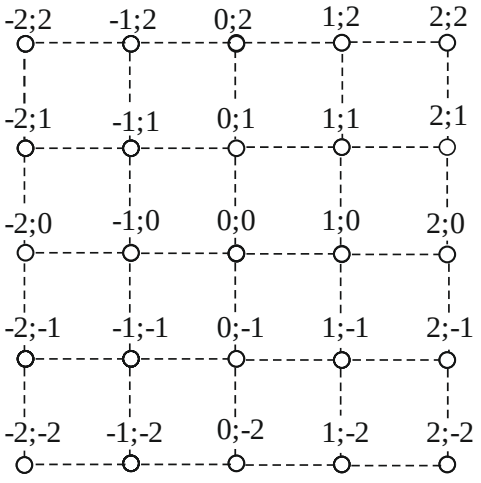


Рис.13.27.

Така ж дифракційна картина отримається, якщо замість двох окремих решіток взяти одну прозору пластину, на яку нанесені дві системи взаємно перпендикулярних штрихів. Подібна пластинка є *двомірна періодична структура* (звичайна гратка – одномірна структура). Вимірявши кути θ_1 і θ_2, які визначають положення максимумів, і знаючи довжину хвилі λ, можна знайти за формулами (13.31) і (13.32) періоди структури d_1 і d_2.

Дифракційну картину, аналогічну тій, що зображена на рис.13.27, дають любі двомірні періодичні структури, наприклад система невеликих отворів або система непрозорих маленьких кульок.

Для виникнення дифракційних максимумів необхідно, щоб період структури d був більший λ. В іншому випадку, умови (13.31) і (13.32) можуть задовільнятись тільки при значеннях m_1 і m_2, які дорівнюють нулю (модуль $\sin\theta$ не може перевищувати одиницю).

Дифракція спостерігається також на трьохмірних структурах, тобто просторових утвореннях, які мають періодичність по трьох напрямках, що не лежать в одній площині,. Подібними структурами є всі кристалічні тіла. У випадку кристалів умова $d > \lambda$ виконується для рентгенівських променів. Вперше дифракцію рентгенівських променів спостерігали в 1913 р. в дослідах Лауе, Фрідріха і Кніппінга.

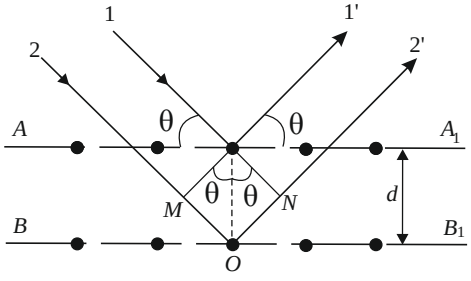


Рис.13.28.

Для рентгенівських променів ідеальними природними дифракційними решітками є монокристали, в яких атоми та іони розмішені на відстані порядку 10^{-10}м. Кристалографічні дослідження показали, що у будь-якому кристалі можна виявити певні площини, де атоми або іони, які утворюють його кристалічну решітку, розміщені найбільш густо. Такі площини відбиватимуть монохроматичне рентгенівське випромінювання, яке може інтерферувати відбившись від різних площин.

На рис.13.28 зображені сусідні площини кристала AA_1 і BB_1. Абсолютний показник заломлення всіх середовищ для рентгенівських променів близький до одиниці.

Тому оптична різниця ходу Δ між двома променями 1' і 2', які відбиваються від площин AA_1 і BB_1, дорівнює

$$\Delta = MO + ON = 2d\sin\theta,$$

де $d-$ відстань між площинами, а кут $\theta -$ кут між площиною AA_1 та падаючими та відбитими променями або кут ковзання. Якщо довжина хвилі рентгенівських променів λ, то

інтерференційні максимуми у відбитих променях спостерігатимуться, коли

$$2d\sin\theta = m\lambda, \quad (m = \pm1, \pm2, \ldots). \quad (13.33)$$

Співвідношення (13.33) є ***формулою Вульфа-Брегга,*** яке означає, що дифракція виникає не в довільному напрямку падіння монохроматичного випромінювання. Для її спостереження треба повертати кристал так, щоб кут ковзання задовольняв умову (13.33). Якщо обертати кристал або проводити експеримент з полікристалічною системою, в якій окремі кристалики орієнтовані довільно, то можна отримати певну систему інтерференційних картин від усіх можливих типів атомних площин певного кристалу.

Дифракція рентгенівських променів є основою ***рентгеноструктурного аналізу,*** який дає можливість вивчати структуру кристалів, аморфних твердих тіл, рідин, а також *рентгеноспектрального аналізу,* що використовується для вивчення рентгенівських спектрів, а також для визначення хімічного складу речовин.

Контрольні питання

1.В чому суть явища дифракції?

2.Сформулювати і пояснити принцип Гюйгенса – Френеля.

3.Пояснити суть методу зон Френеля.

4.Пояснити утворення дифракційної картини від круглого отвору та круглого диску.

5.Застосування методу графічного складання амплітуд для якісного пояснення дифракційної картини.

6.Пояснити явище дифракції Фраунгофера на одній щілині.

7.Що представляє собою дифракційна решітка?

8.Пояснити умову головних фраунгоферових максимумів при дифракції світла на дифракційній решітці.

9.Як визначити максимальний порядок спектру, який може дати решітка?

10.Дати визначення кутової та лінійної дисперсії дифракційної решітки.

11.Як визначається роздільна здатність дифракційної решітки? Чим визначається її величина?

12.В чому особливість дифракції рентгенівського випромінювання?

Розділ 14. Поляризація світла

§14.1. Природне і поляризоване світло. Закон Малюса

Наслідком теорії Максвелла є поперечність світлових хвиль: вектори напруженості електричного $\vec{E}$ і магнітного $\vec{H}$ полів електромагнітної хвилі взаємно перпендикулярні і коливаються перпендикулярно до вектора швидкості $\vec{v}$ поширення хвилі. Оскільки основне значення при дії світла на речовину має вектор напруженості $\vec{E}$ електричного поля – *світловий вектор,* – що діє на електрони в атомах речовини, тому в подальшому будемо розглядати напрямок коливання лише цього вектора.

Розглянемо випромінювання реального джерела світла, яке складається із великої кількості атомів – елементарних випромінювачів (диполів). В основі моделі випромінювання звичайного (нелазерного) джерела світла, лежить *статистична гіпотеза* про те, що у випадку спонтанного випромінювання різні атоми джерела, випромінюють окремі цуги хвиль (§11.6) незалежно один від одного у випадкові моменти часу. Фази коливань електромагнітного поля у випромінюванні

різних атомів не скорельовані одна з одною. Тому виявляється, що розподіл інтенсивності випромінювання всіх атомів джерела в такій не когерентній суперпозиції визначається сумуванням розподілів інтенсивності для індивідуальних атомів.Процес випромінювання окремого атома триває біля 10^{-8} с. За цей час утворюється послідовність цугів хвиль, протяжність яких складає приблизно 3м. «Погаснувши», атом через деякий час «спалахує» знову. Одночасно «спалахує» багато атомів. Збуджені ними цуги хвиль, накладаючись один з одним, утворюють світлову хвилю, яку випромінює тіло. Площина коливань вектора $\vec{E}$ для кожного цуга орієнтована випадковим чином. Тому в результуючій хвилі коливання вектора $\vec{E}$ в різних напрямках представленні з однаковою ймовірністю (рис.14.1). Такий рівномірний розподіл векторів $\vec{E}$ пояснюється великим числом атомарних випромінювачів, а рівність амплітудних значень векторів $\vec{E}$ – однаковою в середньому інтенсивністю випромінювання.

Природним називається світло з усіма можливими рівноймовірними орієнтаціями вектора $\vec{E}$ (і, отже, $\vec{H}$).

Поляризованим називається світло, в якому напрямки коливань вектора $\vec{E}$ якимось чином упорядковані.

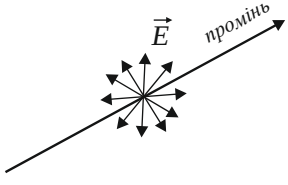


Рис.14.1.

Якщо коливання світлового вектора $\vec{E}$ відбуваються тільки в одній площині, яка проходить через промінь, то світло називається *плоско–* (або *лінійно–*) *поляризованим*. Упорядкованість може бути в тому, що вектор $\vec{E}$ повертається навколо променя, одночасно пульсуючи за величиною. В результаті кінець вектора $\vec{E}$ описує еліпс. Таке світло називається *еліптично-поляризованим*. Якщо кінець вектора $\vec{E}$ описує коло, світло називається *поляризованим по колу.*

Будь-яке світло можна представити, як суперпозицію двох плоскополяризованих у взаємно перпендикулярних площинах хвиль.

Розглянемо дві монохроматичні взаємно перпендикулярні хвилі, що поширюються вздовж додатнього напрямку осі *OX*.

$$E_y = E_{m1}\cos(\omega t - kx), \qquad (14.1)$$

$$E_z = E_{m2}\cos(\omega t - kx + \psi), \qquad (14.2)$$

де $\omega = 2\pi\nu$ – циклічна частота, $k = \omega/v$ — хвильове число, E_{m1} і E_{m2} –амплітуди хвиль, ψ – різниця фаз коливань E_y і E_z.

Щоб знайти траєкторію результуючого коливання світлового вектора при додаванні двох взаємно перпендикулярних коливань (§9.10), визначимо $\cos(\omega t - kx)$ із (14.1): $\cos(\omega t - kx) = \frac{E_y}{E_{m1}}$.

Тоді

$$\sin(\omega t - kx) = \sqrt{1 - \left(\frac{E_y}{E_{m1}}\right)^2}.$$

Оскільки

$$\cos(\omega t - kx + \psi) = \cos(\omega t - kx)\cos\psi - \sin(\omega t - kx)\sin\psi,$$ то

$$\frac{E_z}{E_{m2}} - \frac{E_y}{E_{m1}}\cos\psi = \sqrt{1 - \left(\frac{E_y}{E_{m1}}\right)^2}\sin\psi.$$

Піднесемо до квадрату це рівняння і в результаті отримаємо:

$$\frac{E_y^2}{E_{m1}^2} + \frac{E_z^2}{E_{m2}^2} - 2\frac{E_y E_z}{E_{m1}E_{m2}}\cos\psi = \sin^2\psi. \qquad (14.3)$$

Отримане співвідношення означає, що

результуючий вектор $\vec{E} = \vec{E}_y + \vec{E}_z$ (точніше кінець вектора $\vec{E}$) в кожній точці поля описує еліпс, який лежить в площині, що перпендикулярна до осі *OX* і довільно орієнтований відносно осей *OY* і *OZ*.

Якщо різниця фаз ψ зазнає випадкові хаотичні зміни (тобто хвилі (14.1) і (14.2) некогерентні) і, якщо $E_{m1} = E_{m2}$, то напрямок світлового вектора $\vec{E}$ буде зазнавати скачкоподібні неупорядковані зміни, а величина його буде однаковою у всіх напрямках перпендикулярних до напрямку променя. Таке світло буде природним. Подібне представлення природного світла значно спрощує розгляд його проходження через поляризаційні пристрої.

Якщо хвилі (14.1) і (14.2) некогерентні і різної інтенсивності ($E_{m1} \neq E_{m2}$), то в результаті суперпозиції таких хвиль отримаємо світло *частково поляризоване*. Таке світло можна розглядати як суміш природного і плоскополяризованого.

Якщо хвилі (14.1) і (14.2) когерентні, це означає, що різниця фаз ψ не залежить від часу і ψ =const.

Якщо $\psi = \pm (2m+1)\frac{\pi}{2}$, ($m$ = 0, 1, 2…), то із (14.3) отримаємо рівняння еліпса, орієнтованого відносно осей *OY* і *OZ*:

$$\frac{E_y^2}{E_{m1}^2} + \frac{E_z^2}{E_{m2}^2} = 1. \qquad (14.4)$$

Така хвиля (тобто світло) називається еліптично поляризованою.

При $E_{m1} = E_{m2} = E_m$ еліпс перетворюється в коло. Така хвиля (світло) називається *циркулярно поляризованою (поляризованою по колу)*.

Схематично на рис.14.2 зображено циркулярно поляризовану хвилю. Якщо відносно напрямку променя вектор $\vec{E}$ обертається проти годинникової стрілки, поляризація називається *правою*, в протилежному випадку – *лівою*. Якщо $\delta = \pm m\pi$ (m = 0, 1, 2, ...), то згідно (14.3) еліпс вироджується в пряму:

$$\left(\frac{E_y}{E_{m1}} \mp \frac{E_z}{E_{m2}}\right)^2 = 0, \quad E_y = \pm\frac{E_{m1}}{E_{m2}}E_z. \qquad (14.5)$$

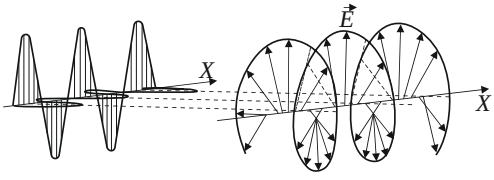


Рис.14.2.

Така хвиля називається *лінійно поляризованою (плоскополяризованою)* (рис.14.3).

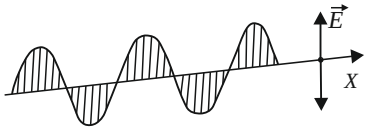


Рис.14.3.

Площина, яка проходить через напрямок коливань електричного вектора лінійно поляризованої світлової хвилі і напрямок розповсюдження цієї хвилі називається *площиною поляризації*.

За міру ступеня поляризації приймають вираз:

$$P = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \qquad (14.6)$$

де I_{max} і I_{min} – відповідно, максимальна та мінімальна інтенсивності світла, що відповідають двом перпендикулярним компонентам вектора $\vec{E}$. Для природного світла $I_{max} = I_{min}$ і P = 0. Для еліптично поляризованого світла поняття ступеня поляризації не застосовується (у такого світла коливання повністю впорядковані).

Поляризацією світла називається виділення лінійно поляризованого або частково

поляризованого світла із природного світла.

Плоско поляризоване світло можна отримати з природного за допомогою приладів, які називаються *поляризаторами*. Ці прилади вільно пропускають коливання вектора $\vec{E}$, які паралельні до площини поляризації, яка називається *головною площиною*, і повністю або частково затримують коливання вектора $\vec{E}$, які перпендикулярні цій площині. В ролі поляризаторів можуть бути, середовища, які анізотропні відносно коливань вектора $\vec{E}$, наприклад, кристали. Одним із природніх кристалів, які використовуються як поляризатори, може бути турмалін, природній і синтетичний монокристал – алюмосилікат. Турмалін оптично анізотропний (подвійне променезаломлення розглядається в подальшому), володіє *дихроїзмом* (грецьке *дихроїзм* означає *двокольорність*).

Прилади, за допомогою яких аналізують ступінь поляризації світла, називають *аналізаторами.*

Нехай на поляризатор падає природне світло (рис.14.4). Виберемо природню хвилю,

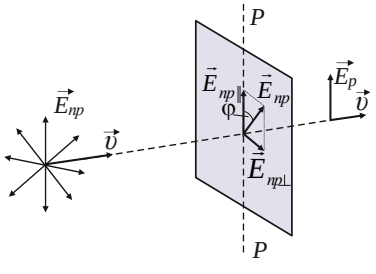


Рис.14.4.

вектор напруженості електричного поля якої $\vec{E}_{np}$ коливається у площині, що утворює з головною площиною поляризатора $P-P$ кут φ. При вході в поляризатор падаючу природню хвилю можна зобразити у вигляді суми двох коливань у взаємно перпендикулярних площинах:

$$E_{np\parallel} = E_{np}\cos\varphi, \quad E_{np\perp} = E_{np}\sin\varphi.$$

Перше коливання пройде через поляризатор, друге буде затримане. Інтенсивність хвилі, що пройшла, пропорційна до $E_{np\parallel}^2 = E_{np}^2\cos^2\varphi$, тобто дорівнює

$$I_P = I_{np}\cos^2\varphi, \qquad (14.7)$$

де I_{np} – інтенсивність коливань з амплітудою природного світла E_{np}.

В природному світлі значення φ рівноймовірні. Тому частка світла, що пройшла через поляризатор, буде дорівнювати середньому значенню $\cos^2\varphi$ тобто $1/2$ і

$$I_P = I_{np}/2. \qquad (14.8)$$

Під час обертання поляризатора навколо напрямку поширення природного світла інтенсивність світла, що проходить поляризатор, залишається однаковою згідно (14.8), а змінюється лише орієнтація його площини поляризації.

Нехай на аналізатор падає лінійно поляризоване світло, отримане за допомогою поляризатора, головна площина якого $P-P$ утворює кут φ з головною площиною аналізатора A-A (рис.14.5).

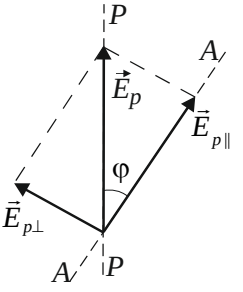


Рис.14.5.

При вході в аналізатор лінійно поляризованого променя, амплітуду його електричного вектора $\vec{E}_p$ можна представити як суму двох лінійно поляризованих променів, площини поляризацій яких взаємно перпендикулярні. Амплітуди вектора

напруженості електричного поля цих променів будуть

$$E_{P\perp} = E_P \sin\varphi \quad \text{і} \quad E_{P\parallel} = E_P \cos\varphi.$$

Оскільки аналізатор пропускає коливання електричного вектора, що відбуваються в площині *A-A*, а інтенсивність пропорційна квадрату амплітуди E_P, то інтенсивність світла після аналізатора буде:

$$I_A = I_P \cos^2\varphi. \qquad (14.9)$$

Якщо аналізатор не абсолютно прозорий, то

$$I_A = k_A I_P \cos^2\varphi, \qquad (14.10)$$

де k_A – коефіцієнт прозорості аналізатора.

Отримані співвідношення (14.9) і (14.10) виражають ***закон Малюса.***

Отже, інтенсивність світла, яке пройшло через аналізатор, змінюється від мінімуму при $\varphi = \pi/2$ до максимуму при $\varphi = 0$.

Якщо пропустити природне світло через поляризатор і аналізатор, площини яких утворюють кут φ, то з першого вийде плоскополяризоване світло, інтенсивність якого $I_P = \frac{1}{2} k_P I_{np}$, де k_P – коефіцієнт прозорості поляризатора, а з другого вийде світло, інтенсивність якого $I_A = k_A I_P \cos^2\varphi$. Отже, інтенсивність світла, яке пройшло через два поляризатори, $I_A = \frac{1}{2} k_P k_A I_{np} \cos^2\varphi$.

Звідси $I_{max} = \frac{1}{2} I_{np} k_P k_A$ ($\varphi = 0$ – поляризатори паралельні) і $I_{min} = 0$ ($\varphi = \pi/2$ – поляризатори схрещені). Нехай еліптично поляризоване світло падає на аналізатор, який пропускає складову $\vec{E}_{np\parallel}$ вектора $\vec{E}_P$ за напрямком площини *A-A* (рис.14.6). Максимальне значення $\vec{E}_{np\parallel}$ досягається в точках 1 і 2. Отже, максимальна амплітуда $\vec{E}_A$ плоскополяризованого світла, що вийде з приладу, буде дорівнювати довжині відрізка 01' або 02'. При повертанні аналізатора навколо напрямку променя спостерігається зміна інтенсивності в межах від I_{Amax}, яка отримується при збігу осі *A-A* з великою віссю еліпса, до I_{Amin}, яка має місце при збігу осі *A-A* з малою віссю еліпса.

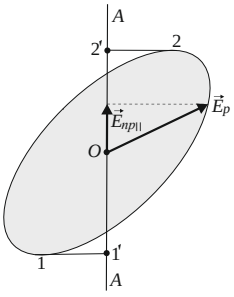


Рис.14.6.

У випадку світла, поляризованого по колу, обертання аналізатора не супроводжується зміною інтенсивності світла, яке пройшло через нього.

§14.2. Поляризація при відбиванні і заломленні

Під час проходження електромагнітної хвилі із одного середовища в інше відбувається її відбивання і заломлення. Виходячи з теорії електромагнітних хвиль Максвелла і враховуючи, що на межі середовищ тангенціальні складові векторів $\vec{E}$ і $\vec{H}$ хвилі задовільняють умові: $E_{\tau 1} = E_{\tau 2}$; $H_{\tau 1} = H_{\tau 2}$, можна отримати формули для інтенсивності відбитих і заломлених хвиль. Вперше такі формули були виведені Френелем у 1823 році при розгляді проходження пружної хвилі через межу двох середовищ. Розглянемо детальніше питання відбивання і заломлення електромагнітних хвиль на межі двох

діелектриків на основі аналізу формул Френеля. Нехай промінь природного світла падає під кутом α до межі двох діелектриків (наприклад скло-повітря),які характеризуються показниками заломлення n_1 і n_2 (чи відповідно діелектричними проникностями, $\varepsilon_1 = n_1^2$, $\varepsilon_2 = n_2^2$ (рис.14.7).

Природне світло зручно представити як суперпозицію двох лінійно поляризованих у взаємно перпендикулярних площинах, некогерентних, однакової інтенсивності хвиль. Світлові вектори (вектори напруженості електричного поля) цих хвиль $\vec{E}_{i\parallel}$ і $\vec{E}_{i\perp}$. Індекс «*i*» (від слова incident – падаючий) означає промінь, що падає, індекс «||» – означає, що вектор $\vec{E}_{i\parallel}$ лежить в площині рисунка, індекс « $\perp$ » – означає, що вектор $\vec{E}_{i\perp}$ пендикулярний до площини рисунка і позначений точкою. Для природного світла $|\vec{E}_{i\parallel}| = |\vec{E}_{i\perp}|$.

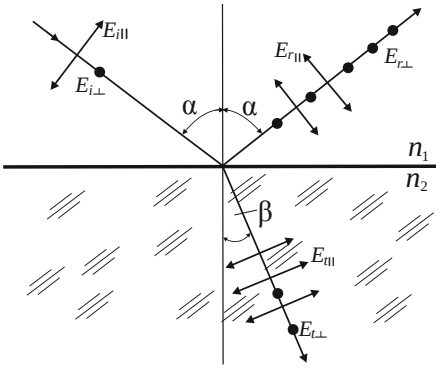


Рис.14.7.

На рис.14.7 індекс «*r*» (від слова reflection – відбивання) означає відбитий промінь, індекс «*t*» (від слова tranzit – проходження) означає промінь, що пройшов (заломлений промінь).

Формули Френеля встановлюють зв'язок між складовими амплітудами падаючою хвилі $E_{i\parallel}$ і $E_{i\perp}$, амплітудами відбитої хвилі $E_{r\parallel}$ і $E_{r\perp}$ і амплітудами $E_{t\parallel}$ і $E_{t\perp}$ хвилі, яка пройшла межу розділу двох середовищ. Ці формули такі:

$$E_{r\perp} = -E_{i\perp}\frac{\sin(\alpha-\beta)}{\sin(\alpha+\beta)}, \qquad (14.11)$$

$$E_{r\parallel} = E_{i\parallel}\frac{\mathrm{tg}(\alpha-\beta)}{\mathrm{tg}(\alpha+\beta)}, \qquad (14.12)$$

$$E_{t\perp} = E_{i\perp}\frac{2\sin\beta\cos\alpha}{\sin(\alpha+\beta)}, \qquad (14.13)$$

$$E_{t\parallel} = E_{i\parallel}\frac{2\sin\beta\cos\alpha}{\sin(\alpha+\beta)\cos(\alpha-\beta)}, \qquad (14.14)$$

де α – кут падіння, β – кут заломлення променя. Співвідношення між кутами α і β задається законом заломлення:

$$n_1\sin\alpha = n_2\sin\beta. \qquad (14.15)$$

Особливу цікависть має випадок, коли $\alpha+\beta=\pi/2$, тобто відбитий і заломлений промені взаємно перпендикулярні. Оскільки $\mathrm{tg}(\alpha+\beta)=\mathrm{tg}\pi/2=\infty$, то із формули (14.12) видно, що $E_{r\parallel}=0$. Тобто відбитий промінь є лінійно поляризований і містить тільки перпендикулярні до площини падіння складові напруженості електричного поля $E_{r\perp}$ (рис.14.8).

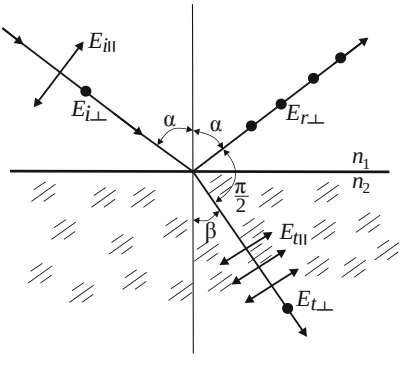


Рис.14.8.

Кут падіння α, при якому $\alpha+\beta=\pi/2$, називається кутом Брюстера

$\alpha_Б$ і закон (14.15) при цьому перепишеться так:

$$\operatorname{tg}\alpha_Б = n_2/n_1 . \qquad (14.16)$$

Він носить назву ***закону Брюстера.***

Таким чином, при падінні природного світла на межу розділу двох середовищ під кутом Брюстера $\alpha_Б$, відбите світло теоретично повинно бути повністю поляризованим. Проте на практиці це не так. Навіть при падінні променя світла під кутом $\alpha_Б$ у відбитому світлі $E_{r||} \neq 0$ і це означає, що воно є частково поляризованим, так як $E_{r\perp} > E_{r||}$. Формули (14.13) і (14.14) показують, що при проникненні у друге середовище фаза хвилі не зазнає скачків (знаки $\vec{E}_{i\perp}$ і $\vec{E}_{t\perp}$, $\vec{E}_{i\parallel}$ і $\vec{E}_{t\parallel}$ однакові).

Фазові співвідношення між відбитою і падаючою хвилями залежать від співвідношення між показниками заломлення n_1 і n_2 першого і другого середовища, а також від співвідношення між кутом падіння α і кутом Брюстера $\alpha_Б$.

Якщо $\alpha < \alpha_Б$, то відбивання від оптично більш густого середовища ($n_2 > n_1$) супроводжується скачком фази на π ("втрата півхилі"−знак мінус у формулі (14.11)); відбивання від оптично менш густого середовища ($n_2 < n_1$) відбувається без зміни фази. У випадку, коли $\alpha > \alpha_Б$, фаза для обох компонентів відбитої хвилі змінюється на π.

Фізичний зміст закону Брюстера.

Хвиля, яка падає, збуджує в середовищі II (рис.14.9) коливання електронів, які стають джерелом вторинних хвиль; ці хвилі і дають відбите світло. Напрямок коливань співпадає з напрямком електричного вектора світлової хвилі, тобто для середовища II він перпендикулярний до *OC*.

Ми можемо представити це коливання як суму двох коливань, одне із яких (*a*) на рис.14.9 лежить в площині *AOC*, а інше (*b*) – до нього перпендикулярне. Іншими словами, ми зображуємо коливання електронів в молекулі як суперпозицію коливань двох елементарних випромінювачів, осі яких напрямлені відповідно по a і b.

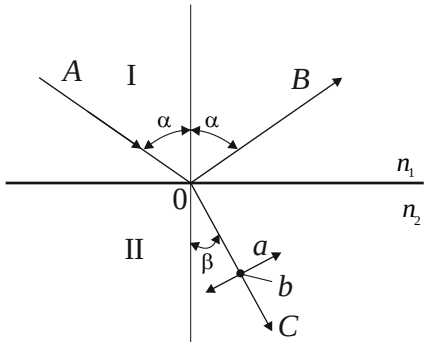


Рис.14.9.

Якщо світло падає під кутом Брюстера, тобто $\alpha + \beta = \pi/2$, то очевидно, що $OB \perp OC$. Значить $OB \parallel a$. Відомо, однак, що електричний заряд не випромінює електромагнітних хвиль вздовж напрямку свого руху. Тому випромінювач типу a вздовж *OB* не випромінює. Таким чином, по напрямку *OB* розповсюджується світло, яке посилає випромінювач типу b, напрямок коливань яких перпендикулярний до *OB*, тобто перпендикулярний до площини рисунка 14.9.

Іншими словами, відбите світло поляризоване і коливання вектора напруженості електричного поля в ньому перпендикулярні до площини падіння (закон Брюстера).

Якщо кут падіння відрізняється від кута Брюстера, то вздовж *OB* може розповсюджуватись хвиля, яка містить поряд з компонентою b і компоненту a, доля якої буде тим більша, чим більший кут між напрямком a і

напрямком відбитої хвилі. Таким чином, відбите світло буде частково поляризоване і ступінь поляризації зростає в міру наближення до кута Брюстера.

Проаналізуємо коефіцієнти відбивання для $E_\perp$ і $E_\parallel$. Згідно формули (14.11) і (4.12) коефіцієнт відбивання для $E_\perp$

$$r_\perp^2 = \left(\frac{E_{r\perp}}{E_{i\perp}}\right)^2 = \left(\frac{\sin(\alpha-\beta)}{\sin(\alpha+\beta)}\right)^2, \quad (14.17)$$

для $E_\parallel$

$$r_\parallel^2 = \left(\frac{E_{r\parallel}}{E_{i\parallel}}\right)^2 = \left(\frac{\mathrm{tg}(\alpha-\beta)}{\mathrm{tg}(\alpha+\beta)}\right)^2. \quad (14.18)$$

Для випадку нормального падіння світла на границю розділу двох середовищ ($\alpha = \beta = 0$) із формул (14.11) і (14.12), розкриваючи невизначеність, знаходимо амплітудні коефіцієнти відбивання:

$$r_\perp = \frac{E_{r\perp}}{E_{i\perp}} = -\frac{n_2-n_1}{n_2+n_1}, \qquad r_\parallel = \frac{E_{r\parallel}}{E_{i\parallel}} = \frac{n_2-n_1}{n_2+n_1}. \quad (14.19)$$

Для n_2=1,5, n_1=1 (скло-повітря) знаходимо, що

$$r_\perp^2 = r_\parallel^2 = 1/25 = 4\%.$$

Відбивання світла від багатьох поверхонь, навіть при падінні близькому до нормального, може значно послабити інтенсивність світла, що необхідно враховувати при побудові складних оптичних систем.

При куті падіння $\alpha = \alpha_Б$ $\mathrm{tg}(\alpha+\beta) \to \infty$ амплітудний коефіцієнт відбивання $r_{||}$ для $E_{r\parallel}$ дорівнює нулю, а $r_\perp$ визначається за формулою (14.11):

$$r_\perp = \frac{n_2^2-n_1^2}{n_2^2+n_1^2}. \quad (14.20)$$

Таким чином, при $\alpha = \alpha_Б$ відбите світло лінійно поляризоване в площині, яка перпендикулярна площині падіння, а коефіцієнт відбивання $r_\perp^2 = \left(\frac{n_2^2-n_1^2}{n_2^2+n_1^2}\right)^2 = 15\%$ (скло-повітря).

Якщо $\alpha \to \infty$ (ковзаюче падіння), то

$$r_\parallel^2 = r_\perp^2 = 1,$$

тобто має місце повне відбивання світла. Цим пояснюються яскраві зображення предметів в спокійній воді (береги річок, ліхтарі, місячна доріжка, сонце, яке заходить і т.п.).

На рис.14.10 зображені графіки залежності $r_\perp^2$ і $r_\parallel^2$ (криві I і III) від кута падіння α для n_2=1,52 і n_1=1. Кут Брюстера при цьому дорівнює 56°40′. Крива II відповідає коефіцієнту відбивання для неполяризованого світла. В цьому випадку $E_{i\parallel}^2 = E_{i\perp}^2$ і коефіцієнт відбивання

$$R = \frac{I_r}{I_i} = \frac{1}{2}(r_\parallel^2 + r_\perp^2), \quad (14.21)$$

де I_i – інтенсивність падаючого світла, I_r – інтенсивність відбитого світла. Із (14.21) видно, що коефіцієнт відбивання дорівнює середньому арифметичному із $r_\perp^2$ і $r_\parallel^2$.

Відбивання під кутом Брюстера дає змогу отримати лінійно поляризоване світло, однак його інтенсивність невелика і для скла (n=1,5) дорівнює близько 15%, тобто основна його частина поширюється у напрямку заломленої хвилі, яка поляризована частково.

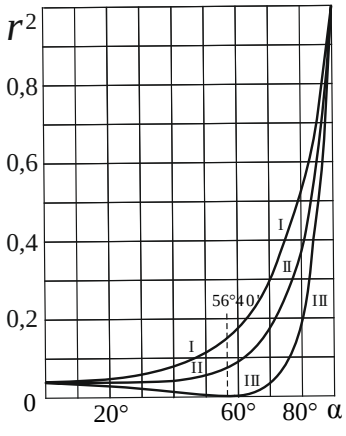


Рис.14.10.

Дійсно із формул (14.13) і (14.14) отримаємо, що для природного світла

$$\frac{E_{t||}}{E_{t\perp}} = \frac{1}{\cos(\alpha-\beta)} = \frac{1}{\sin 2\alpha_Б} > 1. \quad (14.22)$$

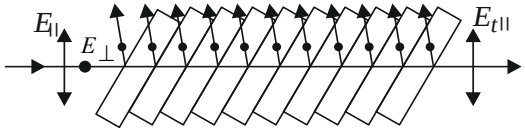


Рис.14.11.

При отриманні (14.22) враховано, що $E_{i\perp} = E_{i||}$ і $\alpha + \beta = \pi/2$. Формула (14.22) показує, що $E_{t\parallel} > E_{t\perp}$. Для збільшення ступеня поляризації заломлених хвиль їх треба пропустити крізь стопу скляних пластинок (рис.14.11).

Так, для $\alpha = \alpha_Б$ стопа з десяти скляних пластинок дає змогу отримати майже стопроцентну поляризацію заломлених хвиль.

§14.3. Подвійне променезаломлення

Проходження світла через анізотропну речовину, оптичні властивості якої в різних напрямках не однакові, супроводжуються рядом своєрідних явищ, які мають принципіальне і практичне значення. Особливості оптичних явищ в анізотропних середовищах пов'язані з тим, що індукований електромагнітною хвилею дипольний момент фізично нескінченно малого об'єму середовища, взагалі кажучи, не співпадає по напрямку з електричним полем хвилі. Так відбувається тому, що в анізотропній речовині під дією зовнішньої сили, елементарні заряди зміщуються в одних напрямках легше, ніж в інших.

Фізична природа анізотропії речовини зв'язана з особливостями будови її молекул або особливостями кристалічної решітки, у вузлах якої знаходяться атоми або іони.

При падінні світлової хвилі на границю *ізотропного середовища* в цьому середовищі від границі розповсюджується одна хвиля, яка лежить в площині падіння і яка складає з нормаллю до границі середовищ кут β, який визначається відомим законом заломлення: $n_1 \sin\alpha = n_2 \sin\beta$, де α – кут падіння, n_1 і n_2 – показники заломлення середовищ. Якщо середовище анізотропне, то в ньому в загальному випадку виникають дві хвилі, які розповсюджуються від межі в різних напрямках і з різними швидкостями. Це явище називається *подвійним променезаломленням.* Воно відкрите Бартоліном в 1670 році в кристалах *ісландського шпату* і в подальшому детально досліджене Гюйгенсом. Гюйгенс дав пояснення подвійного променезаломлення на основі гіпотези про те, що хвиля, яка падає на границю середовищ, породжує в кристалі елементарні вторинні хвилі двох видів: сферичні (*звичайні*) і еліптичні (*незвичайні*), швидкість поширення яких залежить від напрямку розповсюдження.

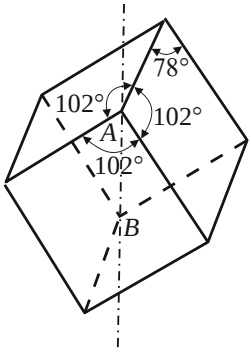


Рис.14.11.Кристал ісландського шпату.

Ісландський шпат – це різновидність вуглекислого кальцію $CaCO_3$ (кальцит), який відноситься до гексагональної системи. Він зустрічається в природі у вигляді досить великих і оптично чистих кристалів. Розколюючи його по певних площинах, кристалу можна надати форму ромбоедру – фігури, яку найлегше всього представити собі

як куб, дещо стиснутий вздовж просторової діагоналі *AB* (рис.14.11). Грані його мають вид ромбів з кутами 78^o і 102^o. У двох протилежних вершинах *A* і *B* сходяться сторони трьох тупих кутів.

Для демонстрації подвійного променезаломлення вузький паралельний пучок світла направляють перпендикулярно до грані природного ромбоедра. Із протилежної грані виходять два пучки, які мають напрямки, що паралельні початковому. Один із них є продовження первинного, а другий зміщений в сторону, тобто для нього кут заломлення відмінний від нуля, не дивлячись на те, що кут падіння дорівнює нулю. Ця обставина дала привід назвати другий пучок *незвичайним* (*e* – extraordinary), а перший, який підпорядковується закону заломлення – *звичайним* (*o*–ordinary). Якщо пучок, який падає, досить вузький, а кристал має достатню товщину, то пучки, що виходять із кристалу, просторово розділені (рис.14.12) і утворюють

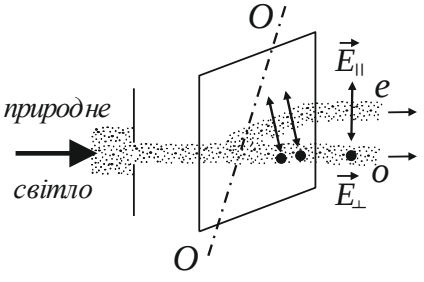


Рис.14.12.

дві плями на екрані. Якщо падаюче світло природне, то обидві плями мають однакову освітленість. При повороті кристалу навколо напрямку падаючого променя одна пляма залишається на місці, а друга обходить навколо неї. За допомогою аналізатора легко переконатись, що пучки, які виходять із кристала *лінійно поляризовані* у взаємно перпендикулярних напрямках (рис.14.12).

Збільшивши переріз падаючого пучка, можна отримати часткове перекриття плям на екрані. При повороті аналізатора частини плям, які не перекриваються, періодично стають світлими і темними, коли у однієї частинки освітленість максимальна, у другої – перетворюється в нуль. Освітленість частини плям, що перекриваються, залишається незмінною.

В кристалі ісландського шпату просторова діагональ *AB* ромбоедра (рис.14.11) є *віссю симетрії*: при обертанні на 120^o навколо цієї осі кристал суміщається сам з собою. Якщо спилити тупі кути у вершинах ромбоедра по площинах, які перпендикулярні осі, і відшліфовавши ці площини, можна дослідити розповсюдження світла в напрямку осі *OO*, яка парельна *AB* (рис.14.13).

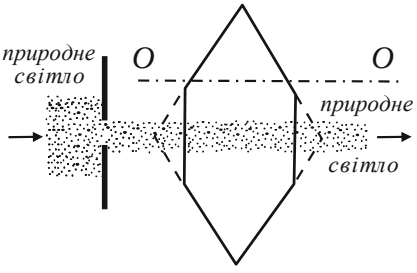


Рис.14.13.

Виявляється, що в цьому випадку подвійне променезаломлення відсутнє: падаючий пучок світла не роздвоюється і стан його поляризації не змінюється. Такий напрямок в кристалі, що має таку властивість, називається його *оптичною віссю*. Відзначимо, що мова йде саме про *напрямок*, а не про окрему пряму, так як пучок не буде зазнавати подвійного заломлення при розповсюдженні вздовж любої прямої, яка паралельна оптичній осі.

Кристали, які мають лише один напрямок, вздовж якого не відбувається

подвійне променезаломлення, називаються одновісними (ісландський шпат, кварц, турмалін та інші). Існують кристали, які мають два таких напрямки. Вони називаються двовісними (слюда, гіпс). У двовісних кристалах обидва промені незвичайні. В подальшому будемо розглядати одновісні кристали, які широко використовуються в оптичних експериментах.

Площина, яка проходить через промінь і оптичну вісь кристала, що перетинає промінь, називається *головною площиною,* або *головним* перерізом кристала, що відповідає цьому променю. Через кристал можна провести нескінченну множину паралельних оптичних осей і нескінченну множину паралельних головних перерізів. Лінія перетину двох довільних головних перерізів завжди є оптичною віссю.

Дослідження звичайного і незвичайного променів показує, що електричний вектор у звичайному промені перпендикулярний до площини головного перерізу і позначається (•), а в незвичайному промені лежить у площині головного перерізу і позначається (↕) (рис.14.12), тобто площина поляризації звичайного променя перпендикулярна до площини головного перерізу, а незвичайного збігається з площиною головного перерізу.

Після виходу з кристала, якщо не брати до уваги поляризацію у двох взаємно перпендикулярних напрямках (площинах), ці два промені нічим один від одного не відрізняються.

Пояснення подвійного променезаломлення

Подвійне променезаломлення пояснюється анізотропією кристалів. У кристалах некубічної системи діалектрична проникність ε є залежною від напрямку дії електричного поля $\vec{E}$. В одновісних кристалах ε в напрямку оптичної осі і в напрямку, який перпендикулярний до неї має різні значення $\varepsilon_{||}$ і $\varepsilon_{\perp}$, які називаються повздовжньою і поперечною діелектричними проникностями кристала. В інших напрямках ε має проміжне значення. А оскільки $n = \sqrt{\varepsilon}$ і $v = \frac{c}{n} = \frac{c}{\sqrt{\varepsilon}}$, то з анізотропії ε випливає, що світловим хвилям з різними напрямками коливання вектора $\vec{E}$ відповідають різні значення показника заломлення n і швидкості світлових хвиль v.

Виділимо в кристалі площину головного перерізу і розглянемо промені, що виходять з деякої точки *C* в різних напрямках в цій площині. На рис.14.14 напрямки, які паралельні до оптичної осі *OO*, зображені

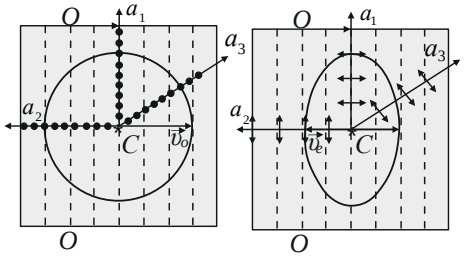


Рис.14.14. Рис.14.15.

пунктирними лініями, а напрямки коливань електричного вектора $\vec{E}$ показані точками. При довільному напрямку звичайного променя, наприклад, C_{a_1}, C_{a_2}, C_{a_3} вектор $\vec{E}$ утворює з оптичною віссю кристала прямий кут і швидкість буде однаковою в усіх напрямках і дорівнюватиме $v_o = \frac{c}{\sqrt{\varepsilon_{\perp}}}$. Геометричним місцем точок, до яких промені, що виходять з центру *C*, будуть доходити за однаковий проміжок часу буде коло (рис.14.14). Якщо розглянути сукупність всіх можливих головних перерізів, що проходять через точку *C*, то

геометричним місцем кінців всіх променів буде сфера.

Коливання вектора $\vec{E}$ в незвичайному промені здійснюються в площині головного перерізу (рис.14.15). Тому для різних напрямків коливання вектора $\vec{E}$ утворюють з оптичною віссю *OO* різні кути α. Для променя C_{a_1} кут $\alpha = \pi/2$ — і швидкість поширення коливань дорівнює $v_o = c/\sqrt{\varepsilon_\perp}$, для променя C_{a_2} кут α = 0 і швидкість $v_e = \frac{c}{\sqrt{\varepsilon_\parallel}}$. Для променя C_{a_3} швидкість має проміжне значення між v_o і v_e. У такому разі хвильова поверхня має вигляд еліпсоїда обертання навколо оптичної осі, переріз якого є однією з головних площин зображених на рис.14.15.

Одновісні кристали характеризуються показником заломлення звичайного променя, який дорівнює $n_o = c/v_o$, і показником заломлення незвичайного променя $n_e = c/v_e$. Значення n_e найбільше відрізняється від n_o для напрямку, який перпендикулярний до оптичної осі.

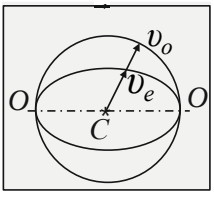


Рис.14.16.
Додатний

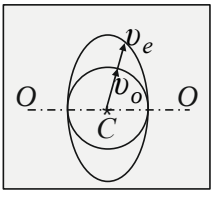


Рис.14.17
Від'ємний

Залежно від того, яка із швидкостей v_o чи v_e більша, розрізняють додатні і від'ємні одновісні кристали.

Якщо $v_o > v_e$ і відповідно $n_o < n_e$, то кристал називається оптично додатним і еліпсоїд вписаний у сферу (рис.14.16). Для від'ємного кристала $v_o < v_e$, $n_o > n_e$. У такому разі еліпсоїд описаний навколо сфери (рис.14.17).

Хід звичайних і незвичайних променів в кристалі можна визначити за допомогою принципу Гюйгенса (рис.14.18).

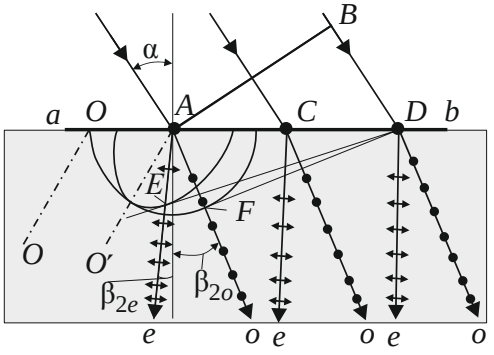


Рис.14.18.

Нехай на плоску поверхню *ab* одновісного оптично додатнього кристалу падає під кутом α плоска неполяризована світлова хвиля. Оптична вісь *OO* з поверхнею кристала утворює гострий кут і лежить в площині рисунка. В момент часу t фронт *AB* падаючої хвилі досяг точки *A* поверхні кристала. За час, поки промінь з точки *B* досягне точки *D* в кристалі з точки *A* виникнуть сферична та еліпсоїдальна хвильові поверхні. З проміжних точок, які лежать на поверхні кристала між *A* і *D*, також утворяться подібні поверхні, але менших розмірів. Згідно з принципом Гюйгенса прямі *DE* і *DF*, які дотикаються відповідно до еліпсоїдальної та сферичної поверхонь, будуть фронтами незвичайної і звичайної хвиль. Промені *Ao* і *Do* – звичайні. Вони нормальні до хвильового фронту *DF*; промені *Ae* і *De* – незвичайні, вони не перпендикулярні до фронту *DE*.

На рис.14.19 зображені три випадки нормального падіння світла на поверхню кристалу, які відрізняються напрямком оптичної осі *OO*. У випадку a промені *o* і *e* розповсюджуються вздовж оптичної осі, і тому

поширюються, не розділяючись. Із рис.14.19,*б* видно, що навіть при нормальному падінні світла на заломну поверхню незвичайний промінь може відхилитись від нормалі до цієї поверхні. На рис.14.19,*в* оптична вісь кристалу паралельна заломній поверхні. В цьому випадку при нормальному падінні світла

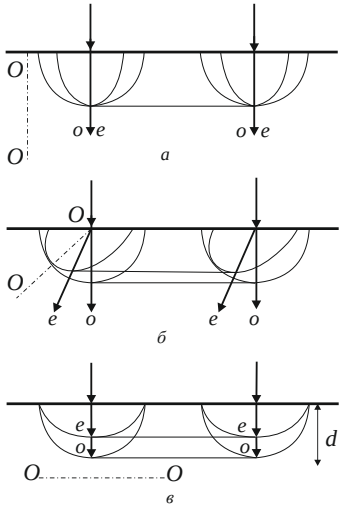


Рис.14.19.

звичайний і незвичайний промені поширюються по одному і тому ж напрямку, але розповсюджуються з різними швидкостями, внаслідок чого між ними виникає зростаюча різниця ходу, а відповідно і різниця фаз. Якщо товщина кристалу d, то оптична різниця ходу

$$\Delta = (n_e - n_o)d, \qquad (14.23)$$

а різниця фаз

$$\delta = 2\pi\frac{(n_e - n_o)d}{\lambda_0}, \qquad (14.24)$$

де λ_0 – довжина хвилі у вакуумі.

Таким чином, якщо пропустити природне світло через вирізану паралельно оптичній осі кристалічну пластинку, із пластинки вийдуть два поляризованих у взаємно перпендикулярних площинах промені o і e, між якими буде існувати різниця фаз (14.24).

Якщо світло, яке падає на пластинку є природним, то o і e промені, некогерентні, тому що вони містять хвилі, які належать різним цугам, що випромінюють атоми джерела світла.

Якщо на пластинку падає плоскополяризоване світло, то в цьому випадку коливання кожного цуга розділяється між звичайним і незвичайним променями в однаковій пропорції, яка залежить від орієнтації оптичної осі пластинки відносно площини коливань в падаючому промені. Тому промені o і e будуть когерентними і можуть інтерферувати.

В залежності від товщини пластинки різниця фаз (14.24) може приймати різні значення і світло на виході із пластинки може мати різну поляризацію (в загальному випадку еліптичну) із-за додавання двох взаємно перпендикулярно поляризованих хвиль.

§14.4.Поляризаційні призми і поляроїди

Явище подвійного променезаломлення лежить в основі роботи поляризаційних пристроїв, які служать для отримання поляризованного світла. Замість окремого кристалу для цієї цілі більш зручними є їх комбінації, які називаються *поляризаційними призмами.*

Призми ділять на два класи:

1)призми, що дають лише плоскополяризований промінь (поляризаційні призми);

2) призми, що дають два поляризовані у взаємно перпендикулярних площинах промені (двозаломні призми).

Найбільш придатний матеріал для їх виготовлення – ісландський шпат, у якого порівняно велика різниця між звичайними і

незвичайним показниками заломлення (n_o=1,658 ÷ 1,486; n_e=1,486) рідше застосовують кварц ($n_o = 1{,}544$; $n_e = 1{,}553$).

Перша поляризаційна призма була винайдена Ніколом в 1828р. Вона виготовлялась із отриманого розколюванням по площинах спайності куска ісландського шпату, довжина якого приблизно в 3,5 рази більша товщини. Торцеві основи *AC* і *BD* (рис.14.20) зішліфовувались під кутом 68° до його довгих ребер (замість 71° у природного кристала). Потім кристал розрізувався по площині, яка перпендикулярна до торцевих поверхонь (розріз *AB* перпендикулярний *AC*), і відшліфовані половинки склеювались в попередньому положенні канадським бальзамом, який утворював між ними тонкий прозорий шар з показником заломлення n=1,549, що має проміжне значення між n_o і n_e. Для звичайного променя канадський бальзам є оптично менш густе середовище, для

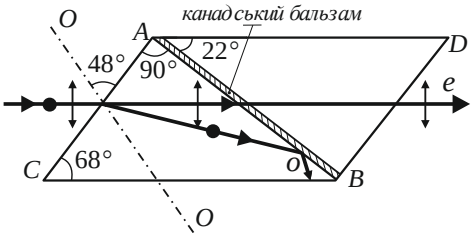


Рис.14.20.

незвичайного – більш густе. Тому звичайний промінь може пройти шар бальзаму тільки тоді, коли кут його падіння на площину розрізу менший граничного кута повного відбивання $\alpha_{гр}$, який визначається із умови $n_o \sin\alpha_{гр} = n \sin(\pi/2)$. В даному випадку $\sin\alpha_{гр} = n/n_o = 1{,}549/1{,}658$ і $\alpha_{гр} = 69^\circ$.

Оптична вісь *OO* спрямована під кутом 48° до вхідної грані *AC*. Промінь при падінні на грань *AC* внаслідок подвійного заломлення поділяється на звичайний промінь o і незвичайний промінь e. При певному виборі кутів призми звичайний промінь падає на шар бальзаму під кутом 76°, який більший за граничний (69°), зазнає повного внутрішнього відбивання, падає на зачорнену грань *CB* і нею поглинається. Незвичайний промінь виходить з призми паралельно до грані *CB*. Площина його поляризації збігається з площиною головного перерізу. Для незвичайного променя повне внутрішнє відбивання від шару канадського бальзаму виключене $(n_e < n)$.

Повне відбиття звичайного променя відбувається тільки тоді, коли кути падіння на вхідну грань призми лежать в певних межах. Кут між крайніми променями падаючого пучка (29°), який задовольняє цій умові, визначає *апертуру повної поляризації призми.*

Існує багато модифікацій призми Ніколя. Для роботи в ультрафіолетовій області спектру канадський бальзам не підходить із-за сильного поглинання світла, тому використовують *призму Фуко* з тонким повітряним прошарком. Повне відбиття звичайного променя в площині розрізу відбувається при менших кутах і тому призма Фуко виходить значно коротша призми Ніколя. Апертура повної поляризації призми Фуко складає всього $8°$.

Деколи необхідно із падаючого світла отримати відразу два промені, які поляризовані у взаємно перпендикулярних напрямках. Призми, які використовуються для цієї цілі, називаються двозаломними.

Прикладом двозаломної призми є *призма Волластона,* яка складається з двох прямокутних призм із ісландського шпату (рис.14.21).

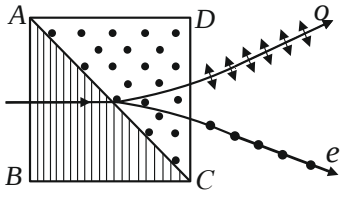


Рис.14.21.

У призмі *ABC* оптична вісь паралельна до катета *AB,* а в призмі *ACD* до ребра *C,* яке перпендикулярне до площини рисунка. Природне світло падає нормально на грань *AB;* обидва промені, звичайний і незвичайний, що виникають в призмі *ABC,* ідуть по одному напрямку, відповідно з швидкостями v_o і v_e. У іншій призмі вони також поширюються в напрямку, який перпендикулярний до оптичної осі, але оскільки оптичні осі в обох призмах взаємно перпендикулярні, то звичайний промінь першої призми перетворюється в незвичайний промінь в іншій і навпаки. Тому звичайний промінь в першій призмі заломиться на межі обох призм з відносним показником заломлення n_e/n_o, а промінь незвичайний – з n_o/n_e . Для ісландського шпату $n_o > n_e$і тому $n_e/n_o < 1$, а $n_o/n_e > 1$. Перший промінь заломиться в бік ребра *C* призми *ACD,* а другий — в бік її основи *AD.* Цим досягається значне розходження променів. Обидва промені плоско поляризовані.

Всі двозаломлені кристали тією чи іншою мірою поглинають світло. Коефіцієнт поглинання неоднаковий для звичайного і незвичайного променів і залежить від напрямку поширення світла в кристалі. Це явище називається *дихроїзмом*. Значний дихроїзм у видимій області спектра мають кристали турмаліну, в якому коефіцієнт поглинання для звичайних променів набагато більший, ніж для незвичайних. Пластинка турмаліну завтовшки 1 мм практично повністю поглинає звичайні промені і світло, яке проходить крізь неї, буде лінійно поляризованим. Але для деяких ділянок спектру незвичайний промінь теж зазнає відчутного поглинання, що обмежує застосування турмаліну як поляризатора.

Дуже зручні у використанні поляризаційні пристрої із синтетичних дихроїчних матеріалів, які відомі як *поляроїди*. Поляроїд – це плівка целлулоїда або іншого прозорого матеріалу, в яку вкраплені певним чином орієнтовані мікроскопічні кристалики сильно дихроїчної речовини, як правило, герапатиту (сульфат йодистого хініну) або інших споріднених йому сполук. Плівка завтовшки 0,1 мм повністю поглинає звичайний промінь видимої області спектру. Таким способом отримують листи великої площі, які забезпечують високий ступінь поляризації світла, яке проходить, при великих апертурах і мають порівняно невелику вартість. Недоліки поляроїдів пов'язані із спектральною селективністю поглинання герапатиту, через що фіолетова частина спектру є поляризована тільки частково, а плівка виявляється неоднаково прозорою для променів різних кольорів.

§14.5.Штучна оптична анізотропія

Подвійне променезаломлення спостерігається в природних анізотропних середовищах. Існують різні способи отримання штучної оптичної анізотропії, тобто надання оптичної анізотропії природно ізотропним речовинам.

Оптично ізотропні речовини стають оптично анізотропними під дією: 1)*одностороннього стиску або*

розтягу (кристали кубічної системи, скла та ін.) (Т.Зеєбек, Д.Брюстер (1815р));

2)*електричного поля* (рідини, аморфні тіла, гази) (Керр);

3)*магнітного поля* (рідини, скло, колоїди) (Коттон, Мутон).

1. Оптичну анізотропію, яка виникає під впливом деформації, можна виявити, якщо помістити досліджуване тіло *T* між поляризатором *P* і аналізатором *A*, які схрещені (рис.14.22). Доки тіло не деформоване, така система світла не пропускає. При односторонньому стиску або розтягу тіла вздовж напрямку *OO* в ньому виникає оптична анізотропія, яка еквівалентна анізотропії одноосного кристалу з оптичною віссю *OO*. Звичайний і незвичайний промені будуть поширюватися в напрямку, який перпендикулярний до *OO*, із різними швидкостями v_e і v_o. Якщо головний переріз поляризатора не паралельний і не перпендикулярний до *OO,* то світло, яке пройшло через деформоване тіло, стане еліптично поляризованим і його не можна погасити аналізатором.

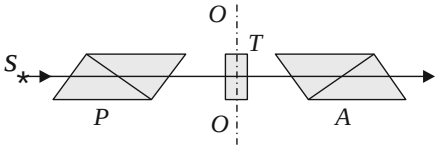


Рис.14.22.

Різниця коефіцієнтів заломлення $n_e - n_o = c/v_e - c/v_o$ може служити мірою анізотропії, що виникла. Досвід показує, що різниця $n_e - n_o$ пропорційна до напруги $\sigma = F/S$ в даній точці тіла (*F*−сила, *S*−площа, до якої прикладена перпендикулярно сила *F*):

$$n_e - n_o = k_1\sigma,$$

де k_1 − коефіцієнт пропорційності, що залежить від властивостей речовини. Різниця фаз, яку матимуть звичайний і незвичайний промені, пройшовши тіло завтовшки l дорівнює

$$\delta = \frac{2\pi l}{\lambda_0}(n_e - n_o) = 2\pi C\sigma l, \quad (14.25)$$

де λ_0 − довжина світлової хвилі у вакуумі, а $C = k_1/\lambda_0$ − новий коефіцієнт. Залежно від роду речовини коефіцієнт *C* може бути позитивним або негативним.

Явище штучної оптичної анізотропії при деформаціях використовують для виявлення внутрішніх залишкових напруг, які можуть виникати у виробах зі скла та інших прозорих ізотропних матеріалів внаслідок порушення технології їх виготовлення. Оптичний метод вивчення на прозорих моделях розподілу внутрішніх напруг у різних напружених частинах машин і споруд широко застосовують у сучасній техніці. Для цього використовують моделі, виготовлені з целулоїду або іншої ізотропної речовини.

Оскільки величина оптичної анізотропії пропорційна до напруги σ, то за виглядом смуг однакового кольору (ізохром), що виникають при спостереженні моделі між схрещеними ніколями, можна зробити висновок про величину напруг.

Оптична анізотропія може виникнути теж і в рідині під впливом зовнішніх динамічних впливів. Виникнення оптичної анізотропії в потоці може бути використана для вивчення властивостей полімерів, що йдуть на виготовлення штучних каучуків і пластмас.

2. У 1875 р. Д. Керр виявив, що рідкий або твердий ізотропний діелектрик, поміщений у дуже сильне однорідне електричне поле, стає оптично анізотропним. Це явище називають *ефектом Керра.* Принципову схему спостереження цього явища в рідинах зображено на рис.14.23, де *P* і *A* − поляризатор і схрещений з ним аналізатор. Між ними

розміщена кювета з конденсатором («комірка Керра»), між пластинами якого знаходиться

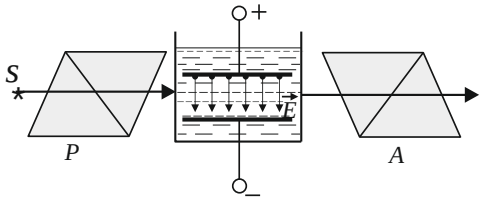


Рис.14.23.

досліджувана рідина. За відсутності електричного поля світло через систему не проходить. Досліди показали, що під дією однорідного електричного поля в плоскому конденсаторі рідина набуває властивостей одновісного двозаломлюючого кристала, оптична вісь якого збігається з напрямком вектора $\vec{E}$ напруженості поля конденсатора. Різниця показників заломлення рідини для звичайного і незвичайного променів монохроматичного світла в напрямку, який перпендикулярний до вектора $\vec{E}$ пропорційна E^2:

$$n_e - n_o = k_2 E^2, \qquad (14.26)$$

де k_2 – коефіцієнт пропорційності.

Такий характер залежності (14.26) пояснюється тим, що для ізотропного середовища (при відсутності поля) величина $n_e - n_o$ не може залежати від напрямку $\vec{E}$, тому розкладання $n_e - n_o$ в ряд по степеням напруженості поля повинно починатись з квадратичного члена. В слабких полях, які використовуються на досліді, в розкладанні досить обмежитись тільки цим членом.

Якщо довжина шляху променів між обкладками конденсатора l, то різниця фаз між звичайними і незвичайними променями дорівнює:

$$\delta = \frac{2\pi}{\lambda_0}(n_e - n_o)l = \\ = \frac{2\pi}{\lambda_0} l k_2 E^2 = 2\pi l B E^2, \qquad (14.27)$$

де $B = k_2/\lambda_0$ – стала Керра, яка залежить від природи речовини, довжини хвилі λ_0, температури і швидко зменшується із її збільшенням. Часто користуються іншою константою Керра K, яка зв'язана з B співвідношенням: $K = B\lambda_0/n$, n – абсолютний показник заломлення речовини за відсутності електричного поля.

У 1930 р. було виявлено існування ефекту Керра і в газах. Складність спостереження цього явища пов'язана з тим, що значення B для газів на кілька порядків менше, ніж для рідин. Для рідин стала Керра B =(10^{-14}÷10^{-12}) м/В^2, для газів B=(10^{-19}÷10^{-16}) м/В^2.

Велике значення (B = 2,2·10^{-12}) м/В^2 має нітробензол ($C_6H_5NO_2$), тому його часто використовують в технічних застосуваннях ефекта Керра.

Для більшості речовин стала Керра додатня, тобто n_e>n_o, що відповідає анізотропії додатного одновісного кристалу. Рідше зустрічаються речовини, у яких B < 0 (етиловий ефір, спирт).

Ефект Керра пояснюється різною поляризацією молекул за різними напрямками. За відсутності поля молекули орієнтовані довільно, тому рідина загалом не виявляє анізотропії. Під дією поля молекули повертаються так, щоб в напрямі поля були орієнтовані або дипольні електричні моменти (у полярних молекул), або напрям найбільшої поляризації (у неполярних молекул). В результаті речовина стає оптично анізотропною.

Ефект Керра практично безінерційний, тобто перехід речовини з

ізотропного стану в анізотропний (і назад) при вмиканні поля становить $\approx 10^{-10}$с . Тому цей ефект може бути ідеальним світловим затвором і застосовується в швидкоплинних процесах (звукозапис, відтворення звуку, швидкісне фото – і кінознімання), в оптичній локації.

3. Штучну анізотропію можна створити теж дією магнітного поля, яка спостерігається в речовинах, молекули яких анізотропні, тобто в парамагнетиках. За відсутності зовнішнього магнітного поля молекули розміщуються хаотично, результатом чого є статистична анізотропія. Якщо таку речовину помістити в досить сильне магнітне поле, то відбудеться напрямлена орієнтація власних магнітних моментів молекул. Це зумовлює анізотропію речовини, що приводить до подвійного заломлення променів. Таке середовище поводить себе як одновісній кристал, оптична вісь якого паралельна до вектора індукції поля $\vec{B}$. Це явище, відкрите в 1905 р., називається ефектом Коттона-Мутона. Різниця показників заломлення середовища при цьому

$$(n_e - n_o) = k_3 B^2, \quad (14.28)$$

де k_3 - коефіцієнт пропорційності, а різниця фаз між звичайним і незвичайним променями становитиме

$$\delta = \frac{2\pi}{\lambda_0}(n_e - n_o)l = \frac{2\pi}{\lambda 0}k_3 lB2 = 2\pi lDB^2, \quad (14.29)$$

де $D = k_3/\lambda_0$ - стала, яка залежить від природи речовини, довжини хвилі світла λ_o і температури.

4. Зміна оптичних характеристик кристала під дією *зовнішнього електричного* поля називається *електрооптичним ефектом Поккельса*. В одновісних кристалах розповсюдження світла вздовж оптичної осі відбувається з однією і тією ж фазовою швидкістю $v_o = c/n_o$ незалежно від напрямку поляризації. Якщо кристал не має центра симетрії, то при прикладанні зовнішнього електричного поля фазові швидкості хвиль з ортогональними напрямками поляризації вздовж осі стають різними. На відмінну від ефекта Керра, *квадратичного* по напруженості зовнішнього електричного поля, в електрооптичному ефекті різниця фазових швидкостей таких хвиль пропорційна напруженості поля (лінійний ефект Поккельса). Безінерційність ефекта Поккельса дозволяє широко використовувати його для створення швидкодіючих оптичних затворів і високочастотних модуляторів світла.

§14.6.Обертання площини поляризації

Явище повертання площини поляризації світлової хвилі на деякий кут при проходженні світла крізь кристалічні тіла і деякі ізотропні рідини, називається обертанням площини поляризації або оптичною активністю.

Якщо речовина не знаходиться у зовнішньому магнітному полі, то оптична активність буде природною.

Природна оптична активність була відкрита в 1811 р. Д. Араго на пластинках *кварцу,* які вирізані перпендикулярно до оптичної осі. Пізніше Ж.Біо виявив обертання напрямку поляризації світла в розчинах цукру. Тепер відомо багато природно-активних речовин, хоч у більшості з них це явище виражене дуже слабо.

Для спостереження оптичної активності шар речовини, яка досліджується, розміщують між встановленими на темноту (схрещеними) поляризатором і аналізатором. Світло, яке при цьому проходить, можна знову

погасити поворотом аналізатором. Це значить, що після проходження через шар активної речовини світло залишається лінійно поляризованим, але напрямок його поляризації повернутий на деякий кут відносно початкового напрямку.

Нехай погляд спостерігача спрямований назустріч падаючому променю. Обертання називають правим (додатним), якщо площина поляризації повертається вправо (за годинниковою стрілкою) для спостерігача, і лівим (від'ємним), якщо вона повертається вліво.

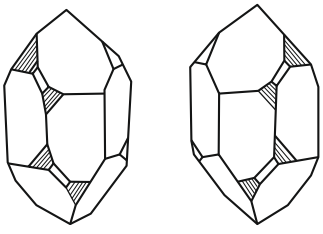


Рис.14.24.

Кристалічний кварц зустрічається у двох модифікаціях: для однієї обертання напрямку поляризації при розповсюдженні вздовж оптичної осі відбувається вправо, для другої – вліво (на той же кут). Для кристалічної гратки кварцу характерна відсутність центру і площин симетрії. Дві модифікації по своїй симетрії і зовнішній формі кристалів (рис.14.24) відрізняються одна від одної, як права рука відрізняється від лівої або як гвинт з правою нарізкою відрізняється від такого ж гвинта з лівою нарізкою. Ніяким переміщенням їх не можна сумістити одна із одною. Одна модифікація є дзеркальним зображенням іншої. Обертальна властивість кварцу зв'язана з його кристалічною структурою, так як плавлений кварц не володіє оптичною активністю.

Кут обертання площини поляризації пропорційний до товщини шару оптично активної речовини:

$$\varphi = \alpha l, \qquad (14.30)$$

де l – довжина шляху променя в оптично активному середовищі, α – коефіцієнт пропорційності , який називають обертальною здатністю або питомим обертанням. Він залежить від природи речовини, від температури та довжини хвилі.

Питоме обертання α дорівнює величині кута, на який повертається площина поляризації монохроматичного світла при проходженні шару завтовшки l=1м.

Далеко від смуг поглинання світла речовиною залежність α від λ задовільняє *закон Біо:* $\alpha \sim 1/\lambda^2$.

Для оптично активних рідин та розчинів Ж.Біо встановив, що кут повороту площини поляризації прямо пропорційний товщині шару і концентрації C оптично активної речовини, тобто

$$\varphi = [\alpha] Cl, \qquad (14.31)$$

де $[\alpha]$ – коефіцієнт пропорційності, який називається *питомим обертанням розчину*. Коефіцієнт $[\alpha]$ залежить від природи оптично активної речовини і розчинника, температури та довжини хвилі світла.

Властивості оптичної активності розчинів дають змогу визначити їх концентрації. Прилади, за допомогою яких проводять такі вимірювання, називаються *поляриметрами.* Оскільки для розчину цукру питоме обертання $[\alpha]$ значне, то поляриметри набули широкого застосування в цукрометрії.

Теорію обертання площини поляризації оптично активними речовинами розробив О.Френель. Він вважав, що це явище зумовлене особливим видом подвійного заломлення променів, при якому швидкість поширення світла в активному середовищі різна для променів, що мають праву і ліву колові

поляризації. Знак кута повертання площини поляризації визначається співвідношенням між швидкостями поширення променів правої циркулярної поляризації $v_{пр}$ і лівої циркулярної поляризації $v_л$. Для $v_{пр}>v_л$ оптично активне середовище буде додатним, а для $v_{пр}<v_л$ – від'ємним.

На вході в оптично активну речовину лінійно поляризоване монохроматичне світло розкладається на дві хвилі тіє ж частоти, але поляризовані по колу у взаємно протилежних напрямках (рис.14.25).

Вектори $\vec{E}_{пр}$ і $\vec{E}_л$ цих хвиль симетричні відносно площини $P-P$ коливань падаючого світла. Якщо $v_{пр}>v_л$, то при виході із оптично активного середовища з товщиною шару l електричний вектор правоциркулярної хвилі буде повернутий на більший кут $\varphi_{пр}$, ніж для лівоциркулярної хвилі $\varphi_л$. Внаслідок цього площина $P'-P'$, відносно якої електричні вектори цих хвиль розміщенні симетрично, буде повернута вправо на кут φ

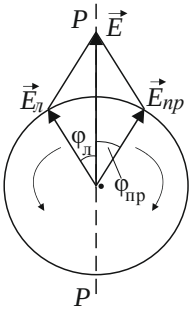


Рис.14.25.

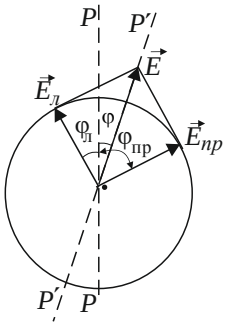


Рис.14.26.

відносно площини поляризації падаючої хвилі (рис.14.26). Кут визначається з умови $\varphi_{пр}-\varphi=\varphi_л+\varphi$ звідки отримаємо, що $\varphi=(\varphi_{пр}-\varphi_л)/2$. Кути повороту електричного вектора правої і лівої хвиль залежать від часу поширення хвилі t і довжини їх шляху в оптично активному середовищі, тобто

$$\varphi_{пр}=\omega\left(t-\frac{l}{v_{пр}}\right),\ \text{а}\ \varphi_л=\omega\left(t-\frac{l}{v_л}\right).$$

Тоді

$$\varphi=\frac{1}{2}\left[\omega\left(t-\frac{l}{v_{пр}}\right)-\omega\left(t-\frac{l}{v_л}\right)\right]=\frac{\omega l}{2}\left(\frac{1}{v_л}-\frac{1}{v_{пр}}\right).$$

Фазові швидкості поширення лівої і правої хвиль виразимо відповідно через коефіцієнти заломлення $n_л$ і $n_{пр}$, $v_{пр}=c/n_{пр}$ і $v_л=c/n_л$. Враховуючи, що $\omega=2\pi\nu=2\pi/T$, з попередньої формули отримаємо:

$$\varphi=\frac{2\pi l}{T\,2}\left(\frac{n_л}{c}-\frac{n_{пр}}{c}\right)=$$
$$=\frac{\pi l}{cT}(n_л-n_{пр})=\frac{\pi l}{\lambda_0}(n_л-n_{пр}),$$

де $\lambda_0=cT$ – довжина світлової хвилі у вакуумі. Отже, при $n_л>n_{пр}$ $(v_л<v_{пр})$ площина поляризації повертається вправо, а при $n_л<n_{пр}$ $(v_л>v_{пр})$ – вліво.

Повертання площини поляризації при проходженні світла через речовину, яка знаходиться в магнітному полі, вперше спостерігав Фарадей у 1846 р. Це був перший експеримент, який продемонстрував зв'язок оптичних і магнітних явищ. Фарадей так охарактеризував своє відкриття: «Мені вдалося намагнітити і наелектризувати промінь світла і освітити магнітну силову лінію» (Літ.№18, стор.619). Для спостереження ефекту Фарадея речовина, яка досліджувалась, розміщувалась між полюсами електромагніту (рис.14.27).

Лінійно поляризоване світло пропускається вздовж магнітного поля через канал, що просвердлений в осерді. Аналізатор A встановлюється так, що при відсутності магнітного поля світло не проходило (схрещенні поляризатор P і аналізатор A). При включені електромагніту світло проходило через аналізатор. Поворотом аналізатора його можна погасити. Це означає, що світло

залишається лінійно поляризованим і тут спостерігається саме поворот напрямку лінійної поляризації.

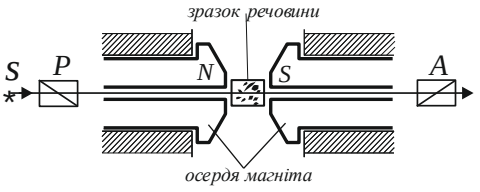


Рис.14.27.

Якщо спостерігач дивиться у напрямку магнітного поля, то повертання праворуч вважається додатним, ліворуч – від'ємним. Досліди М. Фарадея та М. Верде показали, що кут повертання площини поляризації пропорційний довжині шляху l променя у речовині і магнітній індукції B, тобто

$$\varphi = VlB, \qquad (14.33)$$

де V – стала Верде, яка залежить від природи речовини і довжини хвилі світла. Більшість речовин характеризується додатним обертанням.

Ефект Фарадея широко застосовується в наукових дослідженнях. Мала інерційність ефекту (час встановлення менше 10^{-9}с) дозволяє використовувати його для модуляції світла, для створення оптичного затвору і т.п.

Контрольні питання

1.Яке світло називається поляризованим ? Які є види поляризації світла?

2.Що таке ступінь поляризації?

3.Чому природне світло неполяризоване?

4.Як поляризуються відбиті від діелектрика хвилі?

5.Як поляризуються заломлені діелектриком хвилі?

6.Що встановлюють формули Френеля?

7.Що таке кут Брюстера?

8. Як визначити кут Брюстера?

9.Які поляризаційні прилади ви знаєте?

10.Які фізичні явища лежать в основі роботи поляризаційних приладів?

11.Будова і принципи дії призми Ніколя.

12.Будова і принцип дії поляроїдів.

13.Сформулювати і пояснити закон Малюса.

14.В чому суть явища обертання площини поляризації? Як пояснюється це явище?

15.Врахування явища повертання площини поляризації в техніці зв'язку та його використання в наукових дослідженнях.

Розділ15.Взаємодія електромагнітних хвиль з речовиною

§15.1.Дисперсія світла. Області нормальної і аномальної дисперсії

Дисперсією світла називається залежність показника заломлення n *речовини від частоти* ν *(довжина хвилі* λ*) світла або залежність фазової швидкості* v *світла в середовищі від його частоти* ν.

Дисперсію світла представляють у вигляді залежності $n = f(\lambda)$.

Наслідком дисперсії є розклад у спектр пучка білого світла при проходженні його через призму. Такий дослід вперше

виконав в 1672р. Ньютон. Схема досліду наведена на рис.15.1.

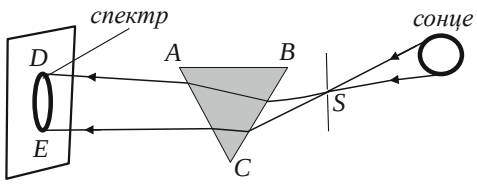


Рис.15.1.

Промінь світла від Сонця проходить через малий круглий отвір *S* у ставні вікна, потім заломлюється в скляній призмі *ABC* і падає на аркуш білого паперу. При цьому кругле зображення отвору *S* розтягується в кольорову смугу *DE,* яку Ньютон назвав спектром.

За допомогою призми, як і за допомогою дифракційної решітки, можна визначити спектральний склад світла.

Розглянемо відмінності в дифракційному і призматичному спектрах:

а) дифракційна решітка розкладає падаюче світло безпосередньо за довжинами хвиль, тому за виміряними кутами можна обчислити довжину хвилі.

Розклад білого світла в призмі відбувається за значеннями показника заломлення, тому для визначення довжини хвилі світла треба знати залежність

$$n = f(\lambda).$$

У дифракційному спектрі кольори розміщуються за порядком зростання довжини хвилі, а в дисперсійному – навпаки. Дифракційною решіткою червоні промені, що мають більшу довжину хвилі, ніж фіолетові, відхиляються сильніше.

На рис.15.2 наведені залежності показника заломлення n від довжини хвилі λ для скла (1), кварцу (2) і флюорита (3).

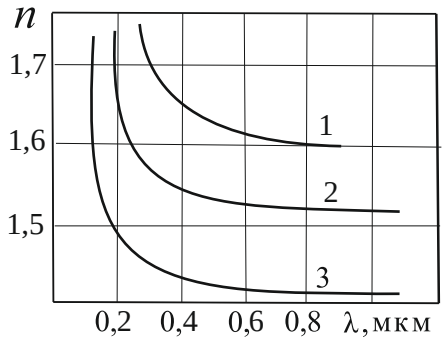


Рис.15.2.

Як видно з рис.15.2, показник заломлення n монотонно зростає із зменшенням довжини хвилі λ. Отже, червоні промені, що мають менший показник заломлення порівняно з фіолетовими променями, відхиляються призмою на менший кут.

б) дифракційний спектр рівномірно розтягнутий на всіх своїх ділянках. Дисперсійний же спектр білого світла стиснутий у червоній області і розтягнутий у фіолетовій, тому що показник заломлення скла в області коротких хвиль при зміні довжини хвилі випромінювання змінюється швидко, а в області довгих хвиль – повільно.

Залежність $n=f(\lambda)$ в оптичній області спектра має складний характер; для тих ділянок спектра, які слабо поглинаються даною речовиною, залежність показника заломлення від довжини хвилі може бути подана у вигляді ***формули Коші***:

$$n = n_0 + \frac{a}{\lambda^2} + \frac{b}{\lambda^4}$$

або

$$n \approx n_0 + \frac{a}{\lambda^2}\cdot \qquad (15.1)$$

Величина, яка показує, як швидко змінюється показник заломлення речовини з довжиною хвилі λ, називається **дисперсією речовини *D***:

$$D = \frac{dn}{d\lambda}. \qquad (15.2)$$

Дисперсія світла в середовищі називається **нормальною**, якщо із зростання частоти ν світла абсолютний показник заломлення n середовища також зростає:

$$\frac{dn}{d\nu} > 0, \quad \left(\frac{dn}{d\lambda} < 0\right).$$

Така залежність показника заломлення n від ν буде в тих областях частот, для яких середовище прозоре. Наприклад, звичайне скло прозоре для видимого світла і в цьому інтервалі частот має нормальну дисперсію.

Дисперсію світла в середовищі називають а**номальною,** якщо із зростанням частоти світла абсолютний показник заломлення середовища n зменшується:

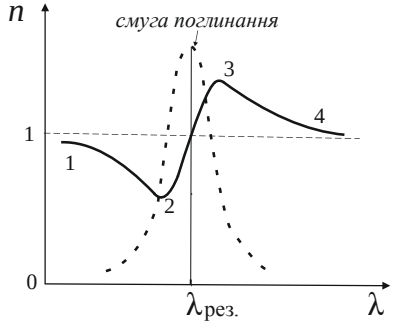


Рис.15.3.

$$\frac{dn}{d\nu} < 0, \quad \left(\frac{dn}{d\lambda} > 0\right).$$

Якщо речовина поглинає частину променів, то в області поглинання і біля неї хід дисперсії проявляє аномалію (рис.15.3). На деяких ділянках дисперсія речовини $dn/d\lambda$ є позитивною. Такий хід залежності n від λ називається *аномальною дисперсією.* Пунктирна крива на рис.15.3 зображує хід коефіцієнта поглинання світла речовиною. Ділянка 1-2 і 3-4 відповідають нормальній дисперсії $dn/d\lambda < 0$. На ділянці 2-3 дисперсія є аномальною $dn/d\lambda > 0$. В області 1-2 показник заломлення менший одиниці.

В різних ділянках спектра дисперсія характеризується тією зміною показника заломлення, яка припадає на одиничний інтервал довжин хвиль. Ця величина $\frac{\Delta n}{\Delta\lambda}$ називається ***середньою дисперсією*** для ділянки спектра $\lambda,\ \lambda + \Delta\lambda$..

У довідникових таблицях показники заломлення речовини даються для жовтої лінії натрію $\lambda_D = 0{,}5893$ мкм і позначаються n_D. Середня дисперсія визначається за синьою $\lambda_F = 0{,}4861$ і червоною $\lambda_C = 0{,}6563$ лініями водню і позначається $n_F - n_C$. Величина $(n_F - n_C)/(n_D - 1)$ називається ***відносною дисперсією,*** обернена їй величина – ***коефіцієнтом дисперсії.***

Оскільки показник заломлення визначається відношенням швидкостей світла у вакуумі і даному середовищі $n = c/v$, то, очевидно, в останньому швидкість червоного світла, яке заломлюється найменше, буде найбільшою, а фіолетового – найменшою. У вакуумі швидкість світла будь-якого кольору однакова.

Зауважимо, що відсутність дисперсії світла у міжзоряному просторі засвідчує, що цей простір можна вважати вакуумом. Коли б це було не так, дисперсія світла виявлялася б, наприклад, у спостереженнях затемнення подвійних зірок. На початку затемнення колір зірки змінювався б від нормального до синього, оскільки відповідне йому світло поширюється у прозорому середовищі з найменшою швидкістю, наприкінці затемнення – від червоного до нормального, оскільки швидкість червоного світла найбільша. Ще Д. Араго (1786 - 1853), спостерігаючи затемнення подвійних зірок Алголь, не виявив цих ефектів.

Із кількісного боку, дисперсія світла

задовільно описується класичною теорією. Суть її зводиться до врахуванні впливу на середовище вимушених коливань електронів, що збуджуються електричним полем світлової хвилі. Йдеться про зовнішні, так звані оптичні електрони атомів; електрони внутрішніх оболонок через великі власні частоти практично не збуджуються світлом.

§15.2.Групова швидкість світла

Оскільки показник заломлення n середовища залежить від частоти коливання ν світла, то швидкість поширення світла теж залежить від частоти коливань електромагнітної хвилі. А яку швидкість поширення світлового сигналу вимірюють дослідним шляхом?

Усі приймачі світла реагують на енергію, тому в дослідах вимірюють *швидкість перенесення енергії світловим сигналом і цю швидкість називають* ***груповою швидкістю.*** Вона відрізняється від швидкості v поширення фази, яка називається ***фазовою швидкістю*** (§10.3, формула (10.7)).

У реальних умовах дістати строго монохроматичну світлову хвилю неможливо. Виділяючи пучок світла одного кольору або модулюючи хвилю для передачі сигналу, фактично дістаємо пучок хвиль близьких, але все-таки різних частот. Тому у раніше розглянутих методах вимірювання швидкості світла ми змушені були оцінювати швидкість поширення групи хвиль різних частот (довжин). Винятком можуть бути лише непрямі вимірювання швидкості світла. Це трапляється, наприклад, коли визначають показник заломлення середовища, а потім за формулою $n = c/v$ знаходять фазову швидкість світла.

Як відомо, у разі накладання хвиль близьких частот виникають биття, тобто пульсації максимумів результуючої амплітуди, й отже, значень енергії. Це наочно можна бачити, якщо зобразити накладання двох хвиль з близькими частотами (довжинами хвиль) (рис.15.4). У точках *A* і *C* хвилі перебувають

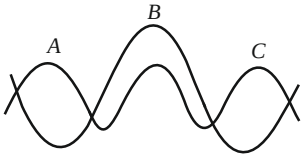


Рис.15.4.

у протифазах, тому результуюча амплітуда й енергія мінімальні; у точці *B* хвилі за фазою збігаються, тому виникає максимум амплітуди й енергії хвильового процесу.

Зауважимо, що у вакуумі внаслідок однакових швидкостей світла різних частот взаємне розміщення хвиль залишається незмінним, тому швидкості поширення фаз і максимуму амплітуди хвильового процесу однакові; фазова й групова швидкості світла теж однакові. У разі переходу світлового пучка у дисперсійне середовище монохроматичні хвилі різних частот поширюватимуться з різними фазовими швидкостями, тому максимум амплітуди хвильового процесу весь час змінюватиме своє положення у смузі складових хвиль. Його швидкість у просторі не збігатиметься зі швидкістю жодної зі складових хвиль. У цьому випадку питання про групову швидкість світла ускладнюється.

Щоб оцінити групову швидкість світла в дисперсійному середовищі, розглянемо накладання двох світлових хвиль із близькими частотами й однаковими амплітудами.

$$E_1 = E_m \cos(\omega_1 t - k_1 x) = \\ = E_m \cos 2\pi(\nu_1 t - x/\lambda_1),$$

$$E_2 = E_m \cos(\omega_2 t - k_2 x) = \\ = E_m \cos 2\pi(\nu_2 t - x/\lambda_2).$$

Їх результуюче зміщення E в деякий момент часу t у точці з координатою x виразиться сумою

$$E = E_m\cos2\pi\left(\nu_1 t - \frac{x}{\lambda_1}\right) + \\ + E_m\cos2\pi\left(\nu_2 t - \frac{x}{\lambda_2}\right),$$

або

$$E = \left[2E_m\cos2\pi\left(\frac{\nu_2-\nu_1}{2}t - \frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}x\right)\right]\cdot \\ \cdot\cos2\pi\left(\frac{\nu_1+\nu_2}{2}t - \frac{\lambda_1+\lambda_2}{2\lambda_1\lambda_2}x\right).$$

Цей вираз можна розглядати як рівняння плоскої хвилі, в якому множник у квадратних дужках є амплітудою хвилі:

$$A = 2E_m\cos2\pi\left(\frac{\nu_2-\nu_1}{2}t - \frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}x\right). \quad (15.3)$$

Якщо в момент часу t_1 максимум амплітуди хвильового процесу знаходиться в точці з координатою x_1, а в момент часу t_2 – в точці з координатою x_2, то групова швидкість світла

$$v_{гр} = \frac{x_2 - x_1}{t_2 - t_1}.$$

Оскільки ми розглядаємо переміщення фіксованого значення амплітуди (15.3), то із (15.3) дістаємо рівняння:

$$\cos2\pi\left(\frac{\nu_2-\nu_1}{2}t_1 - \frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}x_1\right) = \\ = \cos2\pi\left(\frac{\nu_2-\nu_1}{2}t_2 - \frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}x_2\right).$$

Звідки

$$\left(\frac{\nu_2-\nu_1}{2}t_1 - \frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}x_1\right) = \\ = \left(\frac{\nu_2-\nu_1}{2}t_2 - \frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}x_2\right),$$

або

$$\frac{\lambda_1-\lambda_2}{2\lambda_1\lambda_2}(x_2-x_1) = \frac{\nu_2-\nu_1}{2}(t_2-t_1).$$

Знайшовши з останньої залежності $v_{гр}$ і, виразивши далі частоти через відповідні швидкості і довжини хвиль, отримаємо:

$$\begin{cases} v_{гр} = \dfrac{x_2-x_1}{t_2-t_1} = \\ = \dfrac{v_2\lambda_1 - v_1\lambda_2 + (v_1\lambda_1 - v_1\lambda_1)}{\lambda_1-\lambda_2} = \\ = v_1 - \lambda_1\dfrac{v_2-v_1}{\lambda_2-\lambda_1}, \end{cases}$$

або

$$v_{гр} = v_1 - \lambda_1\frac{\Delta v}{\Delta\lambda},$$

де v_1 – фазова швидкість складової хвилі.

Отже, знайдено залежність між груповою і фазовою швидкостями світла, яку записують так:

$$v_{гр} = v - \lambda\frac{dv}{d\lambda}. \quad (15.4)$$

Якщо врахувати, що

$$\frac{dv}{d\lambda} = \frac{d}{d\lambda}\left(\frac{c}{n}\right) = -\frac{c}{n^2}\frac{dn}{d\lambda},$$

то формулу (15.4) можна переписати так:

$$v_{гр} = v\left(1 + \frac{\lambda}{n}\frac{dn}{d\lambda}\right). \quad (15.5)$$

Якщо

$$dn/d\lambda = 0$$

(відсутня дисперсія), то $v_{гр} = v$, тобто фазова і групова швидкості збігаються. Цей випадок має місце для вакууму.

При нормальній дисперсії

$$\frac{dn}{d\lambda} < 0 \ \left(\frac{dn}{dv} > 0\right), \text{тому} \ \ v_{гр} < v.$$

При аномальній дисперсії

$$\frac{dn}{d\lambda} > 0 \ \left(\frac{dn}{dv} < 0\right) \ \text{і} \ \ v_{гр} > v.$$

Поняття групової швидкості застосовне лише за умови, що поглинання енергії хвилі в даному середовищі невелике. При значному згасанні хвилі поняття групової швидкості втрачає сенс. Саме в області аномальної дисперсії поглинання дуже велике, і поняття групової швидкості не можна використовувати.

Наочно картину поширення групи хвиль при нормальній дисперсії можна описати так. Врахуємо, що $\frac{dn}{d\lambda} < 0$. Це означає, що для $\lambda_2 > \lambda_1$, $n_2 < n_1$ і $v_2 > v_1$. Тоді, якби у початковий момент часу максимум амплітуди знаходився посередині смуги хвиль, то у подальшому їхньому русі спостерігалося б поступове ослаблення, а потім і зникнення передніх хвиль; у голову смуги хвиль пік амплітуди хвильового процесу повільно переміщувався б, слабшав і згодом зникав би, одночасно позаду смуги хвиль формувався, повільно переміщувався і виростав би новий пік амплітуди і т.д. Оскільки передні хвилі весь час зникають, а наступний пік амплітуди формується позаду смуги хвиль, то це є свідченням того, що групова швидкість менша за фазові швидкості складових хвиль.

Групова швидкість світла в усіх випадках не може перевищувати швидкості світла у вакуумі.

§15.3.Поглинання світла

Поглинанням світла *називається явище втрати енергії світловою хвилею, яка проходить через речовину, внаслідок перетворення енергії хвилі у інші форми енергії.*

При проходженні паралельного пучка світла крізь шар прозорого середовища його інтенсивність зменшується. Поглинання світла може приводити до нагрівання, іонізації або збудження атомів і молекул речовин, до деформації, поглинання може супроводжуватись розсіянням світла та індуктивним (вимушеним) випромінюванням.

Щоб отримати співвідношення, яке виражає закон поглинання світла, розглянемо шар прозорого середовища завтовшки l, на який падає паралельний пучок променів, інтенсивність якого I_0. Виділимо в середовищі нескінченно тонкий шар dl, який обмежений паралельними поверхнями, що перпендикулярні до напрямку поширення світла (рис.15.5).

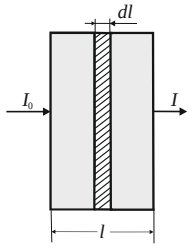


Рис.15.5.

Дослід показує, що зменшення інтенсивності світла шаром середовища dl пропорційне до величини інтенсивності, що входить у цей шар, і товщині шару, тобто

$$dI = -\alpha I dl,$$

де α – коефіцієнт пропорційності, який не залежить від інтенсивності світла і називається коефіцієнтом поглинання. Знак мінус вказує на те, що із збільшенням товщини шару середовища, яке поглинає світло, інтенсивність світла, що проходить крізь нього, зменшується.

Після розділення змінних у рівнянні дістаємо $\frac{dI}{I} = -\alpha dl.$

Проінтегруємо це рівняння:

$$\int_{I_0}^{I} \frac{dI}{I} = -\alpha \int_{0}^{I} dl; \ \ln\frac{I}{I_0} = -\alpha l.$$

В результаті отримаємо:

$$I = I_0 e^{-\alpha l}, \qquad (15.6)$$

де I – інтенсивність світла, що виходить із шару поглинаючого середовища завтовшки l; I_0 – інтенсивність світла, що входить у поглинаюче середовище. При $l = 1/\alpha$ інтенсивність $I = I_0/e$.

Отже, шар, товщина якого дорівнює $1/\alpha$, зменшує інтенсивність світла в $e = 2{,}72\ldots$ рази. Отримане співвідношення встановлене у 1729р. П. Бугером і називається законом Бугера, або ***законом Бугера-Ламберта.*** А. Бер встановив, що поглинання світла розчинами пропорційне молекулярній концентрації C_0 розчиненої речовини, тобто

$$\alpha = \alpha_0 C_0, \qquad (15.7)$$

де α_0 – коефіцієнт пропорційності, який залежить від природи розчиненої речовини і не залежить від її концентрації.

Тоді ***закону Бугера-Ламберта-Бера***, який справедливий для газів і розчинів малих концентрацій, можна надати вигляду

$$I = I_0 e^{-\alpha_0 C_0 l}. \qquad (15.8)$$

Коефіцієнт поглинання α залежить від довжини хвилі λ (або частоти $\omega=2\pi\nu$) і від хімічної природи речовини.

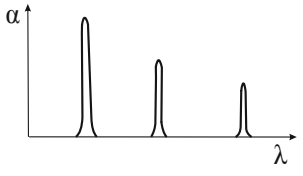


Рис.15.6.

В одноатомних газах і парах металів, у яких атоми розміщені на значних відстанях один від одного і їх можна вважати ізольованими, коефіцієнт поглинання для більшості довжин хвиль близький до нуля і лише для дуже вузьких спектральних областей спостерігаються різкі максимуми. Такий спектр поглинання називається ***лінійчастим*** (рис.15.6). Ці лінії відповідають частотам власних коливань електронів в атомах.

У газів з багатоатомними молекулами спостерігаються системи тісно розміщених ліній, які утворюють ***смуги поглинання.*** Структура цих смуг визначається складом і будовою молекул.

Рідкі й тверді діелектрики мають ***суцільні спектри поглинання,*** що складаються з порівняно широких смуг поглинання, в межах яких коефіцієнт поглинання змінюється плавно. За межами цих смуг $\alpha \approx 0$, тобто діелектрики прозорі.

Метали практично непрозорі для світла. Коефіцієнт α для них має значення порядку $10^2 м^{-1}$, в той час як для скла $\alpha \approx 1 м^{-1}$. Це обумовлено наявністю в металах вільних електронів. Під дією електричного поля світлової хвилі вільні електрони починають рухатись і в металі виникають швидкозмінні струми, що супроводжуються виділенням теплоти. В результаті енергія світлової хвилі швидко зменшується і перетворюється у внутрішню енергію металу.

Структура спектрів поглинання визначається складом і будовою молекул, тому вивчення спектрів поглинання є одним з основних методів кількісного і якісного дослідження речовин.

§15.4.Розсіяння світла

З класичної точки зору процес розсіяння світла полягає в тому, що світло, яке проходить через середовище, викликає коливання електронів в атомах. Ці електрони, що коливаються, збуджують вторинні хвилі, які розповсюджуються у всіх напрямках. Це явище, здавалось би, повинно при всіх умовах приводити до розсіяння світла. Однак вторинні хвилі є когерентними, так що необхідно врахувати їх взаємну інтерференцію.

Відповідний розрахунок дає, що у випадку однорідного середовища вторинні хвилі повністю гасять одна одну у всіх напрямках, крім напрямку розповсюдження первинної хвилі. Тому перерозподіл світла по

напрямках, тобто розсіяння світла, не відбувається.

Вторинні хвилі не гасять одна одну в бокових напрямках тільки при розповсюдженні світла в неоднорідному середовищі. Світлові хвилі, зазнаючи дифракції на неоднорідностях середовища, дають дифракційну картину, яка характеризує досить рівномірний розподіл інтенсивності у всіх напрямках. Таку дифракцію на дрібних неоднорідностях називають ***розсіюванням світла.***

Середовища з явно вираженою оптичною неоднорідністю носять назву ***каламутних*** середовищ. До них належать:

1)дим, тобто наявність в газах дрібних твердих частинок, що висять;

2)тумани – гази, в яких висять дрібні крапельки рідини;

3)суспензії, що утворені твердими частинками, які плавають в рідині;

4)емульсії, тобто наявність дрібних краплинок однієї рідини в іншій, які висять в ній і не розчиняються (прикладом емульсії може бути молоко, яке можна представити як воду, в якій висять крапельки жиру;

5)тверді тіла, як то перламутри, опали, молочне скло і т.п.

Світло, яке розсіяне на частинках, розміри яких значно менші довжини світлової хвилі, є *частково поляризованим.* Це пояснюється тим, що коливання електронів, яке викликане розсіяним світловим пучком, відбуваються в площині, яка перпендикулярна до пучка (рис.15.7).

В результаті розсіяння світла в бокових напрямках інтенсивність в напрямку розповсюдження зменшується швидше, ніж у випадку одного тільки поглинання. Тому для каламутної речовини у виразі (15.6), поряд з коефіцієнтом поглинання α, повинен знаходитись додатковий коефіцієнт α', обумовлений розсіянням:

$$I = I_0 e^{-(\alpha+\alpha')l}. \qquad (15.9)$$

Сталу α' називають ***коефіцієнтом екстинції.***

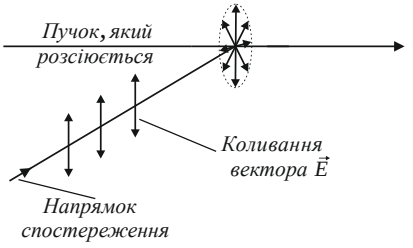


Рис.15.7.

Якщо розміри неоднорідностей малі порівняно з довжиною хвилі (не більші $\sim 0{,}1\lambda$), інтенсивність розсіяного світла I пропорційна четвертій степені частоти або обернено пропорційна четвертій степені довжини хвилі.

$$I \sim \omega^4 \sim \frac{1}{\lambda^4}. \qquad (15.10)$$

Ця залежність носить назву ***закону Релея.*** Її походження легко зрозуміти, якщо врахувати, що потужність випромінювання заряду, який коливається, пропорційна четвертій степені частоти (§11.5, формула (11.41)) і відповідно, обернено пропорційна четвертій степені довжини хвилі.

Якщо розміри неоднорідностей порівняльні з довжиною хвилі, електрони, що знаходяться в різних місцях неоднорідності, коливаються із суттєвим зсувом по фазі. Ця обставина ускладнює явище і призводить до інших закономірностей – інтенсивність розсіяного світла стає пропорційною тільки квадрату частоти (обернено пропорційною тільки квадрату довжини хвилі).

Прояв закономірностей (15.10) легко спостерігати, якщо пропускати пучок білого світла через посудину з каламутною рідиною (рис.15.8).

Внаслідок розсіяння слід пучка в рідині добре видно збоку, причому, оскільки короткі світлові хвилі розсіюються значно сильніше довгих, то цей слід стає голубуватим. Пучок, який проходить через рідину, стає збагаченим довгохвильовим випромінюванням і утворює на екрані E не білу, а червоно-жовту пляму. Якщо поставити на вході пучка в посудину поляризатор *P*, ми побачимо, що інтенсивність розсіяного світла в різних напрямках, які перпендикулярні до первинного пучка, неоднакова. Напрямок випромінювання диполя (див.рис.11.11) приводить до того, що в напрямках, які співпадають з площиною коливань первинного пучка, інтенсивність розсіяного світла практично дорівнює нулю, в напрямках, що перпендикулярні до площини коливань, інтенсивність розсіяного світла максимальна. Повертаючи поляризатор навколо напрямку первинного пучка, ми будемо спостерігати поперемінне підсилення та ослаблення світла, яке розсіюється в даному напрямку.

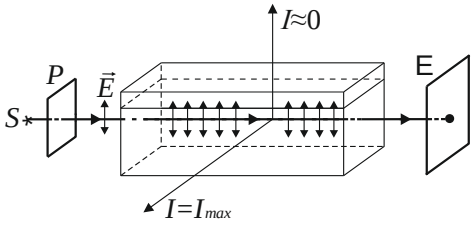


Рис.15.8.

Навіть старанно очищені від сторонніх домішок і забруднень рідини і гази в деякій мірі розсіюють світло. Встановлено, що причиною появи оптичних неоднорідностей є в цьому випадку флуктуації густини (тобто відхилення густини від її середнього значення, яке спостерігається в межах малих об'ємів). Ці флуктуації викликані хаотичним рухом молекул речовини. Тому розсіяння світла, яке обумовлене ним, називається ***молекулярним.***

Молекулярним розсіянням пояснюється голубий колір неба. Місця згущення і розрідження повітря, які безперервно виникають в атмосфері внаслідок хаотичного молекулярного руху, розсіюють сонячне світло. При цьому, згідно закону (15.10), голубі і сині промені розсіюються сильніше, ніж жовті і червоні, чим і обумовлюють голубий колір неба. Коли Сонце знаходиться низько над горизонтом, промені, які розповсюджуються безпосередньо від нього, проходять більшу товщину середовища, яке розсіює, в результаті чого вони стають збагаченими більшими довжинами хвиль. З цієї причини небо на зорі забарвлене в червоні тони.

§15.5.Ефект Доплера в оптичних явищах

Ефект Доплера для звукових хвиль, існування яких можливе в пружних середовищах, розглянутий в §10.9. Перш ніж перейти до розгляду ефекту Доплера, в оптичних явищах зазначимо таке: у класичній фізиці з гіпотези про існування ефіру автоматично випливало, що для світлових хвиль має виявлятися ефект Доплера так само, як і для звукових хвиль. Проте в подальшому було з'ясовано, що ніякого ефіру, який міг би бути абсолютною системою відліку, не існує, що зміст має лише відносний рух джерела і приймача один відносно одного, що системою відліку може бути лише приймач або джерело світла.

З викладеного випливає, що формули, які відображають ефект Доплера в оптичних явищах, мають бути єдиними,

незалежними від того, що саме переміщується – джерело чи приймач. Для знаходження таких формул керуватимемось принципом відносності, згідно з яким закони природи мають однаковий вигляд в усіх інерціальних системах відліку.

Оберемо для системи відліку S приймач, а для системи S' – джерело світла. Розмістимо приймач і джерело у початках відповідних систем координат, осі яких X і X' збігаються зі швидкістю переміщення джерела v (рис.15.9).

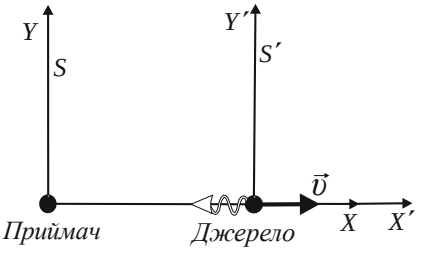


Рис.15.9.

Запишемо в системі S' рівняння плоскої світлової хвилі, що поширюється від джерела до приймача:

$$E' = E'_m \cos\omega'(t' + \frac{x'}{c}), \qquad (15.11)$$

де ω' – циклічна частота хвилі в системі S', c– швидкість світла, однакова у всіх системах відліку.

Згідно з принципом відносності, в системі S рівняння хвилі, що розглядається, матиме вигляд, аналогічний формулі (15.11):

$$E = E_m \cos\omega\left(t + \frac{x}{c}\right), \qquad (15.12)$$

де ω – циклічна частота хвилі, яку фіксує приймач у системі S.

Однак від рівняння (15.11) можна перейти до рівняння (15.12) , якщо за формулами перетворень Лоренца виразити координати x' і t' через x і t:

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Тоді отримаємо:

$$E = E'_m \cos\omega'\left(\frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}} + \frac{x - vt}{c\sqrt{1 - \frac{v^2}{c^2}}}\right),$$

або в іншому вигляді

$$E = E'_m \cos\omega' \frac{1 - \frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}}\left(t + \frac{x}{c}\right). \quad (15.13)$$

Зіставивши рівняння (15.12) і (15.13), знайдемо:

$$\omega = \omega' \frac{1 - \frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} = \omega' \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}}.$$

Переходячи від циклічних до звичайних частот і позначивши частоту ν' в системі S' джерела через ν_0, дістанемо формулу, що описує поздовжній ефект Доплера, який спостерігається під час руху джерела вздовж лінії, що з'єднує його з приймачем.

$$\nu = \nu_0 \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}}. \qquad (15.14)$$

Якщо $v \ll c$, то

$$\nu \approx \nu_0 \frac{1 - \frac{1}{2}\frac{v}{c}}{1 + \frac{1}{2}\frac{v}{c}} \approx \nu_0\left(1 - \frac{1}{2}\frac{v}{c}\right)\left(1 - \frac{1}{2}\frac{v}{c}\right).$$

Обмежившись членами порядку v/c, отримаємо

$$\nu = \nu_0\left(1 - \frac{v}{c}\right).$$

Звідси відносна зміна частоти

$$\frac{\nu - \nu_0}{\nu_0} = \frac{\Delta\nu}{\nu_0} = -\frac{v}{c}. \qquad (15.15)$$

Зазначимо, що при віддаленні джерела і приймача один від одного спостерігається зсув в область довших хвиль ($\nu < \nu_0,\ \lambda > \lambda_0$), так зване червоне зміщення. При наближенні – зсув в область коротших хвиль ($\nu > \nu_0,\ \lambda < \lambda_0$). Крім поздовжнього ефекту, для світлових хвиль існує також ***поперечний ефект Доплера.*** Він полягає в зменшенні частоти, яка сприймається приймачем, і спостерігається в тому випадку, коли вектор відносної швидкості напрямлений перпендикулярно вздовж прямої, що проходить через приймач і джерело (наприклад, джерело рухається по колу, в центрі якого поміщено приймач). У такому випадку

$$\nu = \nu_0 \sqrt{1 - \frac{v^2}{c^2}} \approx \nu_0 \left(1 - \frac{1}{2}\frac{v^2}{c^2}\right) < \nu_0.$$

Відносна зміна частоти при поперечному ефекті Доплера

$$\frac{\Delta\nu}{\nu_0} = -\frac{1}{2}\frac{v^2}{c^2}. \qquad (15.16)$$

Вона значно менша, ніж при поздовжньому ефекті. Для звукових хвиль поперечного ефекту Доплера не існує.

Поздовжній світловий ефект Доплера було експериментально підтверджено дослідами А.А. Бєлопольського у 1905 р. і Б.Б.Голіцина у 1907 р.; поперечний – дослідами Айвса і Стілуела у 1938р. В останніх визначалась зміна частоти випромінювання атомів водню в анодному промінні, де їх швидкість досягла $2 \cdot 10^6$м/с. Ці досліди водночас підтверджували правильність перетворень Лоренца в теорії відносності.

Поздовжній ефект Доплера використовують в астрономії для визначення радіальної швидкості зірок. Вимірявши спектроскопом відносну зміну частоти для будь-якої лінії спектра випромінювання зірки, за формулою (15.15) можна визначити її швидкість v.

Ефект Доплера впливає на спектр випромінювання газів. Внаслідок хаотичного теплового руху молекул частоти випромінювання їх фіксуватимуться у спектрографі в межах від

$$\nu_0 \left(1 - \frac{v}{c}\right) \quad \text{до } \nu_0 \left(1 + \frac{v}{c}\right).$$

Це зумовлюватиме розширення спектральної лінії світного газу. Отже, за розширенням спектральних ліній можна судити про температуру світного газу.

Контрольні питання

1.Дисперсія світла і її природа.

2.Які є види дисперсії?

3.Групова швидкість світла та її зв'язок з дисперсією.

4.Закон Бугера-Ламперта поглинання світла в середовищах.

5.В чому суть явища розсіяння світла? Як це явище проявляється в природі? Які властивості розсіяного світла?

6.Ефект Доплера в оптичних явищах і як він проявляється?

7.Чим відрізняється ефект Доплера в оптичних явищах від ефекта Доплера в акустиці?

Розділ 16. Квантова природа випромінювання

§16.1.Теплове випромінювання. Його характеристики. Абсолютно чорне тіло. Закон Кірхгофа

1.Світло, яке випромінюється джерелом, відносить з собою енергію. В залежності від того, звідки черпається ця

енергія, розрізняють і види світіння.

Випромінювання, яке супроводжується хімічними перетвореннями, називають *хемілюмінесценцією*. Прикладом є світіння фосфору, який повільно окислюється на повітрі. Випромінювання променистої енергії в цьому випадку відбувається внаслідок зменшення внутрішньої енергії тіла при зміні його хімічного складу.

Світіння під дією електричного поля називають *електролюмінесценцією*. В напівпровідниках випромінювання може відбуватись внаслідок рекомбінації носіїв заряду в області *p-n* переходу (інжекційна електролюмінесценція), яка має місце у фотодіодах. Інший приклад – світіння газів або пару в електричному розряді. В цьому випадку необхідна для випромінювання енергія надається атомам або молекулам газу електронами, які прискорюються електричним полем розряду. Бомбування електронним пучком може викликати також світіння твердих тіл, наприклад сірчаного цинку, який нанесений на екран електронно-променевої трубки *(катодолюмінісценція).*

Процеси випромінювання, які викликаються попереднім або одночасним освітленням тіла, об'єднуються під назвою *фотолюмінесценція*. При цьому необхідна для випромінювання енергія доставляється світлом від зовнішнього джерела. В тих випадках, коли необхідна енергія надається нагріванням, тобто підводом теплоти, випромінювання називається ***тепловим*** або ***температурним.*** Серед різних видів світіння воно займає особливе місце. На противагу всім видам люмінесценції *це єдиний вид випромінювання, який може знаходитись в стані термодинамічної рівноваги з тілами*.

Тому фізика *теплового випромінювання є тією ланкою, яка зв'язує термодинаміку й оптику.* Це не тільки розширило границі можливостей термодинаміки і статистичної механіки, але й привело до таких наслідків, котрі виявились революційними в історії фізики взагалі. Велике відкриття Планка – квантова гіпотеза – і є таким наслідком.

У проблемах теплового випромінювання особливо важливе значення має поняття так званого *рівноважного випромінювання*. Для встановлення цього поняття розглянемо оболонку з нерухомими і непрозорими стінками, температура яких підтримується сталою (рис.16.1). Атоми і

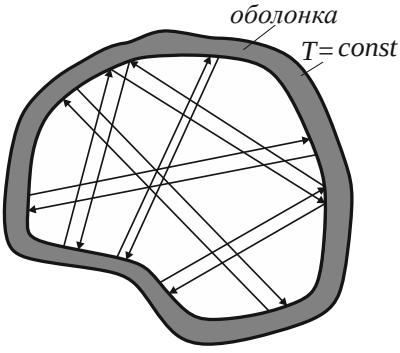


Рис.16.1.

молекули оболонки переходять у збуджені стани внаслідок енергії теплового руху і при обернених переходах в незбуджені стани дають випромінювання, яке заповнює оболонку. Падаючи на стінки оболонки, променистa енергія частково відбивається, частково поглинається. Відбувається зміна напрямку розповсюдження, спектрального складу, поляризації, інтенсивності випромінювання. В результаті всіх цих процесів, як витікає із загальних законів термодинаміки, в оболонці врешті-решт встановиться макроскопічний цілком визначений стан випромінювання, при якому за кожний проміжок часу кількість

променистої енергії певного кольору (частоти), напрямку розповсюдження і поляризації, яка випромінюється, в середньому дорівнює кількості енергії того ж кольору, напрямку розповсюдження і поляризації, що поглинається. *Як і будь-який рівноважний стан, він характеризується тим, що кожному мікропроцесу, що відбувається в системі, з тією ж ймовірністю відповідає мікропроцес, що йде в оберненому напрямку (**принцип детальної рівноваги**).* Завдяки цьому стан випромінювання в оболонці залишається *макроскопічно незмінним* з часом. Перехід в рівноважний стан, як і будь-який статистичний процес, керується *ймовірністними законами*. В оболонці встановлюється хаотичний стан випромінювання, якому відповідає *найбільша ймовірність*. Такий стан випромінювання називається *рівноважним випромінюванням.*

Властивості рівноважного випромінювання: густина променистої енергії, її розподіл по спектру частот (довжин хвиль) і напрямку розповсюдження, а також поляризація випромінювання *зовсім не залежать від форми і матеріалу стінок оболонки*. Ці властивості, як і стан газу в посудині, визначаються *тільки температурою стінок оболонки*. Рівноважне випромінювання *однорідне*, тобто його густина однакова у всіх точках всередині оболонки. Воно *ізотропне і неполяризоване*: всі можливі напрямки розповсюдження випромінювання є рівноймовірними, а напрямки векторів $\vec{E}$ і $\vec{H}$ в кожній точці простору хаотично змінюються з часом. Оскільки випромінювання знаходиться в тепловій рівновазі зі стінками, то можна говорити про температуру не тільки стінок оболонки, а і про *температуру самого випромінювання*, вважаючи за визначенням обидві температури однаковими. Однак, необхідно підкреслити, що температура рівноважного випромінювання є *властивістю самого випромінювання*, а не стінок оболонки, з якою воно знаходиться в тепловій рівновазі. Про температуру випромінювання має сенс говорити і тоді, коли взагалі немає ніяких стінок. Зокрема, наприклад, густина енергії рівноважного випромінювання однозначно визначає і його температуру.

В порожній оболонці зі стінками, що є ідеально дзеркальними, поглинання і випромінювання світла не відбувається. Якщо якимось чином ввести в оболонку випромінювання, то напрямок його розповсюдження буде змінюватись при відбиванні від стінок, але спектральний склад зберігатиметься попередній. Таке випромінювання нерівноважне і нестійке. Скільки завгодно малого відхилення властивостей стінок від ідеальності вже достатньо для того, щоб випромінювання стало рівноважним. Це відбудеться і при ідеальних стінках, якщо внести в оболонку скільки завгодно мале тіло («порошинку»), яке здатне поглинати і випромінювати. Така «порошинка», не впливаючи на енергетичний баланс, переведе випромінювання в оболонці з будь-якого нерівноважного стану в рівноважний. Від властивостей «порошинки» залежить тільки час встановлення рівноважності.

Для експериментального вивчення спектрального складу рівноважного випромінювання можна зробити невеликий отвір в стінках оболонки, які підтримуються при певній температурі. Випромінювання, яке виходить назовні через отвір, має такий самий же спектральний склад, як і всередині оболонки. Від рівноважного воно відрізняється лише тим, що розповсюджується

в межах деякого тілесного кута в одному напрямку, тобто воно не ізотропне.

2.Введемо поняття деяких фізичних величин, що характеризують стан випромінювання в просторі. Ці величини можна застосовувати до довільного випромінювання, а не тільки до рівноважного. Позначимо через *U(T) об'ємну густину енергії* випромінювання, тобто сумарну енергію одиниці об'єму електромагнітних хвиль усіх можливих частот $(0 \le \omega \le \infty)$. Для характеристики розподілу енергії за частотами або довжинами хвиль введемо *спектральну густину* випромінювання $u(\omega, T)$ або $u(\lambda, T)$ так, що величина $u(\omega, T)\, d\omega$ дає енергію одиниці об'єму випромінювання з частотами в інтервалі від ω до $\omega + d\omega$, а $u(\lambda, T)\, d\lambda$ дає енергію одиниці об'єму випромінювання з довжинами хвиль в інтервалі від λ до $\lambda + d\lambda$. Якщо це один і той самий спектральний інтервал, то $u(\omega, T)\, d\omega = u(\lambda, T) d\lambda$. Вираховуючи , що $\omega = 2\pi c/\lambda$, звідси легко виразити $u(\omega,T)$ через $u(\lambda, T)$ або навпаки.

$$u(\lambda, T) = (2\pi c/\lambda^2) u(\omega, T). \qquad (16.1)$$

Очевидно, що

$$U(T) = \int_0^\infty u(\omega, T) d\omega = \int_0^\infty u(\lambda, T) d\lambda. \qquad (16.2)$$

В теоретичній фізиці, як правило, користуються величиною $u(\omega, T)$, в експериментальній фізиці віддають перевагу $u(\lambda, T)$. Для випадку рівноважного випромінювання функції $u(\omega, T)$ і $u(\lambda, T)$ залежать тільки від частоти ω (або довжини хвилі λ) і від температури випромінювання T, але не залежать від форми і матеріалу стінок оболонки. В цьому випадку $u(\omega, T)$ буде *універсальною функцією тільки* ω і T, а $U(T)$ – універсальною функцією тільки T. Знаходження функції $u(\omega, T)$ є *основним завданням* теорії теплового випромінювання.

В той час як спектральний розподіл енергії випромінювання, що виходить з отвору в оболонці, має універсальний характер, для теплового випромінювання з відкритої поверхні тіла це не так: його спектральний розподіл залежить не тільки від температури, але і від матеріалу поверхні. Для кількісної характеристики цього спектрального розподілу вводять поняття ***випромінювальної здатності*** тіла $r(\omega, T)$ або $r(\lambda, T)$, тобто спектральної густини потоку енергії випромінювання, що випромінюється одиничною площадкою поверхні у всіх напрямках (в межах тілесного кута 2π) для одиничного інтервалу частот (довжин хвиль) так, що $r(\omega, T)d\omega$ або $r(\lambda, T)d\lambda$ є випромінювання у відповідному спектральному інтервалі від ω до $\omega + d\omega$ або від λ до $\lambda + d\lambda$.

Аналогічно формулі (16.1)

$$r(\lambda, T) = (\tfrac{2\pi c}{\lambda^2}) r(\omega, T). \qquad (16.3)$$

Інтегральна випромінювальна здатність (повний потік випромінювання для всіх частот) називається *енергетичною світністю R*(*T*) поверхні. Очевидно, що

$$R(T) = \int_0^\infty r(\omega, T) d\omega = \int_0^\infty r(\lambda, T) d\lambda. \qquad (16.4)$$

Підкреслимо, що випромінювальна здатність $r(\omega, T)$ характеризує тільки теплове випромінювання тіла (всі види люмінесценції виключаються). Для даного тіла вид функції $r(\omega, T)$ залежить тільки від його температури. Випромінювальна здатність тіла не залежить від оточуючого середовища, зокрема від того, чи знаходиться тіло в рівновазі з випромінюванням, чи ні.

Випромінювання, яке падає на поверхню тіла, частково відбивається або

розсіюється поверхнею, частково проходить через поверхню. Обмежимось випадком, коли тіло непрозоре, тобто товщина його достатня для того, щоб випромінювання, яке зайшло в нього, встигало повністю поглинутись, не досягнувши протилежної сторони. Тоді можна умовно говорити про поглинання самою поверхнею. Поглинання, як правило, має селективний (вибірковий) характер і залежить від температури поверхні. Так, наприклад, тонкий шар сажі практично повністю поглинає видиме світло, але в значно меншій мірі інфрачервоне випромінювання. Плавлений кварц прозорий в широкому інтервалі довжин хвиль, але починає помітно поглинати світло при температурі близькій 1500 °С. Для характеристики таких властивостей тіл вводять поняття *поглинальної здатності* $\alpha(\omega,T)$або $\alpha(\lambda,T)$ тіла, під якою розуміють безрозмірну величину, що показує, яка доля енергії випромінювання даної частоти (або довжини хвилі), яке падає на поверхню, поглинається нею.За визначенням, $\alpha(\omega,T)\leq 1$. Для тіла, яке при будь якій температурі повністю поглинає випромінювання всіх частот, яке падає на нього, називається ***абсолютно чорним або просто чорним***. Для нього $\alpha(\omega,T)\equiv 1$. Тіло, для якого $\alpha(\omega,T)=$ $=\alpha(T)=\text{const}<1$, називається ***сірим.***

3.Між випромінювальною і поглинальною здатностями будь-якого тіла є зв'язок, який випливає із загальних принципів термодинаміки. Окільки випромінювальна і поглинальна здатності характеризують саме тіло і не залежать від оточення, то для знаходження зв'язку між ними можна розглянути частинний випадок, коли тіло оточене рівноважним випромінюванням з тією ж температурою, наприклад, поміщене в оболонку, стінки якої є ідеально дзеркальними, або підтримуються при тій самій температурі, що й тіло. На рис.16.2 в оболонці розміщено декілька тіл, серед яких і абсолютно чорне (АЧТ). Всередині оболонки вакуум, так що тіла можуть обмінюватись енергією між собою і

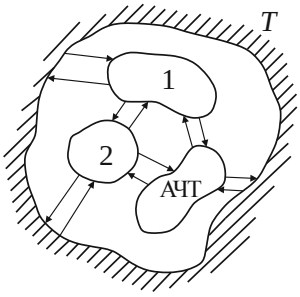


Рис.16.2.

оболонкою тільки шляхом випромінювання і поглинання електромагнітних хвиль. Дослід показує, що така система через деякий час прийде до рівноваги – всі тіла приймуть одну й ту саму температуру, яка дорівнюватиме температурі оболонки *T*. В такому стані тіло, яке має більшу випромінювальну здатність $r(\omega,T)$, втрачає за одиницю часу з одиниці поверхні більше енергії, ніж тіло, що має меншу $r(\omega,T)$. Оскільки температура (а значить і енергія) тіл не змінюється, то тіло, яке випромінює більше енергії, повинно і більше поглинати, тобто мати більший $\alpha(\omega,T)$. Таким чином, чим більша випромінювальна здатність тіла $r(\omega,T)$, тим більша і його поглинальна здатність $\alpha(\omega,T)$.

Звідси витікає співвідношення

$$\left(\frac{r(\omega,T)}{\alpha(\omega,T}\right)_1=\left(\frac{r(\omega,T)}{\alpha(\omega,T}\right)_2=\left(\frac{r(\omega,T)}{\alpha(\omega,T}\right)_{\text{АЧТ}}=\ldots, \quad (16.5)$$

де індекси 1,2, АЧТ і т.д. відносяться до різних тіл. Співвідношення (16.5) є встановлений Кірхгофом в 1859 році закон, який формулюється так: *відношення випромінювальної здатності тіла до його*

поглинальної здатності однакове для всіх тіл і є універсальною функцією частоти (довжини хвилі) і температури:

$$\frac{r(\omega,T)}{\alpha(\omega,T)} = f(\omega,T), \qquad (16.6)$$

де $f(\omega,T)$ – *універсальна функція Кірхгофа.*

Самі величини $r(\omega,T)$ і $\alpha(\omega,T)$ можуть змінюватись дуже сильно при переході від одного тіла до іншого. Відношення ж їх є однаковим для всіх тіл. Закон Кірхгофа є точним кількісним узагальненням правила, яке емпірично встановив Прево в 1809 році. Згідно цього правила, якщо *поглинальні здатності тіл різні, то будуть різними і їх випромінювальні здатності.* Оскільки величина $\alpha(\omega,T)$ не може бути більша одиниці, то із (16.6) витікає, що зі всіх тіл *при одній і тій самій температурі абсолютно чорне тіло має найбільшу випромінювальну здатність.* Із закону Кірхгофа витікає також, що всяке тіло при даній температурі випромінює переважно промені таких довжин хвиль, які воно при тій самій температурі сильніше всього поглинає.

Для абсолютно чорного тіла за визначенням $\alpha(\omega,T) \equiv 1$. Тоді із формули (16.6) витікає, що $r(\omega,T)$ для АЧТ дорівнює $f(\omega,T)$. Таким чином, універсальна функція Кірхгофа $f(\omega,T)$ *є не що інше, як випромінювальна здатність* АЧТ.

Між спектральною густиною енергії рівноважного випромінювання і випромінювальною здатністю АЧТ існує такий зв'язок:

$$u(\omega,T) = \frac{4}{c} r(\omega,T), \qquad (16.7)$$

де c –швидкість світла у вакуумі. Тобто, спектральний розподіл теплового випромінювання АЧТ буде таким самим, як у рівноважного випромінювання при тій самій температурі. Тому рівноважне випромінювання ще називають *чорним випромінюванням.* Сажа і платинова чернь у видимій області спектру

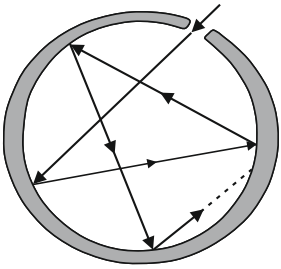


Рис.16.3.

мають поглинальну здатність $\alpha(\omega,T)$ близьку до одиниці. Але в далекій інфрачервоній області спектру $\alpha(\omega,T)$ для них помітно менша одиниці. Абсолютно чорних тіл, як і інших ідеалізованих об'єктів, в природі не існує. Але можна створити пристрій, який за своїми властивостями скільки завгодно близький до абсолютно чорного тіла. Це оболонка, яка показана на рис.16.1, в стінці якої зроблений невеликий отвір (рис.16.3).

Випромінювання, яке падає ззовні оболонки, проникає через отвір всередину, попадає на стінки оболонки, частково поглинається ними, частково відбивається або розсіюється і знову попадає на стінки. Через малі розміри отвору це відбувається багатократно, перш ніж деяка частина випромінювання знову попадає на отвір. Тому практично все падаюче на отвір світло будь-якої частоти повністю поглинається. Матеріал стінок оболонки значення не має. Отвір оболонки по відношенню до випромінювання, яке падає на нього, а також по відношенню до теплового випромінювання, яке виходить із нього, поводить себе як поверхня абсолютно чорного тіла з температурою, що дорівнює температурі стінок оболонки.

Модель АЧТ можна проілюструвати простими демонстраціями. Якщо взяти коробку, яка всередині пофарбована білою фарбою, то малий отвір в ній буде здаватись зовсім чорним навіть на фоні пофарбованої чорною фарбою зовнішньої стінки. Чорними виглядають ззовні відкриті вікна будинків, не зважаючи на те, що в приміщенні стіни і стеля світлі. Якщо розжарити стінки оболонки із матеріалу з малою поглинальною здатністю, наприклад фарфору, то її отвір буде яскраво світитись на тьмяному фоні прямого випромінювання самих стінок.

Джерело у вигляді оболонки з малим отвором, стінки якої підтримуються при постійній температурі, використовується при кількісних вимірюваннях випромінювання чорного тіла. Аналогічним чином побудований стандартний випромінювач, який є еталоном одиниці сили світла (кандели).

§16.2.Закон Стефана-Больцмана. Закон Віна. Формула Релея Джінса

Після встановлення закону Кірхгофа стало очевидним, що першочергове завдання теорії теплового випромінювання полягає у визначенні вигляду універсальної функції Кірхгофа $f(\omega,T)$ ($r(\lambda,T)$ – випромінювальної здатності абсолютно чорного тіла). Однак, спочатку було знайдено залежність інтегральної (повної) випромінювальної здатності (енергетичної світності) $R(T)$ абсолютно чорного тіла від його температури. На основі аналізу експериментальних даних Стефан у 1879 р. прийшов до висновку, що повна енергія, яка випромінюється тілом при тепловому випромінюванні, пропорційна четвертій степені термодинамічної температури. В 1884 році Больцман теоретично отримав формулу для енергетичної світності чорного тіла:

$$R(T)=\int_0^\infty r(\omega,T)d\omega=\sigma T^4. \quad (16.8)$$

В основу доведення формули (16.8) покладена аналогія між термодинамічними властивостями рівноважного випромінювання й ідеального газу. Суть цієї аналогії полягає в тому, що як енергія ідеального газу, так і енергія рівноважного випромінювання залежать тільки від температури; як газ, так і випромінювання здатні створювати тиск при взаємодії з поверхнею тіла. На основі цих уявлень було з'ясовано, що енергетична світність абсолютно чорного тіла пропорційна четвертій степені температури. Формула (16.8) виражає закон Стефана-Больцмана. За сучасними вимірюваннями числове значення постійної σ =5,670373(21)×10^{-8} Вт·м$^{-2}$·К$^{-4}$. Варто зазначити, що закон (16.8) справедливий тільки для чорного тіла. Для нечорних тіл не існує простої аналітичної залежності енергетичної світності від температури. Для реальних тіл енергетична світність $R^*(T)$ може бути визначена на основі закону Кірхгофа:

$$R^*(T)=\alpha(T)\sigma T^4, \quad (16.9)$$

де $\alpha(T)$ – поглинальна здатність реального тіла.

Закон Стефана-Больцмана встановлює залежність інтегральної випромінювальної здатності абсолютно чорного тіла від температури, але не розкриває характеру розподілу енергії випромінювання за довжинами хвиль або частотами, тобто конкретний вигляд функції Кірхгофа $f(\omega,T)$ залишається невідомим. Основним завданням теорії теплового випромінювання є встановлення конкретного вигляду цієї функції.

Криві розподілу енергії в спектрі випромінювання абсолютно чорного тіла вивчав експериментально В.Ленглей у 1866 р. Як абсолютно чорне тіло було використано порожнину з малим отвором, а також чорну сажу. Його досліди показали, що для заданих температур максимуми енергії випромінювання припадають на певні довжини хвиль, причому величина максимуму збільшується з підвищенням температури і зміщується у бік коротких довжин хвиль (більших частот).

Важливий крок в теоретичному вивченні рівноважного випромінювання був зроблений Віном в 1893р. Розглядаючи адіабатичне (без теплообміну) стискування рівноважного випромінювання в оболонці із дзеркальними стінками і враховуючи зміну частоти світла при відбиванні від дзеркала, яке рухається (ефект Доплера), Він показав, що спектральна густина $u(\omega, T)$ повинна мати вигляд

$$u\,(\omega, T) = \omega^3 \varphi\left(\frac{\omega}{T}\right), \qquad (16.10)$$

де φ −деяка функція відношення частоти до температури, конкретний вид якої в рамках електромагнітної теорії і феноменологічної термодинаміки встановити неможливо. Співвідношення (16.10) зводить задачу знаходження універсальної функції $u(\omega, T)$, яка залежить від двох аргументів, до задачі знаходення функції $\varphi(\omega/T)$, яка залежить тільки від одного аргумента (ω/T). Це значить, що якщо відомий спектральний розподіл енергії рівноважного випромінювання при якій-небудь довільній температурі, то з допомогою формули (16.10) можна знайти його при будь-якій іншій температурі.

Закон Віна (16.10) включає в собі як наслідок і закон Стефана-Больцмана. Виражаючи випромінювальну здатність чорного тіла $r(\omega,T)$ через $u\,(\omega,T)$ за допомогою співвідношення (16.7) і, переходячи від інтегрування по частотах до нової змінної $x = \omega/T$, отримаємо:

$$R(T) = \frac{c}{4}\int_0^\infty u\,(\omega,T)d\omega = \frac{c}{4}\int_0^\infty \omega^3\,\varphi\left(\frac{\omega}{T}\right)d\omega = \\ = \frac{c}{4}T^4\int_0^\infty x^3\,\varphi(x)dx = \sigma T^4.$$

Тобто інтегральна світність пропорційна T^4, оскільки інтеграл по x не залежить від *T*.

Вимірювання Люммера і Прінгсгейма в 1899р. розподілу інтенсивності випромінювання із отвору оболонки за довжинами хвиль, стінки якої підтримувались при постійній температурі, показали, що при деякому значенні довжини хвилі λ_m крива $r(\lambda,T)$ має чітко виражений максимум (рис.16.4). Ці досліди підтвердили результати

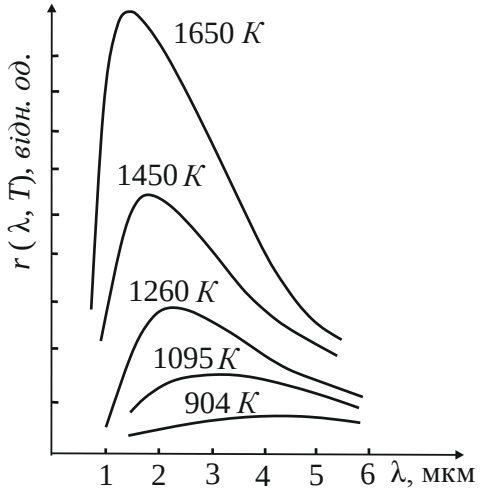


Рис.16.4.

В.Ленглея (1866р.), що із збільшенням температури максимум зміщується в область більш коротких довжин хвиль. Це зміщення відбувається так, що добуток температури на

довжину хвилі, яка відповідає максимуму випромінювання, залишається сталим:

$$\lambda_m \cdot T = b. \qquad (16.11)$$

Значення константи b =2,9 $\cdot 10^{-3}$м $\cdot K$ визначено з дослідних даних. Співвідношення(16.10) отримало назву ***закону зміщення Віна***. Воно безпосередньо випливає із термодинамічної формули Віна (16.10). Переходячи в ній по схемі (16.1) від частот до довжин хвилі, отримуємо

$$u(\lambda,T)= \left((2\pi c)^4/\lambda^5\right))\cdot\varphi\left(\frac{2\pi c}{\lambda T}\right)= \\ = T^5\varphi_1(\lambda T), \qquad (16.12)$$

де $\varphi_1(\lambda T) = (2\pi c)^4/(\lambda T)^5) \cdot \varphi(2\pi c/(\lambda T))$ – нова невизначена функція від аргумента $\lambda \cdot T$. Позначимо $x = \lambda T$. При фіксованій температурі для визначення положення максимуму залежності $u(\lambda, T)$ від λ отримаємо рівняння $\frac{d\varphi_1(\lambda T)}{d\lambda} = T\frac{d\varphi_1(x)}{dx} = 0.$ Або $\frac{d\varphi_1(x)}{dx} = 0,$ яке має деякий корінь $x = x_m$. Це означає, що при будь-якій температурі максимум залежності $u(\lambda, T)$ від λ отримується при одному і тому самому значенні аргумента $x = x_m$, тобто $\lambda_m T = x_m = \text{const}$, що й доказує закон зміщення (16.11). Із (16.12) витікає, що значення спектральної густини $u(\lambda, T)$ в максимумі при ($\lambda = \lambda_m$) пропорційне п'ятій степені термодинамічної температури.

При температурах, які досягаються практично, максимуми випромінювальної здатності чорного тіла лежать в інфрачервоній області спектру. Так, при T =3000 K $\lambda_m \approx$ $\approx$ 0,97мкм. Тільки при $T > 5000K$ максимум попадає в зелену область спектру, до якої око найбільш чутливе. Світність чорного тіла швидко зростає з температурою ($\sim T^4$) і при високих температурах досягає дуже великих значень. При T=6000K (температура поверхні Сонця) з кожного квадратного сантиметра чорного тіла випромінюється потік потужністю 7,4кВт, але тільки невелика частина цієї потужності припадає на видиму область спектру.

Розглянуті вище закономірності теплового випромінювання використовується в оптичних методах вимірювання температури (§16.4).

В 1896 р. Він опублікував роботу «Про розподіл енергії в спектрі випромінювання чорних тіл», в якій він привів доведення формули (16.10) в такому вигляді:

$$u\,(\omega, T) = \frac{\hbar\omega^3}{2\pi^3 c^3}\exp\left\{-\frac{a\omega}{2\pi T}\right\}, \qquad (16.13)$$

де $\hbar$ і a –деякі константи. При отриманні цієї формули Він застосував до випромінювання статистику. Він вважав, що густина випромінювання повинна бути пропорційна числу молекул, які випромінюють і розподілені за швидкостями згідно закону Максвелла, а частота їх коливань пропорційна квадрату швидкості молекул. Формула (16.13) узгоджувалась із вимірюваннями Пашена в короткохвильовій області спектру, які були на той час. Фактично Пашен отримав формулу (16.13) як емпіричну. В 1897р. О.Луммер і Е.Прінсгейм показали, що формула (16.13) добре узгоджується з експериментом тільки в області коротких хвиль (високих частот).

В 1892-1905рр. англійські вчені Д.Релей і Д.Джінс спробували теоретично отримати залежність $u(\omega, T)$ для рівноважного випромінювання, яке знаходилось всередині оболонки. Вони вважали, що таке випромінювання – це система стоячих хвиль. Їхні розрахунки показали, що кількість стоячих хвиль в одиниці об'єму, яке приходиться на одиничний інтервал частот дорівнює

$$n(\omega) = \frac{\omega^2}{\pi^2 c^3}\cdot \qquad (16.14)$$

Тоді

$$u\,(\omega,T)=n(\omega)\,\langle\varepsilon\rangle=\frac{\omega^2}{\pi^2c^3}\langle\varepsilon\rangle,\ (16.15)$$

де $\langle\varepsilon\rangle$ – середня енергія, яка приходиться на одну стоячу хвилю. Формула (16.15) отримана строго , без будь– яких припущень.

Релей і Джінс, виходячи із теореми класичної статистики про рівномірний розподіл енергії за ступенями свободи, припустили, що на кожне електромагнітне коливання приходиться в середньому енергія, яка дорівнює двом половинкам kT, одна половинка на електричну, друга – на магнітну енергію хвилі (за класичними уявленнями на кожну ступінь свободи приходиться в середньому енергія, що дорівнює $\frac{1}{2}kT$). Прийнявши

$$\langle\varepsilon\rangle=kT,\qquad(16.16)$$

Релей і Джінс із (16.15) отримали, що

$$u\,(\omega,T)=\frac{\omega^2}{\pi^2c^3}kT.\qquad(16.17)$$

За допомогою формули (16.7) перейдемо від $u\,(\omega,T)$ до $r(\omega,T)$.

Таким чином, для випромінювальної здатності чорного тіла отримуємо вираз:

$$r(\omega,T)=\frac{\omega^2}{4\pi^2c^2}kT,\qquad(16.18)$$

або

$$r(\lambda,T)=\frac{2\pi c}{\lambda^4}kT.\qquad(16.18)^*$$

Вирази (16.17) , (16.18) і $(16.18)^*$ називають *формулою Релея –Джінса*. Ці формули дуже добре описують розподіл енергії у довгохвильовій частині спектру випромінювання, тобто для малих частот ω (великих довжин хвиль). Для великих частот вони непридатні.

Розбіжність між експериментальною кривою (суцільна лінія) і кривою, одержаною за допомогою формули Релея–Джінса $(16.18)^*$ (штрихова крива), видно з рис.16.5. Крім того виявилось, що спроба вивести закон Стефана-Больцмана за допомогою формули Релея-Джінса призвела до абсурдного результату, а саме:

$$U(T)=\int_0^\infty\frac{\omega^2}{4\pi^2c^2}kT\,d\omega=\infty.$$

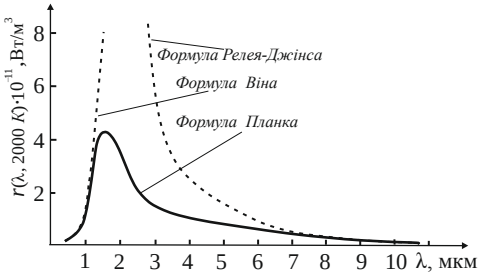


Рис.16.5

Досліди заперечують такий результат. Ця ситуація, що створилась у теорії теплового випромінювання, образно була названа П.Еренфестом «ультрафіолетовою катастрофою». Така назва пов'язана з тим, що зменшення довжини хвилі (збільшення частоти) обумовлює згідно (16.17) необмежене зростання $u\,(\omega,T)$. Проблема отримання функції розподілу енергії у спектрі випромінювання абсолютно чорного тіла, за висловленням У.Кельвіна – це «одна з темних хмар на горизонті класичної фізики, що затьмарила її чисте небо». Як висловився з цього приводу Лоренц, рівняння класичної фізики виявилися нездатними пояснити, чому піч, що згасла, не випромінює синіх променів поряд з випромінюванням променів великих довжин хвиль.

§16.3.Квантова гіпотеза і формула Планка

Планк почав публікувати роботи з теорії теплового випромінювання у 1895 р. Проблема особливо притягувала його тим, що тут мова йшла не про часткову закономірність, а про «абсолютну». Планк розглядав оболонку

(рис.16.3) як набір лінійних, пружних диполів (осциляторів, див.§17.4), які знаходяться в термодинамічній рівновазі з випромінюванням. До 1900р. в нього були два основних результати. Суть першого в тому, що, застосувавши рівняння Максвелла до лінійного пружного осцилятора, який знаходився в оболонці в рівновазі з тепловим випромінюванням, він отримав формулу

$$u\,(\omega,T)=\frac{\omega^2}{\pi^2c^3}\langle\varepsilon\rangle, \qquad (16.19)$$

яка співпадала з формулою (16.14). Порівнявши (16.19) із формулою Віна (16.14), яку Планк розглядав як емпіричну, він отримав, що

$$\langle\varepsilon\rangle=\frac{h\omega}{2\pi}\exp\left\{-\frac{a\omega}{2\pi T}\right\}. \qquad (16.20)$$

Із досліду для коефіцієнта h було знайдено значення $h=6{,}885\cdot10^{-34}$Дж$\cdot$с. Так 18 травня 1899р. народилась постійна Планка, якій судилась така велика доля. Стала $a=h/k$, де k −стала Больцмана. Отримані Планком значення сталих h і k дуже близькі до сучасних: $h=6{,}626\cdot10^{-34}$Дж$\cdot$с; $\hbar=\frac{h}{2\pi}=$ $=1{,}054\cdot10^{-34}$Дж$\cdot$с; $k=1{,}38\cdot10^{-23}$ Дж/К. В нових позначеннях (16.20) можна переписати так:

$$\langle\varepsilon\rangle=\hbar\omega\cdot\exp\left\{-\frac{\hbar\omega}{kT}\right\}. \qquad (16.21)$$

Іншим результатом було те, що на початку 1900р. Планк, як він думав, довів те, на що раніше сподівався, − формула Віна (16.13) є наслідком термодинаміки.

Планк закінчував доведення в момент, коли вимірювання при малих ω поставили під сумнів формулу Віна. Незабаром він відмовився і від доведення, і від самої формули. Тим не менш, 19 жовтня 1900р. Планк знайшов свою знамениту формулу, яка визначає спектральну густину енергії випромінювання для всього спектру частот:

$$u\,(\omega,T)=\frac{\omega^2}{\pi^2c^3}\cdot\frac{\hbar\omega}{\exp\left\{\frac{\hbar\omega}{kT}\right\}-1}\cdot \qquad (16.22)$$

Залишилось знайти її теоретичне обґрунтування. Знайшовши помилку в своєму старому доведенні, Планк виявив, *що з однієї термодинаміки функцію Кірхгофа* $f(\omega,T)$ *визначити неможливо.* Для теоретичного виводу формули (16.22) необхідна гіпотеза, яка корінним чином протирічить уявленням класичної фізики. Рівномірно розподіляти енергію між осциляторами, аналогічно Релею і Джінсу, безперспективно, тому її необхідно розподіляти нерівномірно. Якщо $\langle\varepsilon\rangle$ середня енергія осцилятора, то сумарна енергія N осциляторів буде $N\langle\varepsilon\rangle$. Щоб цю енергію розподілити нерівномірно між N осциляторами, необхідно ввести порцію енергії ε. Таких порцій буде $n=N\langle\varepsilon\rangle/\varepsilon$ і їх потрібно розподілити між осциляторами так, щоб система осциляторів знаходилась в рівноважному, найбільш вірогідному, стані. Така задача була розв'язана Больцманом, розв'язок якої давав формулу (16.26).

Приведемо отримання формули Планка (16.22), коли не енергію розподіляють між осциляторами, а осциляторів за енергіями. Нехай енергія осцилятора може приймати не будь-які, а тільки певні визначені дискретні значення енергії ε_n, які відділені одне від одного кінцевими інтервалами. Перехід осцилятора з одного стану в інший супроводжується поглинанням або випромінюванням кінцевої порції (кванта) енергії випромінювання. В такій системі з дискретним енергетичним спектром середню енергію $\langle\varepsilon\rangle$ в тепловій рівновазі при температурі T вже не можна знаходити за формулою (16.16). Ймовірність p_n того, що

осцилятор знаходиться в стані з енергією ε_n, у відповідності із розподілом Больцмана пропорційна $\exp\{-\varepsilon_n/(kT)\}$, але при розрахунках середніх значень інтеграли замінюються сумами:

$$\langle\varepsilon\rangle = \sum_n \varepsilon_n p_n = \frac{\sum_n \varepsilon_n \exp\{-\varepsilon_n/(kT)\}}{\sum_n \exp\{-\varepsilon_n/(kT)\}}, \quad (16.23)$$

де $p_n = \frac{\exp\{-\varepsilon_n/(kT)\}}{\sum_n \exp\{-\varepsilon_n/(kT)\}}$.

Ще одна гіпотеза необхідна для встановлення значень енергії ε_n дозволених станів осцилятора. Згідно припущення Планка, гармонічний осцилятор має еквідистанційний енергетичний спектр так, що енергія ε_n в стані з номером n складає ціле кратне найменшій порції енергії ε; $\varepsilon_n = n\varepsilon$, де $n = 0,1,2,3,...$ (сучасна квантова механіка дає для рівнів енергії гармонічного осцилятора значення $\varepsilon_n = (n + \frac{1}{2})\,\hbar\omega$). Тоді сума, яка знаходиться в знаменнику формули (16.23), є нескінченна геометрична прогресія зі знаменником $\exp\{-\beta\varepsilon\}$, де $\beta = 1/(kT)$:

$$S(\beta) = \sum_{n=0}^{\infty} \exp\{-\beta n\varepsilon\} = (1-\exp\{-\beta\varepsilon\})^{-1}. \quad (16.24)$$

Легко побачити, що сума, яка знаходиться в чисельнику формули 16.23), дорівнює похідній $\frac{dS(\beta)}{d\beta}$, взятій із протилежним знаком:

$$\begin{cases} \sum_{n=0}^{\infty} n\varepsilon \exp\{-\beta n\varepsilon\} = \\ = -\frac{d}{d\beta}\sum_{n=0}^{\infty} \exp\{-\beta n\varepsilon\} = \\ = (1-\exp\{-\beta\varepsilon\})^{-2} \cdot \exp\{-\beta\varepsilon\} \cdot \varepsilon\,. \end{cases} \quad (16.25)$$

Підставимо (16.25), (16.24) в (16.23) і тоді знайдемо, що

$$\langle\varepsilon\rangle = \frac{\varepsilon}{\exp\left\{\frac{\varepsilon}{kT}\right\} - 1}. \quad (16.26)$$

Цей результат співпадає з класичним виразом (16.16) тільки в граничному випадку при $\varepsilon/(kT) \to 0$. В оригінальній роботі Планка 1900р. середня енергія осцилятора знаходилась шляхом розрахунку середньої енергії на основі співвідношення Больцмана $S = \ln W$ між ентропією S і термодинамічною ймовірністю стану системи W. Підставляючи отримане значення середньої енергії квантового осцилятора (16.26) в (16.19), отримаємо формулу:

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3} \cdot \frac{\varepsilon}{\exp\left\{\frac{\varepsilon}{kT}\right\} - 1}. \quad (16.27)$$

Щоб вона не протирічила термодинамічній формулі Віна (16.10), згідно якої температура може входити в $u(\omega, T)$ тільки в комбінації $\frac{\omega}{T}$, необхідно прийняти, що

$$\varepsilon = \hbar\omega. \quad (16.28)$$

В результаті, з формули (16.27) ми отримаємо формулу (16.22).

14 грудня 1900р. Планк зробив доповідь, в якій приводився повний розв'язок поставленої задачі й отримання формули (16.22). Цю дату можна вважати днем народження квантової теорії. Планк довів, що формулу (16.22) можна вивести тільки в тому випадку, якщо допустити квантування енергії, яке протирічить класичним уявленням. Формула (16.22) в змінних λ, T має вигляд

$$u(\lambda, T) = \left(\frac{8\pi hc}{\lambda^5}\right)/\left(\exp\left\{\frac{hc}{\lambda kT}\right\} - 1\right), \quad (16.29)$$

а для випромінювальної здатності абсолютно чорного тіла

$$r(\lambda, T) = \frac{c}{4}u(\lambda, T) = \left(\frac{2\pi hc^2}{\lambda^5}\right)/\left(\exp\left\{\frac{hc}{\lambda kT}\right\} - 1\right). \quad (16.30)$$

Формула Планка (16.22) (те саме, що і (16.30)) чудово узгоджується з дослідом при всіх частотах і температурах. Для малих частот

(довгих хвиль) і при високих температурах, коли $\hbar\omega/(kT) \leq 1$, експоненту $\exp\{\hbar\omega/(kT)\}$ можна розкласти в ряд по степенях змінної $\hbar\omega/(kT)$. Якщо обмежитись лінійним членом цього розкладу, то формула Планка(16.22) переходить в формулу Релея-Джінса (16.17). Але в області високих частот і низьких температур формула Планка передбачає зовсім іншу поведінку $u(\omega, T)$, ніж класична теорія. При $\hbar\omega/(kT) \gg 1$ експонента $\exp\{\hbar\omega/(kT)\}$ в знаменнику формули (16.22) значно більша одиниці. Тому поведінка $u(\omega, T)$ в області коротких хвиль приблизно описується формулою

$$u\,(\omega, T) \approx \frac{\hbar\omega^3}{\pi^2 c^3} \exp\left\{-\frac{\hbar\omega}{kT}\right\}, \quad (16.31)$$

яка була запропонована Віном в 1896 р. (див формулу (16.13)). Експоненціальний спад спектральної густини з ростом частоти усуває «ультрафіолетову катастрофу» класичної теорії.

Формула Планка (16.30) дає змогу не тільки встановити закон Стефана-Больцмана (16.8), але й знайти величину сталої σ, а саме:

$$\sigma = \frac{2\pi^5 k^4}{15c^2h^3} = (5{,}67032 \pm 0{,}00071) \cdot \cdot 10^{-8}\text{Вт}/(\text{м}^2 \cdot \text{К}^4).$$

З формули (16.30) можна отримати закон зміщення Віна і величину сталої b. Для цього необхідно розв'язати рівняння

$$\frac{dr(\lambda, T)}{d\lambda} = 0.$$

З якого отримаємо, що

$$T\lambda_m = \frac{hc}{4{,}965k} = b = (2{,}897790 \pm 0{,}000090) \cdot \cdot 10^{-3}\text{м} \cdot \text{К} \cong 2{,}9 \cdot 10^{-3}\text{м} \cdot \text{К}.$$

При вирішені проблеми теплового випромінювання виникла постійна Планка $\hbar$ (або $h = 2\pi\hbar$), яка грає в квантовій фізиці таку саму роль, як швидкість світла c в релятивістській фізиці. Ці фундаментальні світові константи визначають границі застосування класичного описання. В масштабах макросвіту числове значення постійної Планка надзвичайно мале. Цим пояснюється широке застосування класичної фізики, в основі якої лежить концепція неперервності при описанні макроскопічних явищ. Значення $\hbar$, які отримані на основі різних фізичних явищ (теплове випромінювання, фотоефект, короткохвильова границя суцільного рентгенівського спектру, ефект Джозефсона в надпровідності та інші), добре узгоджуються одні з одними.

Згідно сучасним уявленням, рівноважне випромінювання в оболонці розглядається як фотонний газ, до якого застосовується статистика Бозе-Ейнштейна. Середнє число фотонів у квантовому стані

$$< n > = \frac{1}{\exp\{\frac{\hbar\omega}{kT}\} - 1},$$

де $\hbar\omega$ – енергія фотона. Число таких станів $n(\omega) = \frac{\omega^2}{\pi^2 c^3} \cdot$ Тому спектральна густина енергії випромінювання

$$u(\omega, T) = < n > \cdot\, \hbar\omega \cdot n(\omega) = = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar\omega}{\exp\{\frac{\hbar\omega}{kT}\} - 1},$$

що співпадає із формулою Планка (16.22).

§16.4. Оптична пірометрія

Для вимірювання температур нагрітих тіл за термометричний параметр можна використати теплове випромінювання. Сукупність оптичних методів визначення температур, в основі яких є закони рівноважного випромінювання, називається оптичною термометрією,а методи вимірювання високих температур називаються *оптичною пірометрією.* Оптичні методи вимірювання температур не потребують безпосереднього

контакту вимірювальних приладів із досліджувальним тілом. Завдяки цьому можна вимірювати високі температури, а також температури досить віддалених тіл. Застосування цих методів не змінює температуру досліджуваних тіл.

Залежно від того, який із законів теплового випромінювання покладено в основу методу вимірювання температур, розрізняють три умовних температури: *енергетичну* або *радіаційну*; *яскравісну*; *кольорову*. Вони функціонально пов'язані із справжньою температурою тіла і його випромінювальною здатністю. Під *радіаційною* розуміють температуру абсолютно чорного тіла T_P, при якій його випромінювальна здатність $R(T)$ однакова з випромінювальною здатністю $R^*(T)$ тіла, температуру якого визначають. Відповідно до визначення радіаційної температури та закону Стефана-Больцмана для сірих тіл (16.9) маємо:

$$R(T) = R^*(T) = \sigma T_P^4 = \alpha(T)\sigma T^4. \quad (16.32)$$

Звідки

$$T_P = T\sqrt[4]{\alpha(T)}. \quad (16.33)$$

Оскільки $\alpha(T) < 1$, то радіаційна температура менша від справжньої температури тіла. Радіаційну температуру вимірюють радіаційним пірометром (рис.16.6), оптична система якого дає змогу

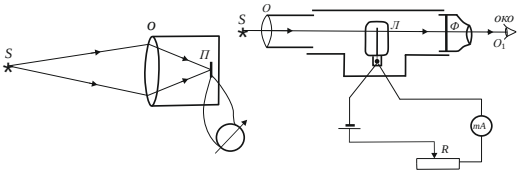


Рис.16.6. Рис.16.7.

сфокусувати чітке зображення джерела світла *S* на приймач *П* так, щоб зображення повністю покрило поверхню приймача. За такої умови енергія випромінювання джерела, що переноситься за одиницю часу на приймач, не залежатиме від відстані між джерелом і приймачем. Приймачами в пірометрах часто є термобатареї або болометри. Пірометри попередньо градуюють за допомогою абсолютно чорного тіла, температура якого відома.

Під яскравісною розуміють температуру абсолютно чорного тіла, при якій його спектральна випромінювальна здатність $r(\lambda, T_я)$ дорівнює спектральній випромінювальній здатності $r^*(\lambda, T)$ досліджуваного тіла для тієї самої довжини хвилі, яка згідно (16.6) дорівнює $\alpha(\lambda, T)r(\lambda, T)$. Тобто

$$r(\lambda, T_я) = r^*(\lambda, T) = \alpha(\lambda, T)r(\lambda, T), \quad (**)$$

де T – справжня температура тіла, $T_я$ – яскравісна температура, $\alpha(\lambda, T)$ – поглинальна здатність тіла.

Оскільки для світлових хвиль видимої частини спектра, які випромінюється нагрітими тілами, температура яких $T{\sim}3000K$, величина $\frac{hc}{kT\lambda} \approx 7$, то $e^7 \gg 1$. Нехтуючи одиницею у формулі Планка (16.30), вираз $(**)$ можна записати так:

$$\exp\left\{\left(\frac{1}{T} - \frac{1}{T_я}\right)\frac{hc}{k\lambda}\right\} = \alpha(\lambda, T). \quad (16.34)$$

Із співвідношенняя (16.34) отримаємо вираз:

$$\frac{1}{T} - \frac{1}{T_я} = \frac{k\lambda}{hc}\ln\alpha(\lambda, T), \quad (16.35)$$

який зв'язує яскравісну і справжню температури тіла. Для нечорних тіл $\alpha(\lambda, T) < 1$; яскравісна температура менша від справжньої; яскравісна температура вимірюється пірометрами із зникаючою ниткою (рис.16.7). Його дія полягає в тому , що за допомогою об'єктива *O* зображення поверхні нагрітого тіла, температуру якого визначають ,

суміщають із площиною нитки розжарення лампи *Л,* яка вмикається в електричне коло. Поверхня тіла і нитка розжарення спостерігаються через окуляр O_1. Перед окуляром знаходиться світлофільтр *Ф*, який пропускає досить вузьку частину спектру (для λ_0=0,66мкм). Під час вимірювань за допомогою реостата змінюють струм через лампу Л доти, поки яскравості нитки розжарення і поверхні тіла не стануть однаковими. Тому $T_я$ називають яскравісною температурою тіла. Міліамперметр попередньо градуюють у градусах за допомогою випромінювання абсолютно чорного тіла. Якщо величина $\alpha(\lambda_0, T)$ випромінюючого тіла відома, то за яскравісною температурою можна визначити справжню температуру тіла .

Під кольоровою розуміють температуру абсолютно чорного тіла, при якій спектральний склад його випромінювання однаковий із спектральним складом випромінювання досліджуваного тіла, тобто відношення спектральних випромінювальних здатностей для двох заданих довжин хвиль однакові. За визначенням *поняття кольорової температури застосовується до сірих тіл*. Для тіл, в яких у значній мірі проявляється селективність випромінювання, поняття кольорової температури не має змісту. Якщо для сірого тіла $r^*(\lambda_1, T)/r^*(\lambda_2, T) = \eta$, то така ж рівність має місце і для АЧТ:

$$r(\lambda_1, T_к)/r(\lambda_2, T_к) = \eta.$$

На основі формули Планка (16.30) це співвідношення запишеться так:

$$\eta = \left(\frac{\lambda_2}{\lambda_1}\right)^5 \cdot \frac{\exp\left\{\frac{hc}{kT_к\lambda_2}\right\} - 1}{\exp\left\{\frac{hc}{kT_к\lambda_1}\right\} - 1}. \qquad (16.36)$$

Звідки, враховуючи, що $\exp\left\{\frac{hc}{kT_к\lambda}\right\} \gg 1$, отримаємо вираз для кольорової температури:

$$T_к = \frac{\frac{hc}{k}\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right)}{\ln\eta - 5\ln\frac{\lambda_2}{\lambda_1}}. \qquad (16.37)$$

Ряд реальних випромінювачів, наприклад, оксиди металів, кераміка, вогнетривкі вироби, графіт та інші практично є сірими тілами.

Кольорова температура $T_к$ вимірюється за червоно-синім відношенням інтенсивностей для $\lambda_1 = 0{,}665$ мкм і λ_2=0,470 мкм. Для несірих тіл здебільшого кольорова температура вища за справжню температуру, тобто $T_к > T$. Кольорова температура більшості твердих тіл і рідин значно менше відрізняється від їх справжньої температури, порівняно з радіаційною або яскравісною.

Отже, за допомогою оптичних пірометрів вимірюють не справжні, а умовні температури: радіаційну, яскравісну і кольорову.

§16.5. Зовнішній фотоефект і його закони. Рівняння Ейнштейна для зовнішнього фотоефекту

Фотоелектричним ефектом *називаються електричні явища, які супроводжуються поглинанням світлового випромінювання в речовині.*

Розрізняють:

1)*зовнішній фотоефект* – виривання електронів з речовини під дією світла у вакуум або в інше середовище. Практичне значення має фотоефект із твердого тіла у вакуум;

2)*внутрішній фотоефект*, при якому відбувається лише перерозподіл електронів за енергетичними станами в конденсованому середовищі внаслідок поглинання електромагнітного випромінювання. В напівпровідниках і діелектриках фотоефект

проявляється в зміні електропровідності середовища та діелектричної проникності;

3)*фотогальванічний ефект*, при якому на границі поділу напівпровідника і металу або на границі поділу двох напівпровідників під впливом опромінювання виникає електрорушійна сила (за відсутності зовнішнього електричного поля);

4)*фотоефект в газоподібному середовищі*, який полягає у фотоіонізації окремих молекул або атомів.

Розглянемо закономірності зовнішнього фотоефекту. Фотоефект відкрив у 1887 р. Г.Герц, який виявив, що при освітленні негативного електрода іскрового розрядника ультрафіолетовими променями розряд відбувається при меншій напрузі між електродами, ніж без такого освітлення.

Перші фундаментальні дослідження фотоефекту виконанні в 1888 – 1889р. О.Г. Столєтовим за допомогою установки, що показана на рис.16.8.

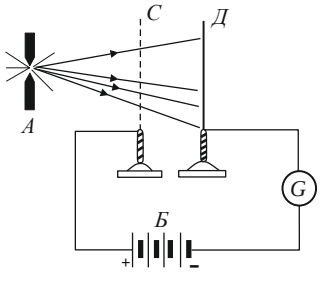


Рис.16.8

Конденсатор, який утворений дротяною сіткою з міді *С* (анод) і суцільною цинковою пластинкою *Д* (катод), був послідовно ввімкнений з гальванометром *G* в коло акумуляторної батареї *Б*. При освітленні негативно зарядженої пластини *Д* світлом від джерела *А* в колі виникає електричний струм, який називається фотострумом. На основі своїх дослідів Столєтов дійшов таких висновків:

1) найбільшу дію чинять ультрафіолетові хвилі;

2) сила струму зростає зі збільшенням освітленості пластини;

3) заряди, які випускаються із пластини *Д* під дією світла, мають від'ємний знак.

У 1898 р. Леонард і Томсон методом відхилення зарядів у електричному і магнітному полях визначили питомий заряд q/m частинок, що вириваються світлом з катода, довівши, що ці частинки є електронами.

Леонард й інші дослідники удосконалили прилад Столєтова, помістивши електроди у вакуумну трубку (рис.16.9). Катод

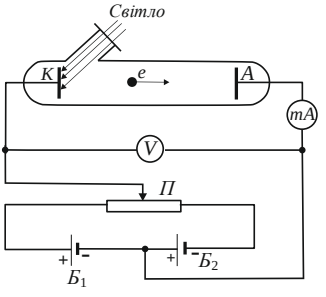


Рис.16.9.

K, покривався металом, який досліджувався, і освітлювався монохроматичним світлом, яке проходило через кварцове вікно. Напругу *U* між катодом *K* і анодом *A* можна регулювати за допомогою потенціометра *П* і вимірювати вольтметром *V*. Дві акумуляторні батареї $Б_1$ і $Б_2$ дають можливість за допомогою потенціометра *П* змінити не лише абсолютну величину, а й знак напруги U. Цей пристрій дав можливість дослідити вольт-амперну характеристику фотоефекту – залежність фотоструму *I* від напруги між електродами (рис.16.10). Така залежність відповідає двом різним освітленостям *E* катода. У міру збільшення напруги U фотострум *I* поступово зростає, тобто все більша кількість фотоелектронів досягає анода. Максимальне

значення фотоструму $I_н$, яке називається фотострумом насичення, відповідає таким значенням U, при яких усі електрони, що вибиваються з катода світлом, досягають анода: $I_н = en$, де n – кількість електронів, які вилітають з катода за 1 с.

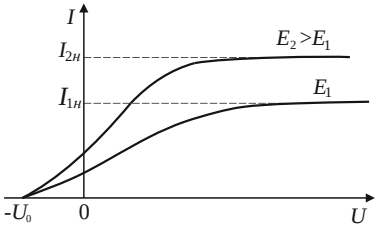


Рис. 16.10.

З вольт-амперної характеристики виходить, що при $U = 0$ фотострум не зникає. Електрони, вибиті з катода світлом, мають деяку початкову швидкість v, а отже і відмінну від нуля кінетичну енергію і можуть досягати анода без зовнішнього поля. Для того, щоб фотострум став нульовим, необхідно прикласти ***затримуючу напругу*** U_0 (на катоді K позитивний потенціал по відношенню до анода A). При $U = U_0$ жоден з електронів не може подолати затримуючого поля і досягнути анода (I=0, рис.16.10). Отже,

$$\frac{m_e v_{max}^2}{2} = eU_0, \qquad (16.38)$$

тобто, вимірявши U_0, можна знайти максимальні значення швидкості v_{max} і кінетичної енергії фотоелектронів, маса яких m_e.

Дослідами встановлено такі основні ***закони зовнішнього фотоефекту:***

1. Закон Столєтова: при фіксованій частоті падаючого світла кількість фотоелектронів, що вириваються з катода за одиницю часу, пропорційна інтенсивності світла (сила фотоструму насичення $I_н$ пропорційна енергетичній освітленості E катода – $I_н = \gamma E$, γ – спектральна чутливість фотокатода).

2.Максимальна початкова швидкість фотоелектронів визначається лише частотою ν світла і не залежить від його інтенсивності. Величина v_{max} зростає зі збільшенням частоти ν.

3.Для кожної речовини існує "червона межа" фотоефекту, тобто максимальна довжина хвилі λ_{max}, при якій ще спостерігається фотоефект. Величина λ_{max} залежить від хімічної природи речовини і стану його поверхні.

Столєтовим була встановлена фактична безінерційність зовнішнього фотоефекту. Проміжок часу між початком освітлення і початком фотоструму не перевищує 10^{-9}с.

Фотоефект – результат трьох послідовних процесів: поглинання фотона і поява електрона з великою (порівняно із середньою) енергією; рух цього електрона до поверхні, при якому частина енергії може розсіятись; вихід електрона в інше середовище через поверхню розділу.

Кількісною характеристикою фотоефекту є квантовий вихід β – число електронів, які вилітають під дією одного фотона, що падає на поверхню тіла. Величина β залежить від властивостей тіла, стану його поверхні і енергії фотона. Біля порогу фотоефекту ($\lambda \approx \lambda_{max}$) для більшості матеріалів $\beta \sim 10^{-4}$ електрон/фотон. Така мала величина β обумовлена тим, що світло проникає в метал на глибину $\sim 10^{-7}$м, і там в основному поглинається. Фотоелектрони при своєму русі до поверхні сильно взаємодіють з електронами провідності, яких в металі багато, і швидко розсіюють енергію, отриману від

випромінювання. Енергію, яка достатня для виконання роботи виходу зберігають тільки ті фотоелектрони, які утворились біля поверхні на глибині, яка не перевищує 10^{-9}м. Крім того, поверхні металів сильно відбивають видиме і ближнє ультрафіолетове випромінювання. Різкого збільшення β і зсуву λ_{max} металів у видиму область спектру досягають за допомогою покриття чистої поверхні металу моноатомним шаром електропозитивних атомів або молекул (Cs, Rb, Cs_2O), які утворюють на поверхні дипольний електричний шар.

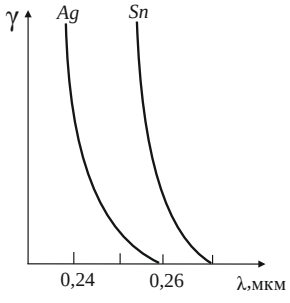


Рис. 16.11

Важливе значення для фотоефекту має залежність спектральної чутливості приймача випромінювання від довжини світлової хвилі (спектральна чутливість приймача γ оптичного випромінювання −це відношення величини, яка характеризує рівень реакції приймача (наприклад, фотоструму), до потоку енергії монохроматичного випромінювання, яке викликає цю реакцію. Вимірюється γ в А/Вт). З рис.16.11 видно, що, починаючи від "червоної межі", із зменшенням λ відбувається зростання чутливості фотокатода.

Фотоефект не можна пояснити з погляду хвильової теорії світла.

А.Ейнштейн в 1905р. показав, що явище фотоефекту і його закономірності можуть бути пояснені на основі запропонованої ним квантової теорії фотоефекту.

Згідно Ейнштейна, світло з частотою ν не лише випромінюється, але і поширюється в просторі і поглинається речовиною окремими порціями, енергія яких $\varepsilon = h\nu$. Поширення світла треба розглядати не як безперервний хвильовий процес, а як потік локалізованих у просторі дискретних світлових квантів, що рухаються зі швидкістю поширення світла у вакуумі. Ці кванти електромагнітного випромінювання отримали назву ***фотонів.***

За Ейнштейном, кожний фотон поглинається лише одним електроном. Тому кількість вирваних фотоелектронів повинна бути пропорційна до кількості поглинутих фотонів, тобто пропорційна до інтенсивності світла (1-ий закон фотоефекту). Безінерційність фотоефекту пояснюється тим, що передача енергії при зіткненні фотона з електроном відбувається майже миттєво.

Енергія падаючого фотона витрачається на виконання електроном роботи виходу Φ з металу і на надання електрону, який вилетів, кінетичної енергії $m_e v_{max}^2/2$.

За законом збереження енергії,

$$h\nu = \Phi + m_e v_{max}^2/2 . \qquad (16.39)$$

Це ***рівняння Ейнштейна для зовнішнього фотоефекту.***

Перевірку формули Ейнштейна з достатньою точністю здійснив в 1916р. Міллікен, який створив прилад, в якому поверхні, які досліджувались, очищались у вакуумі. Після цього вимірювалась їх робота виходу Φ і досліджувалась залежність максимальної кінетичної енергії фотоелектронів від частоти ν світла. Результати вимірювань приведені на рис.16.12. Відповідні графіки залежності $eU_0 = m_e v_{max}^2/2$ приведені для трьох різних металів. Графіки

мають дві цікаві характерні особливості. По-перше, залежність eU_0 від частоти ν – лінійна і, по-друге, графіки для різних металів не співпадають, хоч вони і паралельні один до одного. Лінійна залежність eU_0 від ν підтверджує рівняння Енштейна (16.37). Тангенс кута нахилу прямої, дорівнює сталій Планка h, а відрізок Φ, який відтинається прямою на координатній осі eU_0, – роботі виходу електрона із металу. Стає зрозумілим, чому прямі паралельні (кожна має однаковий нахил h) і чому не зливаються в одну лінію (кожна характеризується своїм певним

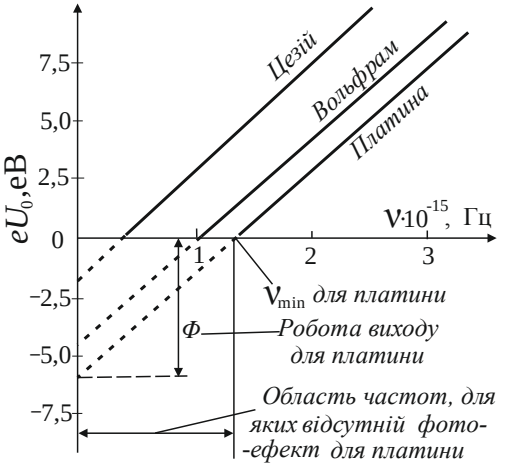


Рис.16.12.

значенням роботи виходу Φ). Коли $eU_0 = 0$, то $\nu = \nu_{min}$. Тут ν_{min} *гранична частота*, тобто мінімальна частота світла, яке падає на метал, при якій ще можливе вибивання електронів із поверхні металу. Цю частоту також називають *червоною межею*. При $\nu = \nu_{min}$ кінетична енергія фотоелектронів буде дорівнювати нулю. Енергії кванта вистачає тільки, щоб вирвати електрон з металу. Отже,

$$h\nu_{min} = h(c/\lambda_{max}) = \Phi. \quad (16.40)$$

Величина λ_{max} залежить лише від роботи виходу електрона, тобто від хімічної природи речовини і стану її поверхні.

Якщо створити поле, яке затримуватиме рух електронів, тобто на катод подати " +", а на анод "−", то фотострум буде зменшуватись і при напрузі U_0 струм дорівнюватиме нулю (рис.16.10). Ця напруга називається затримуючим потенціалом. Тоді формула Ейнштейна матиме вигляд із урахуванням (16.38)

$$h\nu = \Phi + eU_0. \quad (16.41)$$

Якщо замість роботи виходу підставити значення енергії фотона, що відповідає червоній межі фотоефекту, то

$$h\nu = h\nu_{min} + eU_0 \text{ і } eU_0 = h(\nu - \nu_{min}). \quad (16.42)$$

Прямолінійна залежність $eU_0(\nu)$ підтверджується експериментально (рис.16.12).

Прийнято вважати, що фотоефект дає найбільш прямий експериментальний доказ квантової природи випромінювання.

§16.6.Маса та імпульс фотона

Менш ніж через двадцять років після відкриття фотоефекту він виявився одним із явищ, які стали основою квантової теорії світла.

В 1905р. А. Ейнштейн в роботі « Про одну еврістичну точку зору, що торкається виникнення і перетворення світла» довів гіпотезу квантів Планка до логічного кінця і висловив ідею про те, що світло є потік частинок – квантів – фотонів (фотонами назвав кванти світла в 1929р. фізико – хімік Льюїс).

Теорія відносності вимагала, щоб квант випромінювання, який поглинається (або випромінюється), ніс із собою не тільки енергію, але й імпульс, який дорівнює

$$p = \varepsilon/c = \hbar k, \quad (16.43)$$

де $\varepsilon = \hbar\omega$ −енергія кванта, c−швидкість світла у вакуумі, $k = 2\pi/\lambda$. В загальному випадку, $\vec{p} = \hbar\vec{k}$, де $\vec{k}$ −хвильовий вектор.

Звідси уже витікало, що електромагнітне поле не тільки поглинається і випромінюється порціями, але його можна розглядати як сукупність квантів (частинок), які мають енергію і імпульс. В цьому остаточно переконались тільки через два десятиліття – в 1925р., коли фізик Бозе прислав Ейнштейну статтю, в якій він довів, що теплові статистичні властивості випромінювання співпадають із властивостями газу із квантів-фотонів.

Може здатися дивним, що фотон має імпульс, який відрізняється від енергії ε тільки множником c. Однак формули для енергії й імпульсу фотона не потрібно порівнювати з аналогічними формулами для частинки в механіці Ньютона. Справа в тому, що фотон − це частинка, яка летить зі швидкістю світла, а до таких частинок необхідно застосовувати формули теорії відносності, в якій енергія пов'язана з імпульсом співвідношенням

$$\varepsilon^2 = c^2p^2 + m_0^2c^4, \qquad (16.44)$$

де m_0 −маса спокою частинки, тобто маса частинки при швидкості, що дорівнює нулю. Якщо врахувати формулу (16.43) , то із (16.44) витікає, що маса спокою фотона $m_\lambda = m_0 = 0$. Якщо ми схильні розглядати фотони як частинки, то цей результат на перший погляд є дивним. І хоч в квантах Планка реалізувалось передбачення Ньютона, що світло це пучок частинок-фотонів, але вони все-таки не є корпускулами Ньютона: *у фотонів не виявилось маси спокою*. Крім того, вони, на відміну від класичних частинок, могли інтерферувати. Такою властивістю наділила електрони наука XX століття − квантова механіка. Фотон в будь-якій системі відліку має швидкість c і не існує систем відліку, в якій він знаходився б в стані спокою. Рівність нулю маси спокою означає, *що фотон в стані спокою − поняття, яке не має змісту*.

Боротьба за утвердження нового погляду на природу світла була довгою і складною. Її гострота пояснювалась також і тією обставиною, що зміни , які пропонувалися, зачіпали не тільки оптику (світло), але й всю електродинаміку. Із часів Максвелла і Герца ці області не можна було розглядати вже у відриві одна від одної. Якщо погляди на світло зазнавали кардинальних змін, то на закони електродинаміки вже не можна було дивитись по-старому.

Зусиллями багатьох вчених була побудована квантова електродинаміка, яка принципово по-іншому, ніж теорія Максвелла, пояснювала взаємодію зарядів. Але це не значить, що закони класичної електродинаміки були повністю відкинуті. Взаємовідношення квантових і класичних законів визначаються *принципом відповідності*: класичні закони є граничним випадком квантових, і, як наслідок, між ними повинен існувати зв'язок. Одним з проявів цього зв'язку є твердження: *якщо закон Кулона не є точним законом "обернених квадратів ($F \sim 1/r^2$)", то масса спокою фотона* $\boldsymbol{m_\lambda}$ *повинна відрізнятися від нуля.* Варто підкреслити, що питання про рівність нулю m_λ − це питання експериментальне. На цю обставину вперше звернув увагу один із творців квантової механіки Л.де Бройль.

Багато важливих положень сучасної фізики базується на припущені про те, що $m_\lambda = 0$. Тому проблема експериментального підтвердження цього припущення повинна розглядатись як принципіально важлива.

Дослідна перевірка закону Кулона – один із шляхів її розв'язку.

В доквантовий час відхилення від закону Кулона шукали, вважаючи, що потенціал заряду описується не просто формулою $\varphi = q/(4\pi\varepsilon_0 r)$, а з експоненціальним множником:

$$\varphi = \frac{q}{4\pi\varepsilon_0 r}\, e^{r/a}\,, \qquad (16.45)$$

де a – деяка постійна.

Якби потенціал електричного заряду мав вид (16.45), то всередині зарядженої сфери існувало би електричне поле. Саме таке поле всередині зарядженої сфери і намагалися знайти експериментатори. Ніякого поля знайти не вдалось. Результати дослідів виражались в оцінці параметра a. Так із правильності закону Кулона випливало, що по крайній мірі $a > 2 \cdot 10^7$м.

Виникає питання: а яке відношення має вимірювання довжини a до гіпотетичної маси фотона? Зрозуміло, що в класичній фізиці ні про яку масу фотона говорити не можна, і точність закону Кулона можна характеризувати тільки відстанню, в межах якої був перевірений цей закон. Після того як в фізику ввійшла стала Планка, маса частинки m виявилась пов'язана з довжиною a формулою:

$$a = \frac{\hbar}{mc}\cdot \qquad (16.46)$$

Не важко перевірити, що така комбінація фундаментальних величин має розмірність довжини. Теорія прийшла до висновку , що якби $m = m_\lambda \neq 0$, то на відстані a від заряду потенціал електричного поля був би в e=2,72… рази більшим, ніж дає закон Кулона.

Вимірювання поля всередині зарядженої сфери – не найбільш точний спосіб перевірки закону Кулона. Значно точніше було би перевірити закон Кулона, вимірюючи магнітне поле Землі на різних відстанях або, що виявилось ще більш чутливим, вивчаючи магнітне поле Юпітера. Ці вимірювання дали оцінку $a > 0{,}5 \cdot 10^8$м $(m_\lambda < 7 \cdot 10^{-51}$кг).

Однако перевірка закону Кулона – не єдиний спосіб оцінки верхньої межі m_λ. Теоретичний аналіз показує, що якщо $m_\lambda \neq 0$, то повинна спостерігатись дисперсія електромагнітних хвиль в вакумі, тобто швидкість їх поширення v повинна залежати від довжини хвилі λ (відсутність подібної дисперсії—один із наслідків теорії Максвелла). В цьому випадку для групової швидкості електромагнітних хвиль $v_{гр}$ можна було б записати:

$$v_{гр} = \frac{c}{\sqrt{1+(\lambda/(2\pi a))^2}}\,, \qquad (16.47)$$

де c— гранична швидкість руху тіл, яка входить в перетворенна Лоренца (очевидно, що при $m_\lambda \neq 0$ швидкість $v_{гр}$ не співпадає із швидкістю світла c), λ – довжина електромагнітної хвилі, $2\pi a = h/(m_\lambda c)$ – комптонівська довжина хвилі фотона. Із виразу (16.47) витікає, що при $m_\lambda \neq 0$ швидкість синього світла (λ~0,4 мкм) була би більша, ніж червоного (λ~0,7мкм). Той факт, що в дійсності дисперсія електромагнітних хвиль у вакуумі при досягнутій точності вимірювань не виявлена, теж свідчить на користь твердження, що $m_\lambda = 0$. Однак оцінка верхньої межі m_λ, яка отримана із аналізу данних про розповсюдження електромагнітних хвиль, дає значення дещо гірше, ніж досліди з перевірки закону Кулона.

Ще більш точну оцінку маси m_λ отримують із того факту, що в галактиках спостерігаються магнітогідродинамічні хвилі в зарядженій плазмі. В земних умовах такі хвилі

спостерігаються в установках, де створюються умови для термоядерних реакцій. Хвильові процеси в плазмі переносять з собою періодично змінне електромагнітне поле. Існування маси спокою у фотона привело б до «згасання» довгохвильових коливань. В галактиках магнітогідродинамічні процеси захоплюють величезні відстані. Звідси витікає фантастична оцінка $a > 10^{20}$м ($m_\lambda < 3{,}5 \cdot \cdot 10^{-63}$кг), величина достатня для того, щоб не сумніватись в правильності гіпотези про рівність нулю m_λ.

В 1980р. американські фізики Дж.Примак і М.Шер висловили припущення, що при пониженні температури може відбутись фазовий перехід, при якому фотон із безмасової частинки перетвориться в частинку з масою спокою $m_\lambda \approx 2 \cdot 10^{-38}$ кг. Група експериментаторів із Прістонського університету в 1985р. перевірили гіпотезу про «холодний масивний» фотон за допомогою досліду. Ідея нової перевірки закону Кулона залишилась попередньою. Однак виявилось зручним використовувати не систему вкладених куль (метод Кавендиша), а три вкладених коаксіальних циліндри. На два зовнішніх циліндри подавалась змінна напруга. У випадку відхилення закону Кулона від точного закону «обернених квадратів» між двома внутрішніми циліндрами, які зв'язані між собою кінцевою ємністю (≈ 20мкФ), повинен був протікати змінний струм, який експериментатори старались зареєструвати. Основна частина експериментальної установки розміщувалась в кріостаті, де підтримувалась низька температура. Досліди проводились при $T = 4{,}2$ К ($-268{,}8$ ℃) і 1,36К ($-271{,}64$ ℃). Результати аналізу отриманих даних показали, що при переході від 4,2К до 1,36К фазового переходу не спостерігалось і обмеження на можливу масу «холодного» фотона визначалась нерівністю $m_\lambda \leq (1{,}50 \pm 1{,}38) \cdot 10^{-45}$кг (при T=1,36К).

Таким чином, досліди, які виконані в земних умовах, а також оцінки, які зроблені шляхом аналізу астрофізичних даних, не дають підстав вважати масу спокою фотона відмінною від нуля і, отже, закон взаємодії точкових електричних зарядів зберігає статус точного закону «обернених квадратів».

§16.7.Тиск світла

Фотон, подібно до частинки, яка рухається, має енергію $\varepsilon = \hbar\omega$ та імпульс $p = \hbar\omega/c$. Усі ці корпускулярні характеристики пов'язані з хвильовою характеристикою світла – його частотою ω. Одним з підтверджень наявності у фотонів імпульсу є світловий тиск. Тиск світла, який передбачала електромагнітна теорія Максвелла, вперше експериментально виявив і виміряв у 1901 р. П.М.Лебедєв.

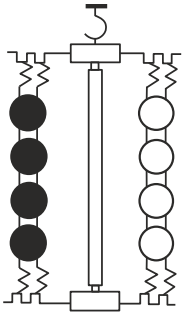


Рис.16.13.

Прилад Лебедєва – це досить чутливі крутильні терези, рухома система яких складалась із легкого каркаса із закріпленими на ньому тонкими кружками, які розміщені симетрично відносно осі підвісу (рис.16.13).

Деякі крильця були дзеркальними, а поверхня інших була зачорнена. Усю цю систему підвішували на тонкій пружній нитці всередині закритого скляного балона, в якому

був створений високий вакуум. Крильця освітлювали вольтовою дугою. Світло від дуги направлялось на них за допомогою системи лінз і дзеркал. Величину світлового тиску на крильця визначали за кутом закручування нитки підвісу.

Тиск світла настільки малий, що для його надійного вимірювання треба було усунути вплив на крильця всіх інших факторів. Вплив конвекційних струмин повітря Лебедєв усунув, створивши в балоні досить високий вакуум. Однак і в такому разі не було усунуто так званий радіометричний ефект. Причина його в тому, що зачорнене крильце нагрівається внаслідок поглинання світла, яке падає на нього, причому температури освітленої і неосвітленої (задньої) поверхонь крильця неоднакові. Молекули розрідженого повітря в балоні відбиваються від нагрітої поверхні крилець з більшою швидкістю, передаючи їм відповідно більший імпульс. Тиск, зумовлений такою дією, значно більший за світловий. Радіометричний ефект може призвести до того, що в досліді тиск на зачорнене крильце буде більший за тиск на дзеркальне крильце тих самих розмірів. Лебедєв усунув вплив радіометричного ефекту, використавши у своїх дослідах дуже тонкі крильця різної товщини від 0,01 до 0,1 мм.

Тиск світла на дзеркальне крильце з коефіцієнтом відбиття R=1 виявився у два рази більший, ніж тиск на зачорнене крильце (R=0).

З точки зору корпускулярних уявлень, тиск світла інтерпретується як результат передачі імпульсів фотонами поверхні, яка поглинає або відбиває їх. При нормальному падінні кожний фотон, який поглинається, передає поверхні імпульс $\hbar\omega/c$, а який відбивається – $\hbar\omega/c$ –($-\hbar\omega/c$)=$2\hbar\omega/c$. На одиничну площину поверхні за одиницю часу падає енергія випромінювання Uc , де U – об'ємна густина енергії (формула (16.2). Тому число фотонів, які падають, буде $N = Uc/\hbar\omega$. Із цього числа NR фотонів відіб'ється, а $N(1-R)$ – поглинеться (R – енергетичний коефіцієнт відбивання поверхні). За визначенням, тиск світла дорівнює повному імпульсу, який передається падаючими фотонами на 1м^2 поверхні за 1с, тобто

$$p = NR \cdot \frac{2\hbar\omega}{c} + \frac{N(1-R)\hbar\omega}{c} = \\ = \frac{(1+R)N\hbar\omega}{c}.$$

Врахувавши, що $U = N\hbar\omega/c$, отримаємо:

$$p = U(1+R), \qquad (16.48)$$

що співпадає із виразом, отриманим на основі електромагнітної теорії.

Часто для визначення тиску світла на поверхню з коефіцієнтом відбивання R використовують формулу:

$$p = \frac{I}{c}(1+R)\cos^2\varphi, \qquad (16.49)$$

де I – інтенсивність світла, яка вимірюється у Вт/м^2, φ – кут падіння світла на плоску поверхню (кут між нормаллю до поверхні і променем світла, який падає на цю поверхню).

Ізотропне випромінювання, яке розповсюджується у всіх напрямках, чинить на поверхню стінки тиск

$$p = U/3\,. \qquad (16.50)$$

Цей вираз можна отримати, розглядаючи випромінювання в замкненій оболонці як ідеальний газ фотонів.

Світловий тиск досить малий. Так, за межами атмосфери Землі інтенсивність сонячного випромінювання дорівнює 1400 Вт/м^2. Відповідний тиск при нормальному падінні світла на поверхню, для якої $R = 0{,}5$, дорівнює $7\cdot 10^{-6}$Па. Ця величина у 10^{10} менша

від атмосферного тиску. Незважаючи на те, що тиск світла малий, його необхідно враховувати у ряді випадків. Так, сила гравітаційної взаємодії частинок пропорційна кубу їх радіуса, а сила світлого тиску пропорційна квадрату радіуса частинки. Для частинок досить малих розмірів ці сили можуть бути однакових порядків. Це дає підставу зробити припущення, що хвости комет зумовлені світловим тиском.

Всередині зірок температура досягає мільйонів градусів. Таким температурам відповідають великі густини енергії випромінювання і тиск світла перешкоджає гравітаційному стисканню зірок. Можливо, цим і пояснюється наявність верхньої межі маси зірок, яка має порядок 10^{30}кг. Крім цього, випромінювання і поглинання його всередині зірки зумовлює швидке перенесення деякої маси з одних областей в інші. Випромінювання, що виходить з центральних областей, має менший момент імпульсу, ніж ті маси периферійних областей, що поглинають випромінювання. Це зумовлює сповільнене обертання зірки.

Нерівномірність освітлення поверхонь штучних супутників Землі викликає небажане їх обертання навколо деякої осі.

Фокусування лазерного пучка у «пляму», радіус якої дорівнює довжині хвилі, дає змогу одержати порівняно великі тиски. За їх допомогою мікроскопічним частинкам можна надати прискорення, які в мільйони разів більші за прискорення вільного падіння і це має практичне застосування.

§16.8.Ефект Комптона

Існують явища, в яких імпульс фотона проявляє себе явно і співвідношення $\vec{p}=\hbar\vec{k}$ допускає експериментальну перевірку. Особливо чітко проявляються корпускулярні властивості світла в явищі розсіяння рентгенівських променів (Про рентгенівські промені детальніше в §17.11). А.Комптон, досліджуючи у 1923р. розсіяння монохроматичних рентгенівських променів речовинами з легкими атомами (парафін, бор), виявив, що у складі розсіяного випромінювання поряд з випромінюванням початкової довжини хвилі λ, спостерігається також випромінювання більшої довжини хвилі λ'.

Схема досліду Комптона зображена на рис.16.14. Пучок монохроматичних рентгенівських променів джерела *R*, пройшовши систему діафрагм *Д*, падає на розсіюючу речовину *P* і після розсіяння на кут θ попадає в рентгенівський спектрограф *C*, де вимірюється довжина хвилі розсіяного випромінювання. Досліди Комптона показали, що довжина хвилі λ' розсіяного випромінювання більша за довжину хвилі λ падаючого випромінювання, причому різниця $\Delta\lambda = \lambda' - \lambda$ не залежить від довжини хвилі падаючого випромінювання і природи розсіювальної речовини, а визначається лише величиною кута розсіяння θ.

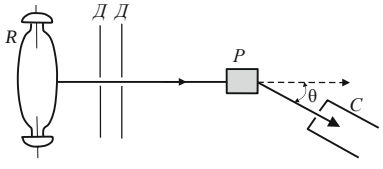


Рис.16.14.

Пояснити ефект Комптона можна на основі квантових уявлень про природу світла. Якщо вважати, що випромінювання є потоком фотонів, то ефект Комптона – це результат пружного зіткнення рентгенівських фотонів з вільними електронами. В атомах легких

елементів енергія зв'язку електрона (енергія іонізації) порядку 10 еВ, що приблизно в тисячу раз менше енергії рентгенівського фотона ($\hbar\omega \approx 10$кеВ). Тому електрони в цих дослідах можна вважати практично вільними. Енергія спокою електрона $m_e c^2 \approx 0{,}5$МеВ, що набагато більше енергії фотона. Тому електрон, який до зіткнення знаходився в стані спокою або рухався в атомі з нерелятивістською швидкістю, і, після зіткнення з таким фотоном, залишається нерелятивістським. Нехтуючи початковою енергією електрона, запишемо закон збереження енергії

$$\hbar\omega = \hbar\omega' + \frac{p^2}{2m_e}, \qquad (16.51)$$

де $\omega = 2\pi c/\lambda$, $\omega' = 2\pi c/\lambda'$, p – імпульс електрона після зіткнення з фотоном, а закон збереження імпульсу

$$\vec{p}_\phi = \vec{p}_\phi' + \vec{p}, \qquad (16.52)$$

де $\vec{p}_\phi$ і $\vec{p}_\phi'$ – імпульс фотона до і після розсіювання (рис.16.15). Згідно теореми косинусів вираз (16.52) переписуємо так:

$$p^2 = p_\phi^2 + (p_\phi')^2 - 2p_\phi p_\phi' \cos\theta. \qquad (16.53)$$

Підставимо в (16.53) вирази для імпульсу фотона через його частоту $p_\phi = \hbar\omega/c$ і $p_\phi' = \hbar\omega'/c$ і квадрат імпульсу електрона із закону збереження енергії (16.51). Тоді (16.53) прийме вигляд:

$$2m_e\hbar(\omega - \omega') = \frac{\hbar^2}{c^2}\{\omega^2 + ((\omega')^2 - -2\omega\omega'\cos\theta.$$

Перепишемо це співвідношення у вигляді

$$\omega - \omega' = \left\{\frac{\hbar\omega}{2m_e c^2}\right\} \cdot \{(\omega^2 + (\omega')^2)/\omega - 2\omega'\cos\theta\} \qquad (16.54)$$

і враховуючи, що $\hbar\omega/(m_e c^2) \ll 1$, бачимо, що зміна частоти $\Delta\omega = \omega' - \omega$ мала порівняно з самою частотою. Тому у правій частині (16.54) можна замінити ω' на ω. Тоді для відносного зсуву частоти $\Delta\omega/\omega$ при розсіянні отримаємо:

$$\frac{\Delta\omega}{\omega} = -\frac{\hbar\omega}{m_e c^2}(1 - \cos\theta) = -\frac{2\hbar\omega}{m_e c^2}\sin^2\frac{\theta}{2}\cdot$$

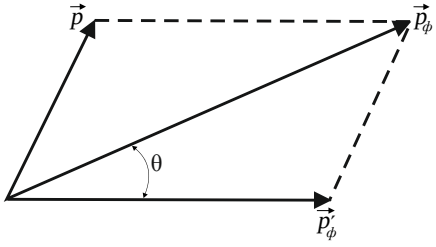


Рис.16.15.

Знак мінус показує, що частота випромінювання при розсіянні зменшується. Це зрозуміло, так як частину своєї енергії фотон передає електрону. Переходячи тепер від частот до довжин хвилі $\lambda = 2\pi c/\omega$, $\Delta\lambda/\lambda = -\Delta\omega/\omega$, отримаємо, що

$$\Delta\lambda = \frac{2h}{m_e c}\sin^2\frac{\theta}{2} = 2\Lambda\sin^2\frac{\theta}{2}\cdot \qquad (16.55)$$

Співвідношення (16.55) дуже добре описує залежність збільшення довжини хвилі від кута розсіяння, яка спостерігається на досліді. Константа Λ, яка входить в (16.55), називається *комптонівською довжиною хвилі* електрона, що виражається у вигляді комбінації трьох універсальних сталих: $\Lambda = h/(m_e c)$ (порівняйте формулу для Λ і формулу (16.46) для a). Підставляючи числові значення сталих, отримаємо $\Lambda = 0{,}24$нм, що співпадає з величиною, яка знайдена із вимірювань зміщення довжини хвилі при розсіянні.

При розсіянні фотона частина його енергії передається електрону. Електрони віддачі грають важливу роль в процесі іонізації газів рентгенівським випромінюванням. Використовуючи камеру Вільсона, яка розміщена в магнітному полі, можна знайти імпульс і енергію електронів віддачі, що

дозволяє ще більш повно перевірити виконання законів збереження. Вимірювання показали, що в кожному елементарному акті розсіяння електрон набуває якраз такі енергію і імпульс, які втрачає фотон.

Досліди Комптона яскраво продемонстрували, що енергія й імпульс фотона дійсно зв'язані формулою (16.43) і що закони збереження енергії і імпульсу виконуються в елементарних актах розсіяння.

Контрольні питання

1.Чим відрізняється теплове випромінювання від інших видів випромінювання?

2.Які є характеристики теплового випромінювання?

3.Які є характеристики поверхні тіла, яка випромінює?

4.Яке тіло вважається абсолютно чорним (АЧТ).

5.Закон Кірхгофа для теплового випромінювання.

6.Закони випромінювання АЧТ.

7.Формула Релея-Джинса та її практичне застосування.

8.Формула Планка та квантування енергії.

9.Безконтактні способи вимірювання температури нагрітих тіл. Які розрізняють типи температур?

10.Зовнішній фотоефект. Рівняння Ейнштейна для фотоефекту.

11.Корпускулярна природа світла: енергія та імпульс фотона.

12.Маса спокою фотона. Чому є принциповим відсутність маси спокою у фотона?

13.Тиск світла на поверхню. Яка природа такого тиску?

14.В чому суть ефекту Комптона? Що цей дослід довів?

Розділ 17.Будова атома

§17.1.Дослід Резерфорда. Планетарна модель атома

Починаючи з Демокріта, вважали, що атом – нескінченно мале абсолютно пружне тіло, що існує вічно, у незмінному вигляді. До початку XX ст. мало у кого, навіть у такого геніального фізика, як Дж. Максвелл, таке твердження породжувало сумніви.

Наприкінці XIX ст. особливо бурхливий розвиток фізики почав гальмуватися проблемами, пов’язаними з внутрішньою структурою атомів та їхніми властивостями. На той час було встановлено, що:

атоми є подільними складними електронейтральними об’єктами, в яких одними із структурних компонентів є електрони;

найістотніша за величиною взаємодія має електричний характер;

атом одного елемента може перетворюватися в атом іншого.

Залишалося нез’ясованим питання про внутрішню будову атома. Адже, за теоремою Ірншоу, ніяка конфігурація позитивних і негативних зарядів, пов’язаних між собою кулонівською взаємодією, не є стійкою. Навіть якщо створити таку стійку систему, то нескінченно мале зміщення будь-якого із зарядів незворотно цю стійкість знищує. Запропонована Кельвіном 1902 року модель, за якою позитивний заряд рівномірно розподілений по всьому об’єму атома, а електрони рівномірно розподілені на тлі позитивного заряду, забезпечувала стійкість атома. Така модель нагадує кекс з родзинками – позитивне желе з вкрапленими в нього електронами.

Англійський фізик Дж. Дж. Томсон доповнив модель Кельвіна припущенням, що електрони в цьому желе рухаються по замкнутих орбітах. Остаточно модель набула назви *модель атома Томсона*. Упродовж 10 років свого існування така модель дала змогу описати складний характер явищ дисперсії і поглинання світла, породила вчення про поляризацію світла, передбачила існування ізотопів тощо.

Англійський фізик Резерфорд піддав сумніву правильність моделі Томсона. Він виконав досліди, в яких аналізували зміни в траєкторіях потоку α – частинок, що проходили через надтонкі металічні пластинки. Відхилення в траєкторіях фіксували спалахами на флуоресцентному екрані в точках попадання відхиленої α – частинки. Маса α – частинки в 7300 разів більша від маси електрона. Тому зіткнення α – частинки з електроном неістотно змінить її траєкторію. Проте зіткнення α – частинки з масивним атомом може відчутно змінити і траєкторію, і її енергію.

α – частинки утворюються під час природного радіоактивного розпаду деяких важких елементів і є позитивно зарядженими частинками з зарядом $2e$ і масою, яка приблизно в чотири рази більша за масу атома водню.

Спрощена схема досліду Резерфорда зображена`на рис.17.1. Джерело α – частинок поміщене всередині свинцевої порожнини з вузьким каналом. Усі α – частинки, крім тих, що рухаються всередині вузького каналу, поглинаються свинцем. Потік α – частинок, пройшовши через вузьку діафрагму *Д*, потрапляє на тонку золоту фольгу *Ф* завтовшки $10^{-6} – 10^{-7}$м, що складається з декількох атомних шарів. При проходженні через фольгу α – частинки відхиляються на різні кути θ і потрапляють на екран *Е*, який можна поміщати в різних положеннях відносно фольги. Екран покритий флюоресцентною речовиною. За допомогою мікроскопа *М* можна спостерігати місця потрапляння α – частинок за свіченням (спалахами) екрану. Поведінку α – частинок після проходження через фольгу вивчали і в камері Вільсона.

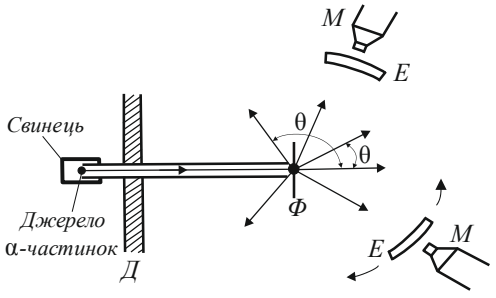


Рис.17.1.

В результаті численних спостережень таких спалахів було встановлено, що переважна більшість α – частинок відхилялась на кути 1-3° щодо їх траєкторії перед пластинкою. Проте траплялися відхилені α – частинки на кути до 150°. Відхилені α – частинки на кути, більші ніж 90°, у середньому, спостерігалися лише одна на 8000. Попри їхню мізерну кількість, вони є важливими з принципового погляду – їх появу ніяк не можна обгрунтувати в рамках моделі Томсона.

З метою пояснення несподіваних результатів експерименту Резерфорд висунув нове припущення про будову атома, за яким додатний заряд атома, на противагу до моделі Томсона, зосереджений у доволі малому за розмірами об'ємі – ядрі. Електрони обертаються навколо ядра, утримуючись на орбітах внаслідок зрівноваження відцентрових сил силою кулонівського притягання до ядра. Така картина нагадує планетарну систему,

подібну до Сонячної, тому така модель будови атома отримала назву планетарної моделі.

Для обґрунтування планетарної моделі будови атома Резерфорд проаналізував процес розсіяння заряджених частинок на атомах речовини з погляду класичної механіки. Коли частинка пролітає біля ядра, на неї діє кулонівська сила відштовхування

$$F = \frac{2Ze^2}{4\pi\varepsilon_0 r^2},$$

де $2e$−заряд α − частинки, Ze−заряд ядра речовини, r − відстань між α − частинкою і ядром. В цьому випадку траєкторія α − частинки є гіпербола. Схема процесу розсіяння зображена на рис.17.2. Тут p – так звана прицільна відстань що, фактично, була б відстанню найбільшого наближення α – частинки до атомного ядра якби вона пролітала без розсіяння. Припускається, що α – частинки, влітаючи в область, зайняту атомом (на рис.17.2 зображене ядро атома, заряд якого $+Ze$) не розсіюються на електронах атома, а лише на ядрі. Можна навести вагомі аргументи

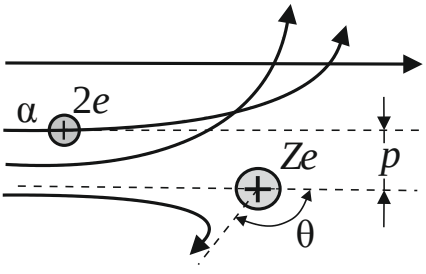


Рис.17.2.

на підтвердження такого припущення. Маса α – частинки, як зазначалося, набагато більша від маси електрона. Тому енергетичні втрати α – частинки у разі однократного пружного зіткнення з електроном будуть дуже малими і це не вплине на її траєкторію. Є також й інші аргументи, які тут не подаємо, бо це вимагає знання квантових понять, які вивчаються пізніше.

Використовуючи закони збереження енергії та імпульсу в замкнутій системі «α – частинка – ядро» та беручи до уваги, що маса ядер значно більша від маси α – частинки, Резерфорд знайшов таке співвідношення між параметрами задачі розсіяння:

$$\operatorname{ctg}\frac{\theta}{2} = W_k \frac{2\pi\varepsilon_0 p}{Ze^2}. \qquad (17.1)$$

Тут $W_k = mv^2/2$ – кінетична енергія α – частинки (v, m – її швидкість та маса, відповідно), а Z – зарядове число. Експериментально перевірити справедливість формули (17.1) неможливо через неможливість виміряти прицільний параметр p (рис.17.2).

Закони динаміки дозволяють розрахувати залежність кута відхилення θ від параметра удару p, а методами теорії ймовірностей можна знайти ймовірність пролітання α – частинки на даній відстані p від частинки і тим самим ймовірність її відхилення на кут θ. Розрахунок показує, що із загального числа N падаючих частинок у пучку, число розсіяних dN_θ α – частинок в тілесному куті $d\Omega$, що міститься між двома конічними поверхнями, твірні яких утворюють кути θ і $\theta + d\theta$ з початковим напрямом руху α – частинок визначається за формулою:

$$dN_\theta = nN\left(\frac{Ze^2}{8\pi\varepsilon_0 W_k}\right)^2 \cdot \frac{d\Omega}{\sin^4(\frac{\theta}{2})}, \qquad (17.2)$$

де n – кількість ядер з розрахунку на одиницю площі поверхні мішені. Це і є *формула Резерфорда* для розсіяння α – частинок. Зауважимо, що формула (17.2) справедлива лише для малих значень прицільних віддалей p, тобто для частинок розсіяних тільки на великі кути. Якщо прицільна віддаль велика, то частинка пролітає поза ділянкою, зайнятою

атомом. Атом, загалом, тепер потрібно розглядати як електро- нейтральний. Характер взаємодії між частинкою, що розсіюється, та атомом змінюється – взаємодія стає відмінною від точкової кулонівської.

Для певної речовини (наприклад, золото), яка розсіює α−частинки при певній енергії і заданій густині потоку, добуток $\frac{dN_\theta}{d\Omega}\sin^4(\frac{\theta}{2})$ повинен бути сталим, що було експериментально підтверджено в дослідах Гейгера і Марсдена. Крім висновку про справедливість планетарної моделі атома, з досліду Резерфорда очевидним був ще один важливий висновок: кількість електронів в атомі збігається з його порядковим атомним числом Z у періодичній системі елементів. Оскільки в нормальному стані атом електронейтральний, то заряд ядра у такому атомі дорівнює $+Z|e|$. До того ж досліди Резерфорда опосередковано свідчили про те, що закон Кулона для взаємодії заряджених частинок залишається справедливим навіть для дуже малих відстаней між ними.

На підставі результатів дослідів з розсіяння α – частинок тонкими фольгами Резерфорд в 1910р. запропонував *ядерну модель атома*.

Згідно з цією моделлю, в центрі атома знаходиться ядро, в якому зосереджені позитивний заряд $+Ze$ і практично вся маса атома. Лінійні розміри ядра $\sim 10^{-14} – 10^{-15}$ м.

Навколо ядра в області з лінійними розмірами $\sim 10^{-10}$м по замкнених орбітах рухаються Z електронів, утворюючи електронну оболонку (*електронний рій*) атома.

Ядерна модель Резерфорда зовні дуже нагадує Сонячну систему: у центрі – ядро, навколо нього по орбітах рухаються електрони. Тому цю модель називають планетарною. Орбіти електронів в атомі стаціонарні, атому властива виняткова стійкість. Стійкість атома не можна погодити з класичним поясненням ядерної моделі.

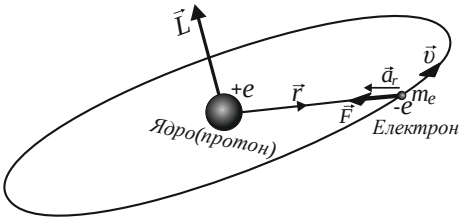


Рис.17.3.

Розглянемо, наприклад, ядерну модель найпростішого атома – атома водню. Для простоти допустимо, що електрон, маса якого m_e і заряд $-e$ рухається рівномірно по коловій орбіті навколо протона (з зарядом $+e$), який займає центральне положення (рис.17.3). В першому наближені рухом протона, маса якого майже в 1836 раз більша маса електрона, можна знехтувати.

Сила, яка утримує електрон на орбіті, є сила кулонівського притягання між протоном і електроном. Ця центральна сила дорівнює $F = \frac{e^2}{4\pi\varepsilon_0 r^2}$, де r – радіус колової орбіти електрона. За другим законом Ньютона запишемо рівняння руху електрона:

$$\frac{m_e v^2}{r} = \frac{e^2}{4\pi\varepsilon_0 r^2}, \qquad (17.3)$$

де $a_r = v^2/r$ – центробіжне прискорення (рис.17.3). На основі рівняння (17.3) кінетична енергія електрона буде:

$$W_k = \frac{m_e v^2}{2} = \frac{e^2}{8\pi\varepsilon_0 r} \cdot \qquad (17.4)$$

Потенціальна енергія системи (електрон − протон)

$$W_p = -\frac{e^2}{4\pi\varepsilon_0 r} \cdot \qquad (17.5)$$

Знак «мінус» означає, що в даній системі діють сили притягання, а не відштовхування. Повна енергія системи

$$W = W_k + W_p = -\frac{e^2}{8\pi\varepsilon_0 r}\cdot \quad (17.6)$$

Знак «мінус» показує, що дана система (електрон-протон) є зв'язаною.

Енергія зв'язку електрона визначається як мінімальна енергія, яка необхідна для повного видалення електрона із атома, або, другими словами, для іонізації атома. Експериментальним шляхом було знайдено, що для атома водню енергія зв'язку дорівнює 13,53 еВ. Підставляючи це значення енергії в (17.6) для W, можна знайти радіус атома r: $r = 0{,}53 \cdot 10^{-10}$ м = 0,53 Å.

Лінійна швидкість v звязана з частотою ω обертання електрона на орбіті співвідношенням: $v = \omega r$.

Тоді із (17.3) можемо знайти, що

$$\omega = \frac{e}{2r}\sqrt{\frac{1}{\pi\varepsilon_0 m_e r}} = 4{,}125 \cdot 10^{16}\text{рад/с}. \quad (17.7)$$

Це значення теж співпадає із значенням ω, знайденим за допомогою інших методів.

Не зважаючи на певні успіхи в поясненні будови атома, які досягнуті в рамках планетарної моделі і класичної теорії, перед фізиками постав ряд непереборних протиріч. Так, згідно законів класичної електродинаміки:

1) заряджена частинка, яка рухається з прискоренням, повинна неперервно випромінювати електромагнітну енергію (§11.5, формула (11.41));

2) частота цього випромінювання повинна дорівнювати частоті обертання частинки навколо ядра. Отже, у відповідності з цією моделлю повна енергія атома повинна зменшуватись (стаючи все більше від'ємною), тоді як частота обертання (формула (17.7)) повинна неперервно зростати. Розрахунок за формулою (11.43) показує, що при цих умовах за час ($\simeq 10^{-8}$с) електрон впаде на ядро і атом перестане існувати (рис.17.4). Крім цього, якщо дана модель правильна, то оптичний спектр випромінювання водню (як і спектри інших

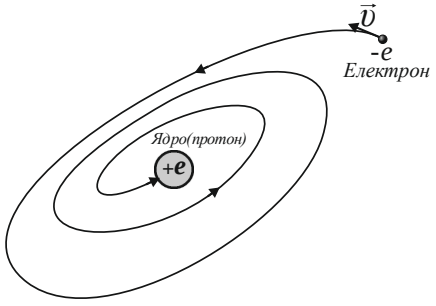


Рис.17.4.

елементів) повинні бути неперервними, а це означає, що всі атоми за дуже короткий час будуть зруйновані. Обидва висновки знаходяться в повному протиріччі з експериментальними даними: атом є дуже стійкою системою, а оптичні спектри газів мають зовсім не неперервний розподіл за частотою, а характеризується набором дискретних частот, тобто мають лінійчасту структуру. Всі ці факти привели до того, що планетарна модель в класичному розумінні швидко була признана такою, що не відбулась.

§17.2.Атомні спектри

Електричний розряд в трубці, яка містить одноатомний газ під низьким тиском, є джерелом світла, аналіз якого за допомогою дифракційної решітки або спектрографа з призмою, дозволяють виявити серію дуже чітких ліній випромінювання. Ці лінії, які характеризують газ, що використовується в трубці, називають лінійчастим спектром даного газу. Якщо в розрядній трубці використовується азот, то отримається так званий смугастий спектр, який містить дискретні групи тісно розташованих ліній. Цей смугастий спектр характеризує двохатомні

молекули N_2 і має інше походження порівняно з лінійчастими спектрами. Біле світло від джерела випромінювання з ниткою розжарення характеризується суцільним спектром і містить неперервний набір довжин хвиль. Якщо таке світло пропускати через одноатомний газ, наприклад водень, то виникає спектр поглинання. При цьому на спектрограмі отримується спектр з темними лініями на світлому фоні. Положення цих ліній відповідає довжинам хвиль яскравих спектральних ліній випромінювання водню, тобто газ поглинає падаюче випромінювання саме з такими довжинами хвиль.

Дослідження атомних спектрів започаткували ще 1860 року, коли була опублікована праця Г. Кірхгофа і Р. Бунзена «Хімічний аналіз за допомогою спостереження спектра». У наступні роки такі дослідження здійснювались доволі інтенсивно. Було встановлено характерну ознаку атомних спектрів – їхню дискретність. Спектри складаються з набору вузьких смуг (ліній) певної частоти. Так, наприклад, у видиму ділянку спектра найпростішого хімічного елемента атомарного водню попадає лише чотири лінії, які прийнято позначати $H_\alpha, H_\beta, H_\gamma, H_\delta$.

Упродовж двох десятків років інтенсивних досліджень атомних спектрів було накопичено величезну кількість експериментального матеріалу, складені докладні таблиці. Проте у великому масиві даних не вдавалося встановити якусь закономірність, хоча було зрозуміло всім дослідникам, що в цих таблицях захована повна інформація про будову атома.

Перші вагомі успіхи у класифікації атомних спектрів, зокрема водню, були отримані 1885 року. Як це не дивно, але певне значення тут мали ідеї античної філософської піфагорійської школи. Послідовники цієї школи стверджували, що гармонія Всесвіту завдячує містичній ролі цілих чисел у природі. Під вливом ідей піфагорійців перебував з юних років і Й.Бальмер (1825 – 1898) – учитель фізики жіночої гімназії м.Базеля (Швейцарія). Впродовж багатьох років він шукав закономірності у сукупності атомних спектрів, вважаючи, що цілі числа мають тут мати не останнє значення.

Докладно аналізуючи спектральні лінії $\{H_\alpha, H_\beta, H_\gamma, H_\delta\}$ атома водню у видимій ділянці, Й.Бальмер зауважив, що вони розташовані не випадково, а утворюють серію довжин хвиль, яку можна описати єдиною формулою.

$$\lambda = \lambda_0 \frac{k^2}{k^2-4}, \qquad (17.8)$$

де $\lambda_0 = 3645{,}6$ Å – емпірично визначена постійна, а k=3, 4, 5, … – цілі числа.

Формула (17.8) настільки точно описувала довжини хвиль ліній у видимій ділянці спектра атома водню, що на неї відразу звернули увагу. Стало зрозумілим, що за цією простою емпіричною формулою криється глибокий фізичний зміст. По суті, з формули Й.Бальмера починається справжня наука про атоми та їхню внутрішню електронну будову.

Незабаром було звернуто увагу на те, що формула (17.8) стає прозорішою і змістовнішою, якщо її переписати не для довжин хвиль, а через частоти, використовуючи співвідношення між частотою і довжиною електромагнітної хвилі $\omega = 2\pi\nu = 2\pi c/\lambda$, тоді

$$\omega = R_\omega\left(\frac{1}{2^2} - \frac{1}{k^2}\right), \qquad (17.9)$$

або

$$\frac{1}{\lambda} = R_\lambda\left(\frac{1}{2^2} - \frac{1}{k^2}\right), \qquad (17.10)$$

де $R_\omega = \frac{8\pi c}{\lambda_0} = 2{,}07 \cdot 10^{16}\text{с}^{-1}$, а $R_\lambda = \frac{4}{\lambda_0} = 1{,}097 \cdot$ $\cdot 10^7\text{м}^{-1}$. Постійні величини R_ω і R_λ називають *сталими Рідберга.* За формулою (17.10) можна розрахувати довжини хвиль спектральних ліній в серії Бальмера. Для k=3 λ=6,563·10^{-7} м (червона лінія H_α), для k=4 λ=4,861·10^{-7} м (голуба лінія H_β), для k=5 λ=4,339·10^{-7} м (голуба лінія H_γ), для k=6 λ=4,101·10^{-7} м (фіолетова лінія H_δ). При k=∞ ми отримаємо границю серії, наприклад, для серії Бальмера $\lambda = 4/R_\lambda$ =3,645·10^{-7} м.

Спектральні серії водню

Табл. 17.1

Серія	Область спектру	Формула серії	Границя серії (k=∞), $\lambda \cdot 10^7$,м
Лаймана	Ультрафіолетова	$\frac{1}{\lambda}=R_\lambda(\frac{1}{1^2}-\frac{1}{k^2})$, k =2,3,4...	0,91127
Бальмера	Видима	$\frac{1}{\lambda}=R_\lambda(\frac{1}{2^2}-\frac{1}{k^2})$, k =3,4,5...	3,6456
Пашена	Інфрачервона	$\frac{1}{\lambda}=R_\lambda(\frac{1}{3^2}-\frac{1}{k^2})$, k =4,5,6...	8,2014
Брекетта	Інфрачервона	$\frac{1}{\lambda}=R_\lambda(\frac{1}{4^2}-\frac{1}{k^2})$, k =5,6,7...	14,580
Пфунда	Інфрачервона	$\frac{1}{\lambda}=R_\lambda(\frac{1}{5^2}-\frac{1}{k^2})$, k =6,7,8...	22,782
Хемфрі	Інфрачервона	$\frac{1}{\lambda}=R_\lambda(\frac{1}{6^2}-\frac{1}{k^2})$, k =7,8,9...	32,806

Пізніше подібна аналітична залежність спостерігалась в інших, ультрафіолетовій та інфрачервоній, ділянках спектра атомарного водню. Всіх їх можна об'єднати в так звану *узагальнену формулу Бальмера.*

$$\omega_{nk} = R_\omega\left(\frac{1}{n^2} - \frac{1}{k^2}\right), \qquad (17.11)$$

$$\frac{1}{\lambda} = R_\lambda(\frac{1}{n^2} - \frac{1}{k^2}), \qquad (17.12)$$

де n=1,2,3,4,… визначає серію, а k=n+1,n+2,n+3,... визначає окремі лінії серії (Табл.17.1) .

Якщо ввести у формулі (17.11) позначення T_n= R_ω/n^2, яке називається термом, то частоту будь-якої лінії спектра водню можна подати як різницю між двома термами, а саме

$$\omega_{nk} = T(n) - T(k). \qquad (17.13)$$

На перший погляд може здатися, що введення поняття терма є тривіальним, не більше ніж перепозначення. Насправді це не так. Подання частоти як різниці двох термів мало вагомі наслідки для подальшого розвитку теорії атомних спектрів. Важливо, що емпірична формула (17.11) записана для атома водню. Проте залишалося нез'ясованим питання про принципову можливість написання подібної простої формули для класифікації спектральних ліній складніших атомів. Враховуючи запис частоти спектральної лінії атома водню у формі (17.13), Й.Рідберг надав спектральним лініям різних елементів форму комбінації термів, застосувавши для терма емпіричну формулу.

$$T(n) = \frac{R}{(n + \alpha + \frac{\beta}{n^2})^2}\cdot$$

Тут α, β – дійсні числа, які мають значення емпірично підібраних параметрів. На підставі такої формули вдалося здійснити класифікацію спектральних ліній багатьох елементів таблиці Менделєєва.

Продовжуючи дослідження Й.Рідберга, 1905р. В.Ритц сформулював комбінаційний принцип, за яким будь-яку спектральну лінію будь-якого атома можна подати як різницю термів. Не важливо чи відомі аналітичні вирази для термів. Набір спектральних ліній атомів, виглядаючи на перший погляд хаотичною сукупністю, розпався на цілком визначені серії із чіткими закономірностями їх розташування. Замість великої кількості спектральних ліній достатньо

знати набагато меншу кількість термів. Правда, комбінуючи терми, можна передбачувати спектральні лінії, які насправді не спостерігалися. Згодом з'ясувалося, що у спектрі відсутні окремі спектральні лінії через так звані правила відбору, які чітко були встановлені та обґрунтовані квантовою теорією.

У результаті виконаної класифікації спектрів стало зрозуміло, що спектральні закономірності мають фундаментальний характер і визначаються внутрішньою будовою атома. Однак встановити зв'язок структури спектра з будовою атома та фізичними процесами на атомному рівні не вдавалося впродовж багатьох років. Лише 1913 року у цій проблемі було досягнуто ясності.

§17.3.Постулати Бора. Теорія атома водню

Як вже зазначалося, планетарна модель атома не узгоджувалася із законами класичної електродинаміки. Рух електрона по замкнутих орбітах навколо ядра є прискореним. За законами класичної електродинаміки пришвидшений рух заряду завжди супроводжується випромінюванням електромагнітних хвиль. Електромагнітні хвилі переносять енергію, тому електрон, втрачаючи енергію на випромінювання, має впасти на ядро. Ніхто не міг спростувати цей важливий наслідок законів електродинаміки. Проте існування стабільних атомів також є незаперечним фактом. Данський фізик Нільс Бор перший дійшов висновку, що протиріччя моделі Резерфорда з класичною електродинамікою можна усунути лише тоді, коли відмовитися від класичного (ньютонівського) опису руху електрона по замкнутих орбітах у планетарній моделі атома.

На початку лютого 1913р. суто випадково Н.Бор зустрів студентського приятеля Ханса Хансена, який розказав йому про спектральні формули Бальмера, Рідберга, Рітца. Н.Бор згадував: «Як тільки я побачив формулу Бальмера, мені все одразу ж стало ясно» Прийнявши гіпотезу Планка про кванти енергії, формулу (17.11) помножимо на сталу планка $\hbar$. Тоді вона перепишеться так:

$$\hbar\omega_{nk} = \frac{\hbar R_\omega}{n^2} - \frac{\hbar R_\omega}{k^2},$$

або

$$\hbar\omega_{nk} = |W_n| - |W_k|, \qquad (17.14)$$

де $\hbar\omega_{nk}$ – квант енергії випромінювання дорівнює різниці двох значень енергії атома водню (W_k –енергія до випромінювання, W_n – енергія після випромінювання).

Для усунення протиріччя планетарної моделі атома із законами класичної електродинаміки Н. Бор припустив, що, рухаючись по орбіті, електрон не випромінює електромагнітні хвилі, а залишається на орбіті достатньо тривалий час. Упродовж 1913 року Нільс Бор опублікував три праці, в яких він сформулював свої знамениті постулати.

З цього приводу американський фізик (один із творців теорії надпровідності) Леон Купер сказав: «Звичайно, було дещо самонадіяно висувати припущення, які протирічили електродинаміці Максвелла і механіці Ньютона, але Бор був молодим».

ПОСТУЛАТ 1 (Про стаціонарні стани).

Атоми можуть перебувати тільки у визначених станах, в яких, попри рухи електронів по замкнутих дискретних орбітах, вони не випромінюють енергію. У цих станах атоми мають енергію, яка утворює дискретний ряд: $W_1, W_2, \ldots, W_n, \ldots$.Ці стани характеризуються стійкістю: будь-яка зміна енергії W_n завдяки випромінюванню чи

поглинанню електромагнітного випромінювання або інших форм обміну атома енергією може відбуватися у разі повного переходу із одного стану в інший.

За цим постулатом електрони під час руху навколо ядра майже весь час перебувають на стабільних, незалежних від часу орбітах, які мають цілком визначену енергію. Такі орбіти називають *стаціонарними*.

Далі, для пояснення спектральних закономірностей, зумовлених процесами поглинання і випромінювання електромагнітних хвиль пропонується такий постулат.

ПОСТУЛАТ 2 (Правило частот).

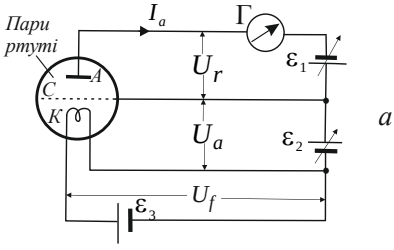
У разі переходу з одного стаціонарного стану в інший атоми випромінюють чи поглинають електромагнітні хвили визначеної частоти. Випромінювання, що висилається або поглинається у разі переходу із стаціонарного стану W_k у стаціонарний стан W_n, є монохроматичним і його частота визначається співвідношенням:

$$\omega_{nk} = \frac{|W_n| - |W_k|}{\hbar}. \qquad (17.15)$$

Теорія Бора, побудована на сформульованих постулатах, блискуче пояснила структуру спектральних ліній атома водню. Власне, Н. Бор і брав за основу ці експериментальні дані під час формулювання своїх постулатів. Тому відразу постало питання про безпосередні, незалежні підтвердження постулатів Бора, про наявність стаціонарних станів та правил переходу між ними.

Дослід Франка і Герца

Першим експериментом, в якому безпосередньо було підтверджено справедливість постулатів Бора став дослід Дж. Франка і Г. Герца (1914р.). Установка, яку вони використовували схематично показана на рис.17.5,*а*. В трубці із гарячим катодом *K*, живлення якого здійснюється від батареї ε_3, із сіткою *C* і анодом *A* знаходяться пари ртуті при низькому тиску при температурі біля 150 °C.



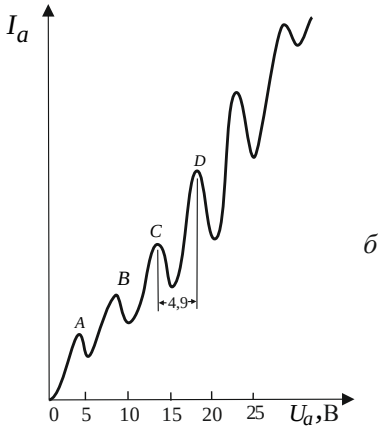


Рис.17.5.

Між катодом і сіткою прикладена прискорююча різниця потенціалів U_a, яку можна змінювати від 0 до 60В. Між анодом і сіткою створюється невелике гальмівне поле з різницею потенціалів U_r ($\approx 0{,}5$В). Дуже чутливий гальванометр Г, який включений, послідовно в коло анода, вимірює анодний струм ($\approx 10^{-9}$А). Характерна крива залежності анодного струму I_a від прискорюючого потенціала U_a приведена на рис.17.5,*б*. Виявилось, що із збільшенням U_a, як і в будь-якій електронній лампі, анодний струм зростає, однак цей ріст не є монотонним, а супроводжується різкими спадами струму

кожний раз, коли U_a збільшується приблизно на 5В.

Немонотонності кривої є результатом взаємодії електрона з атомами ртуті. Маса атома ртуті майже в 400000 разів перевищує масу електрона. Поки енергія бомбуючого електрона не перевищує 4,9еВ, його зіткнення з атомами ртуті є пружними (електрон розсіюється на атомах без зміни енергії, змінюючи лише свою траєкторію). Такий електрон не здатний збудити атом ртуті. Ситуація докорінно змінюється після досягнення електроном енергії 4,9 еВ. Тепер електрон може збудити атом ртуті, передавши йому частину своєї кінетичної енергії. Таке зіткнення електрона з атомом ртуті є непружним. В результаті таких зіткнень енергія електронів наскільки зменшується, що електрони не можуть подолати слабке затримуюче поле, яке створюється різницею потенціалів U_r, і не досягають анода. Анодний струм при цьому різко зменшується. Якщо прискорюючий потенціал U_a збільшити ще на 5В, то електрони можуть зазнати ще одне непружне зіткнення і втратити при цьому повністю свою енергію. Це пояснює другий спад (спостерігається на кривій струму при U_a=10В), який буде більш різким, ніж перший спад, так як відповідає електронам, які зазнали два непружних зіткнення. Третій спад струму I_a відповідає електронам, які зазнають три непружних зіткнення, і т.д. Кожний раз, коли відбувається непружне зіткнення, атом ртуті переходить у збуджений стан і потім, повертаючись в основний стан, випромінює фотон з довжиною хвилі λ=2536Å=0,2536 мкм, що відповідає переходам атомів ртуті із першого збудженого стану в основний стан. Енергія такого фотона дорівнює $\varepsilon = hc/\lambda$ = 4,89 еВ.

Цей результат, а також той факт, що відстань між двома сусідніми спадами струму I_a дорівнює приблизно 4,9В, переконливо доводить існування дискретних енергетичних рівнів в атомі ртуті. Удосконалена методика експерименту дозволила визначити також енергії збудження для інших енергетичних рівнів. І не дивно, що за ці чудові досліди Дж. Франк і Г. Герц в 1925р. отримали Нобелівську премію з фізики.

Покажемо, слідом за Бором, до яких наслідків призводять сформульовані постулати для атома водню. Рівняння (17.3) допускає безмежну кількість замкнутих колових орбіт. Постає питання про існування стаціонарних орбіт серед безмежної кількості можливих. Перший постулат Н. Бора стверджує, що стаціонарні орбіти існують. Тоді відразу таке запитання: «*А як стаціонарні орбіти вирізнити з-поміж безмежної кількості всіх можливих, що задовольняють співвідношенню* (17.3)?». Ось тут Н. Бор зрозумів, а можливо і вгадав, як відповісти на це запитання. І ця відповідь була простою, тому і геніальною.

Уже від 1905 року в науці утвердилися і завоювали визнання квантові ідеї, про що йшлося у попередньому розділі. Нагадаємо, що стала Планка h має розмірність дії і на початку названа Планком квантом дії. Але ж поняття дії у класичній механіці відоме ще з часів Гамільтона і Лагранжа. Розмірність дії має орбітальний момент імпульсу $\vec{L} = \vec{r} \times m\vec{v}$.

$\vec{L}$ ще називають моментом кількості руху. Він перпендикулярний як радіусу-вектору $\vec{r}$ так і вектору кількості руху $m\vec{v}$ (рис.17.3). Величина орбітального моменту імпульсу для колового руху електрона навколо силового центру (ядра) є $L = m_e rv$ (рис.17.3).

Бор відчував, що є глибинний зв'язок між двома "мінімальностями" в природі:

–існуванням найменшої величини для розмірів електронного рою в атомі. Меншою ця величина бути не може;

–існуванням найменшої величини для фізичної дії, менше якої не буває.

Весною 1912р. Бор дійшов висновку, що будова "електронного рою" в резерфордовському атомі управляється квантом дії – сталою Планка "$\hbar$".

Отже, коли фізична величина має розмірність дії, то найменше числове значення, яке вона може набувати дорівнює сталій Планка, а більші значення – кратні сталій Планка. Після тривалих міркувань і розрахунків Н.Бор дійшов висновку, що стаціонарні можуть бути лише ті колові орбіти, для яких орбітальний момент імпульсу кратний $\hbar$:

$$m_e r v = n\hbar\,. \qquad (17.16)$$

Ціле число n називається квантовим числом.

Рівняння (17.13) і (17.16) утворюють систему двох рівнянь з двома невідомими r і v. Розв'язок цієї системи дає, що:

$$v_n = \frac{e^2}{4\pi\varepsilon_0 \hbar n}\,, \qquad (17.17)$$

а

$$r_n = \frac{4\pi\varepsilon_0 \hbar^2}{m_e e^2} n^2\,, \qquad (17.18)$$

де n=1,2,3, … .

Тут у позначені радіуса орбіти введений індекс n, щоб підкреслити залежність радіуса від цього квантового числа.

Радіус r_1, найменшої колової орбіти для атома водню, позначається як a_0:

$$a_0 = \frac{4\pi\varepsilon_0\hbar^2}{m_e e^2} = 0{,}53 \cdot 10^{-10}\text{м} = 0{,}53\ \text{Å}. \quad (17.19)$$

і називається *радіусом Бора*. У певному розумінні радіус Бора a_0 визначає розміри атома водню. Задовго до створення теорії Бора розміри атома водню були оцінені експериментально. Теоретичне значення радіуса Бора (17.19) виявилося доволі близьким до експериментального значення радіуса атома водню. Це було ще одним аргументом на користь теорії Бора.

Енергію електрона у кулонівському полі ядра, заряд якого $|e|$, отримаємо, підставивши значення радіусів орбіт (17.18) у формулу (17.6).

Тоді

$$W_n = -\frac{e^4 m_e}{32\pi^2\varepsilon_0^2\hbar^2}\cdot\frac{1}{n^2} = -\frac{13{,}6}{n^2}\,\text{еВ}. \quad (17.20)$$

Стан атома із найменшим значенням енергії (n=1) називається основним. Для атома водню основному стану згідно (17.20) відповідає енергія $W_1 = -13{,}6$еВ. Відповідно до другого постулату (17.15) для атома водню частоти переходу із стану W_k у стан W_n будуть:

$$\omega_{kn} = \frac{1}{\hbar}(|W_n| - |W_k|) = \\ = \frac{e^4 m_e}{32\pi^2\varepsilon_0^2\hbar^3}\left(\frac{1}{n^2} - \frac{1}{k^2}\right). \qquad (17.21)$$

Порівняння останнього виразу з узагальненою формулою Бальмера (17.11) засвідчує, що вони повністю збігаються, якщо сталу Рідберга вибрати такою, що дорівнює

$$R_\omega = \frac{e^4 m_e}{32\pi^2\varepsilon_0^2\hbar^3} = 2{,}07 \cdot 10^{16}\text{с}^{-1}. \quad (17.22)$$

Отримане значення R_ω співпадає з експериментальним значенням. Це був успіх теорії Бора.

Зобразимо енергетичні рівні (рис.17.6), що описуються формулою (17.20), у вигляді ліній, на яких цифри визначають квантові стани n. За початок відліку прийнятий стан з $n\to\infty$, який позначемо W=0. Електрон з енергією $W > 0$ здійснює необмежений (інфінітний) рух у полі атомного ядра за гіперболічними траєкторіями. За такого руху електрон може мати довільне додатне значення

енергії. Коли $W<0$, то електрон здійснює обмежений (фінітний) рух по замкнутих орбітах, а його енергії можуть мати тільки дискретні значення (17.20).

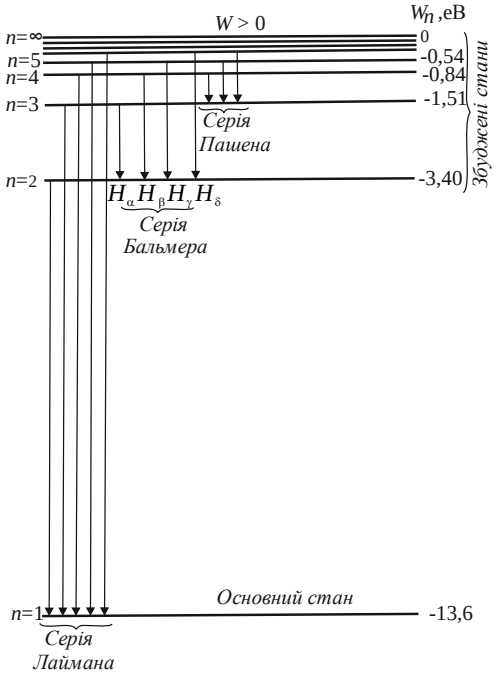


Рис.17.6.

Стан n=1 відповідає основному (нормальному) станові електрона, а всі стани з $n\geq2$ – відповідають збудженим його станам. Вертикальні стрілки, спрямовані догори, відповідають переходам між відповідними станами з поглинанням атомом енергії, а в зворотньому напрямку (рис.17.6) – з випромінюванням енергії за другим постулатом Бора.

Із формули (17.20) і рис.17.6 видно, що якщо атом знаходиться в основному стані (n=1), то для видалення електрона із атома необхідна енергія 13,6 еВ. Отже, енергія зв'язку і енергія іонізації для атома водню, який знаходиться в основному стані, дорівнюють $W_{зв}=W_{іон}$=13,6еВ. За допомогою схеми на рис.17.6 легко пояснити виникнення серіальних спектральних ліній в табл.17.1.

Воднеподібні атоми

Теорію Бора можна з успіхом використати і при розгляді воднеподібних атомів, тобто атомів, які складаються із ядра з зарядом Ze і одного електрона. До таких атомів належать один раз іонізований атом гелію He^+ (Z=2), два рази іонізований атом літія Li^{2+} (Z=3) і т.д.

Для воднеподібних атомів радіуси дозволених орбіт електрона визначаються за формулою

$$r_n = a_0 n^2/Z, \qquad (17.23)$$

яка подібна до формули (17.18). Енергія електрона на орбіті визначається за формулою

$$W_n = -\frac{13{,}6Z^2}{n^2}\ \text{eB}, \qquad (17.24)$$

яка отримується із формули (17.20).

Таким чином, теорія Бора пояснила зміст відомих на той час (1913р.) серіальних формул і передбачила ще серії в інфрачервоній області спектра, які пізніше були виявлені експериментально. Стала Рідберга, яку отримали з теорії Бора, дивовижно добре узгоджується з експериментальним значенням цієї величини. Це говорить про те, що теорія Бора правильно описує будову атома водню.

Теорія Бора була важливим кроком в розвитку теорії атома. Вона показала непридатність класичної фізики для пояснення внутрішньотамних явищ і головну роль квантових законів у мікросвіті. Разом з тим теорія Бора опиралась на класичну механіку і не була ні послідовно класичною, ні послідовно квантовою і могла бути тільки перехідним станом на шляху створення послідовної теорії атомних явищ.

Робота Бора – одне з гідних подиву явищ в історії науки. Тільки геніальним просвітленням можна пояснити появу цієї теорії до того, як з'ясувались хвильові властивості частинок. Саме з цього приводу Ейнштейн сказав: *«…вища музикальність в області теоретичної думки».*

§17.4.Подальший розвиток теорії Бора. Квантування Бора-Вільсона-Зоммерфельда

Постулати Бора, констатуючи наявність в атомі водню стаціонарних орбіт із визначеними значеннями енергії електронів та правил переходу електронів між стаціонарними орбітами, не до кінця пояснюють фізичну природу постулатів, не встановлюють фундаментального зв'язку між стаціонарними станами електронів та їх динамікою по орбітах в кулонівському полі ядра. До того ж Н. Бор для побудови теорії спектрів атома водню застосував правило квантування (17.16). Відразу виникла потреба у фізичному осмисленні цього правила та його узагальненні. Цю працю було проробленo, переважно, В. Вільсоном (1915р.) та, незалежно, А. Зоммерфельдом (1916р.). Продемонструємо це на прикладі гармонічного осцилятора, максимально дотримуючись міркувань і висновків праці А.Зоммерфельда.

Розглядається матеріальна точка маси *m*, що здійснює лінійний рух у потенціальному полі $U(x)= kx^2/2$.

Це може бути тягарець, маса якого *m*, що коливається за допомогою ружини жорсткістю *k*. Повна енергія такого осцилятора:

$$W = W_k + W_p = \\ = \frac{mv^2}{2} + \frac{kx^2}{2} = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2},$$

де $p=mv$ – імпульс точки, x – зміщення від положення рівноваги, $\omega = \sqrt{k/m}$ – власна частота коливання осцилятора, яка не залежить від амплітуда коливання. На рис.17.7,*а* зображена траєкторія осцилятора у фазовому просторі $\{x;p\}$ для заданого постійного

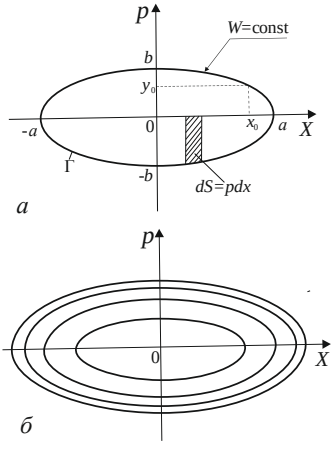


Рис.17.7.Траєкторія гармонічного осцилятора у фазовому просторі для заданого значення енергії W – (*а*): сім'я траєкторій гармонічного осцилятора для різних значень його енергії – (*б*). Площа, обмежена будь-якими сусідніми траєкторіями, однакова за величиною.

значення повної енергії W осцилятора. Такою траєкторією є еліпс з півосями $a = \sqrt{\frac{2W}{m\omega^2}}$ та $b = \sqrt{2mW}$. Класичний опис руху осцилятора на підставі теорії звичайних диференціальних рівнянь другого порядку вказує, що, якою б не була початкова умова $\{x_o; p_o\}$, через цю точку фазового простору обов'язково проходить одна і лише одна траєкторія. Через початкові умови однозначно визначається також енергія осцилятора. Отже, кожному значенню енергії осцилятора відповідає своя траєкторія, яка описується канонічним рівнянням еліпса

$$\frac{x^2}{a^2} + \frac{p^2}{b^2} = 1. \qquad (17.25)$$

Площа еліпса $S = \pi ab$. Підставивши вирази для довжин півосей a та b, знайдемо

$$S = \frac{2\pi W}{\omega} = W/\nu, \qquad (17.26)$$

де $\nu = \omega/2\pi$ – лінійна частота коливання осцилятора.

Тепер до формулювання правила квантування залишається виконати одну, проте доволі сміливу дію. З формули (17.26) бачимо, що площа еліпса у фазовому просторі має розмірність дії: [*S*] = Дж·с. Але ж розмірність дії має також стала Планка – квант дії. Позначимо через ΔW різницю між постійними енергіями осцилятора, фазовими траєкторіями яких є два сусідніх еліпси. Тоді, відповідно до формули (17.26),

$$\Delta S = \frac{\Delta W}{\nu} = h, \text{або } \Delta W = h\nu. \quad (17.27)$$

На рис.17.7,*б* схематично зображено сімейство траєкторій, що задовольняють цю умову. Перенумеруємо еліпси із вказаного сімейства цілими додатними числами 1,2,3, …, *n*, а відповідні значення енергії осцилятора W_1, $W_2, W_3, \ldots, W_n$. Тоді енергія, що відповідає $n-$му еліпсу,

$$W_n = \varepsilon \cdot n\ ,\ n=1, 2, 3, \ldots . \quad (17.28)$$

Квант енергії $\varepsilon = h\nu = \hbar\omega$ – абсолютне значення різниці енергій для двох будь-яких сусідніх фазових траєкторій.

Ґрунтуючись на прикладі гармонічного осцилятора та ще декількох простих моделей, було зроблено узагальнення правила квантування (17.27−17.28) на довільні механічні системи з одним ступенем вільності, що здійснюють рух по замкнутій траєкторії у фазовому просторі. Порівнявши (17.26) і (17.28) знайдемо, що площа еліпса, яка відповідає енергії W_n, буде $S_n = nh$ з однієї сторони, а з другої сторони $S_n = \oint_\Gamma dS = \oint_\Gamma pdx$ (рис.17.7,*а*).

Порівняння цих двох формул дає правило квантування (редакція А. Зоммерфельда)

$$\oint_\Gamma pdx = nh\ ,\ n=1, 2, 3, \ldots . (17.29)$$

Тут n – квантове число, за допомогою якого перенумеровуються дискретні рівні енергії. Інтегрують у (17.29) по замкнутому контуру Г на площині (рис.17.7,*а*).

Після створення сучасної квантової теорії правила квантування Бора-Вільсона-Зоммерфельда, або так звана стара квантова теорія, втратили своє принципове значення для квантової фізики. Проте інколи ця теорія набуває практичного застосування під час знаходження дискретних рівнів енергії квантових частинок, які здійснюють періодичні рухи у потенціальних силових полях. В загальному випадку, правила квантування Бора-Вільсона-Зоммерфельда дають задовільні результати тільки тоді, коли рух квантової частинки є близьким до класичного. Здебільшого це випадки руху частинок у потенціальному полі з великим значенням енергії. Точніше, коли різниця енергій між двома сусідніми дискретними рівнями мала порівняно з енергією цих рівнів*:* $|W_n - W_{n+1}|/W_n << 1$.

Покажемо, що правило квантування (17.16) очевидне з умови квантування (17.29). Якщо електрон рухається по коловій орбіті, радіус якої *r*, як це припускається в теорії Бора, то для обчислення інтеграла у (17.29) зручно перейти до полярної системи координат {*r*,φ} (рис.17.8) за формулами:

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi. \end{cases}$$

Під час руху електрона по колу, радіус якого r (r = const), його просторове положення відносно силового центра (ядра) однозначно визначається тільки однією координатою – полярним кутом φ ($0 \leq \varphi \leq 2\pi$).

Швидкість $v^2 = \dot{x}^2 + \dot{y}^2 = r^2\dot{\varphi}^2$, кінетична енергія $W_k = \frac{m_e v^2}{2} = \frac{1}{2} m_e r^2 \dot{\varphi}^2$, а потенціальна $W_p = U(r) = \text{const}$, так як r = const. Імпульс електрона $p = m_e v = \frac{dW_k}{dv}$, а імпульс, що відповідає координаті φ, $p_\varphi = \frac{dW_k}{d\dot{\varphi}} = m_e r^2 \dot{\varphi}$.

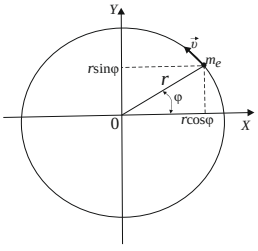


Рис.17.8.

Тоді (17.29) перепишеться так:

$$\oint_\Gamma p_\varphi d\varphi = \oint_\Gamma m_e r^2 \dot{\varphi} d\varphi = \oint_\Gamma m_e r v d\varphi =$$
$$= m_e r v \oint_\Gamma d\varphi = 2\pi m_e r v = n_\varphi h.$$

Звідки отримаємо

$$m_e v r = n_\varphi \hbar\,, \qquad (17.30)$$

що відповідає правилу квантування Бора для колових орбіт (n_φ=1, 2, 3, …).

Розвиваючи теорію Н.Бора і грунтуючись на дослідженнях В.Вільсона, А.Зоммерфельда припустив, що орбіта електрона у атомі водню є еліптичною. В одному з фокусів еліптичної орбіти перебуває ядро атома. Це припущення грунтувалося на поширенні планетарної моделі Сонячної системи на атоми. Це цілком логічно і виправдано, бо планети Сонячної системи і електрони атома рухаються у потенціальних силових полях, які мають схожу математичну структуру. При русі електрона по еліпсу поряд з кутом φ відстань r від ядра теж є змінною величиною (рис.17.9).

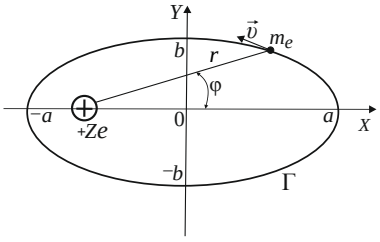


Рис.17.9.

Правила квантування Зоммерфельда для еліптичних орбіт записуються так:

$$\begin{cases} \oint_\Gamma p_\varphi d\varphi = n_\varphi h \\ \oint_\Gamma p_r dr = n_r h, \end{cases} \qquad (17.31)$$

де $n_\varphi -$ азимутальне, а n_r – радіальне квантове число.

Перша квантова умова, як і для колових орбіт, також призводить до квантування орбітального моменту імпульсу (17.16). Обчислення інтеграла у другій умові є технічно значно складнішою задачею. Докладніше про це можна дізнатись з додаткової літератури. Наведемо тут лише основні фізичні висновки.

Розв'язання А. Зоммерфельдом задачі (17.31) для еліптичних орбіт вказує на те, що квантові числа n_r, та n_φ задовільняють умові

$$n = n_r + n_\varphi. \qquad (17.32)$$

Тут азимутальне n_φ та радіальне n_r квантові числа можуть набувати таких значень: n_φ=1, 2,3,…,n; n_r = 0, 1, … , ($n-2$), ($n-1$).

Повна енергія електрона в полі атомного ядра набуває дискретних значень, які визначаються формулою (17.24). Велика піввісь еліпса

$$a = a_0 (n_r + n_\varphi)^2, \qquad (17.33)$$

а мала –

$$b = a_0 (n_r + n_\varphi) n_\varphi, \qquad (17.34)$$

де a_0 – радіус першої боровської орбіти.

Для основного стану n=1 азимутальне квантове число може набувати значення n_φ=1, а n_r = 0. Орбіта електрона в цьому стані є коловою $a = b = a_0$. Для першого збудженого стану n = 2 допустимими є такі набори квантових чисел: {n_φ=2, n_r = 0}, що відповідає $a = b = 4a_0$ і {n_φ=1, n_r = 1}, що відповідає a=4a_0, b=2a_0 (тобто рух по еліпсу). Для n=3 можливими є три орбіти, які визначаються таким набором квантових чисел: {n_φ=3, n_r = 0};{n_φ =2, n_r = 1};{n_φ =1, n_r = 2}. Ці орбіти зображені на рис.17.10. У теорії Зоммерфельда з (17.33) і (17.34) витікає, що відношення півосей еліптичних орбіт задається виразом

$$\frac{b}{a} = \frac{n_\varphi}{n_r + n_\varphi}. \qquad (17.35)$$

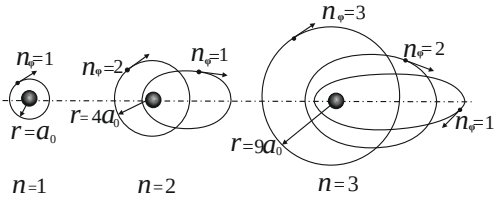


Рис.17.10.Стаціонарні орбіти основного (n=1) та двох збуджених рівнів (n=2;3) атома водню.

Із (17.35) витікає, що чим менше азимутальне квантове число n_φ, то тим більше витягнутий еліпс (рис.17.10). Необхідно зазначити, що при русі різними еліпсами з одним і тим самим квантовим числом n_φ, електрон має однакову енергію. Тому говорять, що має місце виродження по азимутальному квантовому числу n_φ. При наявності, наприклад, зовнішнього електричного поля еліптичні орбіти, з однією енергією, але різної форми, будуть збуджені по-різному і це впливає на енергетичний спектр.

§17.5.Просторове квантування

Ми розглянули рух електрона в площині еліптичної орбіти, тобто при двох ступенях вільності (два квантові числа: азимутальне n_φ і радіальне n_r). Положення електрона в просторі характеризується трьома координатами, тому необхідно враховувати просторову орієнтацію самої орбіти. В атомі всі напрямки рівноймовірні. Однак фізично можна створити напрямки, які відрізняються від решти. Такими напрямками можуть бути напрямки зовнішнього магнітного і електричного полів. Допустимо, що атом з орбітальним електроном знаходиться в магнітному полі. Атом має магнітний момент і у зовнішньому магнітному полі повинен поводити себе як магніт, тобто повинен зайняти положення, при якому магнітний момент буде

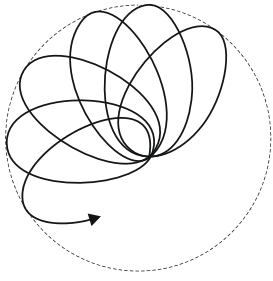


Рис.17.11.

або паралельний або антипаралельний полю (див.§5.8). Однак атом володіє також і моментом кількості руху $\vec{L}$, який перешкоджає орієнтації магнітного момента. Звідси витікає, що в магнітному полі атом повинен поводити себе так, як звичайне веретено (дзига) поводить себе в полі тяжіння, тобто рух електрона по еліпсу супроводжується одночасним обертанням еліпса навколо силової лінії поля (має місце прецесія еліпса). В результаті еліптична орбіта виходить не замкненою, а подібною на розетку (рис.17.11).

Положення площини орбіти електрона по відношенню до виділеного напрямку визначається кутом α між напрямком момента кількості руху $\vec{L}$ і виділеним напрямком (напрямок індукції магнітного поля $\vec{B}$). Цей тип руху в атомі теж квантований. На рис.17.12

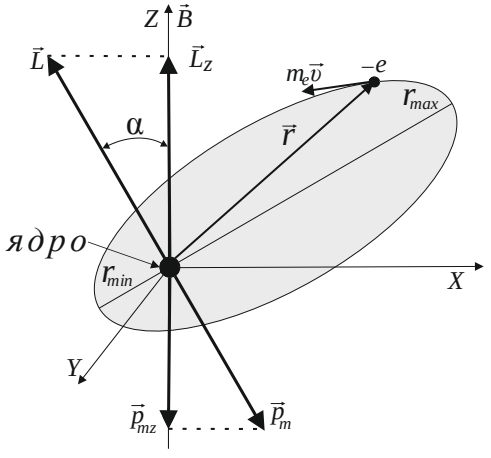


Рис.17.12.

зображено положення орбіти електрона і орієнтація $\vec{B}$ в просторі координат *X,Y,Z.* Згідно рисунку L_z=Lcos α. Враховуючи умову квантування Бора (17.30), отримаємо, що проекція $\vec{L}$ на напрямок $\vec{B}$ (вісь *Z*), буде:

$$L_z = \hbar n_\varphi \cos\alpha.$$

Позначимо

$$n_\varphi \cos\alpha = m\,, \qquad (17.36)$$

і отримаємо

$$L_z = m\hbar\,, \qquad (17.37)$$

де m отримало назву магнітного квантового числа. Так як $-1 \le \cos\alpha \le +1$, то можливі значення $m = n_\varphi \cos\alpha$ знаходяться в межах

$$-n_\varphi \le m \le n_\varphi. \qquad (17.38)$$

Магнітне квантове число m може мати n_φ цілих додатніх, від'ємних і нульове значення, тобто всього ($2n_\varphi$+1) значень:

$$m = 0, \pm1, \pm2, \dots, \pm n_\varphi. \quad (17.39)$$

Таким чином, при дії на атом зовнішнього магнітного поля можуть зреалізовуватись тільки такі орієнтації орбіт, при яких проекція на вісь *Z*, яка співпадає з напрямком магнітного поля $\vec{B}$, момента кількості руху $\vec{L}$ кратна $\hbar$. В цьому і полягає суть *просторового квантування*. Квантування проекції моменту кількості руху L_z на напрямок магнітного поля означає також, що квантується і проекція орбітального магнітного моменту p_{mz} (рис.17.12). Зв'язок між орбітальним магнітним моментом $\vec{p}_m$ і моментом кількості руху $\vec{L}$, як було встановлено раніше (§6.2), задається гіромагнітним відношенням:

$$p_m = -\frac{e}{2m_e}L\,, \qquad (17.40)$$

де m_e – маса електрона. Очевидно, що і для проекції p_{mz} і L_z, згідно рис.17.12, можна записати:

$$p_{mz} = -\frac{e}{2m_e}L_z. \qquad (17.41)$$

Враховуючи умову квантування (17.37), отримаємо, що

$$p_{mz} = -\frac{e\hbar}{2m_e}m = -\mu_Б m. \quad (17.42)$$

Величина $\mu_Б = \frac{e\hbar}{2m_e} = 0{,}93 \cdot 10^{-23}$ Дж/Тл є елементарним магнітним моментом і називається *магнетоном Бора.*

Таким чином, проекція магнітного моменту атома на напрямок магнітного поля кратна елементарному магнітному моменту – магнетону Бора.

Підтвердженням теорії просторового квантування є ефект Зеємана і досліди Штерна і Герлаха.

§17.6.Ефект Зеємана

Ефект Зеємана був відкритий в 1896 р. Суть його в тому, що при дії на атом зовнішнього магнітного поля його енергетичний стан змінюється і стає залежним від магнітного квантового числа m, що проявляться в розщепленні спектральних ліній на декілька компонентів. Причину такого розщеплення легко пояснити наявністю у електрона, який обертається навколо ядра, магнітного орбітального моменту (17.42).

Магнітний момент в магнітному полі має енергію $W = -\vec{p}_m\vec{B} = -p_{mz}B$, де p_{mz} – проекція магнітного моменту на напрямок поля (§5.8). Це означає, що в магнітному полі атом з урахуванням (17.42) отримає додаткову енергію, яка залежить від квантового числа m:

$$\Delta W = \mu_Б Bm, \qquad (17.43)$$

де як відомо $m = 0, \pm 1, \pm 2, \dots, \pm n_\varphi$. Якщо $n_\varphi = 1$, то m приймає значення $-1,0,+1$ (тобто знімається виродження по магнітному квантовому числу m і такий енергетичний рівень розщеплюються на три рівні). При $n_\varphi = 2, m = -2, -1, 0, +1, +2$.

На рис.17.13 показано розщеплення рівнів і спектральних ліній випромінювання. При відсутності магнітного поля спостерігається одна лінія з частотою ω_0. При включені магнітного поля крім лінії ω_0 виникають лінії $\omega_0 + \Delta\omega$ і $\omega_0 - \Delta\omega$. Поява цих ліній (або, як говорять, зміщення компонентів) називається нормальним (або Лоренцовим) зміщенням або нормальним ефектом Зеємана, класичне пояснення якого дав Лоренц. Величину нормального зміщення можна оцінити із (17.43), прийнявши, що $\hbar\Delta\omega = \mu_Б Bm$. Врахувавши, що $\lambda = \frac{2\pi c}{\omega}$ і $|\Delta\lambda| = \frac{2\pi c}{\omega^2}\Delta\omega$, отримаємо:

$$|\Delta\lambda| = \frac{\lambda^2}{c}\frac{\mu_Б Bm}{h}.$$

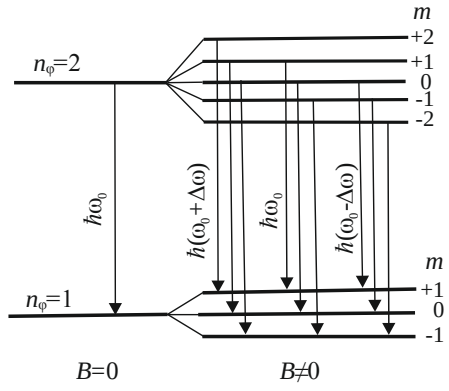


Рис.17.13.

Експериментально спостерігати ефект Зеємана можна в сильних магнітних полях, коли $B \sim 1$ Тл. Для видимої області спектру ($\lambda \sim 0{,}55$ мкм) і $m = +1$, отримаємо: $\Delta\lambda \approx 0{,}14 \cdot \cdot 10^{-10}м = 0{,}14$ Å.

Як бачимо, розщеплення дуже мале і для його спостереження потрібні прилади з великою роздільною здатністю.

Необхідно зазначити, що таким способом розщеплюється невелика кількість ліній спектру, і в більшості випадків розщеплення має більш складний характер, особливо в слабких магнітних полях. Складна картина розщеплення спектральних ліній в слабких магнітних полях отримала назву аномального ефекта Зеємана, який в рамках теорії Бора не має пояснення.

§17.7.Дослід Штерна і Герлаха

Теорія просторового квантування була підтверджена в дослідах Штерна і Герлаха по розщепленню пучків атомів в неоднорідному магнітному полі в 1922р. Прилад, який

використовувався в дослідах Штерна і Герлаха, схематично зображений на рис.17.14. В балоні з вакуумом розміщена піч для випаровування срібла (атоми срібла воднеподібні). Для створення вузького направленого пучка використовувались щілини коліматора. Потім атоми попадають в сильно неоднорідне магнітне поле з особливою формою полюсних наконечників. Неоднорідне магнітне поле створюється з наступних міркувань. Сила, яка

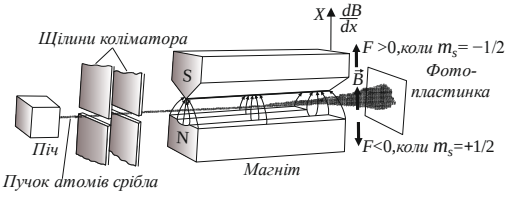


Рис.17.14.

діє на магнітний момент в неоднорідному магнітному полі (див.§5.8), виражається формулою $F_x = p_m \frac{dB}{dx} \cos\alpha$, де p_m – магнітний момент атома, α – кут між напрямками векторів $\vec{p}_m$ і $\vec{B}$, *dB*/*dx* – градієнт індукції магнітного поля. Щоб сила F_x була якнайбільша, для цього і створювалось сильно неоднорідне поле (похідна *dB*/*dx* має велике значення). Коли магнітне поле відсутнє, то на екрані (фотопластинці) пучок атомів залишав слід у вигляді смужки, яка відповідала розмірам пучка атомів (рис.17.15,*a*).

Із класичної фізики відомо, що в магнітному полі площини орбіт відносно напрямку магнітного поля можуть орієнтуватись як завгодно (кут α може приймати довільні значення). А оскільки на магнітні моменти в неоднорідному полі діє сила, то на екрані можна чекати розширення смужки порівняно із випадком, коли *B*=0 (поле відсутнє).

За теорією просторового квантування, площини орбіт атомів срібла можуть орієнтуватись по відношенню до напрямку магнітного поля тільки дискретними способами (17.42). При n_φ= 1 магнітне квантове число m може приймати значення –1,0,+1. Атоми, які летять з магнітними моментами, що орієнтовані по напрямку поля, тобто такі, для яких $\cos\alpha = +1$, повинні відхилятись в одну сторону, атоми з магнітними моментами, які орієнтовані так, що для них $\cos\alpha = -1$, повинні відхилятись настільки ж в протилежну сторону. Атоми, які мають напрямок магнітного моменту, який перпендикулярний полю, тобто для яких $\cos\alpha = 0$, не повинні зазнавати відхилення.

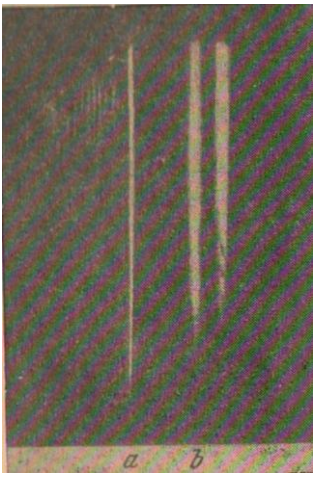


Рис.17.15.Розщеплення пучка парів натрію в магнітному полі: *a* – у відсутності поля; *b* – в присутності поля.

Пучок повинен розпастись на три окремих пучки. Дослід з атомами натрію (рис.17.15,*b*), срібла (а пізніше і з атомами водню) показав, що атомний пучок дійсно розпадається, але не на три, а на два пучки (середній пучок, який не повинен відхилятись, був відсутній) (рис.17.16). На цьому ж рисунку показано розщеплення атомних пучків для різних металів.

Таким чином, ідея просторового квантування підтвердилась, однак механізм розщеплення пучка атомів натрію, срібла та інших елементів був незрозумілий.

Теорія Бора, навіть удосконалена, пояснити цього не могла. Справа в тому, що важливою особливістю атомів першої групи є те, що їхні валентні електрони знаходяться в основному стані атома. В цьому стані у електрона відсутній момент імпульсу L, а значить і магнітний момент. Тому і виникало питання про пояснення результатів дослідів

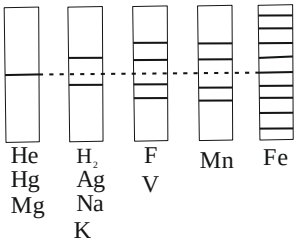


Рис.17.16.

Штерна і Герлаха. Просторове квантування якого магнітного моменту (тобто моменту імпульса) було виявлено в цих дослідах і проекція якого магнітного моменту дорівнює одному магнетону Бора?

§17.8.Власний момент електрона

В спектрах випромінювання лужних металів, особливо в $^{23}_{11}$Na, спостерігається одна різка жовта лінія. Однак детальні дослідження дозволили виявити, що це не проста лінія, а дублет, який складається із двох ліній з довжинами хвиль 5890 Å і 5896Å. Таке розщеплення однієї спектральної лінії на дві різні, які близько примикають одна до одної, відомо під назвою тонкої структури атомного спектру. Для її пояснення навіть удосконалена модель атома Бора вимагала доповнень.

Ще в 1921р. А.Х. Комптон висловив припущення, що електрон повинен бути частинкою, яка подібна до дзиги. Потім в 1925р. С.Гаудсміт і Дж.Уленбек теоретично пояснили розщеплення спектру реальних ліній, припустивши, що електрон здійснює обертальний рух навколо власної осі. У відповідності до цієї гіпотези, електрон, що обертається, повинен мати власний магнітний момент імпульсу $\vec{p}_{ms}$ і механічний момент імпульсу $\vec{L}_s$, який називається спіновим моментом (спін) (англ. spin – обертати). В 1928р. Дірак створив релятивістську квантову теорію електрона, із якої властивості спіна витікали автоматично.

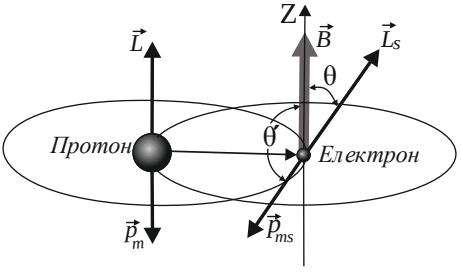


Рис.17.17.

Спостерігач, який рухається разом з електроном по орбіті навколо протона, зазначить, що протон обертається по такій самій орбіті навколо електрона (рис.17.17). При русі протона навколо електрона створюється магнітне поле з індукцією $\vec{B}$, яке показане на рис.17.17 в місці розташування електрона.

На цьому рисунку $\vec{L}$ – орбітальний момент імпульсу електрона, який обертається навколо протона, $\vec{L}_s$ – спіновий момент електрона. При орбітальному русі електрона створюється магнітний момент $\vec{p}_m$, а при власному обертанні електрона по аналогії зі сферою, яка негативно заряджена і обертається, створюється власний магнітний момент $\vec{p}_{ms}$ (рис.17.18).

Магнітне поле, яке створюється в результаті обертання протона навколо

електрона взаємодіє з магнітним спіновим моментом $\vec{p}_{ms}$. Тому до енергії електростатичної взаємодії електрона і протона необхідно добавити додаткову енергію, яка дорівнює

$$\begin{cases} \Delta W_s = -\vec{p}_{ms}\vec{B} = -p_{ms}B\cos\theta' = \\ \quad = -p_{ms}B\cos(\pi-\theta) = \\ \quad = p_{ms}B\cos\theta = p_{msz}B, \end{cases} \qquad (17.43)$$

де p_{msz} – проекція магнітного спінового момента на вісь *Z*.

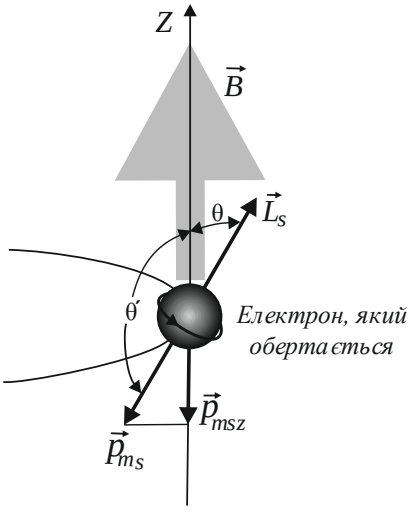


Рис.17.18.

Взаємодія між власними магнітним моментом p_{ms} електрона і магнітним полем, обумовленого орбітальним рухом електрона, називається спін-орбітальною взаємодією. Ця взаємодія є джерелом додаткових енергетичних рівнів, які утворюють тонку структуру спектру.

Примітка. Уявлення про електрон, який обертається навколо своєї осі, використовується тільки для наглядності, але не має буквального змісту. Якщо припустити, що обертання дійсно має місце, і прийняти для електрона розумні розміри, то для отримання відомого значення власного механічного моменту імпульсу електрона довелось би допустити існування лінійних швидкостей обертального руху, які перевищують швидкість світла.

За допомогою магнітного спінового моменту вдалось пояснити аномальний ефект Зеємана, розщеплення пучка атомів натрію, срібла на два пучки в дослідах Штерна і Герлаха, природу жовтого дублета в спектрах випромінювання натрію та інше.

Тому необхідно ввести ще одне квантове число, яке характеризує власний момент імпульсу електрона. Згідно з релятивістською квантовою теорією Дірака, власний механічний момент імпульсу квантований у відповідності з формулою

$$L_s = \sqrt{s(s+1)}\hbar. \qquad (17.44)$$

Спінове квантове число *s*, яке входить в цю формулу, має тільки єдине значення, що дорівнює $1/2$. Тому єдине можливе значення L_s дорівнює:

$$L_s = \sqrt{\frac{1}{2}\left(\frac{1}{2}+1\right)}\hbar = \sqrt{3}\frac{\hbar}{2}.$$

Число *s*, як правило, не використовується разом з іншими квантовими числами, так як має єдине значення (1/2) і не вносить різниці між станами. Власний механічний момент $L_s=\sqrt{3}\,\hbar/2$ є фундаментальною характеристикою електрона, як його маса або заряд. L_s в *Z* – напрямку (напрямок магнітного поля $\vec{B}$) квантований і визначається як

$$L_{sz} = L_s\cos\theta = m_s\hbar\,, \qquad (17.45)$$

де $m_s = \pm 1/2$ і називається магнітним спіновим квантовим числом. Часто $m_s = +1/2$ позначають як «спін-вгору» (↑), а $-\,m_s = -1/2$ як «спін-вниз» (↓) (рис.17.19).

Гіромагнітне відношення для спінового моменту дорівнює:

$$p_{ms} = -\frac{e}{m_e} L_s,$$

і відповідно,

$$p_{msz} = -\frac{e}{m_e} L_{sz}.$$

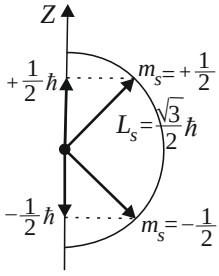


Рис.17.19.

Із урахуванням (17.45)

$$p_{msz} = -\frac{e\hbar}{m_e} m_s,$$

або

$$p_{msz} = \pm\frac{e\hbar}{2m_e} = \pm\mu_Б. \quad (17.46)$$

Таким чином, додаткова енергія спін-орбітальної взаємодії (17.43) буде дорівнювати:

$$\Delta W_s = \pm\mu_Б B. \quad (17.47)$$

Два значення енергії ΔW_s пояснюють жовтий дублет в спектрі випромінювання натрію і аномальний ефект Зеємана.

§17.9.Квантові стани електронів в атомі. Квантові числа

Удосконалена теорія Бора, і навіть наявність у електрона спінового магнітного момента, не могли пояснити розділення пучка атомів срібла в дослідах Штерна і Герлаха на два пучки. Справа в тому, що згідно з теорією Бора-Зоммерфельда, для найменшого із можливих значень азимутального квантового числа n_φ=1, згідно з умовою просторового квантування, магнітне квантове число m повинно приймати три значення (-1,0,1) (§17.5) і пучок атомів повинен розщеплювались на три пучки. Досліди Штерна і Герлаха, як уже зазначалось, пояснюється наявністю в електронів спінового магнітного моменту, проекція якого на напрямок магнітного поля може приймати тільки два значення $\pm\mu_Б$. При цьому вважається, що існує тільки спіновий магнітний момент в електрона, а орбітальний магнітний момент електрона повинен бути відсутнім. Повне пояснення цього явища було отримане за допомогою хвильової (квантової) механіки, основні положення якої ми розглянемо в подальшому. Згідно квантової механіки, в таких атомах як водень, а також воднеподібних атомах, в тому числі і в атомах срібла, орбітальний момент електрона в нормальному стані дорівнює нулю (в планетарній моделі Бора такого бути не може).

Тому одночасно із введенням поняття «спіна» електрона Уленбеку і Гаудсміту довелось ще ввести декілька змін і доповнень в теорію атома. Для пояснення розщеплення пучка атомів срібла (водню) в дослідах Штерна і Герлаха на два пучки (а не на три) азимутальне квантове число n_φ вони замінили на l:

$$l = 0,1,2,\ldots,n-1. \quad (17.48)$$

Це квантове число l називається орбітальним квантовим числом. Орбітальним квантовим числом визначається величина орбітального механічного моменту (орбітального моменту імпульсу L), а також величина орбітального магнітного моменту електрона p_m. В квантовій механіці показується, що орбітальний момент імпульсу, таким чином, виражається через орбітальне квантове число l:

$$L = \sqrt{l(l+1)}\hbar\,. \quad (17.49)$$

Орбітальний магнітний момент із врахуванням гіромагнітного відношення (17.40) визначається так:

$$p_m = -\frac{e}{2m_e}L = -\frac{e\hbar}{2m_e}\sqrt{l(l+1)} = -\sqrt{l(l+1)}\,\mu_Б . \quad (17.50)$$

Магнітне квантове число m, яке характеризує квантування проекції орбітального моменту імпульсу на напрямок магнітного поля (17.39), може приймати значення цілих чисел від $-l$ до $+l$:

$$m = -l, -l+1, \ldots, -1, 0, +1, \ldots, +(l-1), +l, \quad (17.51)$$

тобто $2l+1$ значень.

Табл.17.2

Назва	Визначає	Символ	Приймає значення	Повне число значень
Головне квантове число	Розмір орбіти, енергію електрона	n	1,2,3...	∞
Орбітальне квантове число	Витягнутість (форму) орбіти, частково енергію електрона	l	0,1,2..., n-1	n
Магнітне квантове число	Орієнтацію орбіти в просторі, слабо енергію електрона	m	0,±1,±2,..., $\pm l$	$2l+1$
Спінове квантове число	Орієнтацію спіна, слабо енергію електрона	m_s	-1/2;+1/2	2

Таким чином, стан електрона атома характеризується сукупністю чотирьох квантових чисел n, l, m, m_s, які для зручності зведено в таблицю 17.2.

Замість приведеного набору квантових чисел n, l ,m, m_s, які характеризують стан руху електрона в атомі, можна ввести інший набір квантових чисел n, l, j, m_j, розглядаючи повний момент імпульсу $\vec{L}_j$ електрона в атомі. Повний момент імпульсу визначається геометричною (векторною) сумою орбітального момента $\vec{L}$ і власного моменту імпульсу електрона $\vec{L}_s$:

$$\vec{L}_j = \vec{L} + \vec{L}_s .$$

При цьому моменти $\vec{L}$ і $\vec{L}_s$ можуть складати між собою тільки такі кути , щоб повний момент $\vec{L}_j$ приймав наступні значення:

$$L_j = \sqrt{j(j+1)}\hbar , \quad (17.52)$$

де $j = l + m_s$ повне, або внутрішнє квантове число.

Магнітне квантове число m_j визначає проекцію повного моменту імпульсу на напрямок Z. Так як $m_s = \pm ½$, то при любому $l \neq 0$ квантове число j приймає два значення $j = l \pm ½$; при $l = 0$ квантове число j приймає одне значення $|j| = ½$.

Оскільки можливі значення L_j знаходяться за допомогою векторного додавання, то модель, що розглядається, носить назву векторної моделі. На основі векторної моделі можна пояснити аномальний ефект Зеємана і більш складні випадки в дослідах Штерна і Герлаха.

§17.10.Електронна оболонка атома. Принцип Паулі

Стан кожного електрона в атомі визначається чотирма квантовими числами n, l, m, m_s. Енергія електрона в атомі залежить в основному від головного квантового числа n (формула (17.20)). Квантові числа l, m, m_s в формулу (17.20) не входять. Однак в реальних умовах в атомах завжди існують магнітні і електричні поля, які діють на електрони. Тому енергія залежить в тій чи іншій мірі від квантових чисел l, m, m_s, чого теорія Бора не враховує.

Система електронів у незбуджених атомах намагається зайняти стан з найменшою енергією W_n, що відповідає основному принципу термодинаміки для систем, що

знаходяться в стані рівноваги. Тому електрони повинні розташовуватись на самих низьких доступних для них рівнях. Здавалось би, всі електрони повинні знаходитись на самому низькому (з найменшою енергією) рівні (n = 1, l = 0). Однак практика показує, що це далеко не так.

Влітку 1922р. Вольфгант Паулі приїхав в Копенгаген, щоб працювати разом із Бором над поясненням атомного ефекту Зеємана. Спроби класифікувати рівні енергії електрона, який знаходиться в сильному магнітному полі, поступово підводили Паулі до формулювання принципу Паулі. Остаточно цей принцип визрів після того, як Уленбек і Гаудсміт висунули ідею про існування спіна електрона і було добавлене квантове число m_s.

Закон фізики, відомий під назвою *принципу Паулі* або *принципу заборони*, встановлює, *що не може бути більше, ніж одного електрона в даному стані, якщо цей стан повністю описується чотирма квантовими числами* n, l, m, m_s.

Принцип Паулі можна сформулювати і в іншій формі: *в одному атомі не може бути двох електронів з одним і тим самим набором квантових чисел* n, l, m, m_s. *Принцип Паулі справедливий для мікрочастинок із спіном* $\hbar/2$.

Дія принципу Паулі і обумовлює послідовність заповнення енергетичних рівнів в атомі.

Даному значенню квантового числа n відповідають наступні значення квантових чисел l, m, m_s:

$$l = 0, 1, 2, \ldots, n-1;\ m = 0, \pm 1, \pm 2, \ldots, \pm l;$$

$$m_s = +1/2, -1/2.$$

Електрони з енергією, яка відповідає значенню l, утворюють оболонки, які позначають малими латинськими буквами *s, p, d, f, …* .

l	0	1	2	3
Оболонка	s	p	d	f

Сукупність оболонок, які відповідають квантовому числу n, утворюють шари, які позначають великими латинськими буквами *K, L, M, N, O,…* .

n	1	2	3	4	5	6	7
Шар	K	L	M	N	O	P	Q

Найбільш близько до ядра розташований K – шар, потім L – шар, M – шар, N – шар, ….

Стан із заданими n і l записують так: *ns, np, nd, nf* і говорять, наприклад, так 1*s* – стан, або 1*s* – оболонка (при n = 1), 2*s* – стан, або 2*s* – оболонка (n =2), 3*p* – стан, або 3*p* – оболонка і т. д.

Кожному шару з конкретним значенням n відповідає певне число станів, які визначаються квантовими числами l, m, m_s. Наприклад, при n=1, l=0, m=0, $m_s = \pm 1/2$. Тобто в цьому шарі з оболонкою 1*s* може знаходитись тільки два електрона, які відрізнятимуться квантовими числами m_s = +1/2, $m_s = -1/2$. При n =2 можуть бути дві оболонки: 2*s* (l = 0) і 2*p* (l =1).

На 2*s* оболонці може бути тільки два електрона, які відрізнятимуться квантовими числами $m_s = +1/2$ і $m_s = -1/2$. На оболонці 2*p* може бути шість електронів. Дійсно, при l =1, $m = -1, 0, +1$ і кожному m відповідають два стани, які відрізняються квантовими числами $m_s = +1/2$, $m_s = -1/2$. Таким чином, при n=2 кількість станів дорівнює вісім.

Не важко бачити, що для даної комбінації двох чисел n і l можлива $2(2l+1)$ кількість різних квантових станів, які відрізняються значеннями чисел m і m_s. Для

даного значення головного квантового числа n кількість різних станів дорівнює:

$$N = \sum_{l=0}^{n-1} 2(2l+1) = 2n^2.$$

Нижче складено таблицю 17.3 квантових станів. Для зручності квантові стани $m_s = \pm 1/2$ будемо позначати стрілками (↑↓), які відповідають різним орієнтаціям спіна.

Кількість станів, що зайняті електронами, визначає електронну конфігурацію атома. Наприклад, в атома кисню – 8 електронів, і його електронна конфігурація наступна $1s^2 2s^2 2p^4$ (індекс вгорі показує

Табл.17.3

Шар	n	l	m	m_s	Оболонка	Кількість станів
K	1	0	0	↓↑	1s	2
L	2	0	0	↓↑	2s	8
		1	-1 0 +1	↓↑ ↓↑ ↓↑	2p	
M	3	0	0	↓↑	3s	18
		1	-1 0 +1	↓↑ ↓↑ ↓↑	3p	
		2	-2 -1 0 +1 +2	↓↑ ↓↑ ↓↑ ↓↑ ↓↑	3d	

кількість електронів в даному стані). В 1s стані знаходиться 2 електрона, 2s – стані – 2 електрони, в 2p – стані – 4 електрони, таким чином, всього – 8 електронів.

§17.11.Періодична система

В 1869р. Д. І. Менделєєв розташував хімічні елементи у відповідності з їх атомними масами і фізичними властивостями. Хоч деякі елементи, такі як рідкоземельні, недостатньо добре узгоджувались з його схемою, періодична таблиця Менделєєва виявилась дуже вдалою. Однак закінчений вид таблиця набула тільки після відкриття цілого ряду нових елементів і принципу Паулі.

Прослідкуємо побудову таблиці. Щоб розібратися в розподілі електронів в хімічних елементах, будемо вважати, що кожний черговий елемент утворюється із попереднього шляхом додавання до ядра одного позитивного протона (і відповідного числа нейтронів) і додавання одного електрона, який будемо розміщувати в доступний йому, згідно принципу Паулі, стан з найменшою енергією. При заповненні оболонок електронами будемо користуватись наступним правилом зростання суми $n+l$. Оскільки енергія електрона в атомі залежить і від n, і від l, то заповнення оболонок електронами відбувається в порядку зростання суми $n+l$ (табл.17.4). Із двох однакових сум, наприклад, $n_1+l_1 = n_2+l_2$, перевагу надавати сумі, в якій менше n, оскільки енергія сильніше залежить від n, ніж від l.

Табл.17.4

Шар	Оболонка	$n+l$
K	1s	1
L	2s 2p	2 3
M	3s 3p 3d	3 4 5
N	4s 4p 4d 4f	4 5 6 7
O	5s 5p 5d 5f	5 6 7 8
P	6s 6p 6d 6f	6 7 8 9

Наприклад:

Неон $_{10}\mathrm{Ne} - 1s^2 2s^2 2p^6 -$ інертний газ;

Натрій $_{11}\mathrm{Na} - 1s^2 2s^2 2p^6 3s^1 -$ метал;

Кремній $_{14}\mathrm{Si} - 1s^2 2s^2 2p^6 3s^2 4p^2 -$ напівпровідник;

Аргон $_{18}\mathrm{Ar} - 1s^2 2s^2 2p^6 3s^2 3p^6 -$ інертний газ;

Калій $_{19}\mathrm{K} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d 4s^1 -$ метал;

Нікель $_{28}\mathrm{Ni} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d^8\ 4s^2 -$ метал;

Мідь $_{29}\mathrm{Cu} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^1 -$ метал, хороший провідник;

Германій $_{32}\mathrm{Ge} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^2 -$ напівпровідник.

Для атома міді має місце порушення правила зростання суми $n + l$. Згідно цього правила, електронна конфігурація для міді мала би бути така: $1s^2 2s^2 2p^6 3s^2 3p^6 3d^9 4s^2$. Але енергія електронів буде меншою, якщо $3d$ – оболонка буде повністю заповнена електронами. Тому один $4s$ електрон дозаповнює $3d$ – оболонку і на $4s$ оболонці залишається один валентний електрон. Цей $4s^1$ електрон і обумовлює ряд фізичних властивостей міді, як то висока електропровідність та теплопровідність та інше.

Цезій $_{55}\mathrm{Cs} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4f 5s^2 5p^6 6s^1 -$ рідкоземельний елемент;

Неодим $_{60}\mathrm{Nd} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10} 4f^4 5s^2 5p^6 6s^2 -$ рідкоземельний елемент.

Виходячи з цього, можемо стверджувати, що речовини, атоми яких мають валентний s^1–електрон, є металами (провідниками); p^6 – електрони – є інертними газами; $p^{2\ldots4}$ – електрони – є напівпровідниками; $d^{1\text{-}10}$ – електрони – є речовинами групи заліза; $f^{1\text{-}14}$ – електрони – є рідкоземельними елементами.

Таким чином, повторюваність будови зовнішньої електронної оболонки, яка витікає із принципу Паулі, і пояснює періодичний закон Д. І. Менделєєва в хімії.

§17.12.Рентгенівські спектри

У 1895 році В.Рентген відкрив випромінювання, яке було названо ікс – променями. Рентгенівські хвилі охоплюють широкий інтервал довжини від 0,01 до 10^{-8} мкм. Джерелом рентгенівського випромінювання є рентгенівські трубки. На рис.17.20 наведена схема рентгенівської трубки. У скляній трубці, в якій створено вакуум ~10^{-6} мм.рт.ст., є анод A і катод K (джерело електронів). Між катодом і анодом створюється сильне електричне поле з різницею потенціалів U~10–50 кВ. Електрони, які набули високих енергій у прискорюючому електричному полі, бомбардують анод A, площина якого утворює з напрямом руху

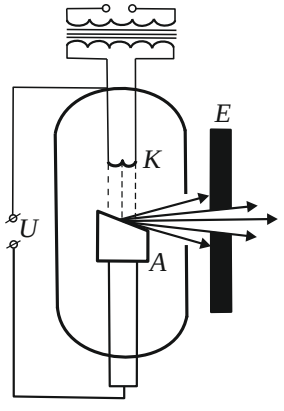


Рис.17.20.

електронів кут 45°. Внаслідок цього рентгенівське випромінювання, що виникає при гальмуванні електронів анодом, виходить

із трубки через отвір захисного свинцевого екрану *E*. Коефіцієнт корисної дії рентгенівських трубок становить близько 1%, тобто 99% кінетичної енергії електронів перетворюється в аноді у теплоту. Тому в тілі анода роблять канали, в яких циркулює охолоджуюча рідина. За принципом отримання електронних пучків рентгенівські трубки поділяють на три типи: іонні, електронні, індукційні. Джерелами рентгенівського випромінювання також можуть бути природні або штучні радіоактивні елементи, ряд небесних тіл.

Рентгенівське випромінювання, яке виникає в результаті гальмування електронів, називається гальмівним. Виникнення гальмівного випромінювання можна пояснити так. Електрони, які рухаються, утворюють навколо себе магнітне поле. Різка зміна швидкості електронів при ударі в анод рівносильна ослабленню і зникненню струму, що приводить до зміни магнітного поля, в результаті чого і виникають електромагнітні хвилі.

Гальмівне випромінювання має суцільний спектр. Це пояснюється тим, що одні електрони гальмуються швидше, а інші повільніше. Це призводить до виникнення електромагнітного випромінювання з різними довжинами хвиль. За квантовою теорією енергія фотона рентгенівського випромінювання

$$h\nu = \frac{m_e v^2}{2} - Q,$$

де $m_e v^2/2$ – кінетична енергія електрона перед ударом в анод, а Q – енергія, яка перетворюється в теплоту. Оскільки величина Q має різні значення, то і енергія $h\nu$ може бути різною.

На рис.17.21 зображені експериментальні криві розподілу інтенсивності рентгенівського випромінювання за довжинами хвиль λ при ударі електронів об вольфрамовий анод при різних значеннях прискорюючої напруги *U*, а на рис.17.22 – при ударі об анод з вольфраму, молібдену і хрому при однакових напругах *U*. Характерною особливістю суцільних спектрів є наявність чіткої короткохвильової межі λ_{min} та її незалежність від речовин анода. При підвищенні напруги інтенсивність випромінювання збільшується, а короткохвильова межа зміщується у бік коротких хвиль (рис.17.21).

Існування короткохвильової межі λ_{min} в рентгенівському спектрі пояснюється так. Енергія електрона при ударі об анод може повністю перейти в енергію випромінювання, тобто

$$\frac{m_e v^2}{2} = h\nu_{\max} = hc/\lambda_{\min}\,.$$

Оскільки $\frac{m_e v^2}{2} = eU$, то

$$eU = \frac{hc}{\lambda_{\min}} \text{ і } \lambda_{\min} = \frac{hc}{eU}\,.$$

Отже, мінімальна довжина хвилі гальмівного випромінювання обернено пропорційна прискорюючій напрузі *U*.

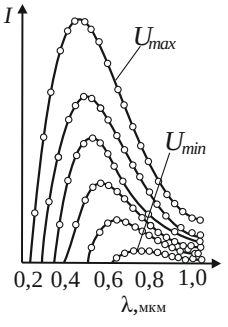


Рис.17.21

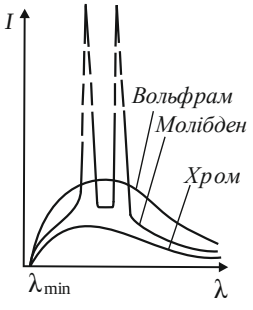


Рис.17.22

При підвищенні напруги на рентгенівській трубці до величини, більшої за деяке критичне значення, що залежить від

речовини анода, на фоні суцільного спектра виникають лінійчасті спектри, інтенсивність яких у тисячі разів перевищує інтенсивність суцільного спектра. Оскільки лінійчастий спектр визначається природою речовини, з якої виготовлено анод, то його називають характеристичним.

З рис.17.22 видно, що для напруги при якій рентгенівська трубка з анодом із вольфраму і хрому випромінює суцільний спектр, трубка з молібденовим анодом, крім суцільного, випромінює ще й лінійчастий спектр.

Існування *характеристичного випромінювання* підтверджує наявність дискретних рівнів енергії атомів.

Кожний хімічний елемент дає характеристичне рентгенівське випромінювання незалежно від того, чи його атоми вільні, чи входять до складу хімічних сполук.

Рентгенівські спектральні лінії утворюють певні серії, які позначають буквами *K,L,M,N,O* і т.д. Кожна серія має невелику кількість ліній, які позначають у порядку зменшення довжини хвилі індексами α, β, γ і т.д. Лінійчасті рентгенівські спектри різних речовин подібні між собою. Це зумовлено подібністю забудови нижніх (найближчих до ядра) шарів електронів в атомах. При збільшенні атомного номеру речовини анода весь рентгенівський спектр лише зміщується в короткохвильову частину спектра, не змінюючи своєї структури. Це можна пояснити тим, що лінійчасті спектри утворюються внаслідок переходів електронів у внутрішніх шарах атомів. Механізм виникнення лінійчастих спектрів можна пояснити так. Якщо під впливом падаючих електронів високих енергій на атоми анода, вибивається один із двох електронів 1*s* – стану *K*-оболонки, то звільнене місце може зайняти електрон з якої-небудь зовнішньої оболонки (*L, M, N* і т.д.). При цьому виникає *K*-серія. Аналогічно виникають й інші серії. Схема виникнення спектральних серій показана на рис.17.23.

Серія *K* обов'язково супроводжується іншими серіями, оскільки при випромінюванні її ліній звільняються рівні в оболонках *L, M,* і т.д., які будуть заповнюватися електронами із більш віддалених оболонок. Частоти ліній зростають при переході від лінії K_α до K_β, K_γ,…, що пов'язано із збільшенням енергії, яка вивільняється при переході електрона на *K*-оболонку з дедалі більш віддалених оболонок. А інтенсивності ліній спадають при переході від лінії K_α до лінії K_β, K_γ, бо ймовірність переходу на $K-$оболонку з $L-$оболонки більша, ніж з $M-$, $N-$ і т.д. Ці закономірності справедливі для $L-$, $M-$ та інших серій.

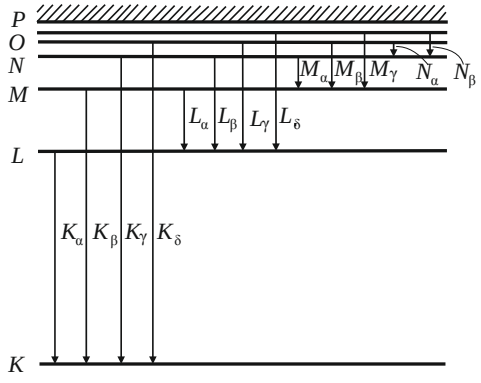


Рис.17.23.

При дослідженні лінійчастих рентгенівських спектрів Г.Мозлі в 1913р. встановив закон, який виражає зв'язок між частотами спектральних ліній ω і атомним номером Z елементу, який випромінює ці лінії. Так, для $K_\alpha -$ лінії ця залежність має вигляд:

$$\omega_{K_\alpha} = R(Z-1)^2\left(\frac{1}{1^2}-\frac{1}{2^2}\right),$$

де R – стала Рідберга. Для лінії K_β –

$$\omega_{K_\beta} = R(Z-1)^2\left(\frac{1}{1^2}-\frac{1}{3^2}\right),$$

а для лінії L_α –

$$\omega_{L_\alpha} = R(Z-7{,}5)^2\left(\frac{1}{2^2}-\frac{1}{3^2}\right)$$

і т.д.

Узагальнений закон Мозлі має вигляд:

$$\omega = R(Z-\sigma)^2\left(\frac{1}{k^2}-\frac{1}{n^2}\right), \qquad (17.53)$$

де k =1, n =2, 3, 4… для ліній K–серії, k =2, n =3,4…для ліній L–серії, k =3, n =4,5,6… для ліній M–серії, σ – стала екранування, яка має певне значення в межах однієї серії.

Зміст сталої екранування полягає в тому, що у важкому атомі, який має Z електронів, на електрон, що здійснює, наприклад, перехід, який відповідає лінії K_α діє не весь заряд ядра Ze, а заряд (Z $-$ σ)e, ослаблений екрануючою дією одного електрона, що залишився на K – оболонці. Для K–серії $\sigma = 1$, для L–серії $\sigma = 7{,}5$.

У загальному вигляді закон Мозлі записують так:

$$\sqrt{\omega} = a(Z-\sigma), \qquad (17.54)$$

де ω – частота, a – стала для певної серії ліній.

Закон Мозлі можна сформулювати так: корінь квадратний із частоти є лінійною функцією атомного номера Z. На рис.17.24 приведені побудовані по експериментальним точкам графіки залежності $\sqrt{\omega}$ від Z для ліній K_α і L_α . По цих графіках можна судити, наскільки точно виконується закон Мозлі. *Закон Мозлі дозволяє за виміряною довжиною хвилі рентгенівських ліній точно встановити атомний номер даного елементу.* Цей закон відіграв велику роль при розміщенні елементів в періодичній системі. На основі цього закону вперше було показано, що не атомна маса, а атомний номер, який дорівнює заряду ядра, визначає хімічні властивості атомів. Мозлі дав

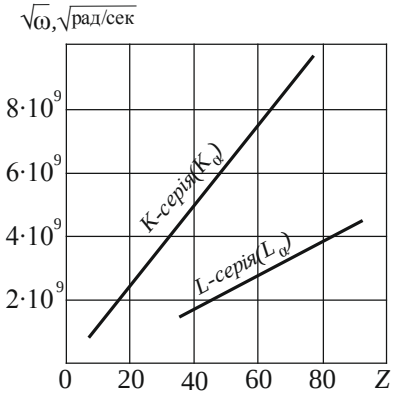


Рис.17.24.

просте теоретичне пояснення знайденого ним закону. Він зазначив, що лінії з частотами, які визначаються формулою (17.53), співпадають з лініями, які випромінюються при переході електрона, що знаходиться в полі заряду $(Z-\sigma)e$, з рівня з номером n на рівень з номером m.

Зауважимо, що формула (17.53) базується на припущенні, що постійна екранування σ для обох термів має однакове значення. Реально екранування, наприклад, для K – терма буде слабше, ніж для L – терма, тому що електрон, який знаходиться в L – оболонці, екранують обидва електрони K – оболонки і, крім цього, частково приймають участь в екрануванні останні електрони L–оболонки, в той час як для електрона K–оболонки екранування здійснюється тільки одним другим K–електроном. Більш строго формулу (17.53) необхідно записати у вигляді:

$$\omega = R\left\{\frac{(Z-\sigma_1)^2}{k^2}-\frac{(Z-\sigma_2)^2}{n^2}\right\}. \qquad (17.55)$$

ПУЛЮЙ ІВАН (1845-1918).

Опублікував у 1896 р. дві статті: "Про походження рентгенівських променів та їх фотографічну дію" і додаток до праці "Про походження рентгенівських променів та їх фотографічну дію". Навів близько десяти досконалих знімків, які ілюструють отримані результати. Рентгенограми відрізнялись від знімків Рентгена високою якістю. Пулюєві належить перша вдала спроба з'ясувати механізм виникнення рентгенівських променів. Зробив перший знімок цілого людського скелета.

Контрольні питання

1.Суть ядерної моделі атома.

2.В чому неспроможність класичної теорії при поясненні ядерної моделі атома?

3.Постулати Бора. Які досліди підтвердили ці постулати?

4.Які є серії спектральних ліній в спектрі випромінювання атома водню?

5.Борівська теорія будови атома водню та воднеподібних атомів.

6.Механічний і магнітний моменти електрона в атомі.

7.Що означає просторове квантування?

8.Що підтверджують досліди Штерна і Герлаха з атомними пучками?

9.Що підтверджують досліди Зеємана?

10.Які квантові числа описують стан електрона в атомі?

11.Пояснити фізичний зміст кожного квантового числа.

12.Сформулювати і пояснити принцип Паулі.

13.Заповнення електронних оболонок атома електронами.

14.Поясніть принцип розміщення хімічних елементів в періодичній таблиці.

15.Рентгенівські спектри. Яка їхня роль в поясненні будови атома?

Розділ 18.Елементи квантової механіки

Квантова механіка (хвильова механіка) – це розділ теоретичної фізики, в якому вивчаються закони руху частинок в області мікросвіту (в масштабах $10^{-8} \div 10^{-15}$м*).* Об'єктами вивчення квантової механіки є: кристали, молекули і атоми, атомні ядра, елементарні частинки. Закони квантової механіки складають фундамент науки, що вивчає будову речовини.

В основі квантової механіки лежать:

1.Уявлення Планка про кванти енергії $W = \hbar\omega = h\nu$.

2.Уявлення Ейнштейна про фотони. Ейнштейн припустив, що світло не тільки випромінюється і поглинається, але і розповсюджується квантами, тобто дискретність властива самому світлу: світло складається із окремих порцій – світлових квантів, які пізніше назвали фотонами. Енергія фотона $W = h\nu$. На основі цієї гіпотези Ейнштейн пояснив закономірності зовнішнього фотоефекту.

3.Дані про існування дискретних значень деяких фізичних величин, які характеризують стан частинки в мікросвіті (квантування: енергії, момента кількості руху, магнітного момента; спектри випромінювання і т.п.).

4.Гіпотеза де Бройля.

§18.1.Протиріччя хвиля – частинка

Із часів філософів античності до І.Ньютона вчені розглядали світло як потік частинок (корпускул), які швидко рухаються. В 1801р. Томас Юнг, здійснивши перші досліди з інтерференції світла, підвів тверду

експериментальну основу під його хвильову теорію. Здавалось, що фізики можуть остаточно відмовитись від корпускулярної теорії світла. Однак подальші дослідження показали, що все більша кількість фізичних властивостей твердих тіл, рідин і газів можна пояснити на основі атомістичних уявлень.

В кінці XIXст. експерименти Дж.Дж. Томсона і роботи Х.А.Лоренца привели до концепції про існування елементарних частинок електрики − електронів. Нарешті, «коло замкнулось», коли в 1900р. Макс Планк встановив закон випромінювання абсолютно чорного тіла, згідно з яким джерело світла випромінює світло не безперервно, а однаковими і кінцевими порціями, які отримали назву квантів. З новою силою протиріччя хвиля – частинка проявилось в 1923р., коли Артур Комптон експериментально довів, що кванти рентгенівського випромінювання поводять себе як частинки з певною енергією і певним імпульсом.

В 1924р. Луї де Бройль у своїй докторській дисертації «Дослідження з теорії квантів» висловив думку про те, що якщо світло в багатьох випадках проявляє корпускулярні властивості, то і частинки, зокрема електрони, в силу симетрії в природі можуть володіти хвильовими властивостями. В 1927р. ідеї де Бройля були експериментально підтверджені К.Дж.Девіссоном і Л.Х.Джермером і незалежно від них Дж.П.Томсоном, які показали, що для пучка електронів, як і для світлового пучка, спостерігається явище дифракції.

§18.2.Хвилі де Бройля

Планк зв’язав енергію частинки світла (фотонів) з частотою світла співвідношенням

$$W = h\nu, \qquad (18.1)$$

де ν – частота фотона, а енергія W співпадає із значенням, яке знаходиться із спеціальної теорії відносності (із формули $W = W_0 + W_k$, де для фотонів енергія спокою $W_0 = m_\phi c^2 = 0$, так як $m_\phi = 0$ (§16.6), і загальна енергія W дорівнює тільки кінетичній енергії W_k. Однак суто корпускулярна теорія не містить в собі жодного елемента, який дозволив би визначити частоту. Щоб розв’язати цю проблему, де Бройль висловив гіпотезу, згідно якої з кожним фотоном зв’язаний хвильовий процес. Це дало можливість пояснити таке суто хвильове явище, як інтерференцію. Комптон показав, що світлові хвилі володіють імпульсом, тобто характеристикою, яку, як правило, пов’язують з частинками. Вираз для енергії фотона можна записати у вигляді

$$W = mc^2 = pc, \qquad (18.2)$$

де $p = mc$ – імпульс фотона. Об’єднавши (18.1) і (18.2), отримаємо співвідношення

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}, \qquad (18.3)$$

яке зв’язує довжину хвилі λ і частоту ν, тобто величини, які характеризують фотон як хвилю, з імпульсом фотона p.

Потім де Бройль висловив чудову ідею: якщо світлові хвилі мають корпускулярну природу, то частинки, наприклад, електрони повинні проявляти хвильові властивості. Таким чином, електрон може володіти певним імпульсом, який дорівнює

$$p = mv = \frac{h}{\lambda}, \qquad (18.4)$$

де m – релятивістська маса, v − швидкість електрона. Із цього співвідношення витікає, що електрону (частинці) приписується певна хвильова характеристика – довжина хвилі λ (частота ν). Тоді із (18.4) отримаємо, що довжина хвилі дорівнює

$$\lambda = \frac{h}{mv} = \frac{h}{p}. \qquad (18.5)$$

Співвідношення (18.5) відображає гіпотезу де Бройля.

В теорії відносності, так само як і в класичній механіці, для замкнутої фізичної системи зберігається імпульс $\vec{p}$ і енергія W. Величина $W^2 - (c\vec{p})^2 = m_0^2c^4$ є інваріантом при перетвореннях Лоренца в релятивістській механіці. В цій формулі m_0 – маса спокою частинки, W – її повна енергія, c – швидкість світла у вакуумі. Тому

$$p = (1/c)\sqrt{W^2 - W_0^2},$$

де $W = mc^2$ – повна енергія, $W_0 = m_0c^2$ – енергія спокою.

Тоді (18.5) перепишемо так:

$$\lambda = \frac{h}{(1/c)\sqrt{W^2 - W_0^2}} = \frac{h}{(1/c)\sqrt{(mc^2)^2 - (m_0c^2)^2}}.$$

Звідки при врахуванні, що релятивістська маса $m = m_0/\sqrt{1 - \frac{v^2}{c^2}}$, отримаємо:

$$\lambda = \frac{h}{m_0c\sqrt{1/\left(1 - \frac{v^2}{c^2}\right) - 1}},$$

а після спрощення

$$\lambda = \frac{h\sqrt{1 - v^2/c^2}}{m_0v} = \frac{h}{mv}.$$

Для нерелятивістського випадку ($v \ll c$) релятивістська маса m дорівнює масі спокою частинки m_0.

Хоч в своїх перших роботах де Бройль розглядав проблему на прикладі електрона, всі отримані співвідношення справедливі для будь-яких матеріальних тіл. М'яч, маса якого 0,14 кг, запускають сильним ударом і він набуває швидкість біля 40м/с. Довжина хвилі де Бройля, яка пов'язана з таким м'ячем, дорівнює:

$$\lambda = \frac{h}{m_0v} = \frac{6{,}625 \cdot 10^{-34}}{0{,}14 \cdot 40}\text{м} = 1{,}2 \cdot 10^{-34}\text{м} = 1{,}2 \cdot 10^{-24}\text{Å},$$

тобто λ є настільки мала величина, що її неможливо визначити з досліду. Рух електрона з такою самою швидкістю супроводжується хвильовим процесом з довжиною хвилі

$$\lambda = \frac{h}{m_ev} = \frac{6{,}625 \cdot 10^{-34}}{9{,}1 \cdot 10^{-31} \cdot 40}\text{м} = 1{,}8 \cdot 10^{-5}\text{м} = 1{,}8 \cdot 10^{5}\text{Å},$$

яка може бути знайдена з експерименту.

Якщо частинка має кінетичну енергію W_k, то замінюючи її імпульс p через енергію W_k за формулою $p = \sqrt{2mW_k}$, можна записати

$$\lambda = \frac{h}{\sqrt{2mW_k}}.$$

Зокрема, для електрона, що прискорюється в електричному полі з різницею потенціалів U, маємо

$$W_k = \frac{m_ev^2}{2} = eU.$$

Тоді для електрона

$$\lambda = \frac{h}{\sqrt{2m_eeU}} = \frac{12{,}25}{\sqrt{U(B)}} \cdot 10^{-10}\text{м} = \frac{12{,}25}{\sqrt{U(B)}}\text{Å}. \qquad (18.6)$$

Для електричних полів з напругою 1−10 В довжини хвиль де Бройля належать інтервалу (10÷0,1)Å, тобто мають довжини хвиль звичайних рентгенівських променів.

Цікавий результат дало порівняння гіпотези де Бройля і теорії атома Бора. Згідно з постулатом Бора: $m_evr = n\frac{h}{2\pi}$. З цього рівняння можна визначити довжину стаціонарної орбіти

$$2\pi r = n\frac{h}{m_ev},$$

де згідно з гіпотезою де Бройля $\frac{h}{m_ev} = \lambda$.

Тоді довжина стаціонарної орбіти

$$2\pi r = n\lambda.$$

Отже, головне квантове число n показує, яке число довжин хвиль де Бройля електрона укладається на довжині борівської колової орбіти.

§18.3.Експериментальне підтвердження гіпотези де Бройля

Для виявлення хвильових властивостей електронів, що рухаються, необхідно користуватись дифракційними гратками, період яких по порядку величини дорівнює довжині хвилі де Бройля, тобто d~1Å. Такою граткою може бути кристалічна решітка кристалів, як у випадку вивчення дифракції рентгенівських променів (§13.8).

В 1925р. в лабораторії фірми «Белл телефон» (США) К.Дж.Девіссоном і К.Х.Кунсманом були проведені експериментальні дослідження явища емісії вторинних електронів, які були пізніше продовжені Девіссоном і Л. Х. Джермером. Схема досліду наведена на рис.18.1. В цих експериментах спостерігалось явище "випаровування" вторинних електронів під дією пучка електронів, який направлявся на зразок із спресованого порошка нікелю.

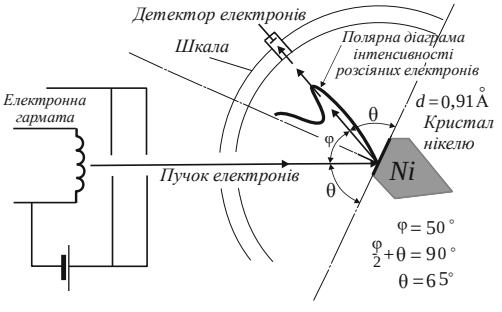


Рис.18.1.

Одного разу одна із пластинок нікеля окислилась і для видалення окисної плівки пластинку відпалили. При повторних дослідах з цією кристалічною мішенню результати різко відрізнялись від тих , які були отримані раніше. Після тривалого прогріву замість дуже дрібних кристаликів утворилось декілька великих монокристалів. Емісія вторинних електронів відбувалась, як і раніше, під будь-якими кутами, але при певних кутах кількість електронів, які розсіювались, різко зростала.

В одному з дослідів пучок електронів з енергією W_k = 54 еВ (U=54В) давав максимум в розподілі інтенсивності розсіяних електронів під кутом φ=50° (рис.18.1). За формулою (18.6) знайдемо довжину хвилі де Бройля

$$\lambda = \frac{12{,}25}{\sqrt{54}}\text{Å} = 1{,}67\text{Å}.$$

В той же час, використовуючи формулу (13.33) для дифракції хвиль від площин в кристалі по методу Брегга (§13.8), для максимуму першого порядку (m=1) дифракції електронів від кристалу нікеля з періодом гратки $d = 0{,}91$Å для $\theta = 65^{\circ}$, за формулою (13.33) отримаємо, що довжина хвилі де Бройля, яка дифрагує, дорівнює

$$\lambda = 2d\sin\theta = 2 \cdot 0{,}91 \cdot \sin 65^{\circ}\,\text{Å} = 1{,}65\text{Å}.$$

Обидва результати добре співпадають, і це переконливо свідчить про те, що електрони проявляють як хвильові властивості, так і властивості частинки.

Функцію $\lambda = f(U)$ (формула (18.6)) К.Девіссон передав на графіку прямолінійною залежністю в координатах $\lambda = f(1/\sqrt{U})$ (рис.18.2), якою підтверджувалась передбачувана де Бройлем обернено пропорційна залежність довжини хвилі де Бройля матеріальної частинки від її швидкості. Розбіжність між величинами, одержаними в дослідах К.Девіссона та Л.Джермера, та підрахованими за теоретичною формулою де Бройля (18.6), не перевищувала прийнятних

2%. Отже, узгодження теоретичної гіпотези та експерименту було переконливо засвідчено.

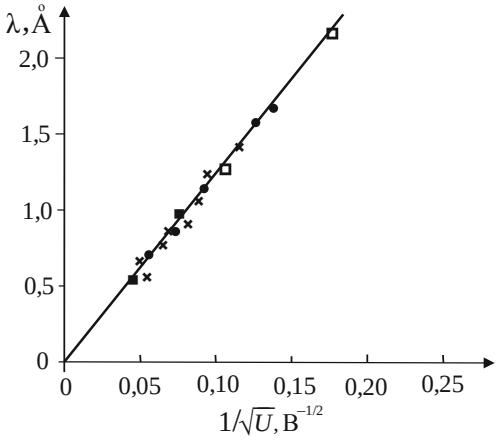


Рис.18.2.Перевірка формули де Бройля $\lambda = h/mv$. На графіку в координатах λ та $1/\sqrt{U}$ зображено для електронів зв'язок їхньої дебройлевої довжини хвилі λ, обчисленої за даними дифракційної картини, з величиною різниці потенціалів U електричного поля (Clinton J. Davisson. The discovery of elektron waves. Nobel Lecture, December 13, 1937. Nobel Lectures, Physics 1922-1941. - p.387-394. 1965. Elsevier Publishing Company, Amsterdam).

В 1927 р. Дж. П. Томсон із своїми студентами здійснив досліди з швидкими електронами і отримав практично ті самі результати. Як і в експериментах по дифракції рентгенівських променів, Томсон досліджував розсіювання швидких електронів на дуже тонкій металевій фользі, яка виготовлялась із полікристалічного матеріалу з хаотичним розташуванням кристалічних осей. Експериментальна установка схематично показана на рис.18.3.

Пропускаючи паралельний пучок електронів (з енергією біля 10^4еВ) через золоту фольгу, товщина якої біля 10^{-5}см = 0,1мкм, Томсон отримав на екрані ряд дифракційних кілець. Кути розсіювання $\theta_1, \theta_2, \theta_3$ і т.д. відповідали значенням, які отримуються із відомого рівняння для дифракції світла, яке проходить через решітку, $d\sin\theta = m\lambda$ (m=1,2,3,…), де θ – кут між дифракційним пучком і напрямком падаючого пучка.

Експериментальна перевірка хвильової природи частинок продовжувалась і в наступні роки. Ось тільки деякі найбільш цікаві експерименти.

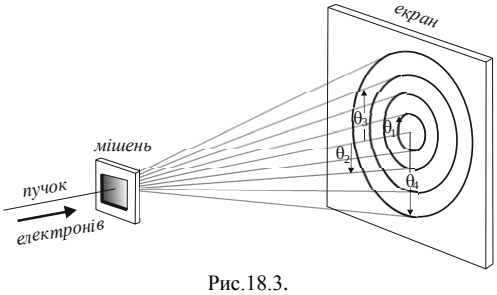


Рис.18.3.

1.Німецький вчений Рупп в 1929 р. виміряв довжину хвилі де Бройля електронів, здійснивши дифракцію пучка електронів за допомогою простої оптичної дифракційної решітки при дуже малих кутах ковзання.

2.В 1931 р. Джонсон продемонстрував дифракцію пучка молекул водню при розсіюванні в кристалі.

3.Естерман, Фріш і Штерн в 1938р. здійснили експеримент по дифракції пучка атомів гелію. Як розсіюючий елемент вони застосували кристали фтористого літію.

4.Протягом тривалого часу залишалось нез'ясованим, чи хвильові властивості є окремо у кожної мікрочастинки, чи їх має лише потік багатьох мікрочастинок. У 1949 р. В.Фабрикант, М. Сушкін, Л. Біберман провели дослід з дифракції потоку електронів настільки незначної інтенсивності, що інтервал часу між пролітанням наступних електронів був приблизно в 30000 разів більший від часу, який необхідний для проходження через увесь прилад, тобто реєструвалось проходження окремих

електронів. При досить тривалій експозиції отримували таку ж дифракційну картину, як і для інтенсивних пучків електронів. Цей дослід переконав, що хвильові властивості притаманні кожному електрону.

Ці експерименти остаточно довели хвильову природу частинок, але разом з тим поставили ряд нових питань. Наприклад, оскільки фотони і електрони проявляють властивості і частинок, і хвилі, то може здатись, що дійсно відсутня будь-яка різниця між частинками і хвилями. Однак, використовуючи тільки на хвильову теорію, не вдається задовільно інтерпретувати такі експерименти, як комптонівське розсіювання, чи фотоефект, а також пояснити той факт, що частинки, наприклад *електрони, не можуть рухатись із швидкістю світла.*

Нільс Бор, намагаючись розв'язати протиріччя хвиля-частинка, ввів так званий ***принцип доповнюваності***. В основі його лежить той факт, що ні хвилі, ні частинки ніколи не проявляють одночасно в одному і тому експерименті і хвильові, і корпускулярні властивості. В кожному випадку – чи то яке-небудь випромінювання, чи потік електронів – для повного описання явища необхідно використовувати і хвильову, і корпускулярну моделі, але кожна з цих моделей має свою власну сферу застосування.

§18.4.Співвідношення невизначеностей як прояв корпускулярно-хвильового дуалізму властивостей матерії

Розглянемо наслідки, що випливають із корпускулярно-хвильової двоїстості електронів, фотонів, протонів та інших об'єктів мікросвіту, і які обмеження накладає цей дуалізм на класичний метод описування їх руху і, зокрема, на можливість одночасного визначення їх імпульсів і координат.

Будемо розглядати одномірний випадок руху частинок. У класичній механіці будь-яка частинка рухається вздовж певної траєкторії так, що фіксовані її координати та імпульси. Сукупність координат x визначає траєкторію руху частинки. В кожній точці траєкторії частинка має конкретні значення координати x і відповідного імпульсу $p_x = m\frac{dx}{dt} = mv_x$. Якщо вважати, що для такої частинки справедливее співвідношення (18.5), то в кожній точці траєкторії буде відповідна довжина хвилі λ, тобто λ буде залежати від координати x. Оскільки довжина хвилі є характеристикою хвильового процесу, то вона не може залежати від координати.

Таким чином, якщо для частинки відома координата x (тобто невизначеність $\Delta x = 0$) і для неї справедливе співвідношення (18.4), то імпульс для такої частинки невизначений, тобто $\Delta p_x = \infty$.

Мікрочастинки внаслідок наявності в них хвильових властивостей відрізняються від класичних частинок. Одна з основних відмінностей полягає в тому, що мікрочастинка не має чіткої траєкторії, і неправомірно говорити одночасно про точні значення її координат та імпульсу.

Як відомо, будь-яка хвиля, незалежно від природи, є процесом, що заповнює більшу чи меншу область простору, а через це не може локалізуватися в одній точці. Наприклад, вздовж осі OX рухається фотон, для якого точно відомий імпульс $p = p_x$, тобто $\Delta p_x \to 0$. Такому фотону відповідає хвиля, довжина якої однозначно визначається величиною імпульсу $\lambda = h/p$.

Така монохроматична хвиля в просторі нічим не обмежена, область її існування – вся вісь *OX*. Це означає, що в ньому випадку просторовий інтервал Δx, в якому замкнено об'єкт з хвильовими властивостями – фотон, дорівнює нескінченності. Іншими словами, якщо $\Delta p_x \to 0$ маємо $\Delta x \to \infty$. Отже, *якщо точно відомий імпульс фотона, то його локалізація стає цілком невизначеною.*

Розглянемо систему хвиль, суперпозиція яких дає огинаючу, яка розповсюджується зі швидкістю, яка відмінна від швидкостей хвиль, які додаються. Цю огинаючу хвиль називають хвильовим пакетом (або групою хвиль). Хвильовий пакет, який побудований відповідною підбіркою великого числа хвиль, можна інтерпретувати як матеріальну частинку. Як просте наближення розглянемо хвильовий пакет, побудований із двох хвиль, які описуються рівняннями:

$$\psi_1 = A\cos(\omega_1 t - k_1 x),$$

$$\psi_2 = A\cos(\omega_2 t - k_2 x),$$

де $\omega_1 \approx \omega_2 = \omega$ і $k_1 \approx k_2 = k$.

Результуюча хвиля буде

$$\psi = \psi_1 + \psi_2 =$$

$$= A_{рез}\cos\left(\frac{\omega_1 + \omega_2}{2}t - \frac{k_1 + k_2}{2}x\right)$$

з результуючою амплітудою

$$A_{рез} = 2A\cos\left(\frac{\omega_1 - \omega_2}{2}t - \frac{k_1 - k_2}{2}x\right).$$

На рисунку 18.4 зображено дві хвилі, при складанні яких утворюється результуюча хвиля. Зазначимо, що в результаті суперпозиції цих двох хвиль виникають биття, або хвильові групи. Кожна хвиля буде розповсюджуватись із своєю власною швидкістю (фазовою швидкістю), але групи хвиль, які отримуються в результаті суперпозиції, переміщуються з іншою швидкістю (груповою швидкістю §§10.5 і 15.2). Так як $\omega_1 \approx \omega_2$ і $k_1 \approx k_2$, то фазові швидкості двох хвиль будуть приблизно однакові. Тоді

$$v_\phi = \nu\lambda = \omega/k. \qquad (18.7)$$

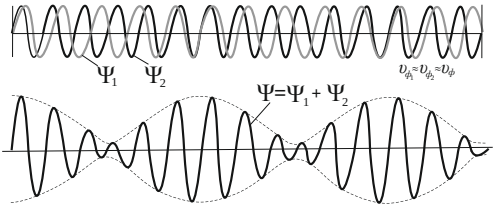


Рис.18.4.

У співвідношенні (18.7) ν і λ можна асоціювати відповідно з частотою і довжиною хвилі де Бройля.

Гіпотеза про хвилі де Бройля набуває ще більш глибокий зміст, якщо врахувати, що швидкість частинки в точності дорівнює груповій швидкості хвиль де Бройля. Швидкість частинки знаходиться із співвідношення $v = dW/dp$.

Групова швидкість, згідно (§10.5, формула (10.28)) буде: $v_{гр} = d\omega/dk$. Це співвідношення можна переписати так:

$$v_{гр} = \frac{d(2\pi\nu)}{d(\frac{2\pi}{\lambda})} = -\lambda^2\frac{d\nu}{d\lambda}. \qquad (18.8)$$

Із співвідношення (18.5) витікає, що $d\lambda = -hdp/p^2$, а із співвідношення релятивістської механіки $W^2 = W_0^2 + p^2c^2$ знаходимо: $dW = \frac{pc^2}{W}dp$. Із формули $\nu = W/h$ отримаємо $d\nu = dW/h = \frac{pc^2}{Wh}dp$. Тоді

$$v_{гр} = \lambda^2\frac{d\nu}{dp}\frac{p^2}{h} = \frac{pc^2}{W} = \frac{mvc^2}{mc^2} = v.$$

Групова швидкість, таким чином, дорівнює швидкості частинки, або, іншими словами, хвилі де Бройля переміщуються разом з частинкою. Вираз (10.30) (§10.5) перепишемо так:

$$v_{гр} = v_\phi - \lambda\frac{dv_\phi}{d\lambda},$$

в якому групова швидкість зв'язана з фазовою швидкістю. Зауважимо, що якщо хвилі розповсюджуються в середовищі без дисперсії, тобто $v_ф = \text{const}$, то $v_{гр} = v_ф$. Що можна

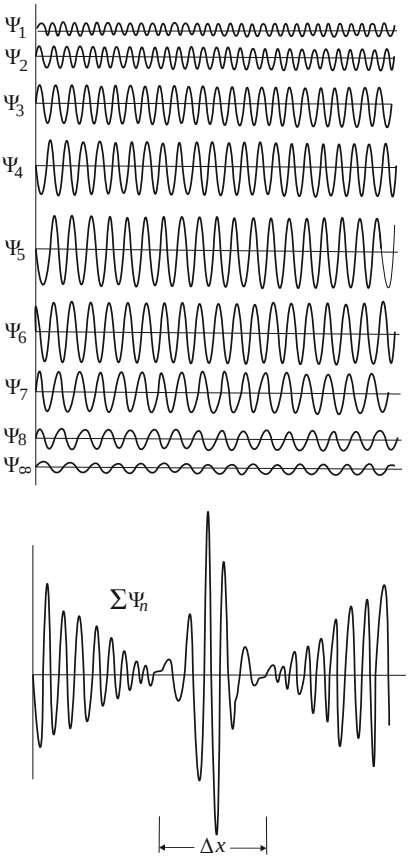


Рис.18.5.

сказати про групову швидкість фотонів? Для фотона $W = pc$ і

$$v_{гр} = \frac{pc^2}{W} = c.$$

Таким чином, ми дістали очікуваний результат – групова швидкість фотонів дорівнює швидкості самого фотона.

На рис.18.5 зображений випадок, коли складається значне число хвиль: $\psi_1, \psi_2, \ldots, \psi_n$. Хвильовий пакет при цьому різко визначається, а його довжина Δx значно скорочується. Оскільки в хвильовому пакеті є набір значень λ, які містяться в деякому інтервалі $\Delta\lambda > 0$, то набір значень імпульсів p, як витікає із (18.4), знаходиться в інтервалі

$$\Delta p = \Delta\left(\frac{h}{\lambda}\right) = \frac{h}{\lambda^2}|\Delta\lambda|.$$

Чим ширший інтервал можливих значень інтерферуючих хвиль $\Delta\lambda$, а разом з ним інтервал Δp_x можливих значень імпульсів, тим вужча область локалізації Δx результуючого пакета. Інакше кажучи, чим більша невизначеність імпульсу фотона, тим точніше можна визначити його координати. При $\Delta p \to \infty$ маємо $\Delta x \to 0$, тобто точне значення координати фотона можливе тільки при повній невизначеності його імпульсу.

Такі висновки справедливі для фотонів, електронів, протонів та інших мікрочастинок.

Щоб усунути протиріччя хвиля-частинка, німецький фізик Вернер Гайзенберг в 1927р. висунув цікаву ідею, ввівши так званий принцип невизначеності (або співвідношення невизначеностей). Цей принцип виражає фундаментальну границю можливості одночасного вимірювання певних пар змінних, наприклад, положення частинки і її імпульсу. На рис.18.6. зображена частинка, яка рухається із швидкістю v. Частинка локалізована всередині хильового пакету, який переміщується із швидкістю $v_{гр} = v$. Співвідношення (18.8) можна записати в такій формі:

$$v_{гр} = \frac{d\omega}{dk} = \frac{d\nu}{d\left(\frac{1}{\lambda}\right)} = \frac{h d\nu}{d\left(\frac{h}{\lambda}\right)} = h\frac{d\nu}{dp}. \quad (8.9)$$

Із рисунка 18.6 і співвідношення (18.9) витікає, що

$$v_{гр} = \frac{\Delta x}{\Delta t} = h\frac{\Delta\nu}{\Delta p}$$

або

$$\Delta p \cdot \Delta x = h \cdot \Delta\nu \cdot \Delta t. \quad (8.10)$$

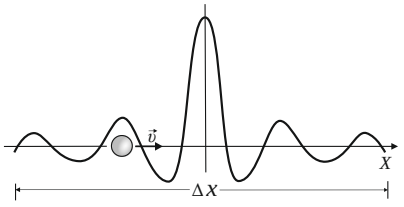


Рис.18.6.

Якщо потрібно виміряти частоту хвилі, то мінімальним часом, який необхідний для вимірювання, буде такий інтервал часу Δt, щоб одна повна довжина хвилі пройшла фіксовану точку. Цей інтервал часу, який повязаний з частотою співвідношенням $\Delta t \geq 1/\Delta\nu$, є час, що відповідає одному повному періоду.

Враховуючи тепер, що $\Delta\nu \cdot \Delta t \geq 1$, із співвідношення (18.10) отримаємо нерівність

$$\Delta p \cdot \Delta x \geq h, \qquad (18.11)$$

яка є одним із співвідношень невизначеностей Гайзенберга. Зміст отриманої фундаментальної границі ($\geq h$) не тільки в тому, що є деяка границя для точності вимірювання координати x і імпульсу p, а і в тому, що при зменшенні області локалізації частинки Δx в просторі збільшується невизначеність в вимірюванні її імпульсу, і навпаки.

Неточність Δp у визначенні імпульсу у співвідношенні (18.11) є неточність у визначенні проекції імпульсу на напрямок осі x і більш правильним буде запис:

$$\Delta p_x \cdot \Delta x \geq h. \qquad (18.12)$$

Такі пари величин (наприклад, p_x і x), для яких невизначеність в значенні однієї величини залежить від точності вимірювання іншої, називаються спряженими величинами. Зауважимо, що не має ніяких обмежень на точність вимірювання проекцій імпульсу на осі Y і Z (Δp_y і Δp_z) і координати x мікрочастинки.

Співвідношення (18.12) може бути встановлене з розгляду, зокрема, наступного прикладу.

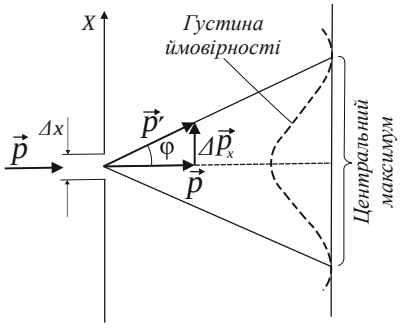


Рис.18.7.

Спробуємо визначити значення координати x мікрочастинки, яка вільно летить, поставивши на її шляху щілину, ширина якої Δx і розташована перпендикулярно до напрямку руху частинки (рис.18.7). До проходження частинки через щілину проекція її імпульсу p_x на вісь X має точне значення, яке дорівнює нулю (щілина за умовою перпендикулярна до імпульсу), так що $\Delta p_x = 0$, зате координата x частинки є зовсім невизначеною. В момент проходження частинкою щілини ситуація змінюється. Замість повної невизначеності координати x виникає невизначеність Δx, але це досягається ціною втрати визначеності значення p_x. Дійсно, внаслідок дифракції є деяка ймовірність того, що частинка буде рухатись в межах кута 2φ, де φ – кут, що відповідає першому дифракційному мінімуму (максимумами вищих порядків можна знехтувати, оскільки їхня інтенсивність мала порівняно з інтенсивністю центрального максимуму). Таким чином, виникає невизначеність

$$\Delta p_x = p \operatorname{tg}\varphi \approx p\varphi.$$

Краю центрального дифракційного максимуму (першому мінімуму), який отримується від щілини шириною Δx, відповідає кут φ, для якого згідно формули (13.16)

$$\sin\varphi \approx \varphi = \lambda/\Delta x.$$

Отже

$$\Delta p_x \sim p\lambda/\Delta x.$$

Звідси з урахуванням (18.5) отримаємо співвідношення

$$\Delta x \cdot \Delta p_x \sim p\lambda = h. \qquad (18.13)$$

Отримання співвідношення (18.12) чи (18.13) не є строгим, і самі вони є приблизними. Найбільш строгі співвідношення невизначеностей отримуються за допомогою методів квантової механіки і записуються для одномірного випадку так:

$$\langle\Delta x^2\rangle \cdot \langle\Delta p_x^2\rangle \geq \frac{\hbar^2}{4}, \qquad (18.14)$$

де $\langle\Delta p_x^2\rangle$ і $\langle\Delta x^2\rangle$ – середні квадратичні відхилення (невизначеності) імпульсу p_x і координати x, ($\hbar = h/2\pi$). (18.14) часто записують в такому виді:

$$\Delta x \cdot \Delta p_x \geq \hbar/2, \qquad (18.15)$$

а для розв'язку задач часто використовують такий запис:

$$\Delta x \cdot \Delta p_x \geq \hbar. \qquad (18.16)$$

Неможливість одночасно точно визначити координату і відповідну складову імпульсу не пов'язана з недосконалістю методів вимірювання та вимірювальних приладів, а є наслідком подвійної корпускулярно-хвильової природи мікрооб'єктів. Співвідношення невизначеностей отримано при одночасному використанні класичних характеристик руху частинки (координати, імпульсу) і наявності у неї хвильових властивостей.

Щоб глибше зрозуміти суть принципу невизначеності, можна здійснити подумки експеримент, який називають "мікроскопом Гайзенберга" (рис.18.8).

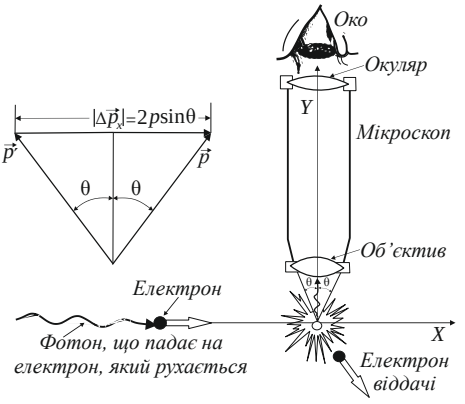


Рис.18.8.

За допомогою уявного мікроскопа з високою роздільною здатністю (рис.18.8) спробуємо визначити одночасно місцерозташування і імпульс електрона. Із геометричної оптики відомо, що даній роздільній здатності мікроскопа місцеположення об'єкта може бути визначене з точністю

$$\Delta x \approx \frac{\lambda}{2\sin\theta}, \qquad (18.17)$$

де θ – кут, який показаний на рис.18.8. Коли відстань між двома точками менша Δx, вони зіллються для спостерігача в одну точку. Отже, величина Δx є мінімальна неточність у визначенні положення електрона, який рухається, координату й імпульс якого намагаємось визначити за допомогою мікроскопа. Із (18.17) витікає, що чим менша довжина хвилі λ світла, яке падає і освітює об'єкт, тим менше Δx і, отже, з більшою точністю може бути зафіксоване положення електрона.

Фотон, що падає, буде взаємодіяти з електроном через механізм ефекта Комптона. Щоб побачити електрон, розсіяний фотон

повинен попасти в мікроскоп в межах кута 2θ. Таким чином, проекція імпульсу фотона на напрямок x (див. рис.18.8) може бути визначена з деякою степенню невизначеності, що дорівнює

$$\Delta p_x \approx 2p \sin\theta, \qquad (18.18)$$

де $p = h/\lambda -$ імпульс падаючого фотона.

У відповідності із законом збереження імпульсу співвідношення (18.18) буде також характеризувати мінімальну неточність у визначенні імпульсу електрона віддачі. Якщо перемножити (18.17) і (18.18) і врахувати, що $p = h/\lambda$, то для електрона віддачі будемо мати:

$$\Delta x \cdot \Delta p_x \geq (2p \sin\theta)\left(\frac{\lambda}{2\sin\theta}\right) = h,$$

тобто отримується співвідношення невизначеностей.

Приклад. Нехай неточність у визначені положення молекули водню, маса якої біля $2 \cdot 10^{-27}$ кг, складає величину порядку її діаметра, тобто десь 10^{-10}м. Тоді неточність у визначенні імпульсу дорівнює:

$$\Delta p_x \geq \frac{\hbar}{\Delta x} =$$
$$= \frac{1{,}054 \cdot 10^{-34}}{10^{-10}}\text{кг} \cdot \text{м/с} \approx 10^{-24}\,\text{кг} \cdot \text{м/с}.$$

Імпульс такої молекули при русі її із швидкістю 2000м/с (швидкість теплового руху при кімнатній температурі) дорівнює:

$$p_x = mv = 2 \cdot 10^{-27} \cdot 2 \cdot 10^3\,\text{кг} \cdot \text{м/с} =$$
$$= 4 \cdot 10^{-24}\,\text{кг} \cdot \text{м/с},$$

а відносна неточність імпульсу:

$$\frac{\Delta p_x}{p_x} = \frac{10^{-24}}{4 \cdot 10^{-24}} = 0{,}25 = 25\%.$$

Таким чином, імпульс цієї молекули не може бути визначений з точністю більшою, ніж 25% від початкового значення. Якщо, однак, розглянути приклад з кулею, маса якої 50 г, швидкість польоту якої дорівнює 1000 м/с і положення відомо з точністю 1 мм, то неточність визначення імпульсу

$$\Delta p_x \geq \frac{1{,}054 \cdot 10^{-34}}{10^{-3}}\text{кг} \cdot \text{м/с} \cong$$
$$\cong 10^{-31}\,\text{кг} \cdot \text{м/с}.$$

Оскільки імпульс

$$p_x = 0{,}05 \cdot 1000\,\text{кг} \cdot \text{м/с} = 50\,\text{кг} \cdot \text{м/с},$$

то відносна неточність імпульсу:

$$\frac{\Delta p_x}{p_x} \cong \frac{10^{-31}}{50} = 2 \cdot 10^{-33}.$$

Це число настільки мале, що ні один із існуючих лабораторних приладів не може його зафіксувати.

Інше співвідношення невизначеностей встановлює зв'язок між енергією та часом і може бути отримане із співвідношення $\Delta\nu \cdot \Delta t \geq 1$, або $\Delta\left(\frac{W}{h}\right)\Delta t \geq 1$. Звідки

$$\Delta W \cdot \Delta t \geq h. \qquad (18.19)$$

Так само як імпульс і координата не можуть бути визначені однозначно з нескінченно високою точністю, так і енергія W і час t (які теж представляють собою пару спряжених змінних) не можуть бути одночасно знайдені з необмеженою точністю. Чим більша точність досягнута нами при вимірюванні часу (тобто чим менше Δt), тим менша точність у визначенні енергії. Співвідношення (18.19) записують в такому виді:

$$\Delta W \cdot \Delta t \geq \hbar. \qquad (18.20)$$

Співвідношення невизначеностей є одним із фундаментальних положень квантової механіки. Вони представляють математичний вираз корпускулярно-хвильового дуалізму частинок, а останнє розглядається як експериментальний факт.

Істинний зміст співвідношень невизначеностей в тому, що в природі об'єктивно не існує станів частинок з точно визначеними значеннями координат і

імпульсів, а значить втрачає зміст поняття траєкторії.

Співвідношення Гайзенберга обмежують застосування понять класичної фізики до об'єктів із мікросвіту. В класичній механіці, щоб повністю визначити рух тіла протягом всього часу, необхідно задати його положення (координату x) і швидкість (v_x) в який-небудь початковий момент часу. Згідно співвідношенню невизначеностей зробити це неможливо. Положення і швидкість (імпульс) не можуть бути визначені одночасно. Доводиться говорити тільки про ймовірність, з якою частинка знаходиться в тому чи іншому місці. Як писав Гайзенберг, «ця невизначеність представляє собою справжню основу для появи статистичних зв'язків у квантовій механіці».

18.5.Статистичне тлумачення хвиль де Бройля

Фізична суть хвиль де Бройля, які зв'язані з рухом частинки, була розкрита не одразу. Спочатку були спроби розглядати самі частинки як утворення з хвиль, що знаходиться в деякій області простору. Інтенсивність хвиль де Бройля розглядалась в цій концепції як величина, що характеризує густину середовища, із якого утворена частинка. Таке розуміння хвиль де Бройля мало зовсім класичний характер. Основою такого розуміння було те, що групова швидкість пакета хвиль співпадала зі швидкістю частинки за законами класичної механіки. Однак з часом група хвиль розпливається, що пов'язано з дисперсією. Окремі хвилі, з яких утворена група, розповсюджується з різними швидкостями, що і є причиною розпливання.

Таким чином, побудована з хвиль де Бройля частинка буде нестійка. Ця нестійкість особливо проявиться, коли частинка рухається в неоднорідному просторі, переходячи з одного середовища в інше. Таким прикладом є класичні досліди по дифракції частинок. Коли пучок електронів проходить через тонку фольгу, то він розділяється на систему конусоподібних дифракційних пучків (рис.18.3). Якщо розглядати електрон як утворення з хвиль, то кожний дифракційний пучок повинен представляти деяку долю електрона. Якщо в один реєстраційний прилад направити тільки перший дифракційний пучок, а в другий прилад – тільки другий дифракційний пучок, то це означає, що кожний із приладів прийме тільки частину електрона. Це і є крайнє порушення атомізму частинок. Дійсно, частинка завжди діє як ціле, і проявляється в приладі вся частинка, а зовсім не її доля. Тому уявлення про частинки, як утворення з хвиль де Бройля протирічить атомізму і повинно бути відкинутим.

Не можна вважати, що самі хвилі є утворенням із частинок або, точніше кажучи, виникають в середовищі, яке утворене частинками. Дослід показує, що дифракційна картина, що виникає на фотопластинці, не залежить від інтенсивності пучка частинок, які падають, а значить, і від густини частинок в одиниці об'єму (§18.3). Щоб отримати одну й ту саму дифракційну картину, можна зменшити інтенсивність, але збільшити експозицію: важливе тільки загальне число частинок, які пройшли. Цей факт однозначно показує, що кожний електрон дифрагує незалежно від інших. Тому існування хвильових явищ не можна зв'язувати з наявністю одночасно великого числа частинок.

Варто підкреслити, що хвильові явища в русі мікрочастинок проявляються і в таких випадках, коли говорити про

середовище, що утворене сукупністю частинок, зовсім не доводиться. Дійсно, такі властивості проявляються в атомах, де число їх зовсім не велике (один електрон у водні, два в гелію і т.д.).

Правильне тлумачення хвиль де Бройля було знайдене М.Борном на зовсім іншому шляху. Щоб зрозуміти основну думку Борна, уявімо собі, що ми досліджуємо дифракцію електронів і реєструємо попадання «дифракційних» електронів на фотопластинці. Кожний із електронів, пройшовши через дифракційний прилад (наприклад, через фольгу), проявиться в якому-небудь місці фотопластинки і викличе там фотохімічну дію. Проходження невеликої кількості електронів дасть на фотопластинці картину, яка подібна на мішень, що простріляну поганим стрілком. Тільки при великій кількості електронів, які пройшли, виявиться регулярність в розподілі електронів на фотопластинці і, нарешті, утвориться розподіл, який відповідає розподілу інтенсивності при дифракції хвиль (§13.4).

Така поведінка частинок привела Борна до статистичного тлумачення хвиль де Бройля, яке дозволило поєднувати атомізм частинок з хвильовими явищами. *Згідно статистичному тлумаченню, інтенсивність хвиль де Бройля в якому-небудь місці простору пропорційна ймовірності виявити частинку в цьому місці.* Так, наприклад, якщо один дифракційний пучок направляється на одну фотопластинку, а другий на іншу, то при великій кількості електронів, що пройшли через апарат, число електронів, які потрапляють на кожну з пластинок, пропорційне інтенсивності хвиль де Бройля, що розповсюджуються в напрямку кожної із фотопластинок.

Якщо фотопластинку розмістити так, що напрямок від дифракційного апарату до пластинки співпадає з напрямком дифракційного мінімуму (в цьому напрямку хвилі гасять одна одну), то частинки зовсім не будуть попадати на таку фотопластинку. Якщо мова йде не про велике число електронів, а про один, то інтенсивність хвиль де Бройля вказує тільки про ймовірність попадання електрона, але зовсім не зобов'язує електрон до тієї чи іншої поведінки.

В такому розумінні, хвилі де Бройля не мають нічого спільного з хвилями, які розглядаються в класичній фізиці. У всіх «класичних» хвилях абсолютне значення амплітуди хвилі визначає фізичний стан. Якщо, наприклад, амплітуда коливань повітря в одному випадку скрізь в два рази більша, ніж в іншому, то це означає вчетверо більшу енергію коливань в першому випадку і, разом з тим, інший фізичний стан середовища.

У випадку хвиль де Бройля їх інтенсивність визначає густину ймовірності місцезнаходження частинки. *Тому важливе тільки відношення інтенсивностей в різних частинах простору, а не сама їхня абсолютна величина.* Це відношення показує, у скільки разів в одному місці простору ймовірніше виявити частинку, ніж в іншому.

Хвилі де Бройля дають, таким чином, статистичний опис руху мікрочастинки. У відповідності з цим ці хвилі необхідно розглядати *як хвилі ймовірності: вони визначають густину ймовірності виявити частинку в даному місці простору в даний момент часу*.

Знаючи хвилю де Бройля, яка описує стан частинки, можна знайти ймовірність не тільки місцезнаходження частинки, але й імовірність будь-якого результату вимірювання

будь-якої механічної величини, яка відноситься до частинки, що вивчається.

§18.6.Ймовірність місцезнаходження мікрочастинки

Позначимо *x, y, z* координати частинки. Згідно викладеного в §18.5, в точний зміст *x, y, z* вкладається наступне: величини *x, y, z* визначаються як координати тієї точки, в якій виявляється частинка. Так, наприклад, це будуть координати плямочки на фотопластинці, яка отримується в результаті попадання на пластинку частинки, або, наприклад, координати, які визначають положення щілини, через яку пройшла частинка, і т.п.

Сформулюємо тепер математично статистичну інтерпретацію хвиль де Бройля. Зауважимо, перш за все, що слово «хвиля» ми вживаємо зараз досить умовно. Тільки в дуже спеціальних випадках стан частинки буде описуватись простими плоскими хвилями. В загальному випадку те, що ми зараз називаємо хвилями де Бройля, може бути досить складною функцією координат частинки *x, y, z* і часу *t*. Тим не менш, і для цих складних випадків використовується термін *хвильова функція* і позначається буквою ψ:

$$\psi = \psi(x, y, z, t). \qquad (18.21)$$

Як було пояснено в §18.5, на основі дослідних *фактів ми приймаємо, що густина ймовірності місцезнаходження частинки визначається інтенсивністю хвиль,* тобто квадратом амплітуди ψ. Враховуючи, що ψ комплексна функція (хвилі ймовірності невидимі), а ймовірність повинна бути завжди дійсною позитивною величиною, будемо брати за міру інтенсивності не ψ^2, а квадрат модуля ψ, тобто величину

$$|\psi|^2 = \psi\psi^*,$$

де через ψ^* позначена величина, яка комплексно спряжена до ψ.

Ймовірність знайти частинку біля точки x, y, z залежить, звичайно, від розмірів області, яку вибираємо. Розглядаючи нескінченно малу область $x, x + dx; y, y + dy; z, z + dz$, можемо вважати ψ всередині неї постійною, а тому ймовірність знайти частинку пропорційна об'єму цієї області. Позначимо цей елемент об'єму через $dV = dxdydz$.

Позначаючи саму ймовірність (нескінченно малу) знайти частинку в елементі об'єму dV біля точки *x,y,z* в момент часу t через $dP(x, y, z, t)$, можемо записати статистичне тлумачення хвиль де Бройля у вигляді наступної рівності:

$$dP(x, y, z, t) = |\psi(x, y, z, t)|^2 dV. \qquad (18.22)$$

Ця рівність дозволяє за відомою хвильовою функцією $\psi(x, y, z, t)$ вирахувати ймовірність місцезнаходження частинки $dP(x, y, z, t)$, в об'ємі dV. Величину

$$\frac{dP}{dV} = |\psi(x, y, z, t)|^2 \qquad (18.23)$$

називають густиною ймовірності.

Ймовірність знайти частинку в момент часу t в об'ємі V, згідно теореми додавання ймовірностей, дорівнює

$$P(V,t) = \int_V dP = \int_V |\psi(x,y,z,t)|^2 dV. \qquad (18.24)$$

Якщо виконати інтегрування по всьому об'єму, то отримаємо ймовірність того, що в момент часу t частинка знаходиться де-небудь. Це – ймовірність вірогідної події. В теорії ймовірностей прийнято ймовірність вірогідної події вважати за 1. Якщо це прийняти, то інтеграл від $|\psi|^2$ по всьому об'єму необхідно прирівняти одиниці:

$$\int_{\infty} |\psi(x,y,z,t)|^2 dV = 1. \quad (18.25)$$

Ця умова називається *нормуванням*, а функція ψ, яка задовольняє цій умові, називається нормованою.

Щоб хвильова функція була об'єктивною характеристикою стану мікрочастинок, вона повинна задовільняти ряд обмежуючих умов:

1.Функція ψ, що характеризує ймовірність виявлення мікрочастинки в елементі об'єму, повинна бути ***скінченною*** (ймовірність не може бути більшою від одиниці), ***однозначною*** (ймовірність не може бути неоднозначною величиною) і ***неперервною*** (ймовірність не може змінюватись стрибком);

2.Похідні $\frac{d\psi}{dx}, \frac{d\psi}{dy}, \frac{d\psi}{dz}, \frac{d\psi}{dt}$ повинні бути безперервні;

3.Функція $|\psi|^2$ повинна бути інтегрованою.

4.Вона не повинна протирічити наступним співвідношенням:

$$\lambda = h/p, \quad \nu = W/h, \quad W=p^2/(2m)+U.$$

5.Якщо $x \to \pm\infty$, $y=\pm\infty$, $z=\pm\infty$, то $\psi(x,y,z,t)$ повинна прямувати до нуля.

Хвильова функція задовільняє принципу ***суперпозиції***: якщо система може знаходитись в різних станах, що описуються хвильовими функціями $\psi_1, \psi_2, \dots, \psi_n$, то вона також може знаходитись в стані ψ, що описується лінійною комбінацією цих функцій:

$$\Psi = \sum_n C_n \psi_n.$$

де C_n – деякі комплексні числа.

За допомогою хвильової функції в квантовій механіці можна обчислити середні значення фізичних величин, які характеризують певний об'єкт, що перебуває в стані, який описується хвильовою функцією ψ.

§18.7.Рівняння Шредінгера. Принцип причинності в квантовій механіці

Статистичне трактування хвиль де Бройля і співвідношення невизначеностей Гайзенберга привели до висновку, що рівняння руху мікрочастинок в різних силових полях, повинно бути рівняння, з якого виходили би хвильові властивості частинок, які спостерігаються під час досліду. Основне рівняння повинно бути диференційним рівнянням відносно хвильової функції $\psi(x,y,z,t)$, оскільки величина $|\psi|^2$ визначає густину ймовірності перебування частинок в момент часу t в об'ємі dV. Оскільки шукане рівняння повинно враховувати хвильові властивості частинок, то воно повинно бути хвильовим рівнянням.

Основне рівняння нерелятивістської квантової механіки, яке сформульоване в 1926р. Е.Шредінгером для одномірного нерелятивістського випадку, має такий вигляд:

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + U(x,t)\psi(x,t) = i\hbar\frac{\partial\psi(x,t)}{\partial t}, \quad (18.26)$$

де m – маса мікрочастинки, $U(x,t)$ – її потенціальна енергія в силовому полі, в якому вона рухається, $i=\sqrt{-1}$ – уявна одиниця. В подальшому будемо розглядати одномірний випадок. Це не зменшує значення тих узагальнень, які витікають із рівняння (18.26).

Не зважаючи на те, що Шредінгер при отриманні свого рівняння користувався класичними уявленнями про «хвилі матерії», це рівняння не можна вивести безпосередньо з фундаментальних законів класичної фізики. Рівняння Шредінгера, подібно, наприклад, до другого закону Ньютона, само по собі є

фундаментальним співвідношенням. Коли потенціальна енергія залежить від часу і координат, то рівняння (18.26) відоме як загальне (або часове) рівняння Шредінгера. Розв'язком цього рівняння і є хвильова функція $\psi(x,t)$, яка служить корисним інструментом, засобом для описання фізичних явищ.

З погляду математики рівняння Шредінгера є лінійним диференційним рівнянням з частинними похідними. З теорії диференційних рівнянь відомо, що кожне лінійне рівняння з частинними похідними має багато розв'язків, причому таких, що лінійна комбінація будь-якої сукупності розв'язків рівняння також буде розв'язком рівняння.

Для значної кількості фізичних явищ, які відбуваються в мікросвіті, наприклад, для опису поведінки електрона в атомі у ряді випадків важливо вміти знаходити стаціонарні розв'язки рівняння Шредінгера, які не містять часу.

Щоб розв'язати цю задачу, необхідно знайти так зване стаціонарне рівняння Шредінгера, в якому виключено залежність ψ від часу. Воно має сенс для тих задач, в яких потенційна енергія U не залежить від часу: $U = U(x)$. Шукатимемо розв'язок рівняння Шредінгера у вигляді добутку:

$$\psi(x,t) = \psi(x)e^{-i\frac{W}{\hbar}t}, \quad (18.27)$$

де $\psi(x)$ є функцією лише координат, а W – повна енергія частинки. Це означає, що стан з певним значенням енергії W (невизначеність $\langle \Delta W^2 \rangle = 0$) гармонійно залежить від часу з частотою ω=$W/\hbar$.

Підставимо вираз для $\psi(x,t)$ (18.27) у рівняння Шредінгера (18.26):

$$-\frac{\hbar^2}{2m}\ddot{\psi}e^{-i\frac{W}{\hbar}t} + U(x)\psi e^{-i\frac{W}{\hbar}t} = \\ = i^2\hbar\psi\cdot(-\frac{W}{\hbar})e^{-i\frac{W}{\hbar}t},$$

де $\ddot{\psi} = \frac{\partial^2\psi}{\partial x^2}$.

Скоротивши на множник $e^{-i\frac{W}{\hbar}t}$, дістанемо

$$\ddot{\psi}(x) + \frac{2m}{\hbar^2}(W - U(x))\psi(x) = 0. \quad (18.28)$$

Це і є ***стаціонарне рівняння Шредінгера.*** Рівняння (18.28) найважливіше співвідношення нерелятивістської квантової механіки. Функції ψ, які задовольняють рівнянню Шредінгера при певному W, називають ***власними функціями.*** В рівнянні Шредінгера як параметр входить повна енергія частинки. В теорії диференційних рівнянь доводиться, що подібні рівняння мають розв'язок не при довільних значеннях параметра, а лише при певних значеннях W. Ці значення ***енергії*** називають ***власними.***

Власні значення W можуть утворювати як безперервний, так і дискретний ряд.

Рівняння Шредінгера дає змогу знайти не тільки конкретний вигляд функції $\psi(x,t)$ у заданому зовнішньому полі $U(x,t)$, а й визначити її зміну з часом $\frac{\partial\psi}{\partial t}$. Отже, рівняння Шредінгера виступає як свого роду «причина» того, який вигляд має ψ – функція в тому чи іншому випадку і як вона змінюється з часом. Знання ж ψ – функції дає змогу знайти всі можливі значення важливих параметрів фізичної системи у будь-який момент часу.

Отже, *рівняння Шредінгера є математичним виразом принципу причинності в квантовій механіці – початковий стан системи (при відомих взаємодіях в системі) повністю визначає всю наступну її еволюцію.*

Але на відміну від класичної фізики, квантова механіка не дає чіткої відповіді на запитання, які є точні значення параметрів фізичної системи у даний момент часу. Ψ –

функція, вказуючи цілий спектр можливих значень параметрів системи, дає змогу обчислити лише ймовірність появи кожного значення під час вимірювання.

§18.8.Рух вільної частинки. Частинка в прямокутній потенційній ямі

Розглянемо декілька порівняно простих прикладів застосування рівняння Шредінгера до руху частинки в конкретних умовах.

1. Рух вільної частинки

Під час руху вільної частинки $(U(x) = 0)$ повна енергія збігається з кінетичною $W = W_к$. Для вільної частинки, що рухається вздовж осі *OX*, $(-\infty < x < +\infty)$, рівняння Шредінгера для стаціонарних станів набирає вигляду:

$$\ddot{\Psi} + \frac{2m}{\hbar^2} W\Psi = 0. \qquad (18.29)$$

Прямою підстановкою можна переконатися в тому, що частинним розв'язком рівняння є функція

$$\Psi(x) = Ae^{ikx} + Be^{-ikx},$$

де $k = 2\pi/\lambda$ – хвильове число, довжина хвилі де Бройля $\lambda = h/p = h/\sqrt{2mW}$, A і B – сталі.

Розв'язок рівняння (18.29) можна переписати ще й так:

$$\Psi(x) = Ae^{\frac{i\sqrt{2mW}x}{\hbar}} + Be^{-\frac{i\sqrt{2mW}x}{\hbar}}.$$

Функція $\Psi(x)$ є тільки координатною частиною хвильової функції $\Psi(x,t)$ стаціонарного стану. Розв'язок повного рівняння Шредінгера отримаємо у формі (18.27):

$$\Psi(x,t) = Ae^{-\frac{i(Wt-\sqrt{2mW}x)}{\hbar}} + Be^{-\frac{i(Wt+\sqrt{2mW}x)}{\hbar}}. \qquad (18.30)$$

Цей розв'язок є суперпозицією двох плоских монохроматичних хвиль однакової частоти $\omega = W/\hbar$. Хвиля з амплітудою A поширюється в додатному напрямі осі *OX,* хвиля з амплітудою B поширюється у протилежному напрямку. Знайдений розв'язок узгоджується із загальним виразом плоскої монохроматичної хвилі $S(x,t) = Ae^{-i(\omega t - kx)}$. Хвильове число k для вільної частинки дорівнює $k = \frac{1}{\hbar}\sqrt{2mW}$, а $\omega = W/\hbar$.

Отже, вільній частинці в квантовій механіці відповідає плоска монохроматична хвиля де Бройля. Вона характеризує ймовірність знаходження частинки в певній точці простору.

Дійсно, взявши лише одну з хвиль, що поширюється в додатньому напрямку осі *OX*, отримаємо

$$|\psi|^2 = \psi\psi^* = |A|^2.$$

Власні значення енергії частинки

$$W = \frac{p_x^2}{2m} = \frac{\hbar^2}{2m\lambda^2} = \frac{\hbar^2 k^2}{2m}. \qquad (18.31)$$

Енергія вільної частинки може набувати довільне значення, тобто її енергетичний спектр неперервний.

2.Частинка в одномірній прямокутній потенційній ямі

Нехай частинка може рухатись лише вздовж осі *OX* і знаходиться всередині прямокутної потенційної ями, яка обмежена нескінченно високими бар'єрами (рис.18.9). В цьому випадку потенційна енергія частинки *U(x)* набуватиме такі значення :

$$U(x) = \begin{cases} \infty, & x < 0; \\ 0, & 0 \le x \le l \\ \infty, & x > l. \end{cases};,$$

де l – ширина ями, а енергія відраховується від дна ями. Рівняння Шредінгера (18.26) у випадку одномірної ями запишемо у вигляді

$$\ddot{\Psi} + \frac{2m}{\hbar^2}(W - U)\Psi = 0. \quad (18.32)$$

За умовою задачі (нескінченно високі «стінки») частинка не проникає за границі ями,

тому ймовірність її виявлення за границями ями дорівнює нулю. На границях ями (при x=0 і x=l) неперервна хвильова функція повинна

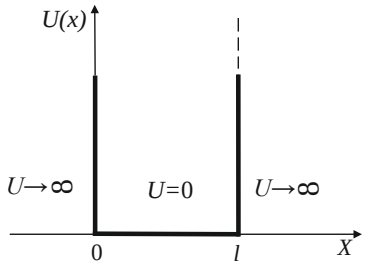


Рис.18.9.

перетворюватися в нуль. Отже, граничні умови мають вигляд:

$$\psi(0) = \psi(l) = 0. \qquad (18.32)$$

В межах ями ($0 \le x \le l$) рівняння Шредінгера матиме вигляд:

$$\ddot{\Psi} + k^2\Psi = 0, \qquad (18.33)$$

де $k^2 = \frac{2m}{\hbar^2}W$.

Загальний розв'язок диференційного рівняння (18.33) будем шукати у вигляді :

$$\Psi(x) = Ae^{ikx} + Be^{-ikx}. \qquad (18.34)$$

Функція (18.34) повинна задовольняти граничним умовами (18.32). Тому отримаємо систему рівнянь:

$$\begin{matrix} A+B = 0, \\ Ae^{ikl}+Be^{-ikl}=0. \end{matrix} \qquad (18.35)$$

Це однорідна система двох рівнянь буде мати нетривіальний розв'язок відносно A і B тоді, коли дискримінант системи дорівнює нулю, тобто

$$\begin{vmatrix} 1 & 1 \\ e^{ikl} & e^{-ikl} \end{vmatrix} = 0 ,$$

або $e^{ikl} - e^{-ikl} = 0$. Врахувавши формулу Ейлера ($e^{i\varphi} = \cos\varphi + i\sin\varphi$), отримаємо:

$$\cos kl + i\sin kl - \\ -\cos(-kl) - i\sin(-kl) = 0.$$

Звідки $\sin kl = 0$ виконується лише при $kl = \pi n$, де n – цілі числа (випадок n=0 не розглядається, обгрунтування цього далі). Тобто необхідно, щоб

$$k = \frac{\pi}{l}\cdot n. \qquad (18.36)$$

Тоді $\frac{\pi^2 n^2}{l^2} = \frac{2m}{\hbar^2}W_n$. Звідки власні значення енергії

$$W_n = \frac{\pi^2\hbar^2}{2ml^2}n^2 = \frac{h^2}{8ml^2}n^2. \qquad (18.37)$$

Рівняння Шредінгера задовольняється лише при значеннях W_n, що залежать від цілого числа n.

Отже, енергія W_n частинки в потенційній ямі з нескінченно високими стінками не може бути довільною, а набуває лише певні дискретні значення, тобто квантується.

Квантові значення енергії W_n називають **рівнями енергії**, а число n, яке визначає енергетичні рівні частинки, називають квантовим числом.

Умова (18.36) має простий фізичний зміст. Оскільки $k_n = \frac{2\pi}{\lambda_n}$, де λ_n – довжина хвилі де Бройля для електрона в ямі, то $\frac{2\pi}{\lambda_n} = \frac{n\pi}{l}$, або $\lambda_n = \frac{2l}{n}$ (l=(λ_n/2)n), тобто на ширині ями повинно вкладатись ціле число півхвиль де Бройля. Знайдемо власні хвильові функції. Із (18.35) видно, що A=−B, тому (18.34) перепишеться так:

$$\Psi(x) = A(e^{ikx} - e^{-ikx}) = 2iA\sin kx.$$

\Сталу A знайдемо з умови (18.25):

$$\int_0^l (2iA\sin kx)(-2iA^*\sin kx)\,dx = 1,$$

або $4|A|^2\int_0^l \sin^2 kx dx = 1$. Врахувавши, що

$$\int_0^l \sin^2 kx dx = \\ = \int_0^l \left(\frac{1}{2} - \frac{1}{2}\cos 2kx\right)dx = \frac{l}{2},$$

отримаємо, що $A = \sqrt{\frac{1}{2l}}$. Тоді власні хвильові функції мають вигляд:

$$\psi_n(x) = i\sqrt{\frac{2}{l}}\sin\frac{n\pi}{l}x, (n = 1, 2, 3 \ldots). \quad (18.39)$$

На рис.18.10,*а* наведені графіки функції $\psi_n(x)$ при $n = 1, 2, 3$, а на рис.18.10,*б*–густини ймовірності знаходження частинки $|\psi_n(x)|^2$ на різних відстанях від "стінок" ями для $n = 1, 2, 3$.

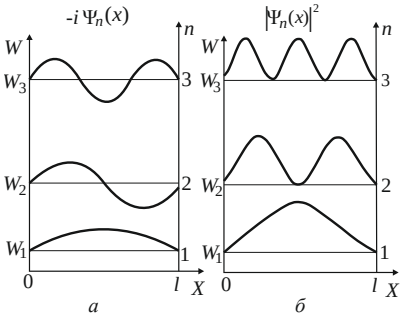


Рис.18.10.

Наприклад, у квантовому стані з n = 2 частинка не може знаходитись посередині ями і в той же час однаково часто може перебувати в її лівій або правій частині. Стан з n = 0 не існує, так як $\psi_n(x)$=0 при n = 0.

Квантовий стан з найменшим значенням енергії (n = 1) називається основним станом. Йому відповідає енергія

$$W_1 = \frac{h^2}{8ml^2}.$$

Оцінимо різницю енергій двох сусідніх рівнів за формулою (18.37):

$$\Delta W_n = \Delta W_{n+1} - W_n = \frac{h^2}{8ml^2}(n+1)^2 - \frac{h^2}{8ml^2}n^2 = \frac{h^2}{8ml^2}(2n+1).$$

Для електрона при розмірах ями $l = 10^{-1}$м (вільні електрони в металі) і $m = m_e = 9{,}1 \cdot 10^{-31}$кг

$$\Delta W_n = (2n+1)\cdot 6\cdot 10^{-36}\text{Дж} = (2n+1)3{,}8\cdot 10^{-17}\text{еВ}.$$

Енергетичні рівні розміщені так тісно, що спектр практично можна вважати неперервним. Якщо область, в якій рухається електрон, порядку атомних розмірів ($l = 10^{-10}$м), то

$$\Delta W_n = (2n+1)\cdot 6\cdot 10^{-18}\text{Дж} = (2n+1)\cdot 3{,}8\cdot 10\text{еВ}.$$

Отже, дискретність енергетичних рівнів в цьому випадку буде досить помітною і отримуємо лінійчастий спектр.

При великих квантових числах ($n \gg 1$)

$$\frac{\Delta W_n}{W_n} = \frac{2n+1}{n^2} \approx \frac{2}{n} \ll 1,$$

тобто сусідні рівні розміщені тим густіше, чим більше n. Якщо n дуже велике, то можна говорити про практично неперервну послідовність рівнів і характерна особливість квантових процесів – дискретність – згладжується.

§18.9.Розв'язок рівняння Шредінгера для потенційного бар'єру

1. Постановка проблеми

Якщо є дві області простору, в яких потенційна енергія частинки менша, ніж на поверхні, яка розділяє ці області, то говорять , що області розділені *потенційним бар'єром.*

Найпростішим прикладом потенційного бар'єру може бути бар'єр в одному вимірюванні, що зображений на рис.18.11. По осі ординат відкладена потенційна енергія *U*(*x*), як функція координати *x*. В точці x_0 потенційна енергія має максимум U_0. Весь простір $-\infty < x < +\infty$ ділиться в цій точці на дві області: $x < x_0$ і $x > x_0$, в яких $U < U_0$. Значення терміну

"потенційний бар'єр" зараз з'ясується, якщо розглянути рух частинки в полі $U(x)$ на основі класичної механіки. Повна енергія частинки W дорівнює

$$W = \frac{p^2}{2m} + U(x), \qquad (18.39)$$

де p – імпульс частинки, а m – її маса. Роз'язуючи (18.39) відносно імпульсу, отримаємо:

$$p(x) = \pm\sqrt{2m(W - U(x))}\,. \quad (18.40)$$

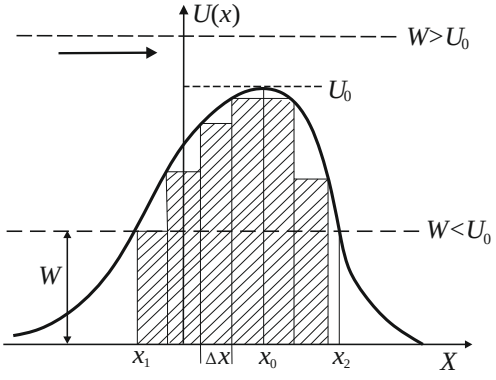


Рис.18.11.

Знаки $\pm$ необхідно вибирати в залежності від напрямку руху частинки. Якщо енергія частинки W більша "висоти" бар'єру U_0, то вона безперешкодно пройде бар'єр зліва направо, якщо початковий імпульс $p > 0$, або в протилежному напрямку, якщо початковий імпульс $p < 0$.

Нехай частинка рухається зліва на право і енергія її W менша U_0. Тоді в деякій точці x_1 потенційна енергія $U(x_1) = W$, $p(x_1) = 0$, частинка зупиниться. Вся її енергія перетвориться в потенційну, і рух почнеться в оберненому напрямку: x_1 є *точка повороту*. При $x > x_1$ імпульс p стає уявною величиною. Тому при $W < U_0$ частинка, яка рухається зліва, не пройде через область максимуму потенціалу $(x = x_0)$ і не проникне в іншу область, де $x > x_0$. Так само, якщо частинка рухається справа наліво з енергією $W < U_0$, то вона не проникне в область за другою точкою повороту x_2, в якій $U(x_2) = W$ (див.рис.18.11). Таким чином, потенційний бар'єр є "непрозорою" перегородкою для всіх частинок, енергія яких менша U_0 (навпаки, він "прозорий" для частинок з енергією $W > U_0$). Цим і пояснюється назва "потенційний бар'єр".

Зовсім по іншому протікають явища біля потенційних бар'єрів, якщо мова йде про рух мікроскопічних частинок в мікроскопічних полях, тобто про рухи, при розгляді яких не можна ігнорувати квантові ефекти (хвильові властивості частинок). В цьому випадку, як побачимо в подальшому, на противагу висновкам класичної механіки, частинки з енергією W, яка більша висоти бар'єру U_0, частково відбивається від бар'єру, а частинки з енергією, яка менша U_0, частково проникають через бар'єр.

2. Розв'язок рівняння Шредінгера для потенційного бар'єру нескінченої ширини

Розглянемо поведінку частинки на границі двох областей I і II, в кожній із яких потенційна енергія частинки стала, але ці потенційні енергії відрізняються на кінцеву величину. Будемо вважати, що на границі областей I і II потенціальна енергія змінюється скачком, як це показано на рис.18.12 і на рис.18.13. Виберемо систему координат так, щоб вісь x була паралельна напрямку руху частинки. При цьому потенційна енергія задається умовами:

область I: при $x < 0 \quad U = 0$,

область II: при $x \geq 0 \;\; U = U_0 = \text{const}$.

Запишемо рівняння Шредінгера (18.28) для області I:

$$\ddot{\psi}_1 + \frac{2m}{\hbar^2} W\psi_1 = 0, \qquad (18.41)$$

для області II:

$$\ddot{\psi}_2 + \frac{2m}{\hbar^2}(W - U_0)\psi_2 = 0. \qquad (18.42)$$

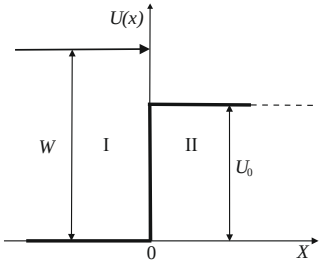


Рис.18.12.

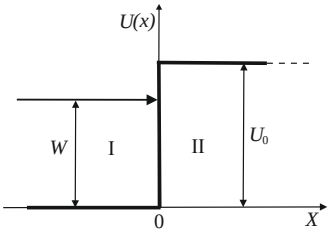


Рис.18.13.

Введемо позначення

$$k_1 = \frac{1}{\hbar}\sqrt{2mW} = \frac{mv_1}{\hbar} = \frac{2\pi}{\lambda_1}, \qquad (18.43)$$

$$k_2 = \frac{1}{\hbar}\sqrt{2m(W - U_0)} = \frac{mv_2}{\hbar} = \frac{2\pi}{\lambda_2}, \qquad (18.44)$$

де λ_1 і λ_2 – довжини хвиль де Бройля в області I і в області II відповідно. Рівняння (18.41) і (18.42) тепер перепишуться у вигляді

$$\ddot{\psi}_1 + k_1^2\psi_1 = 0, \qquad (18.41')$$

$$\ddot{\psi}_2 + k_2^2\psi_2 = 0. \qquad (18.42')$$

Загальні розв'язки цих рівнянь такі:

$$\psi_1 = a_1 e^{ik_1 x} + b_1 e^{-ik_1 x}, \qquad (18.45)$$

$$\psi_2 = a_2 e^{ik_2 x} + b_2 e^{-ik_2 x}. \qquad (18.46)$$

ψ_1 і ψ_2 представлені як суперпозиція плоских хвиль де Бройля, причому $e^{ik_1 x}$ і $e^{ik_2 x}$ характеризують хвилі, які розповсюджуються зліва направо, а $e^{-ik_1 x}$ і $e^{-ik_2 x}$ – хвилі, що розповсюджуються справа наліво.

Розглянемо тепер умови переходу частинки із області I в область II в двох випадках: *а*) коли повна енергія частинки W більше її потенційної енергії U_0 в області II (рис.18.12) і *б*) коли $W < U_0$ (рис.18.13).

а) При $W > U_0$ частинка, яка підпорядковується класичній механіці, обов'язково перейде із області I в область II. Дійсно, якщо ця частинка несе електричний заряд і рухається в області I зліва направо, то на границі I і II вона потрапляє в поле потенціалу *U,* яке її затримує. Але оскільки її повна енергія $W > U_0$, то частинка подолає це затримуюче поле і продовжить свій рух в області II з меншою енергією $W - U_0$.

Частинка, рух якої підпорядковується рівнянню Шредінгера, наприклад електрон, повинен поводити себе зовсім інакше. Це видно навіть із таких якісних міркувань. Рух електрона супроводжується плоскою хвилею де Бройля. На границі двох областей, де відбувається раптова зміна потенціалу, ця хвиля повинна поводити себе так, як веде себе світлова хвиля на границі двох областей з різними показниками заломлення. Це означає, що на границі областей I і II хвиля де Бройля частково відбивається, а частково переходить в область II. Ми можемо також сказати, що, переходячи із однієї області в іншу, електрон має ймовірність відбитись і певну ймовірність пройти далі в область II.

Наша задача тепер знайти ці ймовірності. Будемо вважати, що у (18.45) $|a_1|^2$ інтенсивність хвилі, яка падає, а $|b_1|^2$ – інтенсивність відбитої хвилі. В області II розповсюджується тільки хвиля, що проходить. Тому в (18.46) приймемо $b_2 = 0$. Ймовірність проходження квантової частинки в бар'єр і ймовірність відбиття від бар'єру залежить, таким чином, від коефіцієнтів a_1, b_1, a_2. Ці

коефіцієнти можна визначити, врахувавши так звані граничні умови, тобто неперервність хвильової функції і її першої похідної (яка визначає гладкість функції) при переході із області I в область II :

$$\begin{aligned}(\Psi_1)|_{x=0}&=(\Psi_2)|_{x=0},\\ (\frac{d\Psi_1}{dx})|_{x=0}&=(\frac{d\Psi_2}{dx})|_{x=0}.\end{aligned} \qquad (18.47)$$

Із умов (18.47) отримаємо такі співвідношення:

$$a_1 + b_1 = a_2\ ,$$
$$ia_1k_1 - ib_1k_1 = ia_2k_2$$

або

$$\begin{cases} a_1 + b_1 = a_2 \\ a_1 - b_1 = \frac{k_2}{k_1}a_2. \end{cases} \qquad (18.48)$$

Із системи рівнянь (18.48) отримаємо, що

$$b_1 = \frac{k_1-k_2}{k_1+k_2}a_1, \qquad (18.49)$$

а

$$a_2 = \frac{2k_1}{k_1+k_2}a_1\ . \qquad (18.50)$$

Тепер ми можемо по аналогії з оптикою визначити коефіцієнт відбивання R і коефіцієнт прозорості D. Коефіцієнт відбивання дорівнює відношенню квадратів амплітуд відбитої і падаючої хвиль:

$$R = \frac{|b_1|^2}{|a_1|^2} = \left|\frac{k_1-k_2}{k_1+k_2}\right|^2. \qquad (18.51)$$

При розрахунку коефіцієнта прозорості D необхідно помножити відношення квадратів амплітуд хвилі, що проходить, та хвилі, що падає, на відношення швидкостей частинок в одній та іншій областях з такої причини. Коефіцієнт прозорості дорівнює відношенню потоку, що проходить, до потоку, що падає. Потік частинок дорівнює ρv, де ρ – густина частинок пропорційна квадрату амплітуди хвилі де Бройля, а швидкість $v \sim p \sim \frac{1}{\lambda}$, де λ довжина хвилі де Бройля. Тому коефіцієнт прозорості

$$D = \frac{\rho_2 v_2}{\rho_1 v_1} = \frac{|a_2|^2\lambda_1}{|a_1|^2\lambda_2} = \frac{|a_2|^2 k_2}{|a_1|^2 k_1}.$$

Врахувавши (18.50), отримаємо

$$D = \frac{4k_1k_2}{(k_1+k_2)^2}. \qquad (18.52)$$

Коефіцієнти R і D можемо тлумачити з корпускулярної точки зору так: R є ймовірність частинці зазнати відбивання на границі областей, а D – ймовірність пройти в область II або, як прийнято говорити, ймовірність подолати потенційний бар'єр.

Складаючи (18.51) і (18.52), отримаємо $R + D = 1$, що витікає з теореми додавання ймовірностей, оскільки можна однозначно стверджувати, що на границі областей частинка або відіб'ється, або пройде далі.

Вирахуємо тепер коефіцієнти відбивання й пропускання в залежності від співвідношення між повною енергією W і потенційною енергією U_0.

Підставимо у формули (18.51) і (18.52) значення k_1 (формула (18.43)) і k_2 (формула (18.44)). Тоді отримаємо, що

$$R = \left(\frac{1-\sqrt{1-U_0/W}}{1+\sqrt{1+U_0/W}}\right)^2 \qquad (18.53)$$

і

$$D = \frac{4\sqrt{1-U_0/W}}{\left(1+\sqrt{(1-U_0/W}\right)^2}. \qquad (18.54)$$

В таблиці 18.1 приведені розрахунки декількох числових значень R і D. Із цієї таблиці видно, що коли енергія частинки вдвічі більша висоти "потенційної сходинки" U_0, ймовірність відбивання вже має досить значну величину (близько 3%), а при $U_0 = W$ проникнення в область II взагалі неможливе, тоді як за класичною механікою в цьому випадку, як і в попередніх, частинка достовірно переходить в область II, тільки її кінетична енергія там дорівнює нулю.

Таблиця 18.1.

U_0/W	R	D	W/U_0
0,1	0,007	0,9993	10
0,5	0,0296	0,9704	2
0,8	0,1459	0,8541	1,25
0,9	0,2700	0,2700	1,11
1,0	1,000	0,000	1,0

б) При $W<U_0$ за класичною механікою перехід із області I в область II не можливий, оскільки за цієї умови (потенційна енергія більша повної енергії) кінетична енергія частинки в області II повинна стати від'ємною, а швидкість уявною.

Вирахуємо коефіцієнт відбивання R для цього випадку, скориставшись квантовою механікою. Перш за все необхідно звернути увагу на те, що при $W < U_0$ величина k_2 (формула(18.44)) стає уявною:

$$k_2 = \frac{1}{\hbar}\sqrt{2m(W-U_0)}= $$
$$= i\frac{1}{\hbar}\sqrt{2m(U_0-W)} = ik,$$

де $k = \frac{1}{\hbar}\sqrt{2m(U_0-W)}$. Тоді згідно (18.51)

$$R=\left|\frac{k_1-ik}{k_1+ik}\right|^2 = \left(\frac{k_1-ik}{k_1+ik}\right)\left(\frac{k_1+ik}{k_1-ik}\right) = 1,$$
$$D = 1 - R = 0.$$

Отже, при $W<U_0$ коефіцієнт відбивання дорівнює 1, тобто відбивання є повним. Такого результату й очікували. Однак неочікуваним є те, що, як ми зараз покажемо, хоч відбивання і повне, – є певна ймовірність знайти частинку в області II. Іншими словами, відбивання не обов'язково відбувається на самій границі областей, а деякі частинки заходять в область II для того, щоб потім повернутись в область I. Дійсно, оскільки при $W< U_0$ коефіцієнт k_2 стає суто уявним, то розв'язок рівняння Шредінгера (18.46) за умови $b_2 = 0$ набуває вигляду $\psi_2 = a_2 e^{ik_2 x} = a_2 e^{-kx}$, а ймовірність знайти частинку на одиничній довжині буде:

$$\psi_2\psi_2^* = |\psi|^2 = |a_2|^2 e^{-2kx} = $$
$$= |a_2|^2 e^{-\frac{2}{\hbar}\sqrt{2m(U_0-W)}\cdot x}. \qquad (18.55)$$

Це означає, що є певна ймовірність знайти частинку в області II. Правда, ця ймовірність експоненційно зменшується зі збільшенням x, але вона відмінна від нуля: мікроскопічні частинки можуть проникати в області, що "заборонені" для частинок макроскопічних. Тлумачення отриманих результатів з хвильової точки зору не має труднощів, оскільки випадок $W < U_0$ аналогічний випадку повного внутрішнього відбивання в оптиці. З точки зору геометричної оптики при переході світла з більш оптично густого середовища в менш оптично густе, коли кут падіння більший критичного, світло не може проникати в менш оптично густе середовище. Однак хвильова оптика показує і дослід підтверджує, що дійсно в цьому менш оптично густому середовищі, навіть і при кутах падіння, більших від критичного, є хвильове поле з експоненційно спадною амплітудою. Розрахунки показують, що, не зважаючи на проникнення світла в друге середовище, середнє значення нормальної складової вектора Умова–Пойтинга (потік енергії) до поверхні розділу за достатньо великий проміжок часу дорівнює нулю. Це означає, що відсутній рух енергії з першого середовища в друге. Детальний аналіз цього випадку, виконаний А.А Ейхенвальдом, показав, що при повному внутрішньому відбиванні лінії вектора Умова-Пойтинга виявляються кривими; вони заходять в друге середовище і знову виходять в перше; відбивання продовжує бути повним, не зважаючи на встановлення поля в другому середовищі. У відповідності з цим, і в нашому

випадку коефіцієнт відбивання R дорівнює 1, а коефіцієнт пропускання дорівнює нулю: частинки "заходять"в область II, проникають на деяку глибину і знову повертаються в область I.

У квантовій механіці, однак, не має ніякого парадоксу в тому, що частинку можна виявити в області, де $W<U_0$. По-перше, необхідно завжди пам'ятати, що для мікроскопічних частинок, в силу співвідношення невизначеностей, одночасні точні значення x і p_x неможливі. Тому не має сенсу говорити про одночасні, конкретні значення потенційної (функція x) і кінетичної (функція p) енергії.

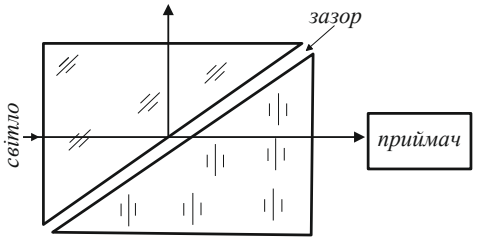


Рис.18.14.

Явище заходження електромагнітної (світлової) енергії в друге, оптично менш густе середовище, спостерігав експериментально на досліді Квінке. На рис.18.14 показана схема досліду. Дві призми повного внутрішнього відбиття розташовані так, що між ними утворюється зазор дуже малої товщини (порядку довжини падаючої хвилі) рис.18.14. При великій товщині зазору світло не проникає через повітряний прошарок і приймач не реєструє енергію. Однак, якщо товщина зазору менша глибини проникнення хвилі при повному внутрішньому відбитті, то енергія, яка проникає, проходячи через другу призму, попадає в приймач. Змінюючи товщину зазору, можна регулювати кількість енергії, яка пройшла. На цьому принципі побудований один із світлових модуляторів. Товщину зазору змінюють під дією звукових хвиль (мовлення). Приймаючи модульоване світло на фотоелемент, можна отримати змінний струм і відтворити таким чином звук (світловий телефон).

Цікаво зазначити, що в даному досліді товщина зазору залежить від довжини хвилі. Так, наприклад, в УКХ – діапазоні вона в 10^5 раз більша, ніж в оптичному діапазоні.

3. Розв'язок рівняння Шредінгера для потенційного бар'єру кінцевої ширини. Тунельний ефект

Нехай частинка, що рухається в додатньому напрямку осі OX зустрічає на своєму шляху прямокутний потенційний бар'єр, висота якого U_0 і ширина l (рис.18.15). Отже, величина U вздовж осі OX змінюється так:

$$U(x)=\begin{cases}0, & x<0;\\ U_0, & 0\le x\le l;\\ 0, & x>l.\end{cases}$$

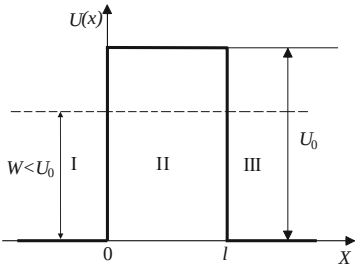


Рис.18.15.

За даних умов задачі класична частинка, яка має енергію W, або пройде над бар'єром при $W> U_0$, або відіб'ється від нього при $W<U_0$ і буде рухатись у зворотній бік.

Для квантової частинки навіть при $W> U_0$ є відмінна від нуля ймовірність того, що частинка відіб'ється від бар'єру і буде рухатись

у зворотній бік. При $W<U_0$ є також відмінна від нуля ймовірність того, що частинка проникає через бар'єр і виявиться в області з $x>l$.

Рівняння Шредінгера (18.28) для кожної з виділених областей має вигляд:
для областей I і III

$$\ddot{\psi} + \frac{2m}{\hbar^2} W\psi = 0, \qquad (18.56)$$

для області II

$$\ddot{\psi} + \frac{2m}{\hbar^2}(W - U_0)\psi = 0. \quad (18.57)$$

Загальні розв'язки цих диференційних рівнянь будемо шукати в такому вигляді:

$$\begin{cases} \text{I: } \psi_1 = a_1 e^{ik_1 x} + b_1 e^{-ik_1 x}, \\ \text{II: } \psi_2 = a_2 e^{ik_2 x} + b_2 e^{-ik_2 x}, \\ \text{III: } \psi_3 = a_3 e^{ik_1 x} + b_3 e^{-ik_1 x}, \end{cases} \quad (18.58)$$

де

$$k_1 = \frac{1}{\hbar}\sqrt{2mW}, \qquad (18.59)$$

$$k_2 = \frac{1}{\hbar}\sqrt{2m(W - U_0)}. \qquad (18.60)$$

В області III є лише хвиля, що пройшла через бар'єр і поширюється зліва направо. Тому коефіцієнт $b_3 = 0$. При $W<U_0$ в області II k_2 – уявне число, тобто

$$k_2 = ik = i\frac{1}{\hbar}\sqrt{2m(U_0 - W)}. \quad (18.61)$$

У випадку широкого бар'єру, незалежно від енергії частинки, як ми бачили, є певна ймовірність проникнути під бар'єр. У випадку, який розглядається, потенціальний бар'єр має кінцеву ширину, тому є ймовірність не тільки проникнення, але і проходження частинки через бар'єр в область III.

Цей випадок відрізняється від попереднього тим, що відбивання частинок має місце як на границі областей I і II, так і на границі областей II і III. Для розрахунку коефіцієнтів R і D необхідно знайти коефіцієнти b_1, b_2, a_2, a_3 (будемо вважати, що a_1=1). Для цього скористаємось умовами неперервності функції ψ і її першої похідної на границях областей I і II, II і III, тобто при x=0 і $x = l$.

Випишемо ці умови:

$$\begin{aligned} (\psi_1)_{|\,x=0} &= (\psi_2)_{|\,x=0}, \\ \left(\frac{d\psi_1}{dx}\right)_{|\,x=0} &= \left(\frac{d\psi_2}{dx}\right)_{|\,x=0}, \end{aligned} \quad (18.62)$$

$$\begin{aligned} (\psi_2)_{|\,x=l} &= (\psi_3)_{|\,x=l}, \\ \left(\frac{d\psi_2}{dx}\right)_{|\,x=l} &= \left(\frac{d\psi_3}{dx}\right)_{|\,x=l}. \end{aligned} \quad (18.63)$$

Із врахуванням виду хвильових функцій (18.58) ці умови дають таку систему рівнянь (враховано, що $a_1 = 1$, $b_3 = 0$).

$$\begin{cases} 1 + b_1 = a_2 + b_2, \\ k_1 - k_1 b_1 = k_2 a_2 - k_2 b_2. \end{cases} \quad (18.64)$$

$$\begin{cases} a_2 e^{ik_2 l} + b_2 e^{-ik_2 l} = a_3 e^{ik_1 l}, \\ a_2 e^{ik_2 l} - b_2 e^{-ik_2 l} = a_3 \frac{k_1}{k_2} e^{ik_1 l}. \end{cases} \quad (18.64)$$

Із цієї системи рівнянь можна визначити коефіцієнти b_1, b_2, a_2, a_3. Однак коефіцієнти b_1, b_2, a_2, а також коефіцієнт відбивання R визначати не будемо, оскільки це нічого принципово нового порівняно з попереднім розглядом не дає. Для коефіцієнта a_3 отримується наступний вираз:

$$a_3 = \frac{i4k_1 k_2 e^{ik_1 l}}{(k_1+k_2)^2 e^{-ik_2 l} - (k_1-k_2)^2 e^{ik_2 l}} \cdot \quad (18.65)$$

Коефіцієнт прозорості бар'єру D буде:

$$D = \frac{|a_3|^2}{|a_1|^2} = |a_3|^2 = a_3 a_3^* . \quad (18.66)$$

Із врахуванням (18.61) і (18.65), можна отримати такий вираз для D

$$D = \frac{4k_1^2 k^2}{(k_1^2+k^2)^2 \mathrm{sh}^2 kl + 4k_1^2 k^2}, \quad (18.67)$$

де гіперболічна функція

$$\mathrm{sh}kl = \frac{e^{kl} - e^{-kl}}{2} \cdot$$

В більшості випадків, які розглядаються, можна вважати, що $\mathrm{sh}^2 kl \cong \frac{1}{4} e^{2kl}$. Тоді

$$D = \frac{4}{\frac{1}{4}\left(\frac{k_1}{k} + \frac{k}{k_1}\right)^2 e^{2kl} + 4} \cdot \quad (18.68)$$

Доданком 4 в знаменнику можна знехтувати, оскільки k_1 і k – одного порядку і $e^{2kl} \gg 4$. З точністю до несуттєвого множника

$$D \sim e^{-2kl} = e^{-\frac{2}{\hbar}\sqrt{2m(U_0-W)}\,l}. \quad (18.69)$$

Ця формула показує, що проникність бар'єру дуже залежить від його ширини l.

Для електрона при $U_0 - W = 2$еВ розрахунок дає наступні значення D: Як видно з таблиці, для ширини бар'єру в 1Å (атомні розміри) проникність досить велика (десятки відсотків), але вже при l=10Å вона стає нехтовно малою.

l, Å	1	2	3	5	10
D, відн. один.	$2{,}3\cdot10^{-1}$	$5{,}5\cdot10^{-2}$	$1{,}3\cdot10^{-2}$	$7\cdot10^{-4}$	$5\cdot10^{-7}$

Ми розглянули випадок проходження частинки через потенційний бар'єр досить спрощеної прямокутної форми. Це ідеалізований випадок. В дійсності потенційний бар'єр має більш складну форму, зображену на рис.18.11. Для такого бар'єру коефіцієнт прозорості приблизно виражається такою формулою:

$$D = D_0 \exp\left\{-\frac{2}{\hbar}\int_{x_1}^{x_2}\sqrt{2m(U-W)}\,dx\right\}, \quad (18.70)$$

де $U = U(x)$, $D_0 = \left(\frac{4k}{k_1}\right)^2/(1+\left(\frac{k}{k_1}\right)^2)^2$, x_1 і x_2 – поворотні точки, в яких $U(x_1)=U(x_2)=W$. k_1 і k визначаються за формулами (18.59) і (18.61). D_0 – константа, яка приблизно дорівнює одиниці, а інтеграл в показнику експоненти формули (18.70) характеризує начебто товщину бар'єру на рівні W. Формула (18.69) є частинний випадок формули (18.70).

Явище проходження частинки через потенційний бар'єр називають *тунельним ефектом*. Така назва дана тому, що частинка проходить через бар'єр, начебто через тунель.

На рис.18.16 приведений якісний вигляд функцій $\Psi_1(x)$, $\Psi_2(x)$, $\Psi_3(x)$.

Важливо зазначити, що частинки виходять із бар'єру з тією ж енергією, з якою на нього і падають ($W = \hbar\omega$). Це означає, що не змінюється частота хвилі де Бройля (при дифракції світла теж не має зміни частоти).

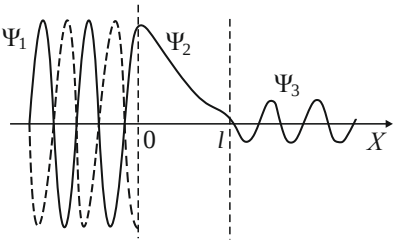


Рис.18.16.

Тунельний ефект є специфічним квантовим ефектом. Проходження мікрочастинки крізь область, в яку, згідно з законами класичної механіки, вона не може проникнути, можна пояснити співвідношенням невизначеностей. Невизначеність імпульсу Δp на відрізку $\Delta x = l$ дорівнює $\Delta p > \frac{\hbar}{l}$. Пов'язана з цим розкидом в значеннях імпульсу невизначеність кінетичної енергії $\Delta W = \frac{p}{m}\Delta p = \sqrt{\frac{2W}{m}}\Delta p = \sqrt{\frac{2W}{m}}\frac{\hbar}{l}$ може виявитися достатньою для того, щоб повна енергія $W+\sqrt{\frac{2W}{m}}\frac{\hbar}{l}$ мікрочастинки виявилася більшою за потенційну U_0.

Проходження частинок через потенційний бар'єр експериментально підтверджено в явищі холодної емісії електронів з металу. Тунельний ефект відіграє основну роль у явищах радіоактивного α – розпаду, перебігу термоядерних реакцій.

Тунельний ефект лежить в основі відомого ефекту Джозефсона, відкритого в 1963р. Суть його в тому, що при проходженні

тунельного струму через тонкий шар діелектрика між двома надпровідниками, відбувається випромінювання електромагнітних хвиль НВЧ (надвисоких частот). Ефект Джозефсона знаходить широке застосування в логічних пристроях, в параметричних перетворювачах та інших електронних приладах.

На основі тунельного ефекту працюють тунельні діоди, які широко використовуються в радіотехніці високих частот, обчислювальній техніці.

§18.10.Холодна емісія електронів із металу

Внаслідок тунельного ефекту, стає можливою емісія з не підігрітого (холодного) катода під впливом електричного поля (E~$10^8 \div 10^9$ В/м). Це так звана *холодна* або *автоелектронна емісія.*

Метал можна розглядати як газ вільних електронів, які можуть переміщуватись в будь-якій області кристалу. Границі кристалу при цьому є начебто своєрідними стінками, які обмежують об'єм, що заповнений електронним газом. Якщо умовно вибрати за початок відліку енергію вільного електрона, що знаходиться в стані спокою у вакуумі, то енергія електрона, який рухається у вакуумі, буде позитивною, а енергія електрона в металі – від'ємною. Таким чином, електрон в металі є начебто в потенційній ямі (рис.18.17) і на його шляху до виходу з металу існує потенційний бар'єр.

Як було показано в §18.8, енергія електрона в потенційній ямі дискретна й вільні електрони в металі (їх концентрація в металі n~10^{28}м$^{-3}$) згідно принципу Паулі заповнюють по два на кожний енергетичний рівень. Навіть при кімнатній температурі верхній із заповнених рівнів утворює різку границю між заповненими і вільними рівнями.

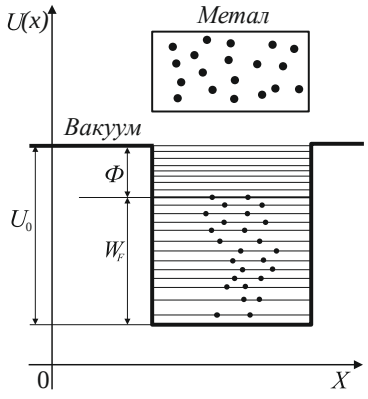


Рис.18.17.Точками позначені електрони.

Цей верхній заповнений рівень називається границею Фермі, а його енергія – енергією Фермі W_F (рис.18.17). Очевидно, що для звільнення електрона з металу необхідно надати йому енергію

$$\Phi = U_0 - W_F,$$

де U_0 – глибина потенційної ями, Φ – робота виходу електрона з металу (входить в рівняння Ейнштейна для фотоефекту). Електрон, однак, може бути звільнений з металу не тільки під дією світла, але і внаслідок нагрівання (термоелектронна емісія). Якщо до металу прикласти електричне поле, як показано на рис.18.18, то всередині металу і на його поверхні енергія електрона не зміниться (електричне поле не проникає в провідник (Розділ 3)). Тому енергію електрона, який вийшов із металу і знаходиться на відстані x від його поверхні в електричному полі, можна записати так:

$$U(x) = U_0 - eEx, \qquad (18.71)$$

де $(-eEx)$ – потенційна енергія електрона в однорідному електричному полі, напруженість

якого E. Залежність (18.71) на рис.18.18 зображена прямою лінією *MN*. Прозорість

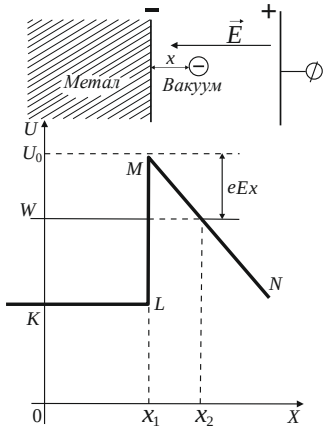


Рис.18.18.

бар'єру *KLMN* визначимо за формулою (18.70), прийнявши $x_1 = 0$:

$$D = D_0 \exp\left\{-\frac{2}{\hbar}\int_0^{x_2}\sqrt{2m_e(U(x)-W)}dx\right\}.$$

Будемо вважати, що $W = W_F$, тоді

$$\int_0^{x_2}\sqrt{U_0 - eEx - W_F}\,dx = \int_0^{x_2}\sqrt{\Phi - eEx}\,dx =$$

$$= -\frac{2}{3eE}(\Phi - eEx_2)^{3/2} + \frac{2}{3eE}\Phi^{3/2} = \frac{2\Phi^{3/2}}{3eE},$$

за умови, що поворотна точка x_2 знаходиться з умови $U_0 - W_F = eEx_2$. Тоді

$$D = D_0 \exp\left\{-4\,\frac{\sqrt{2m_e}\Phi^{3/2}}{3e\hbar}\cdot\frac{1}{E}\right\},$$

або

$$D = D_0 \exp\left\{-\frac{\alpha}{E}\right\}. \qquad (18.72)$$

Струм холодної емісії можна записати так:

$$j = j_0 \exp\left\{-\frac{\alpha}{E}\right\}, \qquad (18.73)$$

де j_0 і $\alpha = 4\,\frac{\sqrt{2m_e}\Phi^{3/2}}{3e\hbar}$ – константи, які характеризують метал, E – напруженість поля. Саме така залежність "холодного струму" від напруженості поля і була знайдена експериментально.

Контрольні питання

1.В чому суть гіпотези де Бройля і корпускулярно–хвильового дуалізму матерії?

2.Чи у всіх випадках руху матеріальних тіл необхідно враховувати гіпотезу де Бройля?

3.Що означає координатно-імпульсний метод описання руху частинок?

4.Що визначають співвідношення невизначеностей Гайзенберга?

5.Хвильова функція та її фізичний зміст. Які властивості хвильової функції?

6.Розв'язок рівняння Шредінгера для нескінченно глибокої потенційної ями. Що означає квантування енергії?

7.Розв'язок рівняння Шредінгера для потенційного бар'єру кінцевої і нескінченої ширини. Яка природа тунельного ефекту?

8.Холодна емісія і причини її виникнення.

Розділ 19.Основи зонної теорії твердих тіл

§19.1.Тверде тіло. Аморфні і кристалічні тверді тіла

Тверде тіло – агрегатний стан речовини, який характеризується стабільністю форми і характером теплового руху атомів, котрі здійснюють малі коливання навколо положень рівноваги. Розрізняють кристалічні та аморфні тверді тіла. Кристали характеризуються просторовою періодичністю в розташуванні рівноважних положень атомів. В аморфних тілах атоми коливаються навколо хаотично

розташованих точок. Аморфний стан (від грецького amorphous – *безформенний),* твердий стан речовини, який характеризується ізотропією властивостей і відсутністю точки плавлення. При підвищенні температури аморфна речовина розм'ягчується і переходить в рідкий стан поступово. Ці особливості обумовлені відсутністю у речовин в аморфному стані строгої періодичності в розташуванні атомів, іонів, молекул і їхніх груп на великих відстанях, яка є в кристалах. Кристали мають певну температуру (точку) плавлення.

Стійким станом (з мінімальною внутрішньою енергією) твердого тіла є кристалічний. З термодинамічної точки зору аморфне тіло знаходиться в метастабільному стані і з часом повинно кристалізуватись. Всі речовини в природі (за винятком рідкого гелію) тверднуть при атмосферному тиску і температурі $T > 0$ К.

Дослідження властивостей твердих тіл об'єднались у велику область – фізику твердого тіла, розвиток якої стимулюється потребами техніки. Фізика твердого тіла – джерело нових матеріалів, нові фізичні ідеї, які народжуються у фізиці твердого тіла, проникають в ядерну фізику, астрофізику, біофізику й інші області науки. Властивості твердого тіла можна пояснити, виходячи зі знань про його атомно-молекулярну будову і закони руху його атомних частинок (атомів, іонів, молекул), а також субатомних частинок (електрони, атомні ядра).

§19.2.Структура кристалів

Кристалами називаються тверді тіла, атоми (іони) в яких просторово упорядковані й утворюють трьохмірну періодичну структуру, яка називається кристалічною решіткою. Багато властивостей твердих тіл пояснюється тією періодичністю, з якою розміщені в просторі їхні структурні елементи.

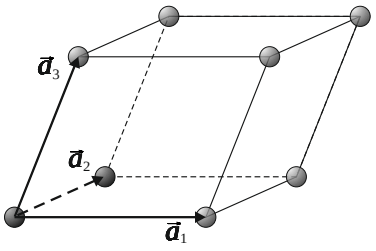


Рис.19.1

Тому таке важливе поняття – кристалічна решітка. Відстань між атомами в кристалічній решітці (близько 10^{-10} м = 1Å) є різною в різних напрямках. В кожній кристалічній решітці можна виділити основний елемент – *елементарну комірку* (рис.19.1).

Елементарна комірка – це той мінімальний за об'ємом елемент решітки (групи атомів), послідовним переміщенням (трансляцією) якого можна побудувати всю решітку кристалу. Тобто для кристалічної решітки характерна періодичність. Виділяючи таку комірку, вдається описати положення атомів чи іонів в речовині, а значить, з її допомогою можна характеризувати структуру кристалів. Положення, які займають атоми чи іони, відповідають точкам (вузлам) решітки. Розглядаючи структуру кристалів, будемо абстрагуватись від теплового руху атомів, який носить характер коливань біля вузлів просторової решітки. Амплітуда цих коливань мала (десь близько 0,1 міжатомної відстані), тому тепловий рух в середньому не порушує симетрії кристалів.

Таким чином, для описання структури кристалу досить задати його елементарну комірку. Сама елементарна комірка, очевидно, геометрично визначається заданням трьох основних векторів $\vec{a}_1, \vec{a}_2, \vec{a}_3$ (рис.19.1). В

напівпровідниках елементарна комірка складається всього лише з декількох атомів. Просторово вона дуже мала і може бути розміщена всередині куба зі стороною $5 \cdot 10^{-10}$м $= 5$Å.

Серед всіх можливих видів решіток можна виділити кубічну (одну з найпростіших) решітку, яка має декілька різновидностей.

1.*Проста кубічна решітка*. В кожній вершині такої решітки розташований один атом, який належить одночасно восьми сусіднім елементарним коміркам (рис.19.2,*а*), яких є вісім. В такій формі кристалізується тільки полоній (Po).

2.*Кубічна об'ємно-центрована решітка.* Тут крім атомів у вершинах кубів є ще один атом в центрі (рис.9.2,*б*). До такого типу відносяться кристалічні решітки молібдену і вольфраму.

3.*Кубічна граніцентрована решітка.* Має шість атомів в центрах граней і, крім того, вісім атомів у вершинах кубів (рис.19.2,*в*). В такій формі кристалізується алюміній і низка інших хімічних елементів.

4.*Решітка типу алмазу.* Може розглядатись як дві вкладені одна в одну кубічні граніцентровані решітки, які зміщені на відстань четвертини діагоналі куба (рис.19.2,*г* темними і світлими кружечками зображені одні й ті самі атоми). В таку форму кристалізується вуглець, кремній, германій і сіра модифікація олова.

5.*Решітка типу арсеніда галію* отримується з решітки типу алмазу в тому випадку, коли атоми галію співпадають з вузлами однієї граніцентрованої решітки, а атоми миш'яку – з вузлами іншої (на рис.19.2,*г* атоми Ga позначені світлими кружками, а атоми As – темними).

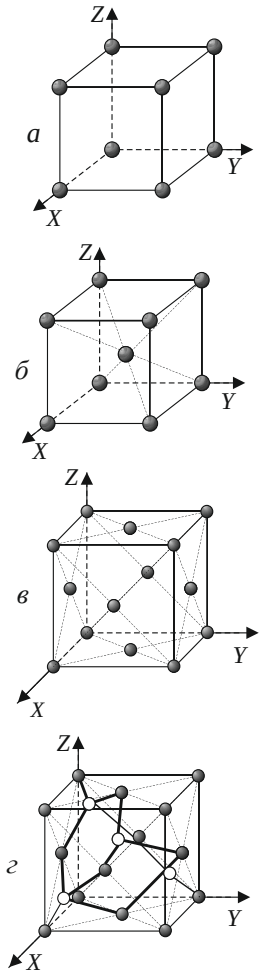


Рис.19.2. Елементарні комірки кубічних решіток: *а*—проста кубічна; *б*—кубічна об'ємно-центрована; *в*—кубічна граніцентрована; *г*—тип алмазу і арсеніду галія.

Щоб знайти число точок, які належать одній елементарній комірці, необхідно розглянути граничні точки комірки. Якщо говорити про елементарну комірку простої кубічної решітки, то тут у вершинах куба є вісім точок, однак кожна з цих восьми точок одночасно належить восьми кубам, які сходяться у вибраній вершині; таким чином, отримаємо $8 \cdot (1/8) = 1$ точку, яка приходиться на елементарну комірку. Для решітки об'ємно – центрованої маємо вісім

точок у вершинах, кожна з яких приходиться на вісім кубів, і крім того, всередині кожного куба міститься одна точка. Звідси витікає, що в одній елементарній комірці міститься $8/8 + 1/1 = 2$ точки. Для решітки граніцентрованої маємо шість точок, які розташовані на гранях комірок, що дотикаються, а також по одній точці у вершинах куба. В результаті знаходимо, що число точок, які належать елементарній комірці, складає тут $8/8 + 6/2 = 4$.

Структура кристалів досліджується, головним чином, наступними методами: дифракції рентгенівських променів (рентгенографія), дифракції електронів (електронографія), дифракції нейтронів (нейтронографія). Дифракція променів або частинок з відповідною довжиною хвилі відбувається на просторовій кристалічній решітці.

Направлені на кристал промені або частинки розсіюються атомами, які розташовані у вузлах кристалічної решітки. Хвилі, які розсіяні окремими атомами, інтерферують між собою. В результаті тільки в деяких напрямках в просторі розповсюджуються підсилені взаємною інтерференцією хвилі (рис.13.28).

Умова появи дифракційного максимуму при розсіюванні рентгенівських променів виражається рівнянням Вульфа-Брегга (13.33). Якщо відомі кути падіння θ променів на кристал (або ж кути, під якими реєструються дифраговані промені) і довжина хвилі λ випромінювання, можна визначити міжплощинну відстань d в даній системі паралельних атомних площин кристалу. Змінюючи кут падіння первинного пучка паралельних променів на кристал, можна досягти виконання умови (13.33) для інших сукупностей паралельних атомних площин і визначити міжплощинну відстань для них. В результаті вдається визначити тип решітки (її симетрію) та її параметри. Рівняння (13.33) справедливе для дифракції на кристалічній решітці всіх випромінювань з відповідною довжиною хвилі, хоч техніка експерименту у випадку використання рентгенівських променів або прискорених електричним полем електронів зовсім різна.

§19.3.Основні типи міжатомної взаємодії в твердих тілах

Сили притягання між частинками, з яких складається тверде тіло, наскільки великі, що атоми твердого тіла не можуть рухатись вільно один відносно одного. Однак ці атоми зовсім не нерухомі. Кожен із них весь час коливається під дією пружних сил, які діють між частинками.

В 1912 р. Фрідріх, Кніппінг і Лауе в дослідах по дифракції рентгенівських променів показали, що кристали дійсно побудовані з атомів, які симетрично розташовані відносно один одного. Брегг в 1913р. визначив кристалічні структури KCl, NaCl, KBr і KI в експериментах по дифракції рентгенівських променів. Ці дослідники вперше пояснили фізичний зміст геометричної правильності кристалічних твердих тіл, яка відповідає їхній атомній будові. Структура і жорсткість твердих тіл визначаються відмінністю в електростатичних силах, які зв’язують разом атоми, що складають тіло.

Іонний зв’язок та іонні кристали

Типовим представником іонних кристалів є кам'яна сіль і, як правило, іонний зв’язок в кристалах розглядають на цьому прикладі. Електронна структура ізольованого

атома натрію $_{11}\mathrm{Na} - 1s^2 2s^2 2p^6 3s^1$, атома хлору $_{17}\mathrm{Cl} - 1s^2 2s^2 2p^6 3s^2 3p^5$. При зближенні цих атомів, наприклад, в процесі конденсації парів, виникає ймовірність переходів електронів зі станів, які належать одному атому, в стан, який належить другому атому (а фактично в цьому випадку вже утворюється система зі спільними станами).

Природно, що перехід зі стану $3s$ електрона натрію у частково пустий стан $3p$ хлору енергетично вигідний, оскільки при цьому обидва атоми набувають на зовнішньо валентній електронній оболонці структуру стійкого октету (вісім електронів). Атом натрію при цьому перетворюється на позитивний іон Na^+, атом хлору – на негативний іон Cl^-. Між ними виникають кулонівські сили притягання, які на близьких відстанях врівноважуються силами відштовхування. Кулонівська взаємодія існує між всіма іонами, які входять в дану конденсовану систему.

В результаті встановлюється деякий рівноважний для даних зовнішніх умов розподіл іонів в просторі, який утворює решітку іонного кристалу (рис.19.3). Найбільш характерні властивості іонних кристалів такі:

– висока (порядку 10^3 К) температура плавлення;

– відсутність електронної і наявність іонної провідності;

– велика механічна міцність, яку, однак, практично не можна використати із-за великої крихкості іонних кристалів;

– іонні кристали – це діамагнетики (магнітна проникність $\mu < 1$, магнітна сприйнятливість $\chi < 0$), якщо вони не містять іонів феромагнітних елементів.

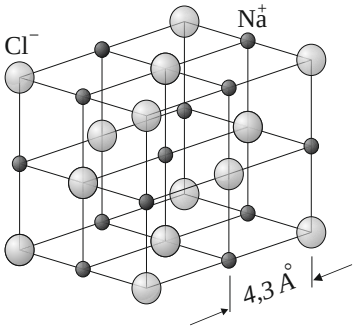


Рис.19.3.

Ковалентний зв’язок

Ковалентний (гомеополярний) зв’язок виникає при зближенні атомів внаслідок перекриття хвильових функцій електронів, що призводить до виникнення при цьому виключно квантової обмінної взаємодії між електронами. Природа обмінної взаємодії обумовлена тотожністю електронів, тобто, принципово не можна розрізнити електрони один від одного. Класичним прикладом обмінної взаємодії є теорія гомеополярного зв’язку в молекулі водню. Як показує детальний квантово-механічний розгляд, у випадку, коли спіни в системі двох електронів антипаралельні, енергія системи, яка складається зі «звичайної» кулонівської енергії і кулонівської енергії обмінної взаємодії, має мінімум при деякій відстані між центрами ядер атомів водню, які входять до складу молекули. Це означає, що утворюється стійка система, в даному випадку – молекула водню H_2. Це приклад найсильнішого хімічного зв’язку в природі, який обумовлює той факт, що атомарного водню в природних умовах не існує.

Ввівши поняття ковалентного зв’язку, розглянемо конкретно кристал кремнію. Атом кремнію $_{14}\mathrm{Si}$ має такий розподіл електронів за станами: $1s^2 2s^2 2p^6 3s^2 3p^2$. В $3p$ – стані є

лише два електрони і чотири вакантних місця. Кожний атом кремнію в твердому стані старається притягнути по одному електрону від кожного з чотирьох найближчих сусідніх атомів, щоб дозаповнити $3p$ оболонку. Виникає так званий алмазоподібний тип кристалічної решітки.

На рис.19.4 приведене двомірне зображення зв'язків в решітці кремнію. Тут у вузлах решітки знаходиться іон кремнію, якому належать чотири валентних електрони $3s^2 3p^2$. Валентні електрони забезпечують ковалентний зв'язок (на рис.19.4 вони зображені у вигляді чорних точок).

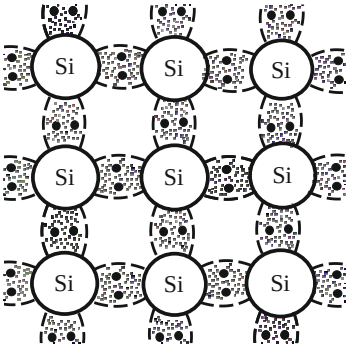


Рис.19.4

Кристали з ковалентним зв'язком часто називають валентними кристалами, основні властивості яких такі:

– висока температура плавлення (10^3К);

– висока твердість і пружність, велика крихкість;

– за електричними властивостями вони належать до напівпровідників і діелектриків (у ряді металів теж зустрічається значна доля ковалентного зв'язку).

Міжатомна взаємодія в металах. Властивості металів

Для металів характерним є наявність в кристалах майже вільних (слабо зв'язаних з іонами гратки) електронів. Цим пояснюється висока електро – і теплопровідність металів, типовий металічний блиск, велика пластичність. Металічний зв'язок – це граничний випадок ковалентного зв'язку. Обмінна взаємодія, яка здійснюється валентними електронами, в цьому випадку відбувається в межах всього кристалу і наявність електронного газу (вільних електронів) визначає ненаправлений характер металічного зв'язку. Цим, зокрема, обумовлена висока пластичність металів: переміщення атомів в решітці під дією зовнішніх сил менше змінює його енергетичний стан, ніж у випадку, коли при такому переміщенні розриваються направлені іонні або ковалентні зв'язки. З точки зору якісних уявлень про утворення стійких октетів на зовнішніх електронних оболонках атомів виникнення металічного зв'язку обумовлене «позбавленням» атомами всіх валентних електронів, в результаті чого кожний атом набуває електронної структури попереднього інертного газу. Наприклад, атом натрію $1s^2 2s^2 2p^6 3s^1$, віддавши в склад електронного газу свій єдиний валентний $3s$ електрон, набуває електронної структури неона $1s^2 2s^2 2p^6$. Це приводить до енергетичного виграшу і до виникнення сил притягання між атомними залишками; кулонівська взаємодія з атомними залишками утримує колективізовані електрони в об'ємі кристалу.

Таким чином, структура металу уявляється як іонний позитивний кістяк, що занурений в електронний газ, який компенсує сили відштовхування між іонами і зв'язує їх внаслідок обмінної взаємодії в тверде тіло (рис.19.10). Більшість металів мають щільно

упаковану решітку: кубічну граніцентровану (рис.19.2,*в*) і гексагональну (рис.19.5).

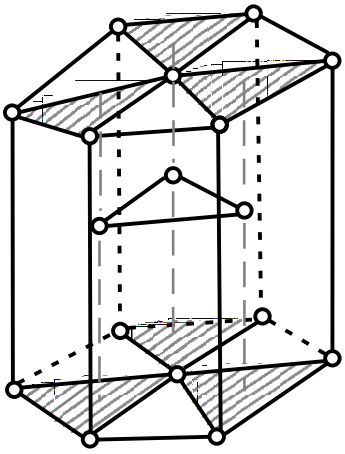


Рис.19.5.Гексагональна комірка металу.Світлі кружечки – іони металу.

Перехідні тугоплавкі метали, такі як Cr, Mo, W, V, Ta, мають кубічну об'ємно-центровану решітку. Кубічну об'ємно-центровану решітку має при температурі до 910 ℃ і залізо. У них, крім звичайного металічного зв'язку в кристалах, також діє ковалентний зв'язок, обумовлений електронами внутрішніх незаповнених оболонок. Більшість металів – парамагнетики ($\mu \geq 1, \chi > 0$). Перехідні метали Fe, Co, Ni і Gd – феромагнетики ($\mu \gg 1$).

§19.4.Рідкі кристали. Сили Ван дер Ваальса

Рідкі кристали – це особливий стан деяких органічних речовин, в якому вони володіють реологічними властивостями рідини – *плинністю*, але зберігають певну впорядкованість в розташуванні молекул і анізотропію ряду фізичних властивостей, яка характерна для твердих кристалів. Рідкі кристали називають ще *мезофазами* (від грецького mezos – середнє).

Рідкі кристали відкриті в 1889р. австрійським ботаніком Ф.Рейніцером, який синтезував незвичайні кристали. При їхньому нагріванні отримувалась рідина, яка в залежності від температури була то каламутною, то прозорою, то набувала синюватий відтінок. Німецький фізик О.Леман

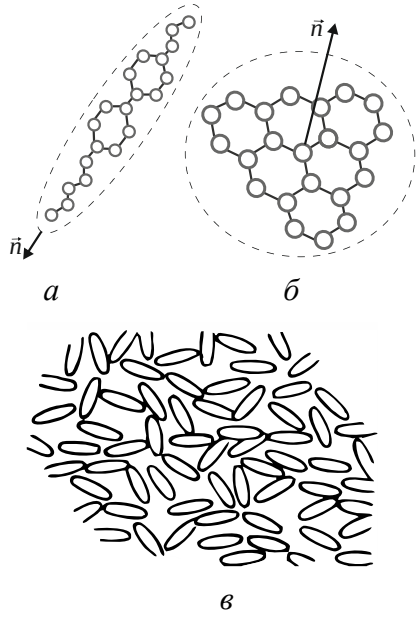
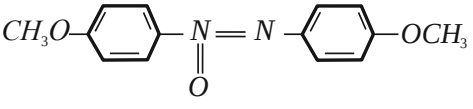


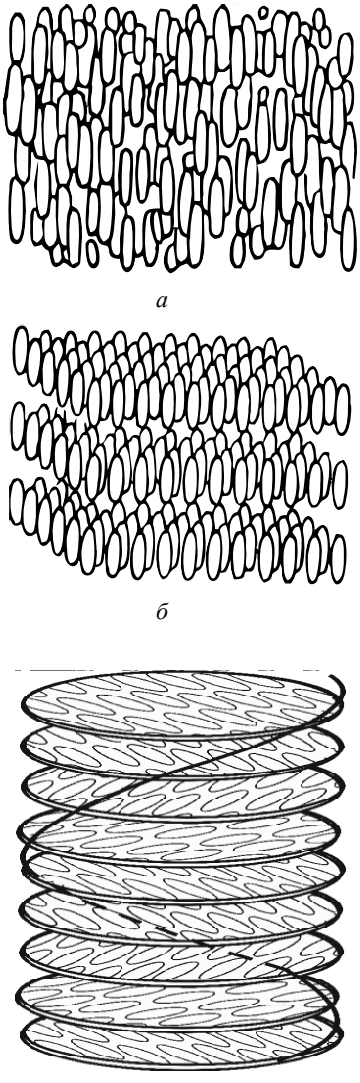
Рис.19.6

почав систематичне вивчення таких речовин і встановив, що відкрито особливий стан, який належить багатьом органічним речовинам. Число хімічних сполук, для яких знайдені рідкі кристали, складають декілька тисяч. Рідкі кристали утворюються при нагріванні деяких твердих кристалів (мезогенних): спочатку відбувається фазовий перехід в рідкий кристал, а подальше плавлення рідкого кристалу у звичайну ізотропну рідину. Кожна мезофаза існує в певному температурному інтервалі (термотропні рідкі кристали). Рідкі кристали

утворюють речовини, молекули яких мають видовжену палицеподібну форму, часто з чергуванням лінійчастих (рис.19.6,*а*) і циклічних плоских атомних груп (рис.19.6,*б*), які паралельні одні одним. На рис.19.6,*в* приведене розташування таких молекул в рідині. Така форма молекул визначає приблизну паралельність їх взаємної укладки, що є основною ознакою рідкого кристалу.

Розрізняють три основні типи рідких кристалів: нематики (рис.19.7,*а*), смектики (рис.19.7,*б*) і холерики (рис.19.7,*в*). Найменшу впорядкованість мають нематичні рідкі кристали (назва «нематичний» походить від грецького слова νημα – нитка). В рідких кристалах в мікроскопі видно тонкі рухомі нитки, які є дефектами структури. В ідеальному рідкому кристалі таких ниток не має. В нематиках молекули паралельні, але зсунуті вздовж своїх осей одна відносно одної на довільні відстані (рис.19.7,*а*). В смектичних рідких кристалах молекули паралельні одна одній і розташовані шарами (рис.19.7,*б*). Структура холестеричних рідких кристалів відрізняються від них додатковим закручуванням молекул в напрямку, який перпендикулярний до їхніх довгих осей (рис.19.7,*в*). Крок такої спіральної «надструктури» може бути досить великим і досягати декількох мікрометрів. Прикладом нематичного рідкого кристалу є параазоксианізол, який існує в мезофазі в інтервалі 116-136 °С. Його молекула має таку структуру:





а

б

в

Рис.19.7.Типи рідких кристалів: *а*–нематики; *б*–смектики; *в*–холерики.

Знайдені смектичні рідкі кристали деяких сполук і їхніх сумішей в інтервалі від –40 до +80 °С, що розширює можливості їхнього практичного використання. Прикладом холестеричних рідких кристалів є ефіри холестерину. Розглянемо тепер сили, які діють в нематичній рідині. Ці сили – електричного походження. Сила притягання

виникає між двома атомами або молекулами, які самі по собі є електрично нейтральними. Вперше такі сили став враховувати голландський фізик Я.Д.Ван дер Ваальс (1873р.) для пояснення властивостей реальних газів і рідин, і з тих пір вони отримали назву сил Ван дер Ваальса. Він припустив, що на малих відстанях r між молекулами діють сили відштовхування, які зі збільшенням відстані змінюються силами притягання. Спробуємо розібратись, як це відбувається. Уявімо собі, що через якусь випадкову причиниу в атомі відбулось зміщення негативно зарядженої електронної хмарини відносно позитивно зарядженого ядра, тобто виникає електричний диполь (рис.19.8,*а*). Біля атома−диполя виникає електричне поле. Напруженість цього поля швидко зменшується при віддаленні від атома, але біля атома воно достатньо велике. Якщо в поле атома I $\vec{E}_1$ потрапляє нейтральний атом II (рис.19.8,*б*), то електричне поле атома I повинно змістити заряди електронів і ядра атома II так, як показано на рис.19.8,*б*, тобто атом II поляризується. Поле атома II $\vec{E}_2$ підтримує розділення зарядів в атомі I. Із рис.19.8,*б* видно, що різнойменно заряджені

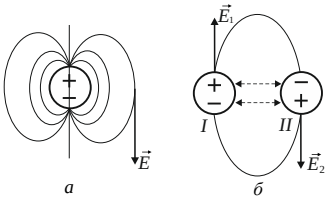


Рис.19.8.

частинки атомів повинні притягуватись, а значить, і атоми в цілому повинні притягуватись один до одного. При зближенні атомів між ними починають діяти сили відштовхування. На відстані, що приблизно дорівнює розміру атома, сили взаємодії між атомами дорівнюють нулю.

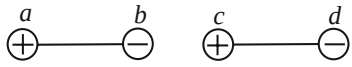


Рис.19.9.

Такі-самі міркування можна привести й у відношенні двох молекул, які складаються з декількох десятків атомів. Нейтральні молекули повинні притягуватись одна до одної внаслідок виникнення електричних диполів. Сила притягання між двома полярними молекулами максимальна в тому випадку, коли їхні дипольні моменти розташовуються вздовж однієї лінії (рис.19.9) і залежить від їхньої взаємної орієнтації. Різнойменні заряди b і c взаємодіють сильніше, ніж однойменні заряди a і c (а також b і d), що знаходяться на більшій відстані один від одного. Сила взаємодії таких молекул $F{\sim}r^{-7}$ зменшується з відстанню значно швидше, ніж кулонівська сила взаємодії заряджених частинок $F_{\text{кул}}{\sim}r^{-2}$. Через ці причини більша частина атомів одної молекули прагне наблизитись до атомів іншої молекули, оскільки тільки в цьому випадку сили взаємодії між молекулами перетворюються в нуль. Але така ситуація можлива тільки тоді, коли довгі осі молекул або площини молекул паралельні одні до одних. Так виникає певний порядок в орієнтації молекул і з'являється виділений напрямок. Цей напрямок можна характеризувати одиничним вектором $\vec{n}$ (рис.19.6). Зрозуміло, що таке паралельне розташування виділених осей (або площин) молекул можливе тільки при достатньо низькій температурі, коли теплові поштовхи не наскільки сильні, щоб зруйнувати орієнтаційний порядок в системі молекул. При підвищенні температури обов'язково настає момент, коли хаотичний тепловий рух стає переважаючим і нематичний порядок руйнується.

Таким чином, система таких особливих молекул може мати два стани: звичайний (ізотропний) рідкий – при високих температурах (рис.19.6,*в*) і анізотропний рідкий – при низьких температурах (рис.19.7).

Теорія рідких кристалів базується на поєднанні принципів симетрії кристалів (у відношенні найближчих сусідів) з законами статистичної фізики.

Рідкі кристали мають анізотропію пружності, електропровідності, магнітної сприйнятливості і діелектричної проникності, оптичну анізотропію, сегнетоелектричні властивості та інше. Наприклад, холестеричні рідкі кристали володіють великою оптичною активністю (в $10^2 - 10^3$) більшою, ніж в органічних речовинах і твердих кристалах.

Найбільші враження створюють оптичні властивості рідких кристалів, які зробили їх такими популярними. Такі оптичні властивості дуже незвичні для рідин і властиві твердому кристалу. Зрозуміло в чому тут справа: в нематичній рідині, як і в кристалі є особливі напрямки. Наявність цих напрямків обумовлює ряд чудових властивостей кристалів і зокрема, оптичні властивості. Тому такі особливі напрямки називають оптичними осями. Напрямок оптичних осей в рідкому кристалі співпадає з вектором $\vec{n}$ (рис.19.6,*а,б*). В твердих кристалах оптичні осі жорстко закріплені. А в рідких кристалах напрямки оптичних осей можна змінювати за допомогою найрізноманітніших впливів, в тому числі й електричними або магнітними полями.

Ефект зміни напрямку орієнтації молекул в нематичній рідині під дією електричного поля спостерігав В. Фредерікс. Назва цього ефекту тепер носить його ім'я. Користуючись популярними електричними годинниками і калькуляторами з екранами на рідких кристалах, ми спостерігаємо саме явище ефекту Фредерікса. Рідкі кристали мають широке практичне застосування, особливо в системах обробки і відображення інформації, в яких використовуються електрооптичні властивості рідких кристалів.

§19.5. Класифікація твердих тіл за величиною питомої електропровідності та її залежності від температури

Класифікація твердих тіл за величною питомої електропровідності наведена в розділі 2 (§2.1). Питомий опір в срібла менший, ніж у тефлона, в 10^{24} (!) раз. Жодна фізична характеристика не змінюється так сильно при переході від речовини до речовини, як питомий опір. Виникає питання, чому ж опір цих речовин в твердому стані відрізняється так сильно. Більше того, атоми однієї і тієї самої речовини, наприклад вуглецю (С), в залежності від того, в яку кристалічну решітку твердого тіла вони з'єднанні, можуть утворювати і дуже хороший провідник – графіт, і прекрасний ізолятор – алмаз. Олово – здавалось би, типовий метал. Однак, якщо всім знайоме біле блискуче олово довго потримати при температурі мінус 30℃, то його кристалічна структура зміниться, а зовнішньо воно посіріє. І це олово – його так і називають «сіре олово» – напівпровідник. А свинець завжди метал.

Якщо почати змішувати різні елементи, то картина зовсім ускладнюється. Візьмемо, наприклад, і сплавим два метали – індій і сурьму в пропорції один до одного. Отримаємо напівпровідник InSb, який широко застосовується в техніці.

Приклади з атомами вуглецю, олова наводять на думку, що тверде тіло буде

металом чи діелектриком залежить не тільки і, може бути, навіть не настільки від властивостей атомів, які складають кристал, скільки від того, які відношення складаються у кожного атома з сусідами, тобто від типу зв'язків, котрі об'єднують атоми в кристалічну решітку твердого тіла. На рис.19.10 схематично зображена кристалічна решітка металу.

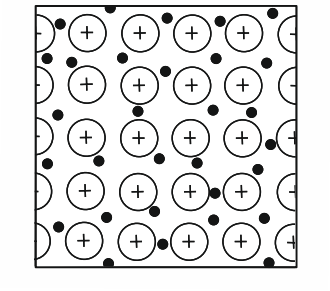


Рис.19.10

Правильна решітка позитивно заряджених іонів занурена в «газ» із вільних електронів (темні кружечки), не зв'язаних жорсткими зв'язками з окремими іонами. Електрони, як прийнято говорити, колективізовані кристалом і можуть вільно переміщуватись під дією зовнішнього електричного поля. В кожному кубічному сантиметрі металу міститься близько 10^{22} вільних носіїв струму – електронів. Тому метали є чудовими провідниками.

На рис.19.4 схематично зображена решітка (фактично діелектрика) кристалічного кремнію. Кожний атом зв'язаний з найближчими сусідами міцними електронними зв'язками. Щоб розірвати електронні зв'язки між атомами решітки і створити в такому кристалі вільні електрони, які здатні створювати електричний струм, необхідно затратити енергію. Звідки може взятись ця енергія? Якщо її величина не дуже велика, а кристал досить сильно нагріти, то електронні зв'язки можуть бути розірванні внаслідок теплових коливань решітки. Світло з відповідною довжиною хвилі теж може розірвати електронні зв'язки і створити в кристалі вільні електрони. Але при досить низькій температурі (і у відсутності світла) кристал, який показаний на рис.19.4 є ідеальним ізолятором: вільних носіїв заряду в ньому немає.

Величина енергії, яка необхідна для того, щоб розірвати електронні зв'язки між атомами і створити в неметалічному кристалі вільні електрони, позначається символом W_g і називається енергією активації. Вона є важливою характеристикою кристалу. Значення W_g залежить від структури кристалічної решітки, від властивостей атомів, які її складають, і у різних матеріалів її значення лежить у межах від нуля до декількох десятків електрон-вольт. Якщо значення W_g велике, то нагрівання навіть до дуже високої температури не створює в кристалі скільки-небудь помітного числа вільних електронів. При досить великому значенні W_g кристал може навіть розплавитись раніше, ніж в ньому внаслідок теплового руху виникатимуть вільні носії заряду. Такі матеріали з великим значенням W_g є типовими *діелектриками*.

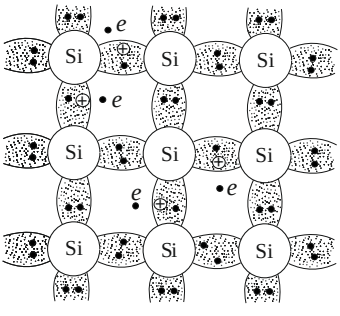


Рис.19.11.Виникнення носіїв заряду в напівпровіднику. Позначення: ● – електрони, ⊕ – дірки.

При невеликому значенні W_g нагрівання навіть не до дуже високих

температур призведе до розриву значного числа електронних зв'язків і появі відносно великого числа вільних носіїв заряду (рис.19.11). Кристал набуває здатності проводити електричний струм. Тому матеріали з невеликим значенням W_g будуть займати «як би середнє місце між провідниками і непровідниками». Адже «здатність проводити електрику» (провідність) пропорційна концентрації вільних носіїв заряду. В провідниках (металах) концентрація вільних носіїв заряду дуже велика. Це відповідає ситуації, коли всі електронні зв'язки розірвані і всі валентні електрони вільні. В «непровідниках» (діелектриках) вільних носіїв заряду практично не має. Зрозуміло, що питомий опір матеріалу з невеликим значенням W_g може дуже сильно змінюватись з температурою. В таких матеріалах концентрація вільних носіїв заряду пропорційна кількості розірваних електронних зв'язків буде дуже різко зростати зі збільшенням температури, а питомий опір різко зменшуватись. Ці матеріали називаються *напівпровідниками*. Зазначимо, що провідність напівпровідників збільшується не тільки при нагріванні, але і при освітленні, при опроміненні ядерними частинками; вона змінюється при прикладанні електричних і магнітних полів, при змінні зовнішнього тиску і т. п.

Тому, *напівпровідники – це речовини, які мають при кімнатній температурі питомий опір в інтервалі від* 10^8 *до* 10^{-6} Ом ·м *(провідність від* 10^{-8} *до* 10^6 $(\text{Ом}\cdot\text{м})^{-1}$*), який в сильній степені залежить від зовнішніх умов* (наприклад від температури):

$$\rho \sim \exp\left\{\frac{W_g}{2kT}\right\}\cdot \qquad (19.1)$$

Більш коротко, напівпровідники – це неметалічні матеріали з відносно невеликою величиною енергії W_g. Значення W_g для найбільш поширених напівпровідників приведено в таблиці при кімнатній температурі Ці напівпровідники знайшли широке практичне застосування. Так, із InSb виготовляють приймачі інфрачервоного випромінювання ($\lambda \cong 10$ мкм), які лежать в основі роботи приладів нічного бачення. Із германію були виготовленні перші транзистори і діоди. Кремній – основний матеріал сучасної

Напівпровідник	InSb	Ge	Si
W_g, еВ	0,17	0,72	1,1
Напівпровідник	GaAs	GaAlAs	SiC
W_g, еВ	1,4	1,4 - 2,17	3

напівпровідникової електроніки. Арсенід галію (GaAs) використовується для створення складник сполук (наприклад, GaAlAs), які використовуються для створення напівпровідникових світлодіодів і лазерів. Із карбіду кремнію (SiC) створенні найбільш надійні і стабільні світлодіоди, що здатні працювати при дуже високій температурі.

Перераховані тільки декілька найбільш важливих напівпровідникових матеріалів. Всього ж синтезовано, вивчено і використовується на практиці сотні напівпровідникових сполук.

В металах концентрація електронів не залежить від температури. Навіть при абсолютному нулі «вільні» електрони залишаються вільними і зберігають здатність проводити струм. У хімічно чистих металах питомий опір з ростом температури збільшується пропорційно абсолютній температурі:

$$\rho = \rho_0(1 + \alpha t) = \rho_0\, T/T_0, \quad (19.2)$$

де ρ_0 – питомий опір даного металу при 0 °C, α – термічний коефіцієнт опору, який дорівнює $1/273\ \text{K}^{-1}$, T_0=273 K.

Таким чином, метали – це такі речовини, які проводять електричний струм при будь-якій температурі.

§19.6.Енергетичні зони в кристалах

Чому одні тіла є провідниками, а інші напівпровідниками і ізоляторами – на це питання не можна відповісти в рамках моделі газу вільних електронів. Необхідно врахувати взаємодію атомів між собою і з електронами.

Якісно задача розв'язується так. Із окремих атомів будується кристалічна решітка і з'ясовується, як впливають сусідні атоми кристалу на поведінку електронів в кожному окремому атомі, тобто досліджується вплив поля сусідніх атомів на енергетичні рівні електронів в атомі. Розмістимо *N* атомів натрію у вигляді просторової решітки, яка властива кристалу натрію, але на великих віддалях один від одного, щоб взаємодією між ними можна було знехтувати. В цьому випадку енергетичні стани електронів в кожному атомі можна вважати такими, як і в окремо взятому ізольованому атомі. На рис.19.12 наведена енергетична схема двох таких атомів. Кожний із них зображений у вигляді потенційної ями, всередині якої проведені рівні 1*s*, 2*s*,2*p*, 3*s*…... Рівні 1*s*,2*s* і 2*p* у натрію заповнені повністю електронами, рівень 3*s* – наполовину: рівні, які розміщені вище 3*s* – вільні.

Ізольовані атоми відокремлені один від одного потенційними бар'єрами завтовшки $r \gg a$, де a – стала решітки кристалу натрію. Для електронів рівня 1*s* - висота бар'єру U_1 , для електронів $2s$ - U_2 , для електронів $2p$ - U_3 . Потенційний бар'єр перешкоджає вільному переходу електронів від одного атома до іншого. У верхній частині рис.19.12 показана якісна картина розподілу густини ймовірності $\rho = 4\pi r^2 \psi\psi^*$ виявлення електронів на даній відстані від ядра. Максимуми кривих відповідають приблизно положенню боровських орбіт для цих електронів. Тепер будемо повільно зближувати атоми решітки, не порушуючи її симетрію. На рис.19.13 наведена картина, яка відповідає такому зближенню. З цього рисунка видно, що потенційні криві, які відділяють сусідні атоми (показані пунктиром), частково накладаються одна на одну і дають результуючі криві, які проходять нижче нульового рівня 00 (на рисунку вони показані суцільними лініями).

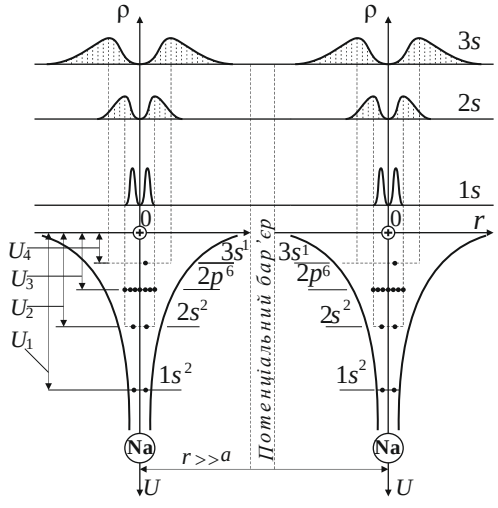


Рис.19.12.

Отже, зближення атомів зменшує товщину потенціального бар'єру до значення порядку a і зменшує його висоту. Для електронів рівня 1*s* - висота бар'єру стає U_1', для електронів $2s$ - U_2', для електронів $2p$ - U_3'.

Для електронів $3s$ рівня висота бар'єра виявляється нижчою від початкового положення рівня $3s$ в атомі натрію. Тому валентні електрони цього рівня отримують можливість практично без перешкод

переходити від одного атома до іншого. Про це свідчить і характер електронних хмар валентних електронів, вони перекриваються настільки сильно, що створюють результуючу хмару рівномірної густини ρ (рис.19.13). Це відповідає стану повної їх колективізації в решітці. Такі колективні електрони називаються вільними, а їх сукупність – електронним газом.

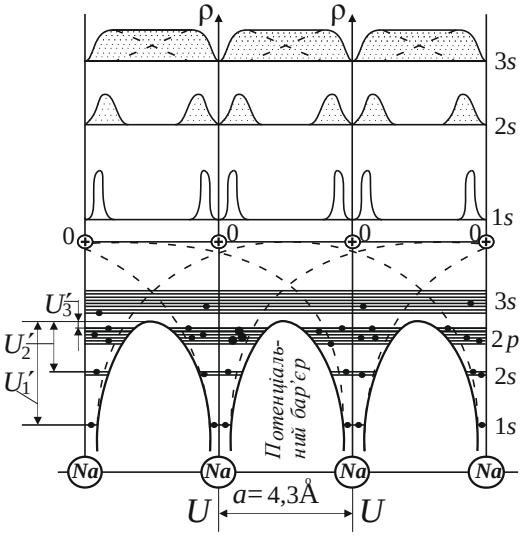


Рис.19.13.

Внаслідок різкого зменшення висоти і товщини потенційного бар'єру при зближені атомів можливість переміщення в кристалі отримують не лише валентні електрони, але й електрони, що розміщенні на інших рівнях атомів. Переміщення відбувається шляхом тунельного переходу електронів через бар'єри, що відокремлюють сусідні атоми. Якщо колективні електрони зберігають досить сильний зв'язок з атомами, то їх потенційну енергію можна подати у такому вигляді:

$$U = U_a + \delta U,$$

де U_a – потенційна енергія електрона в ізольованому атомі. Для кристала вона є періодичною функцією з періодом, що дорівнює постійній гратки, оскільки енергія електрона повторюється при переході його від одного атома до іншого (рис.19.13); δU – доданок, який враховує вплив сусідніх атомів на цю енергію.

Якщо знехтувати доданком δU, тобто розглядати так зване нульове наближення, то в ролі хвильової функції і енергії електрона в кристалі слід взяти хвильову функцію ψ_a і енергію $W_a(n,l)$ електрона в ізольованому атомі.

Відмінність між кристалом і атомом полягає в цьому випадку в наступному: в той час як в ізольованому атомі даний енергетичний рівень $W_a(n,l)$ є єдиним, то в кристалі, що складається з N атомів, він повторюється N разів. Кожний рівень ізольованого атома в кристалі стає $N-$кратно виродженим.

Якщо врахувати доданок δU в потенційній енергії, то в міру зближення ізольованих атомів і утворення з них кристалічної решітки кожний атом потрапляє у поле своїх сусідів, яке зростає. Така взаємодія приводить до зняття виродження. Тому кожний енергетичний рівень, не вироджений в ізольованому атомі, розщеплюється на N близько розміщених один від одного підрівнів, що утворюють енергетичну зону. На кожному підрівні згідно з принципом Паулі може знаходитись не більше двох електронів з антипаралельними спінами. Якщо енергетичний рівень мав в атомі $(2l+1)$ – кратне виродження, то відповідна йому енергетична зона буде складатися із $N(2l+1)$ підрівнів. Так, $s-$рівень дає $s-$зону, яка складається із N підрівнів і здатна вмістити $2N$ електронів; $p-$рівень дає $p-$зону, яка складається із $3N$ підрівнів і здатна вмістити $6N$ електронів, і т.д. Найбільший вплив поле решітки чинить на

зовнішні валентні електрони атомів. Тому стани цих електронів в кристалі відчувають найбільші зміни, а енергетичні зони, які утворені з енергетичних рівнів цих електронів, виявляються найбільш широкими.

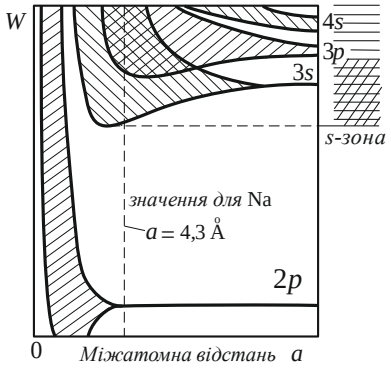


Рис.19.14.Утворення енергетичних зон в кристалі із дискретних рівнів при зближенні атомів.

Внутрішні ж електрони, які сильно зв'язані з ядром, відчувають лише незначні збурення від інших атомів, внаслідок чого їх енергетичні рівні в кристалі залишаються практично такими ж вузькими, як і в ізольованих атомах. На рис.19.14 наведена схематична картина утворення енергетичних зон в кристалі із дискретних атомних рівнів.

Отже кожному енергетичному рівню ізольованого атома в кристалі відповідає зона дозволених рівнів: рівню $1s$ – зона $1s$, рівню $2p$ – зона $2p$ і т. д. Зони дозволених енергій розділені областями заборонених енергій – забороненими зонами. Із збільшенням енергії електрона в атомі ширина дозволених зон збільшується, ширина заборонених – зменшується.

Дозволені енергетичні зони в твердому тілі можуть бути різним чином заповнені електронами. У граничних випадках вони можуть бути цілком заповнені або цілком вільні. Електрони в твердих тілах можуть переходити з однієї дозволеної зони в іншу. Для переходу електрона з нижньої зони в сусідню верхню зону необхідно затратити енергію, що дорівнює ширині забороненої зони, яка розміщена між ними.

Для внутрішніх переходів електронів в межах зони необхідна дуже мала енергія, тому що відстані між сусідніми енергетичними рівнями в зоні дорівнюють приблизно 10^{-22}еВ. Наприклад, для цього достатньо енергії $10^{-4} - 10^{-8}$ еВ, що отримує електрони в металі під дією електричного поля на довжині вільного пробігу при звичайних різницях потенціалів. Під дією спеціального збудження електронам може бути надана енергія, яка достатня як для внутрішньозонних, так і для міжзонних переходів.

Зонна теорія твердих тіл дозволила з єдиної точки зору пояснити існування металів, діелектриків і напівпровідників, пояснюючи відмінності в їхніх електричних властивостях неоднаковим заповненням електронами дозволених зон і різною шириною заборонених енергетичних зон.

Ступінь заповнення електронами енергетичних рівнів у зоні визначається заповненням відповідного атомного рівня. Якщо, наприклад, якийсь рівень атома повністю заповнений електронами відповідно до принципу Паулі, то утворена з таких рівнів енергетична зона також повністю заповнена.

Найвища зона, яка цілком заповнена електронами при $T = 0$ К, називається *валентною зоною.* Зона, яка заповнена електронам частково або вільна при $T = 0$ К, називається *зоною провідності.*

Залежно від ступеня заповнення енергетичних зон електронами і ширини забороненої зони можливі чотири випадки

(рис.19.15). У першому випадку найвища зона, що містить електрони, заповнена лише частково, тобто є вакантні рівні (рис.19.15,*а*). У такому разі електрон, отримавши малу енергетичну добавку за рахунок теплового руху або електричного поля, зможе перейти на вищий енергетичний рівень самої зони, тобто стати вільним і брати участь у провідності.

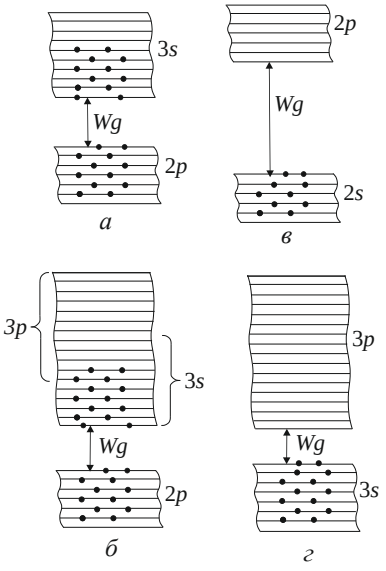


Рис.19.15.Позначення: темні кружечки - електрони.

Внутрішній перехід можливий, оскільки, навіть при $T = 1$ К енергія теплового руху $kT \approx 10^{-4}$ еВ більша від різниці енергій між сусідніми енергетичними рівнями зони (10^{-22} еВ). Отже, тіло буде провідником електричного струму. Саме це властиве *металам.*

Тверде тіло є провідником електричного струму і в тому випадку, коли валентна зона 3*s* перекривається з вільною від електронів зоною 3*p*, що призводить до неповного заповнення зони (рис.19.15,*б*). Це має місце для лужно-земельних елементів, що утворюють другу групу періодичної системи елементів Менделєєва. У даному випадку утворюється так звана «гібридна» зона, яка заповнюється валентними електронами лише частково.

Можливий також перерозподіл електронів між зонами, що виникають із рівнів різних атомів, який може привести до того, що в кристалі виявиться одна цілком заповнена (валентна) зона і одна вільна зона (зона провідності). Тверді тіла, у яких енергетичний спектр електронних станів складається лише з валентної зони і зони провідності є діелектриками або напівпровідниками залежно від ширини забороненої зони W_g.

Якщо ширина забороненої зони кристалу порядку декількох електрон-вольт, то тепловий рух не може перевести електрони з валентної зони в зону провідності і кристал є *діелектриком*, залишаючись ним при всіх реальних температурах (рис.19.15,*в*). Для типових діелектриків $W_g > 3$ еВ. Так, для алмазу $W_g \approx 5{,}2$ еВ, для нітриду бору $W_g \approx 4{,}6$еВ, для Al_2O_3 $W_g \approx 7$ еВ і т. д.

Якщо заборонена зона досить вузька $W_g \leq 1$ еВ, перехід електронів з валентної зони в зону провідності може бути здійснений порівняно легко: або тепловим збудженням, або за рахунок зовнішнього джерела, здатного передати електронам енергію W_g (рис.19.15,*г*). Такий кристал буде *напівпровідником.* При температурах близьких до 0 К, напівпровідники ведуть себе як діелектрики, оскільки переходів електронів у зону провідності не відбувається. Для типових напівпровідників $W_g \leq 1$ еВ.

Зонну структуру енергетичного спектру для електрона в кристалі можна отримати в результаті розв'язування рівняння Шредінгера

$$-\frac{\hbar^2}{2m_e}\Delta\psi(\vec{r}) + U(\vec{r})\psi(\vec{r}) = W\psi(\vec{r}),$$

де $U(\vec{r})$ – потенційна енергія одного електрона не лише в періодичному полі ядер гратки, але і в усередненому періодичному полі всіх інших електронів, $\Delta=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$ – оператор Лапласа, W – повна енергія електрона в кристалі, m_e – маса електрона. Через математичні складнощі розв'язування цього рівняння розглядати не будемо.

Для кращого розуміння виникнення енергетичних зон в кристалах розглянемо ще такий підхід. Рух електрона в атомі можна співставити з будь-якою коливальною системою, наприклад з маятником. Нехай у нас є два зовсім однакових маятники. Поки вони не взаємодіють один з одним, частота коливань обох маятників одна й та сама. Введемо тепер взаємодію між ними – зв'яжемо їх, наприклад, м'якою пружинкою. І зразу ж замість однієї

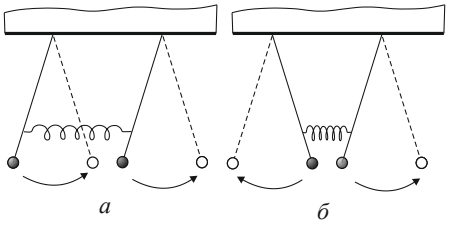


Рис.19.16.

частоти з'явиться дві. Подивимось на рис.19.16: зв'язані маятники можуть коливатись синфазно (рис.19.16,*а*), а можуть на зустріч один одному (рис.19.16,*б*). Очевидно, що в другому випадку рух маятників буде більш швидким, тобто частота коливань такої системи вища власної частоти коливань одного маятника. Таким чином, зв'язок приводить до розщеплення частот. Якщо зв'язати три маятники, то стане вже три власних частоти, у системи з чотирьох зв'язаних маятників чотири власних частоти і так далі до нескінченності.

Поведінка будь-яких коливальних систем подібна. Якщо ми замінимо маятники, наприклад, на електричні коливальні контури, то при введенні зв'язку між ними їхні власні частоти теж розщеплюються. Електрони в атомі – це теж своєрідна коливальна система. Як і маятник, електрони мають масу, є сила Кулона, яка повертає їх в положення рівноваги, і цим визначається рух електронів в атомі, який характеризується згідно квантової механіки власною частотою. Для електронів включення взаємодії при взаємному зближенні атомів приводить до того, що частоти, які були до того однаковими, стають дещо різними. Будемо пам'ятати, що в квантовій механіці є прямий зв'язок між енергією і частотою коливань, який виражається формулою $W = h\nu$.

Розглянемо один з найпростіших елементів, який утворює метал – літій. Порядковий номер Li – три. Це означає, що атом літію має три електрони і його електронна конфігурація $1s^2 2s^1$. Два електрони утворюють заповнену $1s$-оболонку, яка найближча до ядра і тому сильно зв'язана з ним. Третій електрон знаходиться на $2s$-підоболонці, хоч на ній могли би знаходитись два електрони. Всі інші дозволені електронні підоболонки вільні і електрони на них потрапляють тільки при збудженні атома (наприклад, при сильному нагріванні парів літію). Схема рівнів в атомі літію показана на рис.19.17, 1*атом*. При зближенні двох атомів літію кожний із рівнів, які показані на рис.19.17,2*атоми*, розщеплюються на два. Кожному новому рівню енергії буде відповідати своя електронна оболонка тепер вже не окремого атома, а «молекули». Рівні енергії заповнюються електронами по тому же правилу, що й у атома – по два електрони на кожному рівні. Та пара оболонок, яка отримується із самого нижнього рівня, буде повністю заповнена електронами. Дійсно на них можна розмістити чотири

електрони, а їх у двох атомів літію – шість. Залишаються два електрони, які тепер розташуються на нижньому рівні другої пари. Зауважимо, що відбувся якісний скачок: раніше ці два електрони займали два із чотирьох станів, які мали однакову енергію. Тепер у них з'явилась можливість обирати і вони розташувались так, щоб їх сумарна енергія була

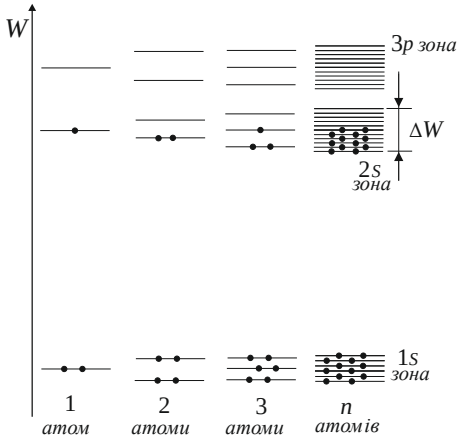


Рис.19.17

найменша. Неважко уявити, що відбудеться при добавленні наступних атомів: для трьох атомів кожний вихідний рівень розщепиться на три (рис.19.17, 3 *атоми*). Дев'ять електронів розташуються так: шість на нижній тріаді рівнів, які виникають із рівня найближчої до ядра внутрішньої заповненої оболонки атома, ще два електрони – на нижньому рівні наступної тріади; електрон, що залишився, – на середньому рівні тієї ж тріади. Ще одне місце на цьому рівні залишається вільним, а верхній рівень повністю пустий. Якщо взяти n атомів ($n \gg 1$), то кожен рівень розщепиться на n тісно розташованих рівнів, які утворюють смугу або зону дозволених значень енергії. В нижній зоні всі стани будуть зайнятими, а в другій – тільки половина – тільки ті стани, енергія яких найменша. Наступна зона – повністю пуста.

Відстань між сусідніми рівнями в зоні легко оцінити. Можна вважати, що при зближенні атомів зміна енергії електронів атома приблизно дорівнює теплоті випаровування речовини, розрахованої на один атом. Для металів вона складає декілька електронвольт, а значить, і повна ширина зони ΔW (рис.19.17), яка визначається взаємодією сусідніх атомів, повинна мати той же порядок, тобто $\Delta W \approx 1$ еВ. Для відстані між рівнями отримаємо $\delta W \approx \Delta W/n$, де n – число атомів в зразку. Об'єм, який приходиться на один атом, дорівнює $a^3 \sim 10^{-22}$см^3 (наприклад, для натрію (рис.19.14) $a = 4{,}3 \cdot 10^{-8}$см). Для зразка, об'єм якого 1см^3, $n \approx 10^{22}$. Тому $\delta W \cong 10^{-22}$ еВ.

§19.7.Статистика електронів в кристалах

Однією з проблем, яка виникає при створенні приладів електронної техніки, є отримання (генерація) достатньої кількості рухомих носіїв заряду, які можуть забезпечити необхідні електронні процеси.

1.Функція розподілу електронів

Єдиною можливістю аналізувати поведінку рухомих носіїв заряду є дослідження статистичних закономірностей, які спостерігаються в цьому колективі, тобто дослідити функцію розподілу електронів за енергіями, за швидкостями і т. п.

Функція розподілу електронів за енергіями $N(W)$ задається як відношення кількості електронів в одиниці об'єму dn, що мають енергію в інтервалі від W до $W + dW$, до величини цього інтервалу:

$$N(W) = \frac{dn}{dW} \cdot \qquad (19.3)$$

Ця функція має зміст густини розподілу електронів за енергіями, тобто кількості

електронів в одиничному енергетичному інтервалі в одиниці об'єму. Функція $N(W)$ має розмірність: кількість електронів/$(\text{м}^3 \cdot \text{Дж})$, але, як правило, у фізиці твердого тіла використовується розмірність: кількість електронів/$(\text{см}^3 \cdot \text{еВ})$. Функцію $N(W)$ розглядають як добуток двох функцій $g(W)$ і $f(W)$, де

$$g(W) = \frac{dn_{рівн}}{dW}, \qquad (19.4)$$

$$f(W) = \frac{dn}{dn_{рівн}} \cdot \qquad (19.5)$$

Функція $g(W)$ характеризує густину станів (рівнів), тобто число енергетичних рівнів в одиниці об'єму кристалу, які приходяться на одиничний енергетичний інтервал біля значення енергії W. Тому $dn_{рівн}$ кількість енергетичних рівнів в одиниці об'єму кристалу в інтервалі енергій від W до $W + dW$ (рис.19.18,*а*). Функція $f(W)$ характеризує ймовірність заповнення енергетичних рівнів

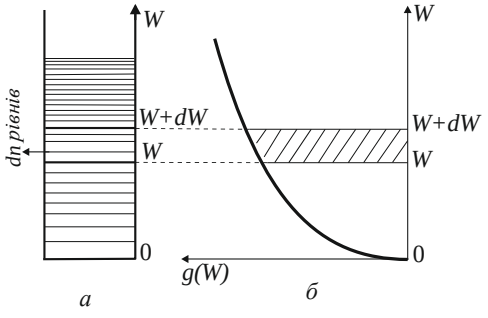


Рис.19.18. Енергетична діаграма зони провідності (*а*) і графік функції розподілу густини рівнів (*б*).

електронами. Зручність такого представлення функції $N(W)$ в тому, що задача знаходження функції розподілу електронів за енергіями розпадається на дві: квантову задачу визначення характеру розподілу енергетичних рівнів і статистичну задачу знаходження ймовірності заповнення цих рівнів.

Функція густини станів (рівнів) для металів при виборі початку відліку енергії від нижнього краю зони провідності (рис.9.18,*а*) має вид:

$$g(W) = \frac{4\pi}{h^3}(2m_n^*)^{3/2}W^{1/2}, \qquad (19.6)$$

де m_n^* – ефективна маса електрона в зоні провідності металу. Для розв'язування задач приймається, що m_n^* в металах дорівнює масі вільного електрона m_e, тобто $m_n^* = 9{,}1 \cdot 10^{-31}$ кг. Функція $g(W)$ є квантовою, так як містить сталу Планка h

Енергія є важливою характеристикою електрона, який рухається в періодичному полі кристалічної решітки металу або напівпровідника. Вона є функцією хвильового вектора, модуль якого $|\vec{k}| = 2\pi/\lambda$, тобто $W = W(\vec{k})$. Рівняння $W(\vec{k}) = \text{const}$ визначає в просторі хвильового вектора $\vec{k}(k_x, k_y, k_z)$ поверхню однакового значення енергії або ізоенергетичну поверхню, форма якої в основному визначає електричні властивості твердих тіл.

Ефективна маса електрона визначається за формулою:

$$m^* = \hbar^2 \left(\frac{d^2W(k)}{dk^2}\right)^{-1}, \qquad (19.7)$$

де $\hbar = \frac{h}{2\pi} = 1{,}054 \cdot 10^{-34}$ Дж · с –стала Планка.

Квантовий розподіл електронів за енергетичними рівнями визначається статистикою Фермі–Дірака і характеризується функцією

$$f(W) = \frac{1}{\exp\left\{\frac{W - W_F}{kT}\right\} + 1}, \qquad (19.8)$$

де T – абсолютна температура, k – стала Больцмана, W_F – енергія Фермі, W – енергія певного рівня. Одиниця в знаменнику враховує принцип Паулі. Із (19.8) витікає, що $f(W_F) = 1/2$ при любій температурі відмінній від

нуля і це формально означає, що енергія Фермі це енергія такого енергетичного рівня (стану), ймовірність заповнення якого електронами дорівнює $1/2$. Це твердження можна розповсюдити і на ті випадки, коли значення W_F попадають в таку область значень енергії, які заборонені в даній квантовій системі. Графік функції Фермі-Дірака (19.8) наведений на рис.19.19.

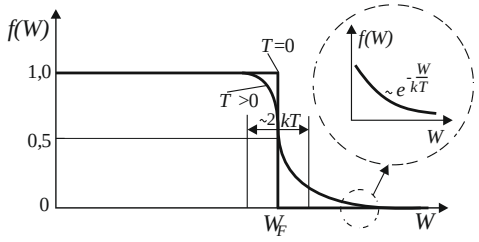


Рис.19.19.

При $T = 0\text{K}$ він має вид розривної ступеневої функції, яка дорівнює 1 в межах значень енергій від 0 до значень W_F і дорівнює 0 при енергіях, які більші W_F. Це означає, що з ймовірністю 1 заповнені електронами всі енергетичні рівні, енергія яких менша величини W_F і не має ні одного електрона в станах, енергія яких перевищує W_F. При підвищені температури частина електронів набуває теплову енергію і повна їх енергія може стати більшою енергії рівня Фермі. На графіку функції розподілу Фермі-Дірака (рис.19.19) це виглядає таким чином, що ступінчастість функції «розмивається» – і тим сильніше, чим вища температура. Аналізуючи (19.8), можна переконатись, що тепловому збудженню піддаються тільки ті електрони, енергія яких відрізняється від W_F на величину порядку kT в більший чи менший бік, тобто на величину порядку теплової енергії.

Вид функції розподілу Фермі-Дірака, а також те, що її значення в точках, які симетричні відносно W_F, в сумі дорівнюють 1, дозволяє представити рівень Фермі, як деяке середнє значення енергії тих електронів, які приймають участь в тепловому русі. Таке уявлення не дуже строге, але часто зручне при якісному розгляді багатьох фізичних явищ. Знаходження функції розподілу електронів в конкретному колективі зводиться фактично до визначення рівня Фермі і його залежності від температури, концентрації носіїв заряду та інших умов. Кількісне значення енергії Фермі може бути знайдене, наприклад, із умови нормування, суть якої в тому, що інтеграл від певної функції розподілу (19.3) по всьому можливому інтервалу енергій дорівнює концентрації електронів, якщо при цьому використовується функція густини енергетичних станів (19.6), яка відноситься до одиниці об'єму, і значення концентрації електронів відоме:

$$n = \int_0^\infty N(W)dW = \\ = \int_0^\infty f(W)g(W)dW. \qquad (19.9)$$

Розрахунок інтеграла (19.9) в явному вигляді можливий тільки в деяких випадках. Як правило приходиться використовувати наближені методи. Розв'язок подібних задач значно спрощується, якщо розглядати область енергії, яка лежить вище рівня Фермі хоч би на декілька одиниць kT. При цьому через сильну залежність (експоненціальну) функції від аргумента $\exp\left\{\frac{W-W_F}{kT}\right\} \gg 1$ і, нехтуючи одиницею в знаменнику функції (19.8), отримаємо розподіл виду

$$f(W) = \text{const} \cdot \exp\left\{-\frac{W}{kT}\right\}. \quad (19.10)$$

Той факт, що в цьому випадку $f(W) \ll 1$, означає, що біля кожного значення енергії кількість станів набагато більша, ніж кількість електронів, які їх заповнюють. Тому особливості, які зв'язані з

квантовою природою частинки (зокрема, з їх спіном), при розподілі за станами не проявляються. Ця ситуація характерна для класичної статистики Максвелла-Больцмана. Електрони, розподіл яких за енергетичними рівнями підпорядковується закону (19.10), називають ***невиродженим газом.*** Відповідно умову $f(W) \ll 1$ можна розглядати як критерій *невиродженості* електронного газу.

Електрони, розподіл яких за енергетичними рівнями підпорядковується закону (19.8) (тобто враховується принцип Паулі), називаються ***виродженим газом.***

2. Статистика електронів в металах

В металічних кристалах всі колективізовані («вільні») електрони енергетично розташовані в зоні провідності. Їх концентрація визначається концентрацією атомів і кількістю колективізованих валентних електронів від кожного атома. Розподіл цих електронів за енергетичними станами зони провідності визначається, по-перше розподілом станів в зоні і, по-друге, ймовірністю їх заповнення. Розподіл станів характеризується функцією густини станів $g(W)$ (19.6).

На рис.19.20 зображена зона провідності металу і функції розподілу $g(W)$, $f(W)$, $N(W) = f(W)g(W)$, які розташовані так, щоб вісь їхнього аргумента співпадала із звичайним напрямком енергії на зонних діаграмах. Із рисунка зрозумілий простий фізичний зміст енергії Фермі, яка для металів відповідає самому верхньому рівню зони провідності, який заповнений електронами при нульовій температурі.

Якщо прийняти, що вид функції (19.6) зберігається не тільки для нижнього краю зони провідності, але і вище, можна скористатись формулою (19.9) і визначити значення енергії Фермі. Це легко зробити для $T = 0$, коли функція $f(W)$ дорівнює 1 для енергій від 0 до W_F і дорівнює 0 для інших значень енергії (рис.19.20,*б*). При цьому підінтегральний вираз

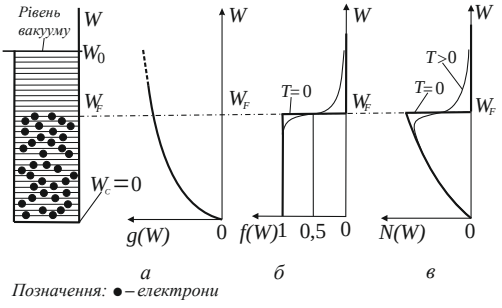


Рис.19.20. Фунцкії: розподілу густини станів (*а*); ймовірності заповнення станів (*б*); густини електронів (*в*) в зоні провідності металів.

у (19.9) значно спрощується і інтеграл вираховується елементарно (дорівнює площі криволінійної трапеції для випадку T=0 на рис.19.20,*в*):

$$n = \int_0^{W_F(0)} g(W)dW = \\ = \frac{4\pi}{h^3}(2m_n^*)^{3/2}\int_0^{W_F(0)} W^{1/2}dW$$

і

$$n = \frac{8\pi}{3h^3}(2m_n^*)^{3/2}\left(W_F(0)\right)^{3/2}. \quad (19.11)$$

Звідси знаходимо значення енергії Фермі при нульовій температурі

$$W_F(0) = \frac{h^2}{2m_n^*}\left(\frac{3n}{8\pi}\right)^{2/3}, \quad (19.12)$$

яка залежить тільки від концентрації електронів. Це стає більш зрозумілим, якщо уявно збільшити кількість електронів в металі. Додаткові електрони будуть розташовуватись на вільних рівнях, які мають енергію більшу W_F, підвищуючи тим самим саме значення W_F. Для типових металів, коли ефективна маса електронів близька до маси спокою електрона

$$W_F(0) = 0{,}365 \cdot 10^{-18} n^{2/3}, \quad (19.13)$$

де n визначається $м^{-3}$, а $W_F(0)-$ в електрон-вольтах і при концентрації електронів порядку $10^{28}-10^{29}\,м^{-3}$ $W_F(0)$ складає величину в декілька електрон-вольт.

При температурах, відмінних від 0, інтегрування (19.9) в загальному вигляді має вид

$$n=\frac{4\pi}{h^3}(2m_n^*)^{3/2}\cdot \cdot\int_0^\infty W^{1/2}(\exp\{\frac{W-W_F}{kT}\}+1)^{-1}dW. \quad (19.14)$$

Це означає, що рівень Фермі залежить від температури (рис.19.21). При низьких температурах поведінка $W_F(T)$ описується приблизним виразом:

$$W_F(T)\cong W_F(0)\left(1-\frac{\pi^2}{12}\left(\frac{kT}{W_F(0)}\right)^2\right), \quad (19.15)$$

який отримується при розкладі підінтегральної функції у (19.14) в ряд.

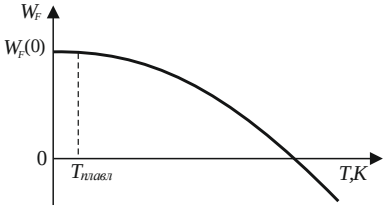


Рис.19.21.

При високих температурах, коли рівень Фермі зміщується в область негативних значень енергій (рис.19.21), експонента в (19.14) стає набагато більшою одиниці, що дає можливість перейти до статистики Максвелла-Больцмана і вирахувати інтеграл в явному виді. Вираз для рівня Фермі, який отримується в цьому випадку

$$W_F(T)=kT\ln\left(\frac{n}{2}\left(\frac{h^2}{2\pi m_n^* kT}\right)^{3/2}\right), \quad (19.16)$$

характерний для невиродженого колективу.

Кількісна оцінка показує, що при температурах біля 1000 К рівень Фермі зміщується всього на величину $(10^{-3}\ldots 10^{-4})\cdot W_F(0)$, а (19.16) стає справедливим при температурах порядку $10^4\ldots 10^5$К. Тому можна вважати, що для металів, які знаходяться в твердому стані, електронний газ є завжди виродженим, а рівень Фермі практично не залежить від температури і може бути розрахований за формулою (19.13) чи при необхідності за більш точною формулою (19.15).

З функцією розподілу (19.8) пов'язана ще одна особливість поведінки електронів в металах, яка проявляється в тому, що їх розподіл за енергіями визначається перш за все не тепловим рухом, а принципом Паулі. Це підтверджується виразом для середньої енергії одного електрона:

$$<W>=W_\Sigma/n= =\int_0^\infty WN(W)dW/\int_0^\infty N(W)dW, \quad (19.17)$$

де $W_\Sigma-$ сумарна енергія всіх "вільних" електронів.

При $T=0$K вираз (19.17) значно спрощується і

$$<W_0>=\frac{\int_0^{W_F(0)}Wg(W)dW}{\int_0^{W_F(0)}g(W)dW}= =3W_F(0)/5. \quad (19.18)$$

В більш загальному випадку, якщо розкласти підінтегральні вирази у (19.17) в ряди, обмежившись членами, які містять T^2, то можна отримати

$$<W>= =\frac{3}{5}W_F(0)\cdot\left(1+\frac{5\pi^2}{12}\left(\frac{kT}{W_F(0)}\right)^2\right). \quad (19.19)$$

Перший доданок у (19.19) є (19.18) і характеризує середню енергію при нульовій температурі і, відповідно, називається «нульовою» енергією, що має чисто квантову природу. Другий доданок у (19.19) залежить від температури і величина

$$< W(T) > \cong \frac{\pi^2 (kT)^2}{4W_F(0)} \qquad (19.20)$$

є середня теплова енергія електронів. Порівняння величин (19.18) і (19.20) показує, що навіть при температурах, які близькі до температури плавлення металу, теплова енергія електронів складає 10^{-3} ... 10^{-4} від «нульової енергії», а відповідна теплова швидкість – одиниці відсотків від «нульової швидкості». Така слабка сприйнятливість електронів в металах до нагрівання витікає із самого характеру функції розподілу електронів. На рис.19.22 приведена функція Фермі-Дірака для електронів зони провідності срібла ($W_F =$ 5,5 еВ) для температури 290К ($kT = 0{,}025$ еВ) і для температури плавлення 1233К ($kT = 0{,}106$ еВ).

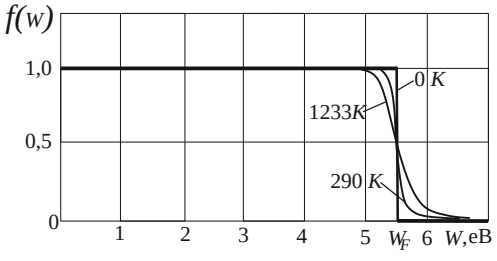


Рис.19.22.

Із рисунка видно, що зміни функції при нагріванні дуже незначні і тільки ті електрони, енергія яких відрізняється від енергії Фермі на величину порядку 0,1 еВ, набувають теплову енергію, приймаючи участь в тепловому русі. Ці електрони складають невелику частку порядку $kT/W_F(0)$ від усіх електронів, тому і теплова енергія всього колективу електронів набагато менша його повної енергії. Ця обставина пояснює, зокрема, малий вклад теплоємності електронного газу в загальну теплоємність металів. Крім того, рис.19.22 показує, що зглажування ступінчастості функції Фермі-Дірака невелике навіть при високих температурах і границя між зайнятими і вільними станами зберігається достатньо різкою. Тому зміст рівня Фермі, який пов'язаний з цією границею, можна приблизно зберегти не тільки для нульової температури.

Звернемо увагу на те, що розподіл електронів за енергіями в зоні провідності металів подібний до розподілу густини молекул рідини з висотою. Процес переходу електронів на більш високі енергетичні рівні при нагріванні аналогічний процесу випаровування молекул із рідини. З цієї точки зору більш правильним для електронів в металі була би назва «електронна рідина» (рівень Фермі електронів відповідає висоті рівня поверхні рідини). Така аналогія характерна для електронного колективу в цілому, а та частина електронів, енергія яких перевищує енергію рівня Фермі на декілька одиниць kT, аналогічна парі над поверхнею рідини і може бути названа "електронним газом", до якого з достатньою точністю може бути застосована статистика Максвелла-Больцмана. Особливу групу складають найбільш «гарячі» електрони, енергія яких на декілька електрон-вольт більша енергії Фермі. Не дивлячись на те, що ця група дуже малочисельна (наприклад, ймовірність заповнення електронами рівнів на 1 еВ вище рівня Фермі при $T = 1160$К не перевищує величину $\exp\{-10\} = 4{,}5 \cdot 10^{-4}$), але саме ці електрони приймають участь в такому, наприклад, явищі, як термоелектронна емісія.

§19.8.Електропровідність металів та її залежність від температури. Надпровідність

Класична теорія

Опір провідника протіканню струму зовсім не така проста річ, як може здатися на перший погляд. Тут не обійтись без поняття кванта. Але спочатку спробуємо обійтись без його допомоги.

В класичній теорії не використовується поняття квант. У багатьох книжках і підручниках написано, що електричний опір металів є результатом зіткнення електронів, які переносять заряд, з іонами кристалічної гратки. Таке розуміння опору металів виражає класична теорія. Проте, вона не може пояснити ряд основних особливостей електричного опору. В класичній теорії безумовно справедливе одне: опір виникає від того, що електрони передають частину своєї енергії і імпульсу (кількості руху) кристалічній решітці. Але яким чином відбувається ця передача – питання зовсім не просте.

Температурна залежність. Опір чистих металів сильно зростає з температурою. Для багатьох із них зростання в основному пропорційне абсолютній температурі. При низьких температурах така проста залежність порушується. Це добре видно із рис.19.23, де

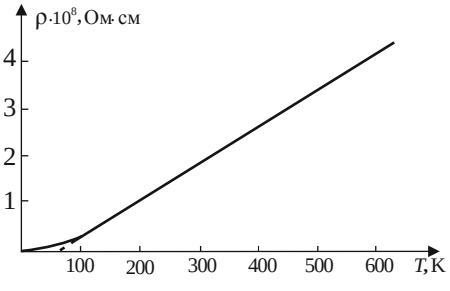


Рис.19.23.Графік залежності питомого опору міді від абсолютної температури.

показано, як змінюється з температурою питомий опір міді, провідника, який найбільш широко використовується для передачі струму.

Якщо пояснювати існування електричного опору зіткненнями електронів з іонами, то температурна залежність опору чистих металів стає зовсім незрозумілою. Справа виглядає так, начебто електрони зіштовхуються тільки з іонами (атомами), які здійснюють тепловий рух, але вільно пролітають мимо нерухомих іонів. Виглядає так, що електрону легше попасти в іон, який рухається. Звичайно, що це не правильно. Таким чином, класична теорія не здатна пояснити залежність $\rho(T)$ для металів

Залишковий опір. Коли вимірюють опір чистих металів при дуже низьких температурах, то виявляється цікавий факт. Продовжуючи криву $\rho(T)$ (аналогічно як на рис.19.23) до абсолютного нуля, отримують опір, який називають залишковим опором. Виявляється, що він сильно змінюється від зразка до зразка. Залишковий опір дуже чутливий до невеликої кількості домішок, до механічної і термічної обробки металу. Очищаючи метал, та обробляючи його так, щоб добитись майже досконалої кристалічної структури, можна зменшити залишковий опір, і не має меж цьому зменшенню. І так досвід підштовхує до такого висновку: ідеальний кристал при температурі абсолютного нуля не повинен мати електричного опору. Іншими словами: електричний опір чистих металів обумовлений порушеннями кристалічної гратки, які викликаються тепловим рухом іонів, домішками, дефектами (неправильностями) кристалічної решітки.

Атомні коридори. Якби електрони рухались за законами класичної механіки, вони повинні були би зіштовхуватись з іонами

(атомами) незалежно від того, розташовані атоми в строгому порядку (як в кристалі), чи по іншому. Звичайно, між правильними рядами атомів є начебто коридори (рис.19.24), але щоб направити електрон по такому коридору, необхідна точна орієнтація кристала по відношенню до електричного поля, яке прикладене до зразка. Але, як правило, мають справу не з одним кристалом, а із великою кількістю дрібних кристалів, які орієнтовані випадковим чином (рис.19.25).

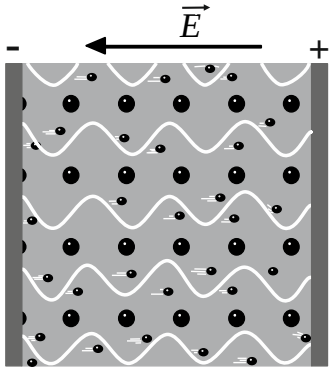


Рис.19.24. Рух електронів і розповсюдження електронних хвиль в ідеальному кристалі. Позначення: ● – іони кристалічної решітки; ● – електрони.

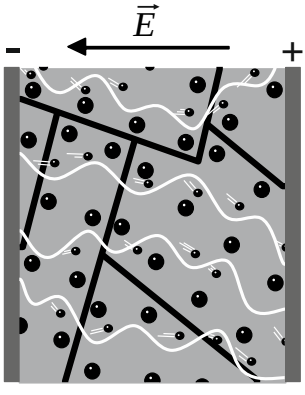


Рис.19.25. Схема руху електронів атомними лабіринтами в полікристалічному зразку. Прямі лінії ▬ – границі мікрокристалів; ● – іони кристалів; ●– електрони.

Одинокий правильний кристал називають монокристалом, а з'єднання багатьох кристалів – полікристалічним агрегатом. Точні вимірювання показують, що опір полікристалічних агрегатів дуже мало відрізняється від опору монокристалів. Особливо важливо те, що опір полікристалічних агрегатів сильно зменшується із зниженням температури. Це не можна пояснити рухом електронів атомними коридорами за законами класичної механіки. При переході із одного кристала в інший атомний коридор різко змінює свій напрямок (рис.19.25) і електрон повинен був би з розгону вдаритись об "стіну". А він продовжує свій рух атомним лабіринтом.

Деяка цікава інформація. Давно відомо, що механічна і термічна обробка суттєво впливає на електричний опір металів. При 20 °С звичайна технічна мідь має питомий опір 17,2 нОм·м. Після холодного протягування питомий опір мідного дроту зростає до 17,7 нОм·м. Навіть намотка проводу на котушку призводить до зростання його опору. Якщо провід відпалити, тобто довгий час нагрівати, а потім знову охолодити до 20°С, то значення опору повертається до нормальної величини. Очевидно, що опір чутливий до невеликих порушень кристалічної структури. Ще більше опір чутливий до мізерної кількості домішок. Старанне очищення зменшує питомий опір міді при t =20°С до 16,9 нОм·м. Ця обставина має велике значення для техніки, так як вона дозволяє зменшити опір проводів і таким чином зменшити непотрібні витрати електроенергії на нагрівання. Якщо до міді додати 1% марганцю, то питомий опір її зростає до 48 нОм·м, тобто майже в 3 рази, і дорівнює опору чистого марганцю (50 нОм·м). Таким же чином впливають добавки заліза, кобальту, іридію та інші. Якби опір виникав від зіткнення електронів з атомами домішки, то ці домішки повинні були би впливати у сто раз

слабше. З точки зору класичної теорії такий сильний вплив домішок на опір пояснити неможливо.

Спеціально розроблені сплави з високим питомим опором: нікелін, манганін, константан, ніхром та інші. Опір цих сплавів в декілька раз більший, ніж у їхніх складових. Так константан, який містить 60% міді і 40% нікелю має питомий опір 440 нОм·м, в той час як у чистої міді він дорівнює 17 нОм·м, а у нікелю −72 нОм·м. У ніхрома питомий опір біля 1000 нОм·м. Температурна залежність опору сплавів значно слабша, ніж у чистих металів. Наприклад, у константана в інтервалі температур від 0° до 400°С питомий опір змінюється тільки від 441 до 448 нОм·м.

З точки зору класичної теорії постійний опір сплавів мало зрозумілий так же, як і пропорційність між опором і температурою для чистих металів. Опір сплавів повинен був би за змістом цієї теорії складатись із опору складових частин за простим правилом змішування, а опір чистих металів пропорційний $\sqrt{T}$.

Квантова теорія

Для теоретичного обґрунтування залежності питомого опору металів від температури візьмемо за основу одне з основних тверджень класичної електронної теорії електропровідності, що

$$\sigma = enu, \qquad (19.21)$$

де $\sigma = 1/\rho$ – питома електропровідність металу (ρ – питомий опір), n – концентрація електронів провідності (так званих «вільних» електронів), u – рухливість електронів.

Вираз (19.21) має дві величини, які принципово можуть залежати від температури – це n і u.

Згідно зонної теорії металів, концентрація електронів провідності (електронів в зоні провідності) визначається при температурі $T = 0$ за формулою (19.11). Для металів енергія Фермі практично не залежить від температури. Тому можна вважати, що концентрація «вільних» електронів n не залежить від температури, тобто є сталою. Залишається вияснити залежність від температури рухливості електронів.

Згідно квантової механіки електрон має одночасно властивості частинки і хвилі (співвідношення де Бройля $\lambda = h/(m_e v)$. Ширина атомних коридорів близька до довжини хвилі де Бройля електрона, і тому при русі такими коридорами хвильові властивості електрона проявляються в повній мірі.

В радіотехніці сантиметрових хвиль широко застосовуються хвилеводи. Це – металічні трубки круглого або прямокутного перерізу. Радіохвилі розповсюджуються по хвилеводу, ідучи за всіма його згинами. Світло – теж електромагнітні хвилі, і його можна передавати світловодами (трубки із дзеркальними стінками). Дія світловода із дзеркальними стінками зрозуміла. Світлові хвилі весь час відбиваються від стінок і, таким чином, залишаються всередині світловода. Радіохвилі теж добре відбиваються від гладких металічних поверхонь. Якщо поверхня шерстка, то хвиля якби розривається об зазубринки і в результаті розсіюється і поглинається. При цьому енергія хвилі переходить в тепло.

У квантовій теорії протікання електричного струму через метал представляється як розповсюдження електронних хвиль атомними коридорами, які грають роль хвилеводів (рис.19.24 і 19.25). Якщо атоми розташовані на площині в

ідеальному порядку, на однакових відстанях один від одного, то така площина повністю відбиває електронні хвилі – як ідеальне дзеркало. Розсіювання і поглинання електронних хвиль відбувається тільки при порушенні строгого порядку в розташуванні атомів. Цей висновок вперше був отриманий досить складним математичним шляхом із рівнянь квантової механіки. Розглянемо якісно причину існування електричного опору. В ідеальному правильному кристалі хвилеводи досконало гладкі. Всякі порушення атомного порядку діють як шерсткість стінок хвилеводу. Електронні хвилі розсіюються порушеннями гратки. Частина енергії електронів поглинається і переходить в тепло. Це і є *квантовий механізм електричного опору*.

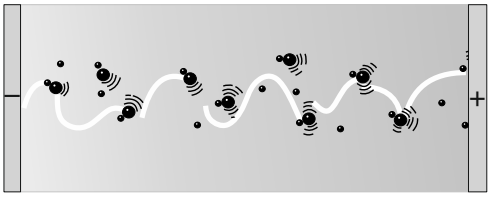


Рис.19.26. Порушення правильності атомного хвилеводу тепловим рухом атомів. Позначення: ● – іони кристалічної решітки; ● – електрони.

Правильність атомних хвилеводів порушується тепловими коливаннями, на яких розсіюються електронні хвилі (рис.19.26). Звідси виникає теплова частина електричного опору, яка залежить від температури. Охолоджуючи метал, її можна зробити скільки завгодно малою. При цьому залишається залишковий, або структурний опір, зв'язаний з постійними дефектами (порушеннями) кристалічної структури. Дефекти є у всякого кристала. Додаткові дефекти виникають при холодній механічній обробці – протягуванні проводу і навіть при намотці його на котушку. Від цього і збільшується опір. При тривалому нагріванні (відпалі) атоми повертаються на свої місця (дефекти заліковуються) – опір зменшується.

Проникнення чужеродних атомів викликає серйозні порушення кристалічної решітки. При цьому один атом домішки може вибити з місця сотні атомів кристала – «господаря». Тому – то домішки і підвищують опір. Встановлено, що дія домішки тим сильніша, чим більше її атоми відрізняються від атомів «господаря» за розмірами та іншими факторами, чим сильніше вони порушують правильність кристалічної решітки.

Якщо охолоджувати метал, то теплова (причиною якої є тепловий рух) частина опору зменшується. Коли вона стає значно меншою, ніж структурна частина, яка не залежить від температури, то при подальшому охолодженні опір майже не змінюється. У сплавів з невпорядкованою будовою структурна частина опору настільки велика, що навіть при високих температурах більша від теплової. Цим і пояснюється слабка залежність електричного опору сплавів від температури.

Пояснення температурної залежності опору. Тепловий рух в твердому тілі (металі) можна уявити як коливання атомів (іонів) біля своїх положень рівноваги. Але частинки в твердому тілі не можуть коливатись незалежно. Вони всі міцно зв'язані між собою і здійснюють колективні коливання. Колективні коливання великої кількості частинок речовин – це теж саме, що звукові хвилі. Якщо вивести будь-яку частинку із положення рівноваги, то збурення буде розповсюджуватися по тілу із швидкістю звуку. Виявляється, що тепловий рух у твердому тілі можна розглядати як розповсюдження звукових хвиль. Нагріваючи тіло, ми примушуємо його звучати. На щастя для нас, частота цих коливань в міліарди раз вища тих, які може чути наше вухо.

Так з'являється зв'язок між електричним опором і звуком. Електрони розсіюються на звукових хвилях, які збуджуються тепловим рухом іонів кристалічної гратки. Але всяка хвиля представляється як *кванти.* У звукової хвилі це фонони, які здатні існувати тільки всередині речовини (у вакуумі фононів бути не може). Якщо звукову хвилю описувати як потік фононів, то і електрони можна вважати частинками, забувши про їх хвильові властивості. Тепер ми повертаємося до простого пояснення опору: електронам заважає вільно рухатися те, що вони зіштовхуються з іншими частинками – звуковими квантами (фононами).

Будемо вважати для простоти, що всі фонони мають однакову частоту ν. Це припущення неправильне, але помилка із-за нього невелика. Енергія фонона, як і всякого кванта, дорівнює $h\nu$, де h – стала Планка, а ν – частота хвилі. Тоді число фононів буде

$$N = \frac{W}{h\nu} = \frac{kT}{h\nu}, \qquad (19.22)$$

де $W=kT$ енергія теплового руху, яка пропорційна абсолютній температурі T, k – стала Больцмана.

Число фононів в твердому тілі пропорційне абсолютній температурі. Значить рухливість електронів у формулі (19.21) буде обернено пропорційна числу фононів. Тобто:

$$u \sim \frac{1}{N} \sim \frac{1}{T}.$$

Це означає, що $\sigma \sim 1/T$, а питомий опір металів буде попорційний абсолютній температурі. Тому можемо записати, що

$$\rho = \rho_0(1 + \alpha t) = \alpha\rho_0 T, \quad (19.23)$$

де t – температура в градусах Цельсія (℃), ρ – питомий опір при температурі t, ρ_0– питомий опір при $t = 0$ ℃, $T=(t + 273)K$, α – температурний коефіцієнт опору, що дорівнює:

$$\alpha = \frac{1}{\rho}\frac{d\rho}{dt}. \qquad (19.24)$$

α *характеризує відносну зміну питомого опору при зміні температури на один градус.*

Зазначимо, що цей результат є справедливий тільки при достатньо високих температурах. Виходячи із змісту квантової теорії число фононів не може бути меншим одиниці. Подивимось, при якій температурі число фононів дорівнює одиниці. Цю температуру називають температурою Дебая і позначають грецькою буквою θ. Прийнявши $N = 1$, із формули (19.22) знаходимо, що

$$\theta = h\nu/k. \qquad (19.25)$$

Температура Дебая (або дебаєвська температура) відіграє дуже важливу роль у фізиці твердого тіла. За її допомогою число фононів, яке нас цікавить, виражається зовсім просто:

$$N = T/\theta. \qquad (19.26)$$

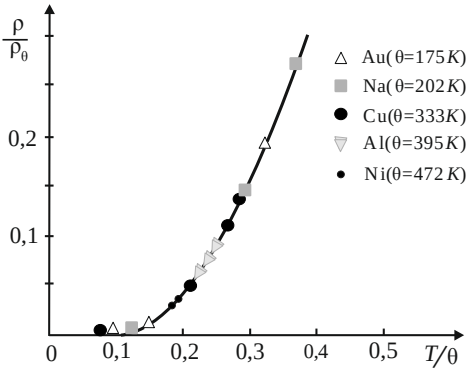


Рис.19.27. Узагальнена крива температурної залежності питомого опору чистих металів.

При температурах, нижчих дебаївської, число $N < 1$. Це означає, що можна говорити тільки про середнє число фононів і розрахунок $\rho(T)$ ускладнюється. Але величина

T/θ, як і раніше, залишається мірою числа фононів в твердому тілі, а значить, і числа зіткнень електронів з фононами.

Виходячи з цього, була висунута наступна гіпотеза: температурна залежність електричного опору чистих металів виражається універсальною функцією від міри числа фононів T/θ. Математично ця гіпотеза записується так:

$$\frac{\rho}{\rho_\theta} = F\left(\frac{T}{\theta}\right), \qquad (19.27)$$

де ρ – опір при температурі T, ρ_θ – питомий опір при температурі θ, F – функція, вид якої може бути знайдений із досліду. Важливо те, що функція $F\left(\frac{T}{\theta}\right)$ – універсальна для різних чистих металів. Як добре ця гіпотеза підтверджується на практиці видно із рис.19.27.

Надпровідність

У 1911р. Г. Камерлінг-Оннес відкрив явище стрибкоподібного перетворення в нуль електричного опору деяких металів при досить низьких температурах. Вперше це було виявлено в дослідах із ртуттю при температурах $T \leqslant 4{,}2$K. Сучасні методи вимірювання дають підставу зробити висновок, що питомий опір провідника в такому стані менший від 10^{-25}Ом·м. Явище перетворення в нуль опору деяких металів при низьких температурах називають *надпровідністю*. Температуру $T_{кр}$, нижче якої спостерігається перехід металів у надпровідний стан, називають *критичною.*

Опір до переходу в надпровідний стан може бути різним. Перехід у надпровідний стан відбувається завжди досить різко. Тепер відомо, що майже половина металевих елементів може переходити у надпровідний стан. Понад тисячу сплавів і сполук металів також здатні переходити в надпровідний стан за певних умов. Речовини, які перебувають у надпровідному стані, мають ряд специфічних властивостей. Основна з них – це відсутність електричного опору. В 1914 р. Г.Камерлінг-Оннес виявив, що надпровідний стан руйнується магнітним полем, коли магнітна індукція $\vec{B}$ перевершує деяке критичне значення $\vec{B}_{кр}$. Величина $\vec{B}_{кр}$ залежить від природи надпровідника та його температури. Критичне поле, яке руйнує надпровідність, може бути створене і самим струмом у надпровіднику. Тому існує критичне значення сили струму, при якому напровідність зникає. Крім цього, в 1933р. В.Мейсснер і Р. Оксенфельд встановили, що при температурі, нижчій від критичної, відбувається виштовхування магнітного поля з надпровідника. Це явище називають ефектом Мейсснера. Виштовхування магнітного поля із надпровідника означає, що в ньому магнітна індукція $\vec{B}$ дорівнює нулю. Фізично ефект Мейсснера зв'язаний з тим, що в надпровіднику, розміщеному в слабкому магнітному полі, у поверхневому шарі завтовшки $(100 \div 1000) \cdot 10^{-10}$м створюються такі колові незгасаючі струми, які повністю компенсують зовнішне магнітне поле.

Явище надпровідності частіше спостерігається в сплавах, ніж у чистих речовинах. Для чистих речовин має місце повний ефект Мейсснера. У сплавах відбувається часткове виштовхування магнітного поля з об'єму. Речовини, в яких проявляється повний ефект Мейсснера, називають *надпровідниками першого роду*, а речовини, в яких не повністю має місце ефект Мейсснера, *надпровідниками другого роду*. В надпровідниках першого роду надпровідність у певному критичному полі зникає раптово, і металевий провідник переходить у нормальний

стан. У надпровідниках другого роду, існування яких передбачив у 1952 р. О.О.Абрикосов, є два критичних магнітних поля – нижнє $H_н$ і верхнє $H_в$. У магнітних полях $H < H_в$ має місце повний ефект Мейсснера. Для полів $H_н < H < H_в$ має місце частково ефект Мейсснера, а при $H > H_в$ надпровідність зовсім зникає.

Мікроскопічну теорію надпровідників першого роду створили Дж. Бардін, Л. Купер, Дж. Шріффер. Теорію надпровідників другого роду розробили О. О. Абрикосов та Л. П. Горьков. Явище надпровідності можна пояснити тільки в рамках квантової механіки.

Відповідно до теорії Бардіна, Купера і Шріффера виникнення надпровідності зумовлене взаємодією електронів провідності з кристалічною решіткою. Надпровідність є колективний квантовий ефект електронів, що взаємодіють між собою внаслідок посередницької ролі решітки. Окремий електрон при своєму русі поляризує (деформує) решітку. Інший електрон, потрапляючи в поле поляризованої решітки, змінює свій стан. Внаслідок цього змінюється його взаємодія з першим електроном. Поляризація решітки рухомим електроном зумовлює її коливання. Індукована коливаннями решітки взаємодія між електронами в деяких випадках має характер притягання електронів, які перебувають в станах близьких до рівня Фермі. Два таких електрони, притягуючись, утворюють своєрідний стан, який називають *куперівською парою*. Між електронами провідності діють сили електростатичного відштовхування, тому утворення куперівських пар можливе у тому випадку, коли притягання електронів внаслідок електрон-фононної взаємодії переважає їх кулонівське відштовхуваня. Істотним є те, що куперівська пара утворюється з електронів з рівними і протилежно спрямованими імпульсами і антипаралельними спінами. Спін такої пари дорівнює нулю. Зазначимо, що при температурі $T < T_{кр}$ існує не одна локалізована пара і не система невзаємодіючих пар, а сукупність взаємодіючих одна з одною пар електронів. Зв'язок між електронами буде тим сильнішим і стійкішим, чим частіше відбуватимуться процеси обміну фононами між електронами. Але при такому процесі обміну електрони, які утворюють пару, змінюють свої імпульси і, щоб утворити пару з нульовим імпульсом, повинні «підшуковувати» собі нових «партнерів». Це означає, що пари електронів неперервно обмінюються між собою електронами, тобто весь час відбувається зникнення одних пар і утворення нових. Тому всі електрони в системі зв'язані між собою.

Отже, внаслідок електрон-фононної взаємодії при $T < T_{кр}$ відбувається перебудова станів та рівнів енергії системи електронів і утворюється надпровідний стан. Це дало змогу стверджувати, що такий тип взаємодії міг би дати основний стан з меншою енергією, ніж повністю невзаємодіючі електрони, і, що такий надпровідний основний стан має відмежуватись енергетичною щілиною від стану з нормальною провідністю. Енергетична щілина розміщена близько до рівня Фермі. Наявність скінченної енергетичної щілини у спектрі надпровідного стану пояснює явище надпровідності. Зауважимо, що безпосередні вимірювання заряду носіїв струму у надпровідних вказують на те, що він дорівнює заряду двох електронів.

Якщо електрон-фононні взаємодії істотні, то температура $T_{кр}$ пропорційна температурі Дебая θ для ізотопу даного хімічного елемента. Це означає, що $T_{кр}$ залежатиме від маси атома ізотопу *M*, а саме

$T_{кр} \sim M^{1/2}$, що підтверджено експериментально.

Однією із основних проблем у фізиці надпровідності є проблема підвищення температури переходу в надпровідний стан, тобто створення таких матеріалів, температура переходу яких у надпровідний стан була б значно вищою за температуру рідкого азоту (77К). Майже за 60 років з часу відкриття явища надпровідності вдалося підвищити температуру переходу в надпровідний стан від 4,2 до 23,2К (Nb_3Ge). У 1986р. виявлено надпровідність у системах $Ba-La-Cu-O$ при температурах $30 \div 35$ К. У 1987р. групою фізиків США на кераміці $Y-Ba-Cu-O$ досягнута температура переходу в надпровідний стан порядку 92К. У 1988р. синтезовані нові вісмутові і талієві сполуки з критичною температурою 100 і 125К. Нині провідні лабораторії світу ведуть пошуки шляхів створення нових матеріалів, які задовольняли б вимогам високотемпературної надпровідності. Успішне розв'язання цієї проблеми матиме надзвичайно важливе значення для науково-технічного прогресу.

§19.9.Статистика електронів і дірок в напівпровідниках

В напівпровідниках колективізовані електрони приймають участь в утворенні ковалентних зв'язків між атомами і при цьому, будучи зв'язані обмінними силами з кристалічною решіткою, не можуть приймати участь в електропровідності. Щоб електрони стали рухливими в електричному полі, вони повинні подолати сили зв'язку, отримавши певну енергію (рис.9.11). На мові енергетичних діаграм це означає, що колективізовані електрони знаходяться в валентній зоні і, щоб стати рухливими, вони повинні перейти в зону провідності, в якій є вільні незаповнені енергетичні стани. Для цього електронам необхідно надати енергію не меншу, ніж величина ширини забороненої зони. Цю енергію вони можуть отримати від різних джерел. Зараз ми будемо розглядати генерацію рухливих носіїв заряду (електронів і дірок), які виникають під дією теплового руху.

Розрахувати концентрації електронів в зоні провідності і дірок в валентній зоні можна, знаючи функції розподілу густини електронів $N_n(W) = g_n(W) \cdot f_n(W)$ і дірок $N_p(W) = g_p(W) \cdot f_p(W)$ для зони провідності і валентної зони, відповідно.

Функція густини дозволених енергетичних рівнів для електронів в зоні провідності напівпровідника визначається за формулою

$$g_n(W) = \frac{4\pi(2m_n^*)^{3/2}}{h^3} \cdot (W - W_c)^{1/2}, \quad (19.28)$$

а – для дірок у валентній зоні –

$$g_p(W) = \frac{4\pi(2m_p^*)^{3/2}}{h^3} \cdot (W_v - W)^{1/2}. \quad (19.29)$$

В цих формулах m_n^* – ефективна масса електронів в зоні провідності, m_p^* – ефективна маса дірок у валентній зоні, $h = 6{,}62 \cdot 10^{-34}$ Дж·с – стала Планка, W_v – енергія «стелі» валентної зони, W_c – енергія «дна» зони провідності (рис.19.28,*а,б*).

Зауваження. За початок відліку енергії в напівпровідниках прийнято «стелю» валентної зони. Тому будемо вважати, що $W_v = 0$.

Вид функції розподілу Фермі-Дірака на рис.19.28,*в* визначається температурою і рівнем Фермі, положення якого підлягає подальшому уточненню. Функція Фермі-Дірака для електронів $f_n(W)$ визначається за формулою (19.8), а для дірок

$$f_p(W) = 1 - f_n(W) = \frac{1}{\exp\left\{\frac{W_F - W}{kT}\right\} + 1}, \quad (19.30)$$

де W – енергія електрона (дірки) в зоні, W_F –енергія рівня Фермі, k – стала Больцмана, T – абсолютна температура.

Концентрація електронів в зоні провідності може бути розрахована як інтеграл (19.9) в межах від нижнього краю зони провідності до верхнього. Однак для зручності необхідно розширити верхню границю інтегрування до нескінченності, що майже не впливає на точність розрахунків, так як підінтегральна функція під впливом множника $f_n(W)$ дуже швидко зменшується із збільшенням енергії. .

Тому

$$n = \int_{W_c}^{\infty} g_n(W)\, f_n(W) dW, \quad (19.31)$$

$$p = \int_{-\infty}^{W_v} g_p(W) f_p(W) dW. \quad (19.32)$$

Як видно із рис.19.28,*в* для розрахунку концентрації електронів і дірок можна з успіхом користуватися класичною функцією розподілу ймовірності Максвелла-Больцмана (19.10). Це пояснюється тим, що вільних рівнів енергії в зоні провідності і валентній зоні набагато більше, ніж електронів і дірок, які знаходяться там. Тому при розподілі частинок за енергетичними рівнями принцип Паулі можна не враховувати. Функція $N(W)$ для електронів і дірок на рис.19.28,*г* має максимум біля дна зони провідності і стелі валентної зони відповідно. Тому можна вважати, що електрони в зоні провідності в основному знаходяться біля його дна, а дірки – біля стелі валентної зони.

Концентрація носіїв заряду n і p в напівпровідниках чисельно дорівнює площам заштрихованих фігур на рис.19.28,*г*. Для розрахунку n і p зручно перейти до нової змінної $x = \frac{W - W_c}{kT}$ для електронів і $x = \frac{W_v - W}{kT}$ для дірок. Врахувавши, що інтеграл $\int_0^{\infty} \sqrt{x} \exp\{-x\} dx = \sqrt{\pi}/2$, отримаємо:

$$n = N_c \exp\left\{-\frac{W_c - W_F}{kT}\right\} \quad (19.33)$$

$$p = N_v \exp\left\{-\frac{W_F - W_v}{kT}\right\}, \quad (19.34)$$

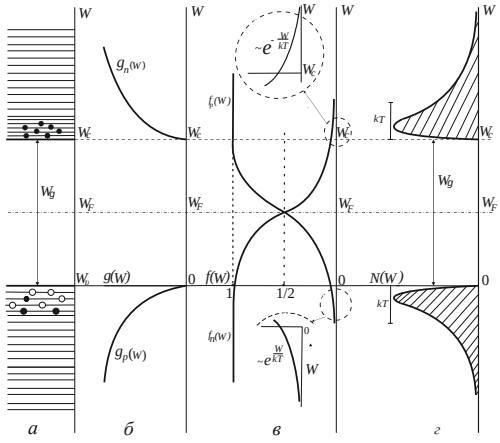


Рис.19.28.*а* – енергетична діаграма напівпровідника, світлими кружечками позначені дірки, чорними – електрони; *б* – графіки функцій розподілу густини рівнів для електронів $g_n(W)$ і дірок $g_p(W)$; *в* – графіки функцій розподілу ймовірності заповнення енергетичних рівнів електронами $f_n(W)$ і дірками $f_p(W) = 1 - f_n(W)$; *г* – графіки функцій $N(W$ розподілу густини електронів в зоні провідності і дірок у валентній зоні. W_g – ширина забороненої зони, W_F – енергія рівня Фермі, W_c – енергія «дна» зони провідності, W_v – енергія «стелі» валентної зони (нуль відліку енергії).

де

$$N_c = \frac{2}{h^3}(2\pi m_n^* kT)^{3/2} = 4{,}83 \cdot 10^{21} \left(\frac{m_n^*}{m_e} T\right)^{3/2}, \quad (19.35)$$

$$N_v = \frac{2}{h^3}\left(2\pi m_p^* kT\right)^{3/2} = 4{,}83\cdot 10^{21}\left(\frac{m_p^*}{m_e}T\right)^{3/2}. \quad (19.36)$$

Тут $m_e = 9{,}1\cdot 10^{-31}$кг – маса вільного електрона, T – абсолютна температура, N_c і N_v визначаються в м$^{-3}$. Простий зв'язок між концентраціями електронів і дірок і положенням рівня Фермі, який виражається формулами (19.33) і (19.34), дуже зручний при аналізі багатьох ситуацій в напівпровідниках.

Величини N_c і N_v називають *ефективними густинами станів в зоні провідності і валентній зоні* відповідно. Цей термін використовується тому, що для класичного (невиродженого) розподілу концентрація електронів в точності дорівнює N_c помноженій на больцманівську функцію розподілу ймовірності (19.10) при енергії рівня $W = W_c$. Хоч в дійсності електрони в зоні провідності, а дірки у валентній зоні розподілені в енергетичному інтервалі, що дорівнює декільком величинам kT, як показано на рис.19.28,*г*. Таким чином зону провідності (і аналогічну валентну зону) можна представити як зону енергій, що складається із $N_c(N_v)$ рівнів, що мають однакову енергію $W_c(W_v)$.

Із співвідношень (19.33) і (19.34) отримаємо, що

$$\frac{p}{n} = \left(\frac{m_p^*}{m_n^*}\right)^{3/2} \exp\left\{\frac{2\left(\frac{W_c+W_v}{2}-W_F\right)}{kT}\right\}. \quad (19.37)$$

Звідси видно, що співвідношення між концентраціями електронів і дірок залежить від їхніх ефективних мас і особливо сильно (*експоненціально*) від положення рівня Фермі відносно середини забороненої зони $(W_c + W_v)/2$. За рахунок невеликого зміщення рівня Фермі можна отримати зміну співвідношення p/n на декілька порядків. Цей перерозподіл концентрацій рухливих носіїв заряду між зонами може бути здійснений введенням в напівпровідник домішок, так як положення рівня Фермі сильно залежить від типу і концентрації домішок.

Закон діючих мас. Із співвідношень (19.33) і (19.34) отримаємо, що добуток концентрацій рухливих електронів і дірок не залежить від положення рівня Фермі і при даній температурі залежить тільки від ширини забороненої зони $W_g = W_c - W_v$ напівпровідника і ефективних мас електронів і дірок :

$$np = N_c N_v \cdot \exp\{-W_g/(kT)\}. \quad (19.38)$$

Із співвідношення (19.38) видно, що збільшення концентрації рухливих носіїв заряду одного типу можливе тільки за рахунок зменшення кількості рухливих носіїв іншого типу. Прологарифмувавши (19.37) знайдемо положення рівня Фермі у напівпровіднику:

$$W_F = \frac{1}{2}(W_c+W_v) - \frac{1}{2}kT\ln\frac{p}{n} + \frac{3}{4}kT\ln\frac{m_p^*}{m_n^*}. \quad (19.39)$$

§19.10. Електропровідність напівпровідників та її залежність від температури

В загальному випадку електропровідність напівпровідника визначається за такою формулою:

$$\sigma = enu_n + epu_p, \quad (19.40)$$

де n – концентрація електронів провідності, p – концентрація дірок провідності, u_n і u_p – рухливості електронів і дірок, відповідно, e – заряд електрона за модулем.

Власний напівпровідник. Це напівпровідник з ідеальною кристалічною решіткою, яка не містить ні домішок, ні яких-небудь дефектів. Енергетична діаграма такого напівпровідника приведена на рис.19.29. При

температурі $T \neq 0$ носії заряду: електрони в зоні провідності (ЗП) і дірки у валентній зоні (ВЗ) появляються внаслідок «переходу» електронів із ВЗ в ЗП (процеси 1 на рис.19.29 – процеси генерації).

Одночасно мають місце і зворотні процеси – процеси рекомбінації електронів провідності і дірок (електрони заповнюють розірвані зв'язки, процеси 2 рис.19.29). При сталій температурі між процесами 1 і 2 встановлюється рівновага.

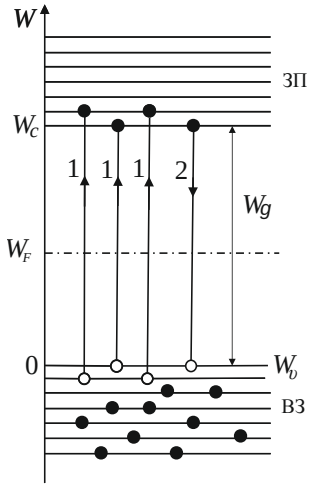


Рис.19.29.Енергетична діаграма власного напівпровідника. Світлі кружечки – дірки, темні кружечки електрони. Температура $T \neq 0$. ЗП – зона провідності, ВЗ – валентна зона. W_c – енергія дна зони провідності, W_v – енергія стелі валентної зони, прийнята за початок відліку енергії, W_F – енергія рівня Фермі, $W_g = W_c - W_v$ – ширина забороненої зони.

Таким чином, з'являється рівноважна концентрація електронів провідності n_i і дірок провідності p_i, які однакові між собою:

$$n_i = p_i. \qquad (19.41)$$

Цей простий зв'язок (індекс i підкреслює приналежність до власного напівпровідника) є умовою електричної нейтральності речовини і дозволяє визначити положення рівня Фермі. Для цього достатньо у (19.39) відношення p/n прирівняти одиниці і тоді

$$W_F = W_i + \frac{3}{4}kT\ln\frac{m_p^*}{m_n^*}. \qquad (19.42)$$

Тут через $W_i = (W_c + W_v)/2$ позначена середина забороненої зони. Отриманий результат показує, що з ростом температури рівень Фермі лінійно зміщується від середини забороненої зони при $T = 0$ в сторону більш легких носіїв. Рівень Фермі відіграє роль «центру мас» між електронами і дірками. Рівень Фермі буде знаходитись ближче до тих носіїв заряду, яких більше.

Формула (19.40) для власного напівпровідника може бути записана так:

$$\sigma = en_i(u_n + u_p), \qquad (19.43)$$

де згідно з (19.38) $n_i^2 = np$ і

$$n_i = \sqrt{N_c N_v}\exp\left\{-\frac{W_g}{2kT}\right\}. \qquad (19.44)$$

У формулі (19.44) W_g – ширина забороненої зони напівпровідника, k – стала Больцмана, N_c, N_v – ефективні густини станів біля дна ЗП і біля стелі ВЗ, відповідно. Вони визначаються за формулами (19.35) і (19.36).

Рухливості u_n і u_p різні за величиною внаслідок різниці ефективних мас m_n^* і m_p^* і часу вільного пробігу електрона і дірки (розгядається в подальшому), які залежать від механізмів розсіювання електронів і дірок в кристалічній решітці.

Можна виділити декілька основних механізмів розсіювання носіїв заряду:

1)на теплових коливаннях структурних елементів кристалічної решітки;

2) на іонізованих домішках (іони домішки);

3) на нейтральних домішках (атом домішки);

4) на дефектах кристалічної решітки;

5) на носіях заряду.

Четвертим і п'ятим видами розсіювання в теорії напівпровідників, як правило, нехтують. При високих температурах (кімнатній і вище) перший вид розсіювання відіграє основну роль. Оскільки абсолютно чистих з ідеальною кристалічною граткою напівпровідників не існує, то на відміну від металів, рухливості

$$u_n \sim T^{-3/2} \text{ і } u_p \sim T^{-3/2}. \qquad (19.45)$$

Тоді формула (19.43) із урахуванням (19.45) і (19.44) перепишеться так :

$$\sigma_i = \sigma_{0i} \exp\left\{-\frac{W_g}{2kT}\right\}, \quad (19.46)$$

де $\sigma_{0i} = e\sqrt{N_c N_v}(u_n + u_p)$ – стала, яка слабо залежить від температури. Якщо формулу (19.46) прологарифмувати, то

$$\ln \sigma_i = \ln\sigma_{0i} - \frac{W_g}{2k} \cdot \frac{1}{T} \cdot \qquad (19.47)$$

Тобто залежність $\ln\sigma_i$ від оберненої температури $1/T$ є лінійною (при умові, що W_g не залежить від температури), що дає можливість визначити коефіцієнт $W_g/(2k)$, який містить такий важливий параметр напівпровідника, як ширина забороненої зони W_g.

Домішковий напівпровідник. Найбільш цікаві результати технічного застосування отримали напівпровідники, електричними властивостями яких можна управляти, вводячи в них хімічні домішки. Так введення в кристалічну решітку германію і кремнію незначної кількості домішкових атомів (на 100.000 атомів кристалічної решітки один атом домішки) різко змінює електричні властивості напівпровідників.

Атоми домішок називають донорами, якщо, знаходячись в кристалічній гратці напівпровідника, вони створюють заповнені електронами рівні енергії, які розташовані в забороненій зоні і з яких електрони можуть перейти в зону провідності. Класичним прикладом донорного напівпровідника є германій $_{32}\text{Ge} - 1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^2$ і кремній $_{14}\text{Si} - 1s^2 2s^2 2p^6 3s^2 4p^2$, в яких частина атомів напівпровідника заміщена п'ятивалентними домішковими атомами (фосфору $_{15}\text{P} - 3s^2 3p^3$, миш'яку $_{33}\text{As} - 4s^2 4p^3$, сурьми $_{51}\text{Sb} - 5s^2 5p^3$). Чотири валентних електрони кожного домішкового атома приймають участь в утворенні хімічних зв'язків із сусідніми атомами решітки, п'ятий валентний електрон залишається зв'язаний із

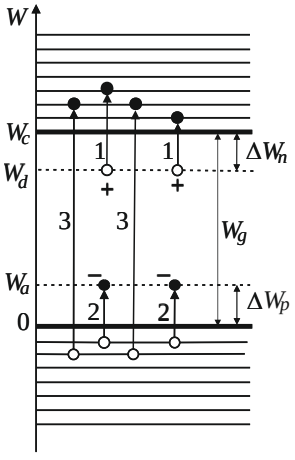


Рис.19.30. Механізм теплової генерації рухливих носіїв заряду в домішкових напівпровідниках. Світлими кружечками позначені дірки, чорними – електрони. Знаками "+" і "–" позначені іонізовані донори і акцептори. ΔW_n– енергія активації донорних домішок, ΔWp – енергія активації акцепторних домішок, W_d *і* W_a– енергія донорних і акцепторних рівнів, відповідно.

своїм атомом. Однак енергія цього зв'язку через «відволікаючий» вплив інших атомів всередині кристалу набагато менша, ніж для вільного атома, і складає величину порядку 10^{-2}еВ, яку називають енергією іонізації (активації) донора і позначається як ΔW_n (рис.19.30). В результаті якої-небудь дії домішковий атом може бути легко іонізований і п'ятий електрон стає вільним

(рис.19.30,процеси 1). На мові енергетичної моделі напівпровідника процес іонізації означає утворення рухливого електрона в зоні провідності.

Подібна ситуація спостерігається в напівпровідниках четвертої групи при заміщені деякої частини їх атомів атомами третьої групи (алюміній $_{13}\mathrm{Al} - 2s^2 2p^1$, галій $_{31}\mathrm{Ga} - 4s^2 4p^1$, індій $_{49}\mathrm{In} - 5s^2 5p^1$). Три валентних електрони домішкового атома приймають участь в утворенні хімічного зв'язку з чотирма сусідніми атомами напівпровідника, при цьому один валентний зв'язок залишається вільним. Ця вакансія під дією якого-небудь збудження може бути заповнена електроном із другого валентного зв'язку, в результаті чого в системі зв'язків утворюється вакантний стан, який може переміщуватись під дією, наприклад, електричного поля. Цей рух зручно інтерпретувати як рух деякої квазічастинки, яка називається ***діркою***. На мові енергетичної моделі такий процес відповідає переходу електрона із валентної зони на більш високий вакантний стан, зв'язаний з домішковим атомом (рис.19.30, процеси 2). В результаті домішкий атом стає негативно зарядженим іоном (енергія активації (іонізації) такої домішки позначається як ΔW_p на рис.19.30), а в валентній зоні утворюється дірка. Вище сказане пояснює, чому такий атом називають *акцептором*, а напівпровідник з такою домішкою *акцепторним*. Значення енергії іонізації (активації) донорних і акцепторних атомів в германії і кремнії приведені в табл.19.1.

Таким чином, в домішкових напівпровідниках рухливі носії заряду можуть появитись не тільки в результаті порушення валентних зв'язків в основному матеріалі (процеси 3 на рис.19.1), але і за рахунок іонізації домішкових атомів. Іонізація донорних атомів приводить до утворення вільних електронів в зоні провідності, а іонізація акцепторних атомів до утворення дірок в валентній зоні. На енергетичній діаграмі (рис.19.30) ці процеси зображені як

Табл.19.1.

Акцептор	Енергія іонізації, еВ	
	Германій	Кремній
B	0,0104	0,045
Al	0,0102	0,057
Ga	0,0108	0,065
In	0,0112	0,16
Tl	0,0131	

Донор	Енергія іонізації, еВ	
	Германій	Кремній
P	0,0120	0,044
As	0,0127	0,049
Sb	0,096	0,039
Bi	0,0127	0,039

переходи електронів з донорних локальних рівнів в зону провідності (процеси 1), переходи електронів із валентної зони на акцепторний рівень (процеси 2), а також міжзонні переходи (процеси 3), при яких генеруються і дірки, і електрони. Інтенсивність різних переходів залежить, перш за все, від температури і змін енергії, яка необхідна. Ймовірність іонізації домішкових атомів, які мають енергію іонізації порядку 10^{-2}еВ, стає помітною при температурах в одиниці і десятки градусів Кельвіна, а інтенсивність міжзонних переходів, які залежать від ширини забороненої зони, суттєві при сотнях градусів. Тому для домішкових напівпровідників розглядають умовно область низьких температур ($kT < \Delta W_n, \Delta W_p$) і область високих температур ($kT > \Delta W_n, \Delta W_p$).

Область низьких температур ($kT < \Delta W_n, \Delta W_p$)**.** Донорному напівпровіднику або напівпровіднику n-типу і відповідним величинам будемо присвоювати індекс n. При нульовій температурі всі стани валентної зони і всі домішкові стани донорних атомів (при відсутності акцепторної домішки) зайняті. Підвищення температури приводить до того, що вже при абсолютній температурі в декілька градусів появляється помітна ймовірність іонізація донорних атомів з енергією активації $\Delta W_n \sim 10^{-2}$еВ. Цей процес викликає появу рухливих носії тільки одного виду – електронів в зоні провідності. Їх концентрація в цьому випадку буде дорівнювати концентрації нерухомих позитивно іонізованих донорів:

$$n_n = N_d^+. \qquad (19.48)$$

Генерацією електронно-діркових пар (процеси 3 на рис.19.30) можна знехтувати. Вважаючи, що концентрація дірок $p_n \ll N_d^+$, (19.48) можна вважати рівнянням електронейтральності. Така якісна картина процесів дозволяє зрозуміти, чому в донорному напівпровіднику електрони називають *основними* носіями, а дірки – *неосновними.*

В акцепторному напівпровіднику *основними* носіями будуть дірки, які утворюються в валентній зоні внаслідок переходів електронів на домішкові акцепторні стани (іонізація домішки) і в зону провідності, а *неосновними* – електрони в зоні провідності.

Концентрація основних носіїв заряду в цьому випадку визначається за формулами: для n – напівпровідника

$$n_n = \frac{2N_d}{\left(1+\sqrt{1+\frac{8N_d}{N_c}\exp\left\{\frac{\Delta W_n}{kT}\right\}}\right)}; \qquad (19.49)$$

для p – напівпровідника

$$p_p = \frac{2N_a}{\left(1+\sqrt{1+\frac{8N_a}{N_v}\exp\left\{\frac{\Delta W_p}{kT}\right\}}\right)}. \qquad (19.50)$$

Для області дуже низьких температур замість (19.49) і (19.50) можна записати більш прості вирази:

$$n_n = \sqrt{\frac{N_c N_d}{2}}\exp\left\{-\frac{\Delta W_n}{2kT}\right\}, \qquad (19.51)$$

$$p_p = \sqrt{\frac{N_v N_a}{2}}\exp\left\{-\frac{\Delta W_p}{2kT}\right\}, \qquad (19.52)$$

де N_d і N_a – концентрації донорних і акцепторних атомів, відповідно.

Провідності для області низьких температур для $n-$ і $p-$напівпровідників будуть:

$$\sigma_n = \sigma_{0n}\exp\left\{-\frac{\Delta W_n}{2kT}\right\}, \qquad (19.53)$$

де

$$\sigma_{0n} = e\sqrt{\frac{N_c N_d}{2}}u_n$$

і

$$\sigma_p = \sigma_{0p}\exp\left\{-\frac{\Delta W_p}{2kT}\right\}, \qquad (19.54)$$

де

$$\sigma_{0p} = e\sqrt{\frac{N_c N_a}{2}}u_p.$$

В цих формулах σ_{0n} і σ_{0p} – сталі, які слабо залежать від температури.

Якщо формулу (19.53) прологарифмувати, то

$$\ln\sigma_n = \ln\sigma_{0n} - \frac{\Delta W_n}{2k}\cdot\frac{1}{T}\cdot \qquad (19.55)$$

Тобто $\ln\sigma_n$ лінійно залежить від оберненої температури $1/T$. Це дає змогу визначити коефіцієнт лінійної залежності $\frac{\Delta W_n}{2k}$ і таким чином визначити енергію активації домішки. Вище сказане в рівній мірі відноситься і для

напівпровідника *p*–типу.

Область високих температур ($kT > \Delta W_n$, $kT > \Delta W_p$). В цій області температур всі домішкові атоми іонізовані, і рівняння електронейтральності для напівпровідника $n-$ типу набуває вигляду:

$$n_n = N_d + p_n\,, \qquad (19.56)$$

Із закону діючих мас (19.38) можемо записати, що

$$n_n p_n = n_i^2\,. \qquad (19.57)$$

Розв'язуючи систему двох рівнянь (19.56) і (19.57) відносно невідомих n_n і p_n, отримаємо :

$$n_n = \frac{N_d}{2}\left(1+\sqrt{1+\left(\frac{2n_i}{N_d}\right)^2}\right), \qquad (19.58)$$

а

$$p_n = n_i^2/n_n\,. \qquad (19.59)$$

Для напівпровідника *p* – типу по аналогії запишемо відповідні формули:

$$p_p = \frac{N_a}{2}\left(1+\sqrt{1+\left(\frac{2n_i}{N_a}\right)^2}\right), \qquad (19.60)$$

$$n_p = n_i^2/p_p. \qquad (19.61)$$

Примітка. Концентрація власних носіїв заряду n_i для конкретного напівпровідника знаходиться за табличними даними.

Практичний інтерес має випадок сильного легування, коли $N_d \gg n_i$ або $N_a \gg n_i$. Тоді концентрації основних носіїв заряду згідно (19.58) і (19.60) дорівнюють:

$$n_n \approx N_d, \qquad (19.62)$$

$$p_p \approx N_a. \qquad (19.63)$$

Згідно закону діючих мас, концентрація неосновних носіїв заряду (дірок) в донорному напівпровіднику буде дорівнювати

$$p_n \approx n_i^2/N_d, \qquad (19.64)$$

а, відповідно, концентрація неосновних електронів в акцепторному напівпровіднику

$$n_p \approx n_i^2/N_a. \qquad (19.65)$$

Положення рівня Фермі в домішковому напівпровіднику знайдемо для випадку сильного легування (цей випадок має практичний інтерес). Для напівпровідника *n* – типу, підставляючи (19.62) в (19.33) , отримаємо:

$$N_d = N_c \exp\left\{-\frac{W_c - W_{F_n}}{kT}\right\}.$$

Звідси знайдемо, що

$$W_{F_n} = W_c - kT\ln(N_c/N_d)\,. \qquad (19.66)$$

Для напівпровідника *p* – типу, підставляючи (19.63) в (19.34) отримаємо, що

$$N_a = N_v \cdot \exp\left\{-\frac{W_{F_p} - W_v}{kT}\right\}.$$

Звідси знайдемо ,що

$$W_{F_p} = W_v + kT\ln\left(\frac{N_v}{N_a}\right). \qquad (19.67)$$

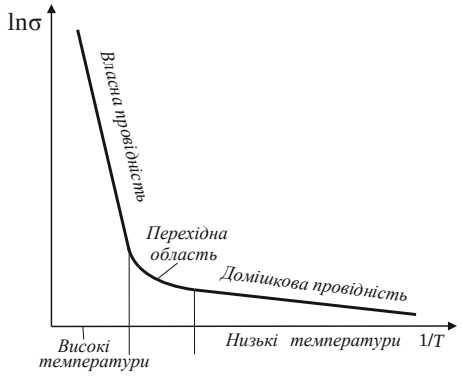


Рис.19.31.

Електропровідність домішкового напівпровідника. В загальному випадку електропровідність напівпровідника *n* – типу буде

$$\sigma = \sigma_{0i}\exp\left\{-\frac{W_g}{2kT}\right\} + \sigma_{0n}\exp\left\{-\frac{\Delta W_n}{2kT}\right\}, \qquad (19.68)$$

де перший доданок власна провідність в області високих температур, а другий доданок домішкова провідність в області низьких

температурах. На рис.19.31 приведена схематична залежність lnσ від оберненої температури $1/T$ для домішкового напівпровідника.

Оскільки $\Delta W_n < W_g$, то домішкова та власна провідності, як правило, реалізуються при різних температурах, оскільки електрони використовують енергію теплового руху. Домішкова провідність реалізується при більш низьких температурах, при яких енергія руху кристалічної гратки вже достатня для генерації домішкових носіїв, а тому власною провідністю (перша складова у формулі (19.68)) можна знехтувати.

Навпаки, при більш високій температурі, коли домішкові центри вже себе вичерпали (повністю іонізовані), починає діяти механізм генерації власних носіїв. Оскільки власних атомів завжди більше, ніж домішкових ($10^6 : 1$), то при більш високих температурах власних носіїв утворюється значно більше, ніж домішкових, і тепер можна знехтувати домішковою провідністю (друга складова у формулі (19.68)). При проміжних температурах можуть працювати обидва механізми провідності (рис.19.31, перехідна область).

§19.11.Внутрішній фотоефект

Світло має корпускулярно-хвильову природу: з одного боку йому притаманні хвильові властивості, що обумовлюють явище інтерференції, дифракції, поляризації, з іншого – корпускулярні властивості. Одне з явищ, в якому проявляються корпускулярні властивості – це фотоелектричний ефект. Він виникає в результаті взаємодії квантів електромагнітного випромінювання з електронами речовини.

Залежно від наслідків такої взаємодії розрізняють 2 види фотоефекту:

1.Зовнішній фотоефект (фотоелектронна емісія), при якому поглинання електромагнітного випромінювання супроводжується випусканням електронів речовиною (§16.5).

2.Внутрішній фотоефект, при якому через поглинання електромагнітного випромінювання відбувається лише перерозподіл електронів за енергетичними станами без їхнього випускання конденсованим середовищем.

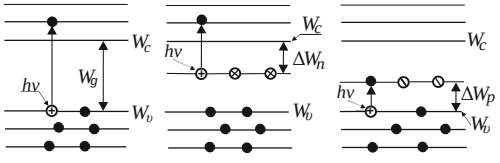
Зовнішній фотоефект має місце, якщо енергія квантів електромагнітного випромінювання більша або дорівнює роботі виходу електронів з поверхні речовини: $h\nu \geq \Phi$ (формула (16.38)). Електрон, отримавши енергію, може досягнути поверхні тіла, подолати потенціальний бар'єр, створений внутрішнім електричним полем, і вийти у вакуум або в інше середовище.

Внутрішній фотоефект відбувається у тих випадках, коли енергія фотонів недостатня для вибивання електронів з речовини. Якщо фотон спроможний лише збудити атом, то це може викликати зміну діелектричної проникності середовища – *фотодіелектричний ефект*. Якщо енергія фотонів достатня для генерації електронів і дірок провідності в напівпровіднику, тоді виникає явище фотопровідності. *Зміна електричного опору напівпровідника під дією випромінювання називається внутрішнім фотоелектричним ефектом, або фоторезистивним ефектом.* Додаткова провідність, яка обумовлена носіями заряду, що створюються випромінюванням, носить назву фотопровідності.

В неоднорідних напівпровідниках, в області p-n–переходів, в результаті

внутрішнього фотоефекту виникає електрорушійна сила (*фото е.р.с*) – *вентильний ефект*. Розглянемо детальніше механізм виникнення та основні закономірності явища фотопровідності в напівпровідниках.

Якщо температура напівпровідника $T > 0$, то завдяки енергії теплового руху навіть в неосвітленому зразку завжди існує деяка кількість вільних електронів та дірок, які знаходяться в термодинамічній рівновазі з кристалічною решіткою. Їх концентрація n_0 і p_0 визначає електропровідність напівпровідника, коли він неосвітлений (темнова електропровідність) – $\sigma_т$. При освітленні напівпровідника можуть виникнути додаткові носії заряду не обумовлені термодинамічною рівновагою з кристалічною решіткою. Такі носії називаються *нерівноважними*. Виникнення цих носіїв проявиться в зміні електропровідності напівпровідника, оскільки він отримає додаткову електропровідність $\Delta\sigma_ф$. На рис.19.32 показана схема утворення електронів



а *б* *в*

Рис.19.32.Схема утворення фотоносіїв у напівпровіднику: *а*)власному, *б*)донорному, *в*)акцепторному. W_g – ширина забороненої зони, ΔW_n – енергія активації донорів, ΔW_p – енергія активації акцепторів. Позначення: ● – електрон, ⊕– дірка, ⊗ – атом донора, ∅ – атом акцептора.

та дірок провідності у напівпровідниках різного виду. При освітленні напівпровідника деяка частина фотонів поглинається кристалічною решіткою і вільними носіями, в результаті чого збільшується їхній тепловий рух, але нові носії не утворюються. Деяка частина фотонів віддає свою енергію електронам валентної зони. Якщо енергія фотона дорівнює або більша ширини забороненої зони ($h\nu \geq W_g$) , тоді електрон, що одержав цю енергію, переходить у зону провідності і може приймати участь у створенні фотоструму. Одночасно з переходом електрона у зону провідності у валентній зоні створюється незаповнений рівень, тобто виникає дірка. Таким чином, один поглинутий фотон з енергією ($h\nu \geq W_g$) створює пару електрон – дірка (рис.19.32,*а*). Це процеси генерації фотоносіїв.

Реальними носіями струму в більшості напівпровідників є не електрони і дірки одночасно, а лише ті з них, у яких більший час життя. В одних випадках це можуть бути електрони, в інших – дірки.

Утворені світлом додаткові носії струму деякий проміжок часу (час життя τ порядку $10^{-3}\text{с} \div 10^{-7}\text{с}$) знаходяться в зоні провідності, а потім знову повертаються на одне з місць (дірки), залишене іншими електронами або приєднуються до надлишкових атомів домішки. Це процеси рекомбінації фотоносіїв. При неперервному освітленні напівпровідника процеси генерації та рекомбінації відбуваються одночасно. В результаті встановлюється така динамічна рівновага, при якій постійно існують деякі додаткові концентрації електронів Δn і дірок Δp. Так виникає фотопровідність.

Надлишкові електрони Δn і дірки Δp можуть мати енергію значно більшу, ніж середня енергія рівноважних носіїв заряду. Однак в результаті взаємодії з фононами (коливаннями) і дефектами кристалічної решітки енергія нерівноважних носіїв заряду стає однаковою з середньою тепловою енергією рівноважних носіїв заряду. Цей процес відбувається за час порядку $10^{-10} \div 10^{-12}\text{с}$.

Оскільки час життя нерівноважних носіїв заряду значно більший цієї величини, то на протязі всього часу життя кінетична енергія нерівноважних носіїв заряду відповідає середній тепловій енергії рівноважних носіїв заряду. *Тому можна вважати, що розподіл за енергіями нерівноважних носіїв заряду в зонах є таким же, як і рівноважних*. Значить і рухливості нерівноважних носіїв заряду не відрізняються від рухливості рівноважних.

Для створення додаткових носіїв в домішкових напівпровідниках фотон повинен мати енергію $h\nu \geq \Delta W_n$ (ΔW_p) (ΔW_n (ΔW_p) – енергія активації атомів домішки (донорів або акцепторів)). У напівпровідниках з донорними домішками фотон викликає перехід електронів з донорних рівнів у зону провідності (рис.19.32,*б*), внаслідок чого концентрація "вільних" електронів у напівпровіднику підвищується і електропровідність збільшується. У напівпровідниках з акцепторними домішками фотон викликає перехід електронів з валентної зони на домішкові рівні (рис.19.32,*в*) з утворенням негативно заряджених іонів домішки. У валентній зоні при цьому утворюються дірки, що забезпечують додаткову фотопровідність.

Таким чином, якщо

$$h\nu \geq W_g \qquad (19.69)$$

для власних напівпровідників і

$$h\nu \geq \Delta W_n\ (\Delta W_p) \qquad (19.69)^*$$

для домішкових напівпровідників, то при освітленні напівпровідника в ньому з'являються додаткові носії заряду і його провідність збільшується.

При енергії фотона

$$h\nu < W_g \qquad (19.70)$$

для власних і

$$h\nu < \Delta W_n\ (\Delta W_p) \qquad (19.70)^*$$

для домішкових напівпровідників фотопровідність не спостерігається. З умови (19.69) виходить, що повинна існувати *довгохвильова межа* (*граница*) фотопровідності – така довжина хвилі світла λ_0, що світло з більшою довжиною хвилі фотопровідність не викликає.

Для власних напівпровідників:

$$\lambda_0 = hc/W_g. \qquad (19.71)$$

Для домішкових напівпровідників:

$$\lambda_0 = hc/\Delta W_n, \quad \lambda_0 = hc/\ \Delta W_p. \qquad (19.72)$$

Іноді довгохвильову межу фотопровідності називають червоною межею (границею). Назва ця умовна, оскільки вперше виміряні значення величини λ_0 для зовнішнього фотоефекту припадали на червону область видимого спектру.

Для власних напівпровідників ширина забороненої зони W_g порядку одиниць електрон-вольт. Тому для них довгохвильова межа фотопровідності може належати до видимої області спектру. В домішкових напівпровідниках енергія активації домішок ΔW_n (ΔW_p) складає десяті і навіть соті долі електрон-вольта. Довгохвильова межа фотопровідності для них знаходиться в інфрачервоній частині спектру.

Розглянемо деякі основні закономірності, які визначають явище фотопровідності і проаналізуємо їх. Зробимо це на прикладі власного напівпровідника.

До освітлення напівпровідника його електропровідність визначається концентрацією та рухливістю електронів і дірок у рівноважному стані і дорівнює:

$$\sigma_0 = en_0u_n + ep_0u_p, \qquad (19.73)$$

де u_n і u_p – рухливості; n_0 і p_0 – рівноважні концентрації електронів і дірок відповідно.

При освітленні напівпровідника,

внаслідок поглинання випромінювання, відбуваються міжзонні переходи, що збільшує концентрацію електронів на величину Δn, а дірок – на величину Δp. У загальному випадку може змінитись і рухливість носіїв, але вплив зміни рухливості буде помітним лише при дуже низьких температурах. Тому будемо вважати, що фотопровідність викликана змінами концентрації носіїв.

Електропровідність напівпровідника при освітленні буде

$$\sigma = e(n_0 + \Delta n)u_n + e(p_0 + \Delta p)u_p \ , \quad (19.74)$$

а зміна електропровідності або фотопровідність визначиться як

$$\Delta\sigma_\phi = \sigma - \sigma_0 = e\Delta n u_n + e\Delta p u_p. \quad (19.75)$$

Якщо власний напівпровідник не містить будь-яких домішок і дефектів, а швидкість рекомбінації електронів і дірок однакова, тоді можна вважати, що $\Delta n = \Delta p$ і (19.75) запишеться так:

$$\Delta\sigma_\phi = e\Delta p\left(u_n + u_p\right). \quad (19.76)$$

Тут під $\Delta\sigma_\phi$ і Δp розуміють стаціонарні значення величин, тобто розглядається $\Delta\sigma_\phi$ і Δp після завершення всіх перехідних процесів і встановлення стаціонарної фотопровідності. Як видно із (19.76), величина фотопровідності буде визначатися концентрацією нерівноважних носіїв заряду $\Delta n = \Delta p$, які в свою чергу залежать від процесів генерації та рекомбінації.

Нехай на напівпровідникову пластину, товщина якої d, перпендикулярно до поверхні падає монохроматичне випромінювання, частота якого ν, а інтенсивність I (рис.19.33). Під інтенсивністю розуміють енергію, яка падає за одиницю часу на одиницю площі поверхні зразка. Якщо *I(x)* – інтенсивність всередині зразка, тоді за одиницю часу в одиниці об'єму буде поглинатися енергія, яка дорівнює $W = \alpha I(x)$, де $\alpha(\nu)$ – коефіцієнт поглинання (розмірність $[\alpha] = м^{-1}$). Поділимо цю енергію на енергію одного кванта

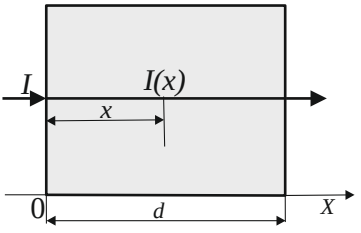


Рис.19.33.

випромінювання.Отримаємо кількість квантів світла (фотонів), що знаходяться в одиниці об'єму зразка в одиницю часу. Будемо мати на увазі, що не кожен фотон взаємодіє з електронами валентної зони, тому вводять поняття квантового виходу. Тоді швидкість генерації електронно-діркових пар, тобто кількість електронно-діркових пар, що генеруються за одиницю часу в одиниці об'єму, буде дорівнювати:

$$G = \frac{\alpha I(x)}{h\nu}\eta, \quad (19.77)$$

де $\eta(\nu)$ – *квантовий вихід* внутрішнього фотоефекту, який визначає число пар носіїв заряду (або число носіїв заряду при домішковій фотопровідності), що утворюється одним фотоном, який поглинувся, якщо інтенсивність світла вимірюється числом квантів енергії $h\nu$, що падає за секунду на одиницю поверхні. Підкреслимо, що η і α залежать від частоти (довжини хвилі) випромінювання, а I різна залежно від товщини пластини. Квантовий вихід $\eta(\nu)$ враховує, що частина фотонів може передати свою енергію решітці або вільним носіям зразка, не утворюючи нових електронів та дірок. Тому η буде меншим за одиницю. Іноді можливі випадки, коли $\eta > 1$. Наприклад, якщо енергія квантів випромінювання дуже

велика ($h\nu >> W_g$), тоді вільний електрон, який появився внаслідок поглинання енергії фотона, може мати велику кінетичну енергію і сам буде створювати нові носії заряду завдяки ударній іонізації.

Величина $I(x)$ у загальному випадку змінюється за складним законом через велику кількість відбивань від різних поверхонь і можливих інтерференційних ефектів.

У випадку тонкої плівки ($\alpha d << 1, d -$ товщина плівки) можна вважати, що $I(x) = I$. Тобто, випромінювання буде поглинатись у всьому зразку рівномірно, а швидкість генерації:

$$G = \frac{\eta\alpha I}{h\nu}. \qquad (19.78)$$

Розглянувши динаміку утворення нерівноважних носіїв (фотоелектронів та фотодірок), можна показати, що у стаціонарних умовах їх концентрація визначається співвідношенням:

$$\Delta n = \Delta p = G\,\tau_p = \frac{\eta\alpha I\tau_p}{h\nu}, \qquad (19.79)$$

де $\tau_p = \tau_n$ – час життя дірок і електронів. Тоді фотопровідність (з урахуванням (19.76) і (19.79)) можна записати так:

$$\Delta\sigma_\phi = e\frac{\eta\alpha I\tau_p}{h\nu}\left(u_n + u_p\right). \qquad (19.80)$$

Тому фотострум (з урахуванням закону Ома $I_\phi = jS = \Delta\sigma_\phi ES$) буде дорівнювати:

$$I_\phi = e\left(u_n + u_p\right)bd\frac{\eta\alpha I\tau_p}{h\nu}IE, \qquad (19.81)$$

де b і d – ширина і товщина зразка відповідно, E – напруженість електричного поля вздовж зразка (рис.19.34). Формула (19.81) дозволяє проаналізувати основні закономірності явища фотопровідності. Вона має дві величини, які залежать від інтенсивності світла I: по-перше сама інтенсивність світла I, по-друге, час життя нерівноважних пар τ_p. Час життя нерівноважних пар у загальному випадку обернено пропорційний концентрації носіїв: $\tau_p \sim \frac{1}{n_0 + p_0 + \Delta p}$. Тому при малих інтенсивностях світла $\Delta p << p_0$ і τ_p буде сталою величиною. Тоді фотострум I_ϕ повинен

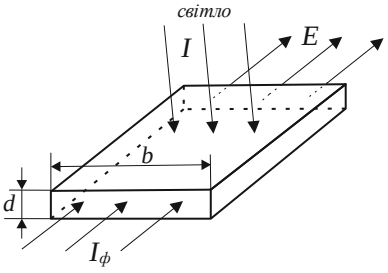


Рис.19.34.

лінійно залежати від інтенсивності світла (рис.19.35,область*а*). При великих інтенсивностях світла, коли $\Delta p \geq p_0$, час життя τ_p змінюється із зміною I, а залежність $I_\phi = f(I)$ стає нелінійною (рис.19.35, область *б*). З формули (19.81) випливає, що величина фотоструму повинна залежати від частоти (або

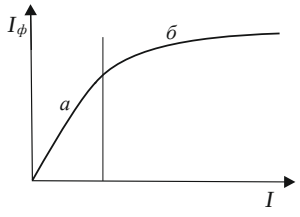


Рис.19.35. Залежність фотоструму I_ϕ від інтенсивності світла I: *а*) область малих інтенсивностей (τ_p = const); *б*) область великих інтенсивностей ($\tau_p \neq$ const).

довжини хвилі) світла. Ця залежність має назву *спектральної характеристики*. Вигляд її може бути дуже складним, тому що від довжини хвилі залежить і квантовий вихід η, і коефіцієнт поглинання α, і енергія фотонів $h\nu$. Також на вид залежності $I_\phi = f(\lambda)$ впливає матеріал напівпровідника, якість поверхні та розміри (товщина зразка).

Типовий вигляд спектральної характеристики $I_\phi(\lambda)$ напівпровідникового

зразка приведений на рис.19.36. Спад в довгохвильовій області пов'язаний з існуванням довгохвильової (червоної) межі фотопровідності λ_0. Спад в області малих довжин хвиль визначається збільшенням

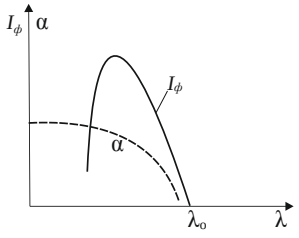


Рис.19.36.

коефіцієнта поглинання α і інтенсивним поглинанням випромінювання в тонкому поверхневому шарі напівпровідника, що призводить до зменшення швидкості об'ємної генерації G і зменшення концентрації нерівноважних носіїв, тобто до зниження фотопровідності. Свій вклад в залежність $I_\phi(\lambda)$ вносить теж і складна залежність $\eta(\lambda)$.

Релаксація фотопровідності

Розглянемо зміну нерівноважних концентрацій Δn та Δp, наприклад, в результаті опромінення напівпровідникового зразка світлом, інтенсивність якого змінюється так, як зображено на рис.19.37,*а*. Тобто періодично стрибком інтенсивність світла зростає від 0 до I_{max} потім залишається постійною деякий час Δt, після чого знову стрибком зменшується до 0. Потім це все багаторазово повторюється.

До опромінення зразка концентрація електронів (або дірок) рівноважна, тобто дорівнює n_0 (або p_0). На початку опромінення зразка починається процес генерації нерівноважних носіїв із швидкістю *G*, яка визначається кількістю носіїв, що утворюються в одиниці об'єму за одиницю часу. Якщо інтенсивність світла I_{max} не дуже велика і населеність енергетичних рівнів далека від насичення, то швидкість генерації носіїв визначається формулою (19.77) і не буде змінюватися із часом оскільки величини, що її визначають, не змінюються з часом.

Концентрація нерівноважних носіїв буде зростати за лінійним законом відносно часу (на рис.19.37,*б* – зображено пунктиром):

$$\Delta n = Gt. \qquad (19.82)$$

Проте одночасно із процесом генерації починається також процес рекомбінації електронів. При малому рівні освітлення, для якого $\Delta n << n_0+p_0$, швидкість рекомбінації

$$R = \frac{\Delta n(t)}{\tau}, \qquad (19.83)$$

де $\tau = \frac{1}{r(n_0+p_0)}$ – час життя фотоносіїв. Тут r – коефіцієнт рекомбінації.

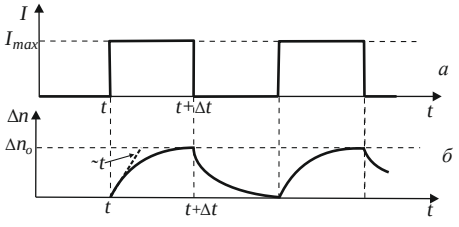


Рис.19.37.

Зміна концентрації надлишкових носіїв заряду за одиницю часу ($\frac{d(\Delta n)}{dt}$) є різниця між швидкостями генерації і рекомбінації носіїв заряду. Тому з урахуванням (19.77) і (19.83)

$$\frac{d(\Delta n)}{dt} = G - R = \\ = \eta\alpha I/(h\nu) - \frac{\Delta n(t)}{\tau}. \qquad (19.84)$$

Розділяючи змінні ($\Delta n(t)$ і t) і інтегруючи з врахуванням початкової умови $\Delta n(0)=0$ при $t = 0$, отримуємо вираз, який характеризує зміну концентрації надлишкових електронів з часом для процесу наростання кількості фотоносіїв:

$$\Delta n = \eta\alpha\tau\frac{I}{h\nu}\left(1-\exp\left\{-\frac{t}{\tau}\right\}\right) = \\ = \Delta n_0\left(1-\exp\left\{-\frac{t}{\tau}\right\}\right). \qquad (19.85)$$

Формула (19.85) пояснює процес наростання фотоносіїв на рис.19.37,*б*.

Розглянемо, як буде змінюватись Δn після вимкнення опромінення, тобто після моменту часу $t + \Delta t$. Тоді $G = 0$ і (19.84) матиме вигляд

$$\frac{d(\Delta n)}{dt} = -\frac{\Delta n}{\tau}\cdot \qquad (19.86)$$

Інтегрування цього рівняння з урахуванням початкової умови, що при $t + \Delta t = 0$ $\Delta n = \Delta n_0$, дає для моменту часу більшого від $t + \Delta t$ розв'язок:

$$\Delta n = \Delta n_0 \exp\{-t/\tau\}, \qquad (19.87)$$

де

$$\Delta n_0 = \eta\alpha\tau\frac{I}{h\nu}\cdot \qquad (19.88)$$

Величина τ має розмірність часу і за своїм значенням в рівнянні (19.87) є характерним часом процесу зменшення нерівноважної концентрації Δn, оскільки при $t = \tau$

$$\Delta n = \Delta n_0/e, \qquad (19.89)$$

де e=2,71828…основа натурального логарифму. Тобто за час τ концентрація фотоносіїв (нерівноважних носіїв) за відсутністю генерації зменшиться в e раз. Отже τ може бути *мірою тривалості життя* нерівноважних носіїв заряду.

Оскільки Δn змінюється за експоненціальним законом (19.87), то за таким же законом буде змінюватись і електропровідність σ та фотострум I_ϕ під впливом світла (рис.19.38) для моменту часу, який більший за момент часу $t + \Delta t$.

Для часу $t' > t + \Delta t$ можемо записати, що струм фотопровідності

$$I_\phi = (I_{max} - I_0)\exp\{-\tfrac{t}{\tau}\}\cdot \qquad (19.90)$$

Розглянутий вище механізм зміни фотопровідності є досить приблизним і умовним, оскільки не враховує багатьох впливових факторів. Зокрема, не врахована можливість різних механізмів рекомбінації, поверхнева рекомбінація і т.п.

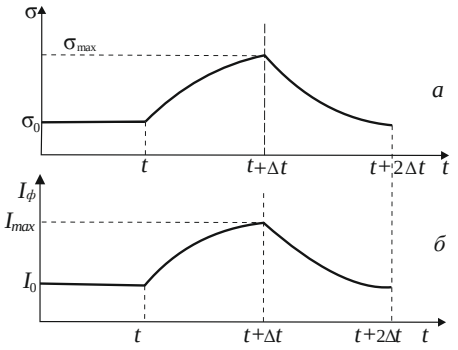


Рис.19.38. Залежність електропровідності (*а*) та сили фотоструму (*б*) через зразок напівпровідника. σ_0 – темнова електропровідність, I_0 – темновий струм, σ_{max} – максимальна провідність, I_{max} – максимальний струм.

Отже, реєструючи залежність фотоструму від часу, можна, застосувавши відповідний метод обробки даних, визначити (правильніше–оцінити) τ – час релаксації фотопровідності, який у випадку міжзонної рекомбінації (19.83), дорівнює часу життя нерівноважних носіїв.

§19.12.Дрейф і дифузія носіїв заряду

Дрейф зарядів в електричному полі

Для вільних електронів і дірок життя – це рух. Перш за все – хаотичний тепловий рух, який існує завжди і відбувається тим енергійніше, чим вища температура кристалу. Вільні електрони і дірки під впливом теплових коливань кристалічної гратки знаходяться в стані безладного хаотичного руху. Середня енергія теплового руху дорівнює $3kT/2$ (k – стала Больцмана, T – абсолютна температура). Прирівнявши величину цієї енергії кінетичній енергії частинки $mv^2/2$, можна знайти середню швидкість v_T хаотичного руху

електронів (або дірок):

$$v_T = \left(\frac{3kT}{m}\right)^{\frac{1}{2}}. \qquad (19.92)$$

При T = 300К і $m = m_e$ = 9,1· 10^{-31}кг, отримаємо для швидкості вільного електрона значення $v_T \sim 10^5$м/с. З такою великою швидкістю носії заряду зазнають $10^{12} - 10^{13}$ раз зіткнень з граткою кристалу і тому знаходяться в тепловій рівновазі з нею. Якщо до напівпровідника (металу) прикладено електричне поле, то на хаотичний тепловий рух вільних носіїв (електронів та дірок) накладається переносний або дрейфовий рух, який є результатом дії електричного поля. Механізм його виникнення такий. Вільні носії заряду при прикладеному електричному полі, як і раніше, з частотою $\sim 10^{12} \div 10^{13}$ раз за секунду зазнають зіткнень з кристалічною решіткою. Як і раніше, неможливо передбачити, куди полетить носій після чергового зіткнення: вперед, назад, вліво, вправо, вгору, вниз – під любим кутом в любому напрямку. Але, куди би він не полетів, електричне поле буде, нехай слабо, тягнути його завжди в одному і тому ж напрямку. Як результат, виникає направлений (дрейфовий) рух.

В електричному полі, рухаючись без зіткнень, носій за час t (згідно другого закону Ньютона $\Delta\vec{p} = \vec{F}\Delta t$), набуває імпульс, $m\vec{v} = e\vec{E}t$ і прискорення $\vec{a} = e\vec{E}/m$, де m – маса носія, $\vec{E}$ – напруженість електричного поля. Щоб правильно вирахувати середню швидкість, яку набуває носій в умовах багатократних зіткнень, необхідно пам'ятати про дві обставини.

По-перше, після зіткнення носій може з однаковою ймовірністю рухатись в любому напрямку. Це означає, що швидкість направленого руху після зіткнення дорівнює нулю.

По-друге, так як процес зіткнення – процес випадковий, носій на проміжку між двома послідовними зіткненнями проводить у «вільному русі» різний час. Обидві ці обставини наглядно видні на рис.19.39, на

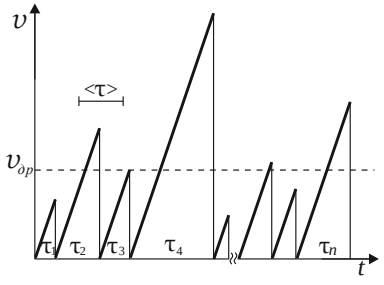


Рис.19.39.

якому показана залежність швидкості направленого руху носія від часу. Середня швидкість направленого руху (швидкість дрейфу) $v_{др}$ дорівнює

$$v_{др} = a\langle\tau\rangle = (e\langle\tau\rangle/m)E, \qquad (19.92)$$

де ‹τ› – середній час між зіткненнями, $a = eE/m$ – прискорення носія. Таким чином, швидкість направленого руху вільних носіїв в кристалі пропорційна напруженості електричного поля.

Коефіцієнт пропорційності

$$u = e\langle\tau\rangle/m \qquad (19.93)$$

носить назву *рухливості носіїв.* Рухливість вільних носіїв є однією із найважливіших характеристик носіїв заряду в кристалі.

Із формул (19.92) і (19.93) витікає, що

$$v_{др} = uE. \qquad (19.94)$$

Чим більша рухливість, тим більша швидкість направленого руху носіїв при тій же напруженості електричного поля *E*. Розмірність рухливості – м^2/(В·с). Із формули (19.94) видно, що рухливість носіїв заряду це їхня дрейфова швидкість в одиничному електричному полі (E=1В/м).

Результатом направленого дрейфового руху електронів є дрейфовий

струм, густина якого пропорційна направленому потоку частинок і їхньому заряду.

$$j_{др} = env_{др} = enuE, \quad (19.95)$$

де n – концентрація носіїв заряду.

Вираз (19.95) виражає закон Ома в диференціальній формі. Коефіцієнт пропорційності між $j_{др}$ і E є питома електропровідність

$$\sigma = enu = e^2 n\langle\tau\rangle/m. \quad 19.96)$$

В слабих електричних полях, коли швидкість дрейфу мала порівняно з тепловою ($v_{др} << v_T$, наприклад, для міді $v_{др} \sim 5 \cdot 10^{-4}$м/с), наявність чи відсутність поля ніяк не впливає на характер зіткнень носіїв з кристалічною граткою. При цьому рухливість u є постійною величиною, яка не залежить від напруженості поля E. Цей випадок відповідає виконанню закону Ома (19.95).

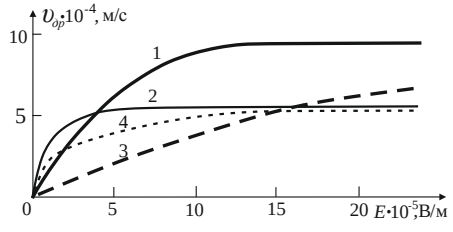


Рис.19.40. Залежність швидкості направленого руху (дрейфової швидкості) носіїв від напруженості електричного поля в Ge і Si (T=300K). Суцільні лінії – $v_{др}(E)$ для електронів, штрихові для дірок; 1,3-Si; 2,4-Ge.

В сильних полях кінетична енергія носіїв, яку вони набувають в електричному полі, може зрівнятись і навіть стати більшою, ніж енергія хаотичного теплового руху. Частота і характер зіткнень, яких зазнають тепер носії, залежить від їхньої енергії. Тому в сильних електричних полях частота і характер зіткнень електронів і дірок (а значить і рухливість) залежать від напруженості поля. Тому в сильних полях (коли $v_{др} = v_T$) закон Ома не виконується.

У всіх напівпровідниках, які використовуються на практиці, при кімнатній температурі рухливість в сильних полях зменшується з ростом напруженості поля E. В дуже сильних полях $u \sim 1/E$. Це означає, що, згідно (19.94), в цьому випадку $v_{др} = \text{const}$. Залежність дрейфової швидкості від напруженості поля електронів і дірок показана на рис.19.40. Видно, що у всіх випадках в слабих електричних полях швидкість пропорційна напруженості поля. Цю ділянку залежності $v_{др}$ (E) називають омічною або лінійною. З ростом поля дрейфова швидкість зростає все повільніше і повільніше і в досить сильних електричних полях, перестає залежати від напруженості поля E.

Дифузія носіїв заряду

Поняття «дифузія» – від латинського diffusio (розмиття)– застосовується до дуже багатьох явищ в газах, рідинах і твердих тілах.

Якщо на поверхню напівпровідникової пластинки, яку старанно очистити від домішок, нанести шар речовини, що містить домішки у великій кількості, то через деякий час цю домішку можна виявити в об'ємі напівпровідникової пластини "далеко" від поверхні напівпровідника. Домішка попадає в об'єм із-за процесу дифузії. Швидкість дифузії домішки в напівпровіднику дуже сильно зростає із збільшенням температури. Тому при кімнатній температурі потрібні десятиліття, щоб домішка проникла в область напівпровідника, а при високій температурі процес дифузії може зайняти всього декілька годин або навіть хвилин.

Таким чином, суть процесу дифузії – спонтанне (самодовільне), яке відбувається без яких-небудь зовнішніх впливів, переміщення

речовини із місця, де її багато, туди, де її мало. Процес дифузії є прямий наслідок хаотичного теплового руху атомів (або молекул). На рис.19.41 показаний розподіл концентрації частинок n вздовж осі x. Швидкість зміни концентрації з координатою характеризують похідною dn/dx, яка називається градієнтом концентрації і має напрямок в сторону збільшення концентрації n. Внаслідок хаотичного руху частинки з однаковою ймовірністю будуть проходити через одиничну плоску площину S (рис.19.41) зліва направо і навпаки. Але так як концентрація частинок

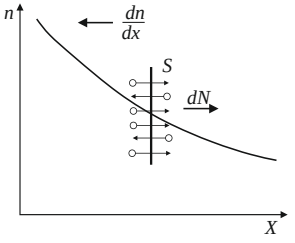


Рис.19.41

зліва від площини більша, то буде переважаючий потік частинок через площину зліва направо, ніж справа наліво. Тому можна вважати, що дифузійний потік частинок (носіїв заряду) буде пропорційний dn/dx:

$$dN = -D\frac{dn}{dx}\cdot \qquad (19.97)$$

Знак мінус в (19.98) означає, що потік частинок направлений в протилежну сторону до напрямку градієнта концентрації dn/dx (рис.19.41). Коефіцієнт пропорціональності D носить назву коефіцієнта дифузії. Встановимо розмірність коефіцієнта дифузії. Потік частинок dN чисельно дорівнює числу частинок, які перетинають за одиницю часу одиничну площадку S. Таким чином, розмірність потоку $[dN]$=1/(м$^2\cdot$с). Розмірність концентрації $[n]$=м$^{-3}$, а градієнта концентрації $[dn/dx]$, очевидно, м$^{-4}$. Тому коефіцієнт дифузії D має розмірність: $[D]$= м^2/с.

Згідно кінетичної теорії газів

$$D = \frac{1}{3}\langle l\rangle v_T, \qquad (19.98)$$

де $\langle l\rangle = \langle\tau\rangle v_T$ – довжина вільного пробігу частинки, $\langle\tau\rangle$ – середній час між зіткненнями частинок, v_T – середня швидкість теплового руху.

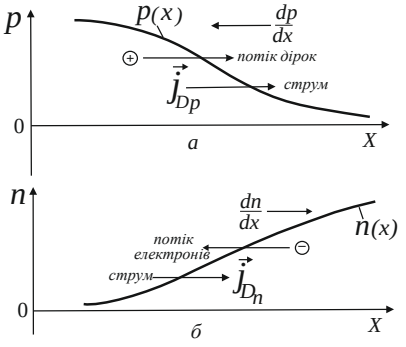


Рис.19.42. Напрямок дифузійних потоків і струмів для позитивних (дірок) (*а*) і негативних (електронів) (*б*) частинок.

Внаслідок дифузії виникає потік дірок і електронів із області з більш високою концентрацією в область з більш низькою концентрацією (рис.19.42). Але направлений потік заряджених частинок є електричний струм. Густина дифузійного струму визначається потоком частинок (19.97) та їхнім зарядом. З урахуванням напрямків вирази для дифузійних струмів дірок і електронів в напівпровідниках мають вид (рис.19.42):

$$j_{D_p} = -eD_p\frac{dp}{dx}, \qquad (19.99)$$

$$j_{D_n} = eD_n\frac{dn}{dx}, \qquad (19.100)$$

де D_p і D_n – коефіцієнти дифузії дірок і електронів, відповідно. Мінус у формулі (19.99) означає, що дифузійний струм і $\frac{dp}{dx}$ мають протилежні напрямки.

При визначенні напрямків дифузійних струмів необхідно враховувати знак носіїв

заряду. Якщо їхній заряд негативний (електрони), то, хоч потік електронів буде направлений в сторону зменшення концентрації носіїв, дифузійний струм електронів буде направлений в протилежну сторону (рис.19.42,*б*), в сторону зростання концентрації. Для дірок (позитивний заряд) дифузійний струм направлений в ту ж сторону, що і потік дірок (рис.19.42,*а*) – в сторону зменшення концентрації. Дифузійний струм – це справжній струм, який нічим не поступається добре знайомому нам струму, який виникає під дією електричного поля *E*. При проходженні дифузійного струму виділяється джоулеве тепло і відхиляється магнітна стрілка. Щоб розрахувати дифузійний струм, необхідно знайти коефіцієнти дифузії електронів D_n і дірок D_p. Між коефіцієнтами дифузії любих частинок та їх рухливістю існує просте співвідношення. Перепишемо (19.98) із врахуванням (19.91) у вигляді:

$$D = \frac{1}{3} v_T^2 \langle\tau\rangle = \frac{kT}{m} \langle\tau\rangle. \quad (19.101)$$

Порівнюючи тепер вираз (19.101) з формулою (19.93) для рухливості частинок, отримаємо

$$D = \frac{kT}{e} u \cdot \quad (19.102)$$

Знаючи величину рухливості при даній температурі за формулою (19.103) легко знайти значення коефіцієнта дифузії. Зв'язок між D і u є наглядним відображенням тієї обставини, що і направленому руху частинок під дією сили, і процесу дифузії заважає один і той самий процес: зіткнення частинок з кристалічною решіткою, що відображається через середній проміжок часу $\langle\tau\rangle$ при середній тепловій швидкості частинок v_T. Співвідношення (19.102) для дірок і електронів може бути записане так:

$$D_p = \frac{kT}{e} u_p, \quad (19.103)$$

$$D_n = \frac{kT}{e} u_n \cdot \quad (19.104)$$

Отримані співвідношення (19.103) і (19.104) називаються *співвідношеннями Ейнштейна* між коефіцієнтом дифузії та рухливістю.

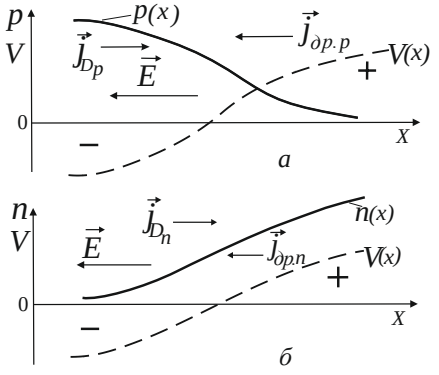


Рис.19.43. Напрямок дифузійних і внутрішніх дрейфових струмів при дифузії для позитивних (*а*) і негативних (*б*) частинок.

Особливістю дифузії заряджених частинок є перерозподіл зарядів в результаті дифузійного процесу. Якщо нейтралізація дифундуючих зарядів не відбувається, виникає об'ємний заряд, який створює електричне поле або потенціальний бар'єр, який стримує процес дифузії. Це явище можна врахувати при розгляді дрейфового струму, викликаного електричним полем (рис.19.43). Нехай дифузійний потік електронів направлений справа наліво (рис.19.43,*б*). В результаті дифузії зліва виникає надлишок електронів, справа нестача, і, відповідно, утворюється об'ємний заряд у вигляді подвійного електричного шару, поле якого направлено, як показано на рис.19.43,*б*. Створений полем потенціальний бар'єр $U(x)$ перешкоджає дифузійному руху, повертаючи назад ті

електрони, енергія яких недостатня для подолання потенціального бар'єру. Саме ці електрони, які повертаються полем і створюють дрейфовий струм. В такому випадку повний струм визначається сумою дрейфового і дифузійного струмів:

$$j = j_{др} + j_D,$$

де $j_{др}$ визначається за формулою (19.95).

Вирази для повного струму електронів і дірок мають вид:

$$j_n = -enu_nE + eD_n\frac{dn}{dx}, \quad (19.105)$$

$$j_p = -enu_pE + eD_p\frac{dp}{dx}\cdot \quad (19.106)$$

У виразах (19.105) і (19.106) струм вважається позитивним, якщо його напрямок співпадає з напрямком осі *X*. Вирази (19.105) і (19.106) справедливі і тоді, коли електричне поле створене не внутрішнім перерозподілом зарядів, а зовнішнім джерелом.

Коли між дифузійними і дрейфовими процесами наступає рівновага, то повний струм j_n і j_p дорівнює нулю. Рівність і взаємна компенсація струмів $j_{др}$ і j_D означає існування зв'язку між параметрами, які характеризують процеси дифузії і дрейфу, тобто між коефіцієнтом дифузії D і рухливістю u. Такий підхід дозволяє іншим способом отримати співвідношення (19.103) і (19.104).

Дифузія нерівноважних носіїв

В деяких типах приладів важливу роль грає особливий вид дифузії – дифузія надлишкових (нерівноважних) носіїв заряду. Особливість цього виду дифузії в тому, що надлишкові заряди, дифундуючи, віддаляються від того місця, де вони введені в напівпровідник, рекомбінують.

Описати процеси, які при цьому відбуваються можна за допомогою рівняння неперервності (19.107), яке виражає закон

$$\frac{\partial \rho}{\partial t} = -\mathrm{div}\vec{j} = -\frac{\partial j_x}{\partial x} - \frac{\partial j_y}{\partial y} - \frac{\partial j_z}{\partial z}, \quad (19.107)$$

збереження заряду: зміна густини об'ємного заряду ρ в даній точці можлива тільки внаслідок розходження струму, а поява і зникнення зарядів неможлива.

На перший погляд може здатись дивним, що рівняння (19.107), не дивлячись на його фундаментальність, не враховує процесів генерації чи рекомбінації носіїв заряду. Але ці процеси означають утворення або зникнення рухомих зарядів, тобто їх перехід із зв'язаного стану у вільний або навпаки, так що нейтральність середовища в цілому не порушується і зміни об'ємного заряду не відбуваються. Явища генерації і рекомбінації повинні бути враховані в частинних рівняннях неперервності для носіїв додатного і негативного зарядів окремо. Так для електронів $\rho_n = -en$, а для дірок $\rho_p = ep$, то (19.107) можна переписати так:

$$\frac{\partial n}{\partial t} = \frac{1}{e}\mathrm{div}j_n + G_n - R_n, \quad (19.108)$$

$$\frac{\partial p}{\partial t} = -\frac{1}{e}\mathrm{div}j_p + G_p - R_p, \quad (19.109)$$

де $G\ i\ R$ – визначають швидкості генерації і рекомбінації відповідно, тобто кількість рухомих носіїв заряду, які появляються або зникають в одиниці об'єму за одиницю часу.

Розглянемо приклад застосування рівняння неперервності для дослідження дифузії нерівноважених носіїв заряду. Допустимо, що біля поверхні донорного напівпровідника якимось чином створюється надлишкова концентрація Δp_0 дірок, які є неосновними носіями заряду для даного напівпровідника (рис.19.44).

Для спрощення задачі розглядаємо одномірний випадок (похідні по *y* і *z* відсутні) і стаціонарний випадок (відсутні похідні по часу). Крім того, будемо вважати, що процеси

нерівноважної генерації відсутні ($G_p = 0$), рекомбінація дірок з рівноважними електронами лінійна (R_p=($p-p_0$)/$\tau_p = -\Delta p/\tau_p$) і напівпровідник має достатньо малий опір,

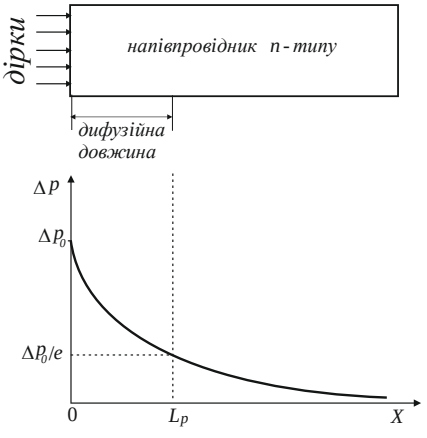


Рис.19.44.Дифузія нерівноважних дірок в напівпровіднику n-типу.

щоб вважати, що всередині напруженість електричного поля $\vec{E}$=0. Останнє припущення означає, що можна знехтувати дрейфовим рухом дірок порівняно з дифузійним. Ці припущення суттєво спрощують вихідне рівняння (19.109):

$$0 = -\frac{1}{e}\frac{d}{dx}\left(eD_p\frac{d(\Delta p)}{dx}\right) + \frac{\Delta p}{\tau_p}$$

або

$$D_p\frac{d^2(\Delta p)}{dx^2} - \frac{\Delta p}{\tau_p} = 0. \qquad (9.110)$$

Розв'язком цього рівняння є вираз

$$\Delta p(x) = \Delta p_0 \exp\{-x/L_p\}, \quad (19.111)$$

де

$$L_p = \sqrt{D_p\tau_p}\,. \qquad (19.112)$$

Результат (19.111) показує, що надлишкові дірки Δp_0 дифундують в об'єм напівпровідника і там рекомбінують з електронами, які є в надлишку в напівпровіднику $n-$типу (рис.19.44). Величина L_p, яку називають *дифузійною довжиною*, відповідає відстані, на якій надлишкова концентрація дірок зменшується в e =2,72… рази і характеризує середню глибину, на яку дифундують дірки за час свого життя τ_p. Рух дірок від поверхні створює дифузійний струм, величина якого визначається згідно (19.99) як:

$$j_{D_p} = -eD_p\frac{d(\Delta p)}{dx} = \frac{eD_p}{L_p}\Delta p = \frac{eL_p}{\tau_p}\Delta p = j_{D_p}(0)\exp\{-x/L_p\}, \qquad (19.113)$$

де $j_{D_p}(0) = \frac{eL_p}{\tau_p}\Delta p_0$. При цьому враховано, що градієнт $\frac{d(\Delta p)}{dx} \approx \frac{\Delta p}{L_p} < 0$. Із (19.113) видно, що дифузійний струм надлишкових носіїв заряду змінюється з координатою x по тому ж закону, по якому змінюється надлишкова концентрація дірок (19.111).

Величина v_{D_p}, яка дорівнює:

$$v_{D_p} = \frac{L_p}{\tau_p} = \frac{D_p}{L_p} \qquad (19.114)$$

називається *дифузійною швидкістю дірок.* Чисельно вона дорівнює швидкості, з якою нерівноважні носії заряду за час життя проходять шлях, який дорівнює дифузійній довжині.

Вище розглянутий випадок в повній мірі відноситься до напівпровідника $p-$типу і надлишкових електронів Δn_0.

Аналогічні явища дифузії відбуваються і в тому випадку, якщо нерівноважні носії створюється освітленням поверхні напівпровідника.

§19.13.Ефект Ганна

Цікавим ефектом, який спостерігається в напівпровідниках в сильних електричних полях, є *розігрів електронного газу* або утворення "*гарячих*" (нерівноважних) електронів. Це явище привело англійського

фізика Джона Ганна на початку шістдесятих років XX століття до відкриття: коли до зразка із напівпровідників арсеніда галію (GaAs) або фосфіду індія (InP) прикласти постійну напругу, яка перевищувала певне значення, струм через напівпровідник починав періодично змінюватись з часом. Інженери відразу зрозуміли, що відкриття Ганна означає появу нового напівпровідникового генератора, який може працювати в області надвисоких частот. Такі генератори були дуже потрібні у багатьох областях радіоелектроніки, зокрема – в радіолокації і системах зв'язку.

Десятки спеціалістів розмірковували над відкритим Ганном явищем. Однак пройшов рік поки його вдалося пояснити.

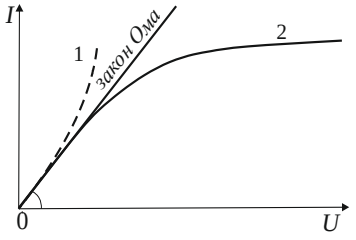


Рис.19.45.

Ганн досліджував вольтамперні характеристики зразків із n –напівпровідників. Поки напруга U, яка прикладена до зразка, невелика, струм, як і очікувалось, був пропорційний напрузі – виконувався закон Ома $I = U/R$, де R – опір зразка. Однак при більших напругах струм починав зростати значно швидше, ніж цього вимагав закон Ома (крива 1 на рис.19.45). Це пояснювалось нагріванням зразка джоулевим теплом. При нагріванні напівпровідника збільшується концентрація носіїв заряду і зменшується опір зразка (зростає провідність напівпровідника §19.10).

Цікаво оцінити, яка дрейфова швидкість електронів (формула (19.94)) при умові, коли джоулеві тепло починає різко збільшувати струм в зразку з ростом напруги. Діапазон значень рухливості електронів в різних напівпровідниках дуже широкий: від $\sim 10^{-5}$ м^2/(В$\cdot$с) до $\sim 10^2$ м^2/(В$\cdot$с). Для Ge $u_n \sim$ 0,4м^2/(В$\cdot$с), для Si – 0,13 м^2/(В$\cdot$с), для GaAS – 1 м^2/(В$\cdot$с). При підвищенні напруженості електричного поля дрейфова швидкість може виявитись співрозмірною із хаотичною швидкістю, що зменшує час вільного пробігу $\langle\tau\rangle$, а значить і рухливість носіїв (формула (19.93)). Так як типове значення рухливості порядку 1м^2/(В$\cdot$с), хаотичної швидкості $v_T \sim 10^5$ м/с, то напруженість електричного поля, при якому $v_{др} \sim 0{,}1 v_T$, не перевищує $\sim 10^4$ В/м.

Важливо підкреслити, що коли внаслідок джоулевого нагріву напівпровідника струм починає різко зростати з ростом напруги, величина рухливості не змінюється. Струм зростає внаслідок того, що розігрів викликає появу «додаткових» електронів. Щоб усунути розігрів зразка, до зразка прикладалась напруга у вигляді коротких імпульсів. Якщо використовувати короткі імпульси (дуже короткочасні) напруги і прикладати їх до зразка так рідко, щоб в проміжку між імпульсами тепло, яке виділяється в зразку, встигало повністю розсіятись, то можна досягти дуже сильних електричних полів в напівпровіднику, не нагріваючи його. Таким чином, вплив розігріву усувався і вольтамперна характеристика знову ставала лінійною ($I \sim U$) і залишалась омічною до значно більших значень напруг, ніж при постійному струмі. Однак, по мірі зростання напруги на зразку знову спостерігається відхилення від закону Ома. Для Ge і Si із збільшенням напруги струм зростає повільніше, ніж за лінійним законом. При дуже великих напругах струм практично перестає залежати від напруги – досягає

насичення (крива 2 на рис.19.45). Проводячи експерименти із зразками різної довжини l, було встановлено, що відхилення від закону Ома і насичення завжди наступає при одних і тих же значеннях напруги електричного поля $E = U/l$. Перші спроби пояснити цей факт зменшенням концентрації n електронів в зразку з ростом поля E виявились невдалими. Експерименти з чистими Ge і Si показали, що при тих полях, при яких спостерігається насичення струму, концентрація n від поля не залежить. Так як густина дрейфового струму (формула (19.95)) при незмінній n залежить від швидкості дрейфу $v_{др}$, то це означає, що в Ge і Si дрейфова швидкість електронів в сильніших електричних полях від напруженості поля не залежить всупереч формулі (19.94). Пояснення цьому факту приведено у §19.12. Можемо тільки добавити, що при великих електричних полях ефективність процесів зіткнення електронів провідності з кристалічною граткою підвищується, і рухливість електронів в Ge і Si з ростом поля зменшується.

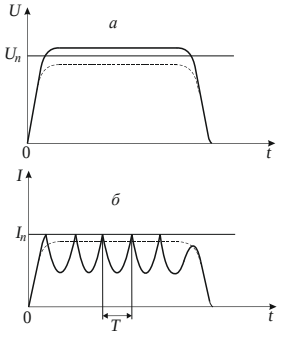


Рис.19.46.

Насичення дрейфової швидкості електронів в Ge і Si в сильних електричних полях було відкрито задовго до того, як Ганн приступив до своїх досліджень. Та обставина, що властивості електронів в сильних полях можуть суттєво відрізнятися від властивостей в слабких електричних полях, теж було добре відомо. Але ефект, який Ганн спостерігав на зразках арсеніду галію ($n-$GaAs) і фосфіду індія (n –InP), був просто вражаючий. Дійсно! До зразка прикладались імпульси напруги (рис.19.46,*а*). Поки амплітуди імпульсів були менші деякої порогової величини U_n (цей випадок показаний на рис.19.46,*а* пунктиром), все було зрозуміло (приблизно виконувався закон Ома). Струм через зразок зростав, поки наростала напруга, залишався стали, коли напруга не змінювалась, і зменшувався, коли напруга спадала (пунктирна крива на рис.19.46,*б*). Коли напруга перевищувала значення U_n (суцільна крива на рис.19.46,*а*), як струм при *постійній напрузі* починав періодично змінюватись з часом (цей випадок показаний на рис.19.46,*б* суцільною кривою). Ганн встановив, що порогова напруга пропорційна довжині зразка l: $U_n = E_n l$, де E_n – приблизно однакове значення напруженості поля для всіх зразків. Для n–GaAS величина E_n складала приблизно $3{,}2 \cdot 10^5$ В/м. Період коливань струму був теж пропорційний довжині зразка і дорівнював приблизно $T \cong 2l/v_n$, де $v_n \cong uE_n$. Намагаючись зрозуміти, що відбувається в зразку, коли виникають коливання струму, Ганн виконав дослідження і встановив, що, коли напруга на зразку перевищує порогове значення U_n, характер протікання струму і розподіл напруженості електричного поля в зразку різко змінюється. При $U < U_n$ струм в зразку утворює потік електронів від катода до анода, а електричне поле розподілене вздовж зразка однорідно (пунктирна лінія на рис.19.47). В цих умовах вольтамперна характеристика зразка відповідає закону Ома. Але як тільки напруга на зразку досягає порогового значення U_n, характер розподілу поля вздовж зразка різко

змінюється. В зразку біля катода виникає вузька область дуже сильного поля – **«домен»** (визначення поняття "домен" приводиться в §2.3). На рис.19.47 показаний домен сильного електричного поля. Хоч ширина домена невелика, але поле в ньому настільки сильне (E_4), що на ньому падає значна частина прикладеної до зразка напруги (рис.19.47). Домен рухається від катода до анода із швидкістю, яка приблизно дорівнює $\upsilon_n/2$. Досягнувши анодного контакта, домен руйнується. Негайно після цього біля катода зароджується новий домен і все повторюється спочатку.

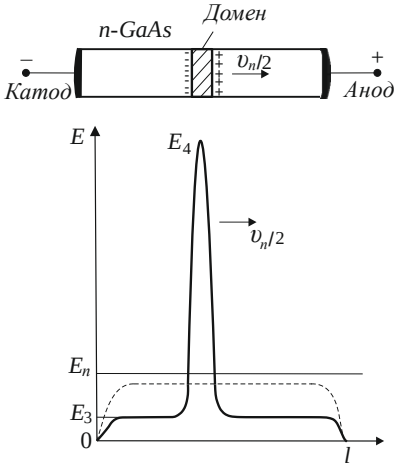


Рис.19.47.

Легко зрозуміти, як пов'язаний періодичний рух домена по зразку з коливаннями струму в ньому. Нехай до зразка приклали напругу, яка дорівнює U_n. В зразку виник домен і значна частина напруги впала на домені. Але це означає, що на долю решти частини зразка, де домена не має, буде приходитись менша напруга; значить, поле зовні домена значно зменшиться (E_3) (порівняти пунктирну і суцільну криві на рис.19.47). Зменшиться і значення швидкості електронів, яке пропорційне напруженості поля, а значить зменшиться і струм, що протікає через зразок.

Поки домен рухається вздовж зразка, струм не змінюється. Коли домен, досягнувши анода, починає руйнуватись, спад напруги на ньому зменшується. При цьому поле в решті частині зразка зростає і, отже, струм зростає.

Як тільки напруженість поля в зразку в районі катода стане дорівнювати пороговому значенню $E_n = U_n/l$, в зразку почне формуватись новий домен і струм знову почне зменшуватись. Поки напруга, яка прикладена до зразка, перевищує порогове значення U_n, цей процес періодично повторюється (рис.19.46,*б*).

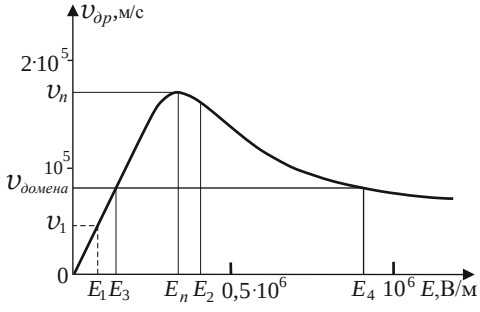


Рис.19.48.

Виникає питання, чому в зразку утворюється домен, в чому фізична причина утворення домена. На це питання в 1984 році зумів відповісти американський фізик Герберт Кремер. Він нагадав, що незадовго до відкриття Ганна на основі теоретичних розрахунків було показано, що в арсеніді галію, фосфіді індію і деяких інших напівпровідниках залежність дрейфової швидкості електронів від напруженості поля повинна мати такий вид, як показано на рис.19.48. В певному діапазоні значень напруженості поля E, які починаються із значення E_n, дрейфова швидкість електронів зменшується із збільшенням напруженості поля. Виходячи із формули (19.94) ми повинні

зробити висновок, що в цьому діапазоні значень E рухливість u електронів зменшується із збільшенням E так різко, що дрейфова швидкість зменшується, хоч напруженість електричного поля зростає.

Як же пов'язані ефекти, що спостерігав Ганн, з таким видом залежності швидкості електронів від напруженості поля? Розглянемо простий приклад. Нехай до зразка прикладена напруга $U_1 < U_n$. При цьому поле вздовж зразка розподілене однорідно і дорівнює $E_1 = U_1/l$. Дрейфова швидкість електронів v_1 дорівнює $v_1 = uE_1$ (рис.19.48) і тоді потік електронів від катода до анода створює струм. Допустимо тепер, що десь в зразку випадково утворилась невелика ділянка, в якому напруженість поля буде трохи більшою, ніж значення E_1 (цьому сприяють неоднорідності в легованому напівпровіднику). За рахунок хаотичного характеру руху електронів в кристалі маленькі відхилення поля від середнього значення E_1 (флуктуації поля) виникають неперервно і тепер прослідкуємо за тим, що відбудеться з такою флуктуацією, коли вона виникає. Якщо на якій-небудь ділянці зразка напруженість поля стала дещо більшою, ніж в решті зразка, то це значить, що на границях цієї ділянки накопичились електричні заряди: негативні зі сторони катода і позитивні зі сторони анода (рис.19.47). Іншими словами, зліва від виділеної ділянки утворився невеликий «надлишок» електронів, а справа – невеликий «дефіцит» електронів. Поле цих «об'ємних» надлишкових зарядів складається із полем зовнішнього джерела і таким чином підсилюється поле всередині виділеної ділянки.

Що ж відбувається з такого флуктуацією в подальшому?

Оскільки ми розглядаємо випадок, коли $E_1 < E_n$, то швидкість електронів пропорційна напруженості поля. Отже, електрони всередині флуктуації рухаються швидше, ніж електрони в решті зразка. Це означає, що флуктуація буде «тікати» від надлишкових електронів зліва і «наздоганяти» ту ділянку справа, де електронів не вистачає. В результаті флуктуація поля розсмоктується.

А тепер розглянемо, що відбудеться з флуктуацією, коли напруженість E_2 поля в зразку дещо більша, ніж порогове значення E_n (рис.19.48). Знову будемо вважати, що спочатку поле розподілене вздовж зразка однорідно; внаслідок флуктуації зарядів виникають об'ємні заряди; в результаті в малій ділянці зразка напруженість поля виявиться більшою, ніж в решті зразка. Але тепер більшому полю відповідає менша дрейфова швидкість (рис.19.48). Значить, флуктуація рухається повільніше, ніж електрони зліва і справа від неї. В результаті зліва до надлишкових електронів, які там уже є, добавляться нові електрони, що доганяють флуктуацію. Справа, де був «дефіцит» електронів, їх стане ще менше: електрони «втечуть» від флуктуації. Таким чином, об'ємний заряд стане більшим і напруженість поля всередині флуктуації збільшиться. Але збільшення напруженості поля при $E > E_n$ відповідає зменшенню дрейфової швидкості. В результаті, флуктуація стане рухатись ще повільніше, ще більше електронів скупчиться зліва від флуктуації і ще менше їх стане справа, а напруженість поля всередині флуктуації стане більшою і т.д. Тобто, при умові, коли дрейфова швидкість електронів зменшується з ростом напруженості поля, однорідний розподіл поля по зразку не стійкий по відношенню до малої флуктуації поля; раз виникнувши, така флуктуація починає нестримно наростати. Ріст флуктуації припиниться, коли швидкість

електронів всередині флуктуації зрівняється зі швидкістю електронів зовні неї. При цьому напруженість поля зовні домена буде E_3, а в домені E_4 (рис.19.47 і 19.48).

Поки надлишкове поле всередині флуктуації мале, на ній падає мала напруга, і її присутність в зразку практично не впливає на швидкості електронів поза флуктуацією. Однак по мірі зростання поля всередені флуктуації, на ній починає падати все більша і більша напруга. При цьому напруга на решті зразка стає меншою (адже повна напруга, яка прикладена до зразка, постійна). Поле поза флуктуацією зменшується. Така флуктуація, яка сильно виросла, і є домен поля. Поза доменом швидкість електронів зменшується із зменшенням напруженості поля, так як поза доменом напруженість завжди менша E_n. Таким чином, по мірі росту домена швидкість електронів поза доменом зменшується. В кінці кінців швидкості електронів всередині і поза доменом зрівняються і ріст домена припиниться. Зразком від катода до анода буде «бігти» незмінна за формою (стабільна) ділянка сильного поля–домен із швидкістю $v_{дp} \approx v_n/2$ (рис. 19.48).

Ці домени і спостерігав Ганн у своїх експериментах. Максимальне поле в домені E_4 може досягати значень $2\cdot 10^6 \div 30\cdot 10^6$ В/м.

Таким чином, вияснення природи існування ефекту Ганна звелось до пояснення залежності швидкості дрейфу електронів $v_{дp}$ від напруженості електричного поля в напівпровіднику, що приведена на рис.19.48. Згідно формул (19.93) і (19.94) можемо записати, що

$$v_{дp} = u_n E = \frac{e\langle\tau_n\rangle}{m_n^*}E, \qquad (19.115)$$

де m_n^* – ефективна маса електронів, $\langle\tau_n\rangle$ – середній час їх вільного пробігу.

Для арсеніда галію зона провідності складається із двох зон, енергетична відстань між якими складає 0,36еВ. Ширина забороненої зони GaAs при T=300К дорівнює 1,4еВ, то енергія дна першої нижньої зони провідності W_{c_1}=1,4еВ, а другої (верхньої) $-$ $W_{c_2} = 1{,}76$еВ. Ефективна маса електрона в першій зоні m_{n1}^* = =0,07 m_e, а в другій зоні $-$ $m_{n2}^* = 1{,}2m_e$, де m_e $-$ маса вільного електрона. Це означає, що рухливість електронів в нижній зоні u_{n1} більша рухливості електронів в верхній зоні u_{n2}, так як рухливість u_n~1/m_n^*. Крім того, густина енергетичних станів ($g(W)$=4π$(2m_n^*)^{\frac{3}{2}}W^{\frac{1}{2}}/h^3$) значно більша у верхній зоні, так як $m_{n2}^* >> m_{n1}^*$.

При звичайних температурах, коли енергія теплового руху $kT \ll$0,36еВ і в слабих електричних полях, електрони заповнюють практично тільки стани нижньої зони і, маючи велику рухливість, забезпечують значну електропровідність і лінійну вольт– амперну характеристику. В сильних електричних полях ($E > 3{,}5\cdot 10^5$В/м) спостерігається розігрів електронів провідності. Це означає підвищення енергії дрейфу і перехід електронів у верхню зону провідності. Більша густина станів у верхій зоні забезпечує можливість переходу туди значної частини електронів зони провідності, а мала їхня рухливість u_{n2} призводить до зменшення швидкості дрейфу $v_{дp}$. Так пояснюється залежність $v_{дp}$ від E, яка приведена на рис.19.48.

Ефект Ганна спостерігається більш ніж в двадцяти напівпровідникових сполуках: антимоніді індія (InSb), антимоніді миш'яку (AsSb), в германії під стискуванням, в кремнії при низьких температурах, в потрійних напівпровідникових сполуках GaAlAs, InGaSb, InAsP, в сполуці InGaAsP і в багатьох інших

напівпровідниках. Основне застосування ганнівських діодів (так прийнято називати зразки, в яких спостерігається ефект Ганна) – генератори надвисокої частоти.

Контрольні питання

1.Дати визначення твердого тіла.

2.Які тверді тіла є кристалами?

3.Приведіть класифікацію твердих тіл за величиною питомої електропровідності і її залежності від температури.

4.Які існують типи зв‘язків атомів в твердих тілах?

5.Чому енергетичний спектр електронів в кристалах повинен бути зонним?

6.Дати визначення зон: валентної, провідності, забороненої.

7.Поясніть електричні властивості металів, напівпровідників і діелектриків з точки зору зонної теорії.

8.Дайте визначення функції густини енергетичних рівнів в зоні. Як ця густина залежить від енергії в зоні?

9.В чому різниця квантового і класичного розподілу електронів за енергетичними рівнями в зоні?

10. Дайте визначення енергії Фермі. Від чого залежить величина енергії Фермі?

11. Дайте визначення функції розподілу густини електронів за енергетичними станами (рівнями).

12.Який фізичний зміст ширини забороненої зони?

13.Як ви думаєте, в чому зміст введення поняття ефективної густини станів для електронів і дірок в напівпровідниках?

14.Поясніть, від яких факторів залежить відношення концентрацій електронів і дірок в напівпровідниках?

15.Поясніть фізичний зміст закону діючих мас.

16.Поясніть, чому концентрація електронів і дірок у власних напівпровідниках сильно залежить від температури?

17.Як ви думаєте, чи можливий такий випадок в напівпровідниках, щоб ефективна маса електронів дорівнювала ефективній масі дірок? Де розташовувався б в цьому випадку рівень Фермі?

18.Чому для електричних властивостей домішкових напівпровідників важливим є: чи це область високих, чи низьких температур?

19.Як ви розумієте вислів: “Сильно легований напівпровідник”?

20.Від яких факторів залежить величина часу життя в напівпровідниках?

21.Поясніть роль основних і неосновних носіїв заряду в напівпровідниках з точки зору електричних і фотоелектричних властивостей.

23.Які існують методи визначення основних фізичних параметрів напівпровідників?

24.Які види струму існують в напівпровідниках? В чому різниця між ними?

25.В чому суть ефекту Ганна?

26.Як пояснюється ефект Ганна за допомогою домена електричного поля?

27.Яке практичне застосування ефекту Ганна?

Розділ 20.Термоелектронна емісія

§20.1.Історія відкриття термоелектронної емісії

Термоелектронну емісію вперше спостерігав у 1883 р. Томас Алва Едісон (1847-1931). Термоелектронна емісія була відкрита випадково. В кінці 1879 р. Едісон в основному закінчив роботу по створенню електричної лампи розжарення і наполегливо працював над проблемою збільшення строку її роботи. Лампа ця по суті мало чим відрізнялась від сучасних ламп розжарення – тільки замість вольфрамової спіралі використовувалась нитка із обвугленого бамбукового волокна. Перші такі лампи (рис.20.1) могли неперервно горіти біля 40 годин, із-за чого для масового застосування вони були не придатні. В пошуках причини швидкого перегорання вуглецевої нитки Едісон звернув увагу на те, що внутрішня поверхня лампового балону покривається темним нальотом із вуглецевих частинок, які мабуть вилітали із розжареної нитки. Надіючись якось усунути це небажане явище, Едісон розмістив в лампі біля нитки додатковий електрод, електричний потенціал якого можна було змінювати. При цьому він зі здивуванням виявив, що, коли потенціал цього електрода вищий потенціала нитки (позитивний по відношенні до потенціалу нитки), між ним і ниткою – через вакуумний проміжок! – протікав електричний струм. При всій своїй технічній винахідливості ніякої практичної користі із свого відкриття Едісон добути не зміг, і воно було надовго забуто. Ефект Едісона вдалось пояснити тільки після відкриття електрона 1897р. Пояснення дав Дж.А.Флемінг, який в 1882–1885 роках працював у Едісона. Флемінг зрозумів, що електричний струм в балоні лампи створюють електрони, які виходять із розжареної нитки і притягуються позитивно зарядженим додатковим електродом.

Продовжуючи роздумувати над явищем термоелектронної емісії, в 1904 році Флемінг запатентував першу в історії техніки радіолампу. Це був вакуумний діод, який повністю базувався на односторонній провідності, зв'язаній з ефектом Едісона, що дозволяла випрямляти змінний струм. Флемінг задумав свою лампу як детектор для радіохвиль.

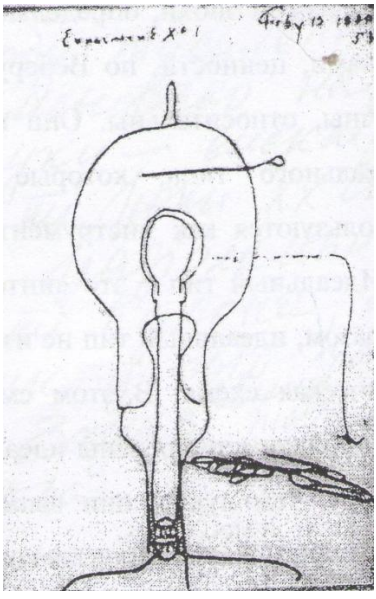


Рис.20.1.Фотокопія сторінки із лабораторного журналу Едісона.

В 1906 році американський інженер Л. де Форест (1873-1961) винайшов трьохелектродну лампу – тріод, історично перший електронний підсилювальний прилад. Відкриття ефективного способу підсилення слабих електричних сигналів ознаменувало народження нової області техніки і науки – електроніки. До появи транзисторів (50-ті роки XX століття) в ній повністю панували

радіолампи, робота яких заснована на ефекті Едісона, ефекті якому знаменитий винахідник у свій час не надав особливого значення.

Зараз явище термоелектронної емісії використовується в сотнях мільйонів джерел світла, в першу чергу – в люмінесцентних лампах.

§20.2.Робота виходу електрона із металу

Суть цього явища в тому, що із підвищенням температури збільшується енергія теплових коливань атомів металу і, відповідно, теплова енергія електронів. Не зважаючи на те, що середня теплова енергія електронів навіть

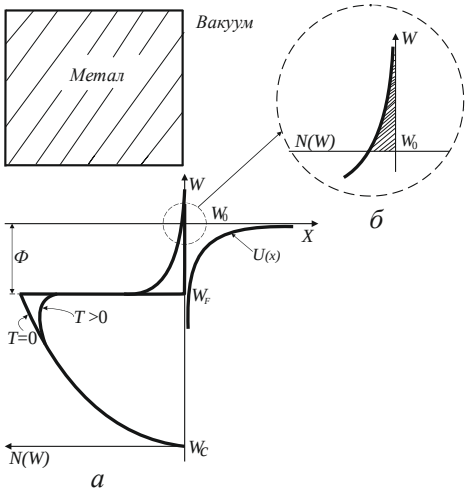


Рис.20.2.Енергетична діаграма електронів біля поверхні металу. $а$— $N(W) = g(W)f(W)$ – функція розподілу густини електронів, де $g(W)$ – функція розподілу густини енергетичних рівнів, $f(W)$ – функція Фермі-Дірака (функція розподілу ймовірності заповнення енергетичних рівнів електронами), W_c – енергія дна зони провідності електронів в металах (приймається за початок відліку енергії), W_0 – енергія вакууму (енергія вільного електрона), W_F – енергія рівня Фермі, $\Phi = W_0 - W_F$ – робота виходу електронів із металу.

при температурах в тисячу градусів є незначною (складає десяті долі електрон-вольта) порівняно із загальною середньою енергією електронів в металі (§19.7), все ж мала частина електронів набуває енергію, яка є достатньою для подолання сил, які перешкоджають виходу електронів із металу.

Хоч таких електронів мало, але саме вони обумовлюють процеси емісії. За своїм фізичним змістом термоелектронна емісія багато в чому аналогічна процесу випаровування молекул із рідини. Вище сказане проілюструємо енергетичною діаграмою електронів провідності в металі на рис.20.2. На рис.20.2 показано, що при $T = 0$ функція розподілу густини електронів $N(W) = 0$ для $W > W_F$. Це означає, що електрони заповнюють енергетичні рівні, які лежать нижче рівня Фермі (квантовий розподіл). Електрони з енергіями $W > W_0$ (електрони поза металом) появляються при підвищенні температури і їхня кількість, яка пропорційна заштрихованій області під кривою $N(W)$ для $W > W_0$ (рис.20.2,*б*), зростає з температурою. Це обумовлено тим, що з ростом температури відбувається «розмиття» функції Фермі-Дірака біля енергії $W > W_F$. Так як $W > W_F \gg kT$, то функція Фермі-Дірака з великою точністю може бути замінена класичною функцією Больцмана (19.20):

$$f(W) \cong \exp\left\{-\frac{W-W_F}{kT}\right\}. \qquad (20.1)$$

Таким чином, електрони з енергією $W > W_0$ (так звані «гарячі» електрони) підпорядковуються класичному розподілу і вони відповідальні за термоелектронну емісію.

Виникає питання, а чому електрони, які вийшли з металу (тобто мають енергію $W > W_0$), не покидають метал зовсім, а утримуються біля поверхні металу. Справа в тому, що при виході електрона за межі металу на нього діє сила зі сторони нескомпенсованого

позитивного заряду (рис.20.3), яка утримує його. Розрахунок сили взаємодії може бути проведений методом дзеркальних відображень, суть якого в тому, що розподілений по поверхні металу позитивний заряд замінюється точковим позитивним зарядом $+e$, який знаходиться по іншу сторону поверхні металу на такій же

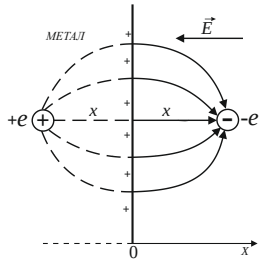


Рис.20.3.

відстані від поверхні, що і електрон, який розглядається (рис.20.3). Згідно з законом Кулона, ця сила буде:

$$F(x) = -\frac{e^2}{16\pi\varepsilon_0 x^2}\ , \qquad (20.2)$$

а потенційна енергія електрона в полі такої сили

$$U(x) = -\frac{e^2}{16\pi\varepsilon_0 x}\cdot \qquad (20.3)$$

Графік цієї функції схематично показаний на

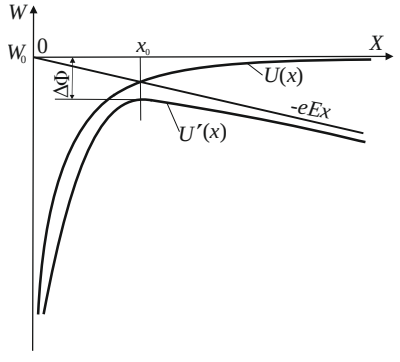


Рис.20.4

рис.20.4. Ця функція описує форму потенційного бар'єру, який перешкоджає виходу електрона із металу. Робота, яка виконується електроном по подоланню сил цього потенційного бар'єру при емісії називається *роботою виходу.* Для металів робота виходу

$$\Phi = W_0 - W_F \qquad (20.4)$$

називається термодинамічною (рис.20.2,*а*).

Сили дзеркального відображення (20.2) можуть бути послаблені зовнішнім електричним полем, напруженість якого $\vec{E}$ (рис.20.3). Сила дзеркального відображення (20.2) і сила електричного поля $e\vec{E}$ направлені назустріч одна одній, а в точці

$$x_0 = \sqrt{e/(16\pi\varepsilon_0 E)} \qquad (20.5)$$

(рис.20.4) взаємно компенсують одна одну. Потенційна енергія електрона буде дорівнювати сумі енергії (20.3) і його енергії в зовнішньому полі $-eEx$ (рис.20.3). Тобто

$$U'(x) = -\frac{e^2}{16\pi\varepsilon_0 x} - eEx. \qquad (20.6)$$

Графіки цих потенційних енергій схематично показані на рис.20.4. Точка x_0 відповідає максимуму потенційної енергії $U'(x)$ і визначає найбільшу висоту потенційного бар'єру. Так як $U'(x)$ відраховується від енергії вакууму W_0, то це означає, що має місце зменшення роботи виходу електрона із металу на величину

$$\Delta\Phi = U'(x_0) = -e\sqrt{\frac{eE}{4\pi\varepsilon_0}}\cdot \qquad (20.7)$$

Таким чином, кількість електронів, які можуть приймати участь в термоелектронній емісії, визначається площею заштрихованої фігури на рис.20.2, величина якої визначається температурою і напруженістю зовнішнього електричного поля $\vec{E}$ (рис.20.3), яка зменшує роботу виходу Φ.

§20.3.Закони термоелектронної емісії

Термоелектронна емісія широко використовується в електровакуумних приладах. Тому для її вивчення може бути використана двох – або трьохелектродна лампа (рис.20.5). Вивчення законів термоелектронної емісії за допомогою таких ламп зводиться до експериментального дослідження залежності анодного струму I_A від анодної напруги U_A, тобто до дослідження вольтамперних характеристик (ВАХ) (рис.20.6).

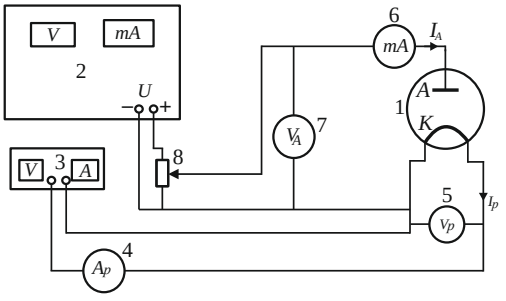


Рис.20.5. Принципова схема установки для дослідження ВАХ діода.

1 – електронна лампа (діод); 2 – універсальне джерело живлення широкого спектру призначення (воно живить анодне коло лампи); 3 – джерело живлення, за допомогою якого регулюють величину струму розжарення, а відповідно, і температуру катоду *K*; 4 – амперметр, який вимірює струм розжарення катоду; 5 – вольтметр, який вимірює напругу на катоді; 6 – міліамперметр, який вимірює струм в анодному колі; 7 – вольтметр, який вимірює напругу між анодом і катодом (анодна напруга); 8 – потенціометр, який дає можливість змінювати анодну напругу в широких межах.

На графіках можна побачити три характерні ділянки: початкова ділянка (I), ділянка зростання (II), ділянка насичення струму (III). Звернемо увагу на те, що вольтамперна характеристика електронної лампи є *нелінійною,* а отже електричний струм у лампі не підпорядковується закону Ома.

Спочатку (ділянка I ВАХ) із збільшенням анодної напруги сила струму зростає повільно, потім значно швидше (ділянка II) (рис.20.6).

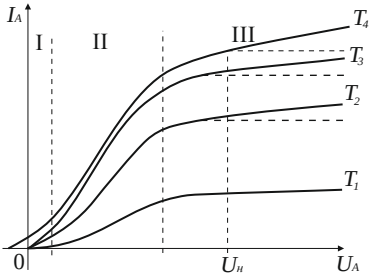


Рис.20.6.Вольтамперні характеристики для різних температур катода ($T_4 > T_3 > T_2 > T_1$).

Починаючи з деякого значення напруги $U_н$ подальше зростання струму практично припиняється (ділянка III) (рис.20.6). Залежність струму діода від напруги легко пояснити. При наявності термоелектронної емісії в просторі між катодом і анодом у будь-який момент часу є електрони, що рухаються від катода до анода, які утворюють хмару негативного заряду (*просторовий заряд*). Цей просторовий заряд змінює розподіл потенціалу в діоді. Якщо катод і анод є плоскими пластинками, паралельними одна одній (рис.20.7), то при відсутності просторового

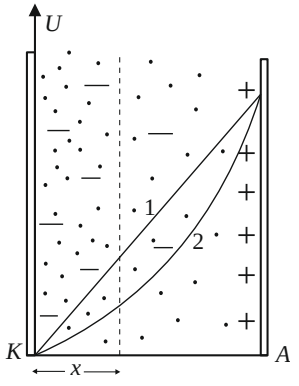


Рис.20.7.Просторовий заряд у діоді і спричинений ним перерозподіл потенціалу: 1– розподіл потенціалу при відсутності просторового заряду; 2 – розподіл потенціалу при наявності просторового заряду.

заряду (при холодному катоді) розподіл потенціалу між катодом і анодом, які утворюють плоский конденсатор, зображується прямою лінією 1. При наявності термоелектронного струму (розжарений катод) між катодом і анодом виникає просторовий заряд і розподіл потенціалу змінюється; він подається тепер кривою 2. При цьому значення потенціалу в будь-якій площині x менше, ніж при відсутності просторового заряду, а отже, і швидкості руху електронів при наявності просторового заряду зменшуються. Із збільшенням анодної напруги концентрація електронів у хмарі просторового заряду зменшується. Тому і гальмівна дія просторового заряду зменшується, і анодний струм зростає.

Зауважимо, що розподіл потенціалу, зображений кривою 2 дістаємо тоді, коли початкові швидкості вилітання електронів з катода досить малі, що, звичайно, й спостерігається. А коли початковими швидкостями знехтувати не можна, розподіл потенціалу має складніший вигляд.

Коли потенціал анода стає $U_A > U_н$ (рис.20.6), то всі електрони, які випускає катод за кожну одиницю часу, потрапляють на анод і струм I_A повинен досягати максимального значення (струм насичення) при певній температурі катода. Струм насичення характеризує емісійну здатність катода, що залежить від природи катода і його температури. Проте в області III насичення анодного струму (рис.20.6) не має. Анодний струм I_A залежить від потенціала анода U_A і ця залежність тим сильніша, чим вища температура катода. Пояснення цієї залежності дається в подальшому.

При напругах, що менші $U_н$ (ділянка II), залежність сили струму I_A від анодної напруги U_A описується законом Богуславського−Ленгмюра:

$$I_A = \alpha U_A^{3/2}, \qquad (20.8)$$

де α – стала величина, яка залежить від конструкції електродів лампи.

Закон (20.8) ще часто називається «законом трьох других». Він отриманий теоретично для ряду спрощень, а саме:

1.Катод і анод площини, що утворюють плоский конденсатор;

2.На поверхні катода ($x = 0$) потенціал дорівнює нулю ($U = 0$), а на аноді $U = U_A$;

3.Напруженість електричного поля ($E = -\frac{dU}{dx}$) на поверхні катода та безпосередньо біля нього дорівнює нулю;

4.Початкова швидкість електронів, що вилетіли з поверхні катода, дорівнює нулю.

Зауважимо, що залежності $I_A = f(U_A)$, які спостерігаються на практиці (рис.20.6), дещо відрізняються від теоретичних. Існують такі розходження:

1.Реальні ВАХ (рис.20.6) не виходять з початку координат. Вони починаються при невеликих від'ємних напругах на аноді, тому при $U_A = 0$ існує деякий невеликий початковий струм. Це можна пояснити тим, що електрони, які вилітають із поверхні катода, мають початкову швидкість відмінну від нуля. Деякі з них, маючи велику швидкість, можуть навіть при нульовій анодній напрузі ($U_A = 0$) подолати гальмівне поле просторового заряду і долетіти до анода, створюючи в анодному колі незначний струм;

2.Реальні ВАХ на ділянці II, як правило, є більш пологими, ніж теоретичні. Крім того, для різних температур катоду ділянки II реальних кривих розходяться, тобто не співпадають. Оскільки, за законом

Богуславського-Ленгюмера, на ділянці II повинна бути єдина крива за будь-яких температур. Це пояснюється таким чином: зі зміною температури змінюється і характер її ж розподілу вздовж катода. При незначному розжарені катод має довші охолоджені кінці, і його робоча довжина, а відповідно, і діюча поверхня, значно менші за геометричні розміри цього електрода. Тому в формулі (20.8) коефіцієнт α і величина струму значно менші від теоретичного значення при заданій U_A. З підвищенням температури розжарення збільшується робоча довжина катода, що призводить до зростання коефіцієнта α і згідно (20.8), струм I_A стає більшим при тому самому значенні U_A.

3.На крутизну ділянки зростання ВАХ впливає спад потенціалу вздовж катода. Струм розжарення, що проходить через катод, створює падіння потенціалу по його довжині. Внаслідок цього потенціали різних точок катода неоднакові. Потенціал катода поступово збільшується від його кінця з меншим потенціалом до його кінця з більшим потенціалом. Завдяки цьому потенціал анода відносно різних ділянок катода різний, а отже не однакова і густина струму на різних ділянках катода.

Ділянка III ВАХ описується законом Річардсона-Дешмана:

$$j_s = AT^2 \exp\left\{-\frac{\Phi}{kT}\right\}, \qquad (20.9)$$

де j_s – густина струму насичення; $A = \frac{4\pi m_e e k^2}{h^3}$ – стала, що не залежить від виду металу; Φ – робота виходу електронів із металу; k – постійна Больцмана; m_e – маса електрона; h – стала Планка; e – заряд електрона.

Формула (20.9) є основним законом термоелектронної емісії. Вона характеризує залежність струму емісії від температури і термодинамічної роботи виходу. Ця залежність експоненційна, тобто струм насичення дуже різко залежить від температури.

Точне виведення формули (20.9) виходить за рамки курсу, але можемо сказати, що струм насичення є пропорційним швидкості і кількості електронів, що виходять за межі металу, тобто площі заштрихованої фігури на рис.20.2,*б*.

Із формули (20.9) випливає, що струм насичення не повинен залежати від напруги між катодом (нагрітий метал) і анодом.

Однак експериментально (в області III рис.20.6) спостерігається залежність струму насичення від анодної напруги, всупереч передбаченням формули (20.9). Цей ефект пояснив Шотткі. Зростання струму насичення зі збільшенням анодної напруги пов'язане зі зменшенням роботи виходу електронів з катода при збільшенні електричного поля між катодом і анодом (формула 20.7).

Тому замість формули (20.9) необхідно записати, що

$$j = AT^2 \exp\left\{-\frac{\Phi - |\Delta\Phi|}{kT}\right\} = j_s \exp\left\{\frac{|\Delta\Phi|}{kT}\right\},$$

або

$$j = j_s \exp\left\{\frac{e}{kT}\sqrt{\frac{eE}{4\pi\varepsilon_0}}\right\}. \qquad (20.10)$$

Формула (20.10) пояснює збільшення струму при $U>U_н$ внаслідок ефекту Шотткі, що проявляється у відсутності насичення струму на ВАХ, особливо при значних температурах катоду. Таким чином, можна теоретично пояснити вигляд основних ділянок ВАХ електронної лампи (діода).

Явище термоелектронної емісії використовують для створення катодів в різних електронних приладах. Для виготовлення катодів найбільше розповсюдження отримав вольфрам, рідше використовується молібден,

тантал, цирконій і деякі інші метали. Завдяки великій тугоплавкості робоча температура вольфрамового катода може бути доведена до 2500К, при якій, не зважаючи на велику роботу виходу (4,5еВ), густина струму досягає $(0{,}3 \dots 0{,}5) \cdot 10^4$ А/м^2. Металічні катоди мають малу ефективність, оскільки вимагають великих затрат на нагрівання до високої температури.

З іншого боку, для практики дуже важливо поможливості зменшити робочу температуру катода електронної лампи, бо при цьому зменшується потужність, що витрачається на розжарювання катода, і збільшити строк служби лампи. Тому тепер поряд з катодами з чистих тугоплавких металів (вольфрам, молібден) широко застосовують катоди складнішої будови.

Велике застосування в техніці має так званий *оксидний катод*. Він має металеву підкладку (керн) (рис.20.8), на яку нанесено

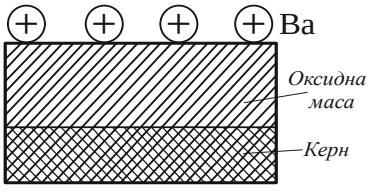


Рис.20.8. Структура оксидного катоду (схематично).

шар окисів лужноземельних металів (BaO, SrO, CaO або їх суміш). Для розжарювання катода через керн пропускають струм («катоди прямого нагрівання») або нагрівають катод, користуючись допоміжною металевою спіраллю ("підігрівні катоди"). Щоб надати катоду високої емісійної здатності, його додатково обробляють (активують); при цьому через електронну лампу при температурі катода близько 1300К протягом деякого часу пропускають термоелектронний струм. Згідно з сучасними поглядами при активуванні катода на його поверхні виникає одноатомний шар позитивних іонів лужноземельного металу (рис.20.8), який значно зменшує роботу виходу і цим збільшує емісійну здатність катода.

При виготовленні оксидних катодів на керн спочатку наносять вуглекислі сполуки лужноземельних металів ($BaCO_3, SrCO_3, CaCO_3$) і потім прожарюють катод перед активуванням у вакуумі. При цьому вуглекислі сполуки розкладаються за реакцією

$$BaCO_3 \rightleftarrows BaO + CO_2$$

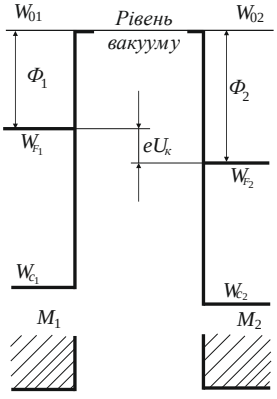
і керн покривається окисами.

Сучасні оксидні катоди мають високі якості. Їх робоча температура дорівнює 1070-1170 К, а іноді і менша. Їх нормальна емісійна здатність досягає 10^4 А/м^2. Для порівняння зазначимо, що робоча температура вольфрамових катодів лежить близько 2500°С, а термоелектронні струми, які знімаються з них, на практиці не перевищують 10^3 А/м^2. При дуже короткочасних струмах (імпульси струму тривалістю $10^{-6} - 10^{-5}$ сек) оксидні катоди здатні давати емісію до 10^6 А/м^2 і вище.

§20.4. Контакт двох металів

Утворення контакту між двома твердими тілами неминуче викликає взаємний обмін рухомими носіями заряду. При однаковій температурі тіл в стаціонарному стані встановлюється термодинамічна рівновага, яка характеризується постійним значенням рівня Фермі, єдиним для всього колективу рухомих носіїв. Встановлення рівноваги супроводжується переходом електронів, який здійснюється в результаті термоелектронної емісії або тунельного проходження, або дифузійних процесів в залежності від якості контакту і властивостей тіл, які контактують.

Розглянемо два метали, енергетичні діаграми яких приведені на рис.20.9,*а*. При достатньо малій відстані (порядку декількох Å),



а

Рис.20.9, *а*.

між ними можливий обмін електронами внаслідок термоелектронної емісії. Якщо робота виходу $\Phi_1 < \Phi_2$, то струм емісії (20.9) j_1 буде перевищувати j_2, і другий метал набуде надлишкового електронного заряду. Зменшення його потенціалу викликає появу в зазорі електричного поля, яке перешкоджає потоку електронів із першого металу в другий. На енергетичній діаграмі (рис.20.9,*а*) це відобразиться взаємним зміщенням енергетичних рівнів і появою додаткового потенціального бар'єру $eU_к$, який зменшує струм j_1 і практично не впливає на струм j_2:

$$j_1 = A_1 T^2 \exp\{-(\Phi_1 + eU_к)/(kT)\},$$
$$j_2 = A_2 T^2 \exp\{-\Phi_2/(kT)\}. \qquad (20.11)$$

Рівновага встановлюється, коли вирівняються зустрічні потоки електронів. Значення рівноважного потенціального бар'єру знаходимо з умови $j_1 = j_2$ (для простоти приймемо, що $A_1 = A_2$):

$$eU_к = \Phi_2 - \Phi_1. \qquad (20.12)$$

Різниця потенціалів $U_к$, яка виникає між металами в рівноважному стані, називається контактною. Зауважимо, що перехід електронів з одного металу в інший

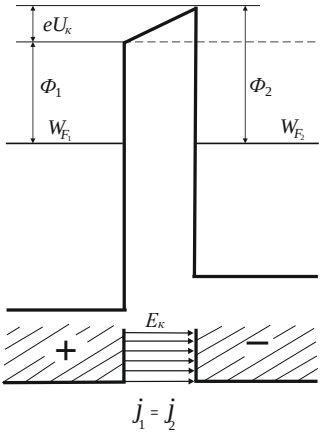


Рис.20.9, *б*.

змінює концентрацію електронів в них, але настільки незначно, що значення енергії Фермі W_F, яка відраховується від дна зони провідності W_c і залежить від концентрації електронів (§19.7, формула (19.22)), практично не змінюється. Тому контактна різниця потенціалів визначається відстанню між початковими положеннями рівнів Фермі (рис.20.9,*а*), або, що теж саме, різницею початкових значень термодинамічних робіт виходу (20.9,*б*).

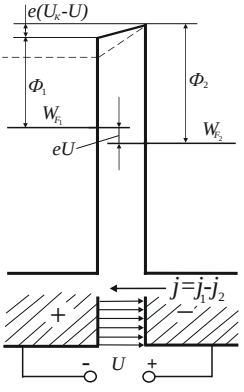


Рис. 20.9, *в*.

Рівновагу металів, які контактують, можна порушити, якщо до них прикласти напругу U від зовнішнього джерела. В залежності від знаку цієї напруги (позитивною будемо вважати напругу, яка створює електричне поле, що має напрямок протилежний до напрямку контактного поля, як показано на рис.20.9,*в*) буде змінюватись висота бар'єру і величина струму j_1:

$$j_1 = A_1 T^2 \exp\{-(\Phi_1 + eU_к - eU)/(kT)\}.$$

Тоді результуючий струм j, який протікає через контакт, дорівнює різниці $j_1 - j_2$. Із врахуванням (20.12)

$$j = j_s(\exp\{eU/(kT)\} - 1), \quad (20.13)$$

де $j_s = A_2 T^2 \exp\{-\Phi_2/(kT)\}$ – так званий струм насичення.

Вираз (20.13) описує вольт-амперну характеристику контакту двох металів. Її вид приведений на рис.20.10 і свідчить про появу

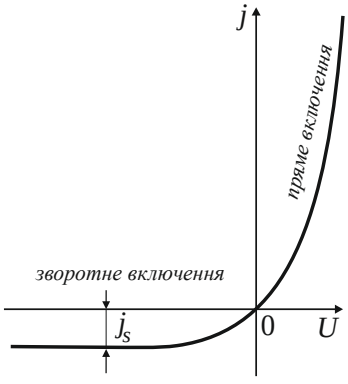


Рис.20.10.Ідеалізована вольтамперна характеристика контакту двох металів, які розділені вакуумним проміжком.

нелінійних властивостей контакту, тобто про різний опір в прямому і зворотному напрямку включенні зовнішнього джерела. Необхідно мати на увазі, що мова йде не про ідеальний контакт, а контакт з досить малим вакуумним зазором між металами. Розглянуті явища відбуваються в приладах, які використовують структури типу МДМ (метал-діелектрик-метал), які можна представляти як контакт двох металів з вакуумним зазором, який замінює тонка діелектрична плівка.

Вольтамперні характеристики МДМ−структур відрізняються від ВАХ, які визначаються за формулою (20.13), оскільки тут не врахована можливість тунельного проходження електронів через бар'єр, а також сили дзеркального відображення, які суттєво змінюють форму бар'єру і викликають ефект Шотткі (формула (20.7)).

Контрольні питання

1.Що таке термоелектронна емісія? Який її механізм?

2.Сформулювати закон термоелектронної емісії Богуславського-Ленгмюра.

3.Сформулювати закон термоелектронної емісії Річардсона-Дешмана

4.Що таке робота виходу електрона із металу?

5.Які фізичні фактори визначають величину роботи виходу?

6.Зобразити ВАХ (теоретичні) та дати пояснення їм.

7.Які причини відхилення теоретичних ВАХ від експериментальних?

8.Що таке ефект Шотткі та як він вливає на вигляд ВАХ?

9.Які основні властивості контакту двох металів?

Розділ 21.Контактні явища

Властивості напівпровідникових пристроїв визначаються параметрами енергетичних бар'єрів, які виникають між двома частинами напівпровідникового кристалу. Найбільш важливими параметрами бар'єрів є їхня енергетична висота, ширина і розподіл напруженості електричного поля всередині бар'єру.

§21.1.Проникнення зовнішнього електричного поля в метал, діелектрик, напівпровідник

Метал. Розмістимо всередині конденсатора металічну пластину, так як показано на рис.21.1,*а*. Як відомо із §3.1 напруженість електричного поля всередині метала дорівнює нулю. Поки електричного поля біля металу не має, в кожному, навіть дуже маленькому об'ємі всередині металу і на його поверхні число позитивних іонів точно дорівнює числу негативно заряджених електронів. Тому будь-яка частина металу електронейтральна.

В металі, який розміщений в електричному полі, електронейтральність біля поверхні порушується. До поверхні, яка повернута до позитивного електроду зовнішнього джерела, притікають електрони. А від поверхні, повернутої до негативного електроду, вони відтікають, залишаючи біля поверхні шар позитивно заряджених іонів (рис.21.1,*а*). Таким чином, заряд на поверхні металу створює електричне поле, яке направлене проти зовнішнього поля, яке його викликало. Зовнішнє поле повністю екранується, і в результаті поле всередині металу дорівнює нулю.

Це можна ще зрозуміти із таких міркувань. Відомо із §1.5, що густина силових ліній пропорційна напруженості електричного поля і лінії починаються на позитивних і закінчуються на негативних зарядах. Тому, який би не був густий потік силових ліній поля, яке створене зовнішнім джерелом біля поверхні металу, внаслідок великої концентрації електронів в металі ($\sim 10^{28}$м$^{-3}$), силові лінії поля завжди замикаються на поверхневих зарядах і не проникають в глибину провідника (рис.21.1,*а*). Але якщо біля поверхні провідника електронейтральність порушується, то що ж означає термін «біля»! На яку все ж таки глибину поле проникає в метал?

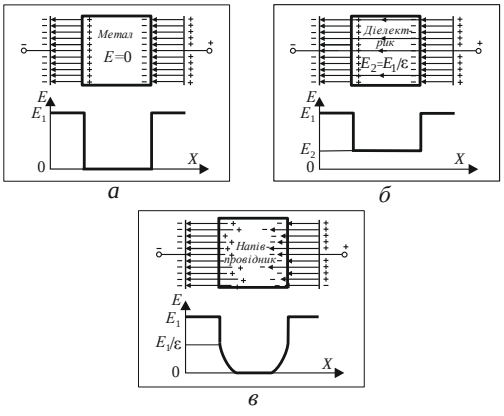


Рис.21.1. Розподіл поля в металі (*а*), діелектрику (*б*) і напівпровіднику (*в*), які розміщені у зовнішньому електричному полі.

Допустимо, наприклад, що під дією зовнішнього поля для того, щоб екранувати його, всі електрони в металічній пластині відтекли із шару, товщина якого Δx дорівнює всього тільки постійній гратки, яка в твердих тілах дорівнює приблизно $5 \cdot 10^{-10}$м. Яке поле створюється при цьому біля поверхні металу ? Або, що теж саме, яке зовнішнє поле буде при цьому екрановане ? Якщо із якої-небудь ділянки металу пішли всі «вільні» електрони, то всередині цієї ділянки виникає позитивний об'ємний заряд з густиною $\rho = |e|N$, де $N-$концентрація позитивно заряджених іонів, яка точно дорівнює концентрації електронів n в металі. Об'ємному

заряду ρ біля поверхні в шарі, товщина якого Δx, відповідає густина поверхневого заряду $\sigma = \rho\Delta x = eN\Delta x$. Тоді згідно (3.1) напруженість електричного поля *E* біля поверхні металу буде:

$$\begin{cases} E = \frac{\sigma}{\varepsilon_0} = \frac{eN\Delta x}{\varepsilon_0} = \\ = \frac{1{,}6\cdot10^{-19}\cdot10^{28}\cdot5\cdot10^{-10}}{8{,}85\cdot10^{-12}}\text{В/м} = \\ = 9\cdot10^{10}\text{В/м}. \end{cases}$$

Отримане значення *E* дуже велике. При $E{\sim}3\cdot10^6$В/м у повітрі виникає електричний пробій. Але щоб екранувати дію навіть такого сильного поля, достатньо щоб із приповерхневого шару металу, товщина якого в один атомний шар, пішла $\frac{3\cdot10^6}{9\cdot10^{10}} \approx 3\cdot10^{-5}$ – стотисячна доля електронів.

Таким чином, навіть дуже сильні електричні поля не проникають в метал на глибину більшу, ніж постійна решітки. Та навіть в межах цієї товщини відносне збіднення електронами зі сторони "мінуса" зовнішнього джерела (і, відповідно, відносне збагачення зі сторони «плюса») не перевищує декількох долей відсотка.

Діелектрик. Проникнення електричного поля в діелектрик розглядалось в розділі 2. Тут тільки зазначимо, що на мові силових ліній діелектрик в електричному полі показаний на рис.21.1,*б*. Частина силових ліній зовнішнього поля залишається на зв'язаних поверхневих зарядах діелектрика. Але ті силові лінії, які не замкнулись на поверхневих зв'язаних зарядах і проникли всередину діелектрика, пронизують його наскрізь. Всередині діелектрика не має вільних зарядів, які могли би перегрупуватись і, змінивши своє положення, екранувати дію електричного поля, яке проникло в діелектрик.

Напівпровідник. В пластині напівпровідника, яка розміщена в електричному полі (рис.21.1,*в*), з одного, боку, як в діелектрику, на поверхні виникнуть внаслідок поляризації зв'язані заряди. Таким чином, в напівпровіднику біля межі з вакуумом (або повітрям) напруженість поля E_2 буде в ε раз менша, ніж в вакуумі. З другого боку, в напівпровіднику, як і в металі, існують вільні носії заряду, які здатні, перегрупувавшись, екранувати дію електричного поля, яке проникло в напівпровідник.

На яку глибину поле проникає в напівпровідник? В чому різниця між напівпровідником і металом? Різниця в тому, що концентрація вільних носіїв і відповідно заряджених іонів в напівпровіднику, як правило, значно менша ніж в металі. Тому, щоб створити достатньо великий позитивний заряд, який здатний екранувати дію зовнішнього поля, електрони в напівпровіднику *n*-типу повинні "відтікти" не із тонкого шару, товщина якого порядку постійної гратки, а із шару досить великої товщини (рис.21.1,*в*). Ясно, що товщина цього шару буде тим більша, чим менша рівноважна концентрація електронів n в напівпровіднику.

Концентрація електронів n практично в багатьох випадках визначається концентрацією донорів N_d. Таким чином, глибина, на яку поле проникає всередину напівпровідника, визначається рівнем його легування.

Рівень легування напівпровідників, які використовуються для виготовлення різних напівпровідникових приладів, лежить в межах від ${\sim}10^{16}$ до ${\sim}10^{26}$м$^{-3}$, тобто може відрізнятись в десятки мільярдів раз. Відповідно і глибина проникнення поля в напівпровідник, або, іншими словами, ширина енергетичних бар'єрів в напівпровідниках може змінюватись в дуже широких межах. В

дірковому напівпровіднику ситуація із проникненням електричного поля в напівпровідник аналогічна як і в електронному напівпровіднику.

§21.2.Розподіл поля в бар'єрі. Ширина бар'єру

Поверхневі стани. Сама поверхня напівпровідникового кристалу є порушенням в ідеальній періодичній решітці. Ідеальний порядок атомів, які розташовані в об'ємі напівпровідника на строго однаковій один від одного відстані, на поверхні раптово порушується. Ясно, що умови зв'язку між валентними електронами, які належать приповерхновим атомам, будуть дещо іншими, ніж в об'ємі. В 1932р. І.Е.Тамм теоретично показав, що обрив кристалічної решітки є причиною виникнення біля поверхні додаткових енергетичних рівнів. Ці поверхневі стани носять назву рівнів Тамма. Крім того, на поверхні адсорбуються чужеродні атоми, на поверхні завжди існують структурні дефекти: спотворення решітки, вакансії, тобто пусті місця в вузлах решітки, які виникають, коли атоми напівпровідника зміщуються із своїх законних місць в вузлах гратки в міжвузловини. Всі ці домішкові і структурні порушення теж є джерелом поверхневих станів.

Густина поверхневих станів на реальних поверхнях знаходиться в межах від $\sim 10^{13}$до $\sim 10^{19}$м$^{-2}$. Поверхневі рівні можуть бути як донорними так і акцепторними.

В залежності від того, який тип поверхневих станів переважає (донорний чи акцепторний), поверхня буде або віддавати електрони в об'єм напівпровідника і заряджатись позитивно, або навпаки, захоплювати електрони із об'єму і заряджатись негативно.

Нехай на поверхні напівпровідника $n-$ типу переважають акцепторні стани. Тоді поверхневі рівні будуть захоплювати електрони із частини об'єму напівпровідника, який прилягає до поверхні. В результаті поверхня заряджається негативно. А в приповерхневій області напівпровідника виникне шар, збіднений електронами, тобто заряджений позитивно. Таким чином, біля поверхні утворюється подвійний заряджений шар, поле якого буде відштовхувати електрони в глибину кристалу. Це поле максимальне біля поверхні і спадає по мірі віддалення від неї в глибину кристалу. Це поле відштовхує електрони від межі кристалу, а позитивно заряджені дірки притягуються до межі (рис.21.2).

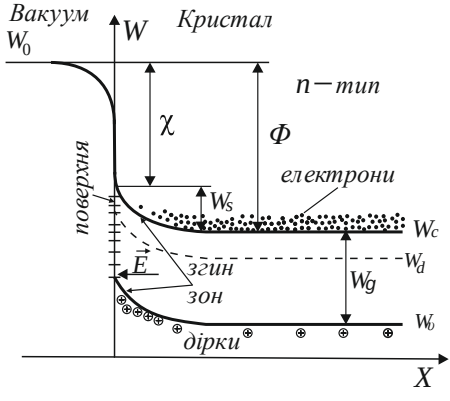


Рис.21.2.

Біля межі всередині кристалу утворився, із-за подвійного зарядженого шару, енергетичний бар'єр, висота якого W_s. Величина $U_s = W_s / e$ називається поверхневим потенціалом. Біля самої межі концентрація електронів провідності буде:

$$n_s = n_n \exp\left\{-\frac{W_s}{kT}\right\}, \qquad (21.1)$$

де $n_n -$ концентрація електронів в глибині кристалу далеко від межі.

Енергія дірок біля межі менша, ніж енергія дірок в об'ємі кристалу на величину W_s. Тому концентрація дірок біля поверхні p_s буде дорівнювати

$$p_s = p_n \exp\left\{\frac{W_s}{kT}\right\}, \qquad (21.2)$$

де $p_n -$ концентрація дірок в глибині кристалу. Розберемося, до чого призводить наявність на межі кристалу поверхневого потенціалу і згину зон, який супроводжує його. Нехай в кремнії n-типу концентрація електронів $n_n = 10^{21}\text{м}^{-3}$. Тоді згідно закону діючих мас (19.48), рівноважна концентрація дірок в такому кристалі при T=300K, $p_n = \frac{n_i^2}{n_n} = \frac{(1{,}1\cdot 10^{16})^2}{10^{21}} = 1{,}21\cdot 10^{11}\text{м}^{-3}$. Електронів в об'ємі кристалу в 10^{10} (10 міліардів) раз більше, ніж дірок.

Нехай величина поверхневого потенціалу $W_s = 0{,}1\text{еВ}$. Тоді при кімнатній температурі (T=300K) kT=0,026еВ і концентрація електронів біля поверхні буде згідно (21.1): $n_s = 10^{21}\cdot\exp\left\{\frac{-0{,}1}{0{,}026}\right\} \approx 2\cdot 10^{19}\text{м}^{-3}$, а концентрація дірок згідно (21.2) $p_s = 5{,}66\cdot 10^{12}\text{м}^{-3}$. Електронів біля поверхні приблизно в 50 раз менше, ніж в об'ємі, дірок в стільки ж раз більше. Але дірок все ще в $3{,}5\cdot 10^6$ раз менше, ніж електронів і їхній вклад в провідність незначний. Тому наявність поверхневого потенціалу в цьому випадку означає виникнення біля поверхні збідненого електронами шару з питомим опором, який значно більший, ніж об'ємний питомий опір. *Саме виникнення на межі кристалу збіднених шарів з високим питомим опором лежить в основі роботи діодів і транзисторів*.

При $W_s = 0{,}2\text{еВ}$ $n_s = 4{,}6\cdot 10^{17}\text{м}^{-3}$, $p_s = 2{,}7\cdot 10^{14}\text{м}^{-3}$. Якісно ситуація залишилась тією самою $n_s/p_s = 1{,}7\cdot 10^3$. При $W_s = 0{,}4\text{еВ}$ виникає якісно нова ситуація. Тепер $n_s = 2\cdot 10^{14}\text{м}^{-3}$, $p_s = 6\cdot 10^{17}\text{м}^{-3}$. Дірок біля поверхні майже в три тисячі раз більше, ніж електронів! На поверхні напівпровідника $n-$типу утворився шар з дірковою провідністю, так званий шар з інверсним типом провідності, або *інверсний шар*.

На рис.21.3 приведена енергетична діаграма для діркового напівпровідника при позитивному заряді поверхневих станів. На рис.21.2 і 21.3 приведене таке позначення: $\chi -$електронна спорідненість напівпровідника і, як видно із рисунків $\chi \approx W_o - W_c$ (для даного напівпровідника є величина стала); $\Phi -$ робота виходу електронів із напівпровідника.

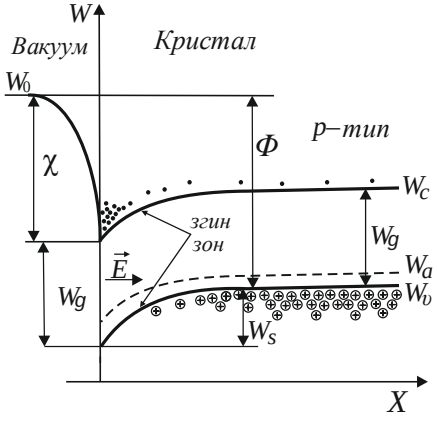


Рис. 21.3

Для $n-$ і $p-$ кремнію, величина поверхневого бар'єру W_s приблизно дорівнює $W_g/2$ (цей випадок зустрічається дуже часто). Тоді для $n-$кремнію рис.21.2 робота виходу дорівнює $\Phi_n \approx \chi + W_s = \chi + W_g/2$, а для $p-$кремнію (рис.21.3) $-\Phi_p \approx \chi + W_g - W_s \approx \chi + W_g/2$. Таким чином, $\Phi_n \approx \Phi_p$. Тобто, робота виходу електрона із напівпровідника $n-$ і $p-$ типу (кремнію) приблизно дорівнює значенню роботи виходу у власного кремнію.

Визначимо розподіл електричного поля і ширину енергетичного бар'єру в області поверхневого потенціалу W_s (рис.21.4). Для цього використаємо рівняння Пуассона (1.38) і деякі припущення про характер бар'єру і напівпровідник.

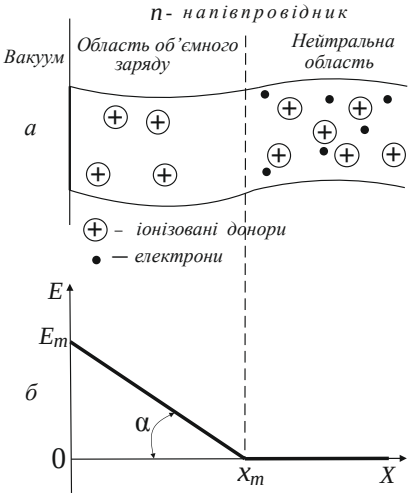


Рис.21.4.Формування поверхневого потенційного бар'єру (напівпровідник *n*-типу): *а* – негативний заряд поверхневих станів спричинив відтік електронів із приповерхневого шару, де залишився нескомпенсований позитивний заряд іонізованих донорів; *б* – поле приповерхневої області напівпровідника лінійно зменшується від величини E_m (на границі вакуум-кристал, $x = 0$) до нуля (на границі область об'ємного заряду – нейтральна область, $x = x_m$), $\mathrm{tg}\alpha = eN_d/(\varepsilon\varepsilon_0)$.

По-перше, будемо вважати, що напівпровідник легований однорідно, тобто концентрація домішки скрізь однакова і не залежить від координати. Якщо це напівпровідник $n-$типу, то концентрація іонізованих донорів N_d скрізь однакова. По-друге, будемо вважати, що висота потенційного бар'єру W_s, з одного боку, достатньо велика: $W_s \gg kT$, але, з іншого боку, недостатня для того, щоб викликати інверсію провідності біля межі кристалу. За цих двох умов розрахунок поля в бар'єрі стає дуже простим.

Дійсно, потенційний бар'єр на межі обумовлений тим, що рухомі носії (в даному випадку електрони) відтекли від межі, залишивши нескомпенсованим позитивний об'ємний заряд іонізованих донорів (рис.21.4,*а*). При $W_s \gg kT$ із приповерхневого шару йдуть практично всі вільні електрони. Тоді густина об'ємного заряду, природно, дорівнює нескомпенсованому заряду іонізованих донорів, які залишились: $\rho = eN_d$, де $N_d -$ концентрація донорів, $e -$ величина заряду електрона за модулем. Тоді згідно рівняння Пуассона (1.38), поле всередині поверхневого шару буде змінюватись за законом

$$\frac{dE}{dx} = \frac{\rho}{\varepsilon\varepsilon_0} = \frac{eN_d}{\varepsilon\varepsilon_0}. \qquad (21.3)$$

Так як $N_d -$ величина постійна, то, у відповідності з формулою (21.3), швидкість зміни поля з координатою – теж величина стала, тобто напруженість поля E лінійно спадає з координатою x (рис.21.4,*б*). Із формули (21.3) і (рис.21.4,*б*) легко отримати, що

$$E(x) = \frac{eN_d}{\varepsilon\varepsilon_0}(x_m - x), \qquad (21.4)$$

де $x_m -$ межа об'ємного заряду.

Якби нам була відома величина $E_m -$ поле на межі кристалу, то із (21.4) ми легко визначили б ширину бар'єру x_m. Із (21.4) видно, що при $x = 0, E = E_m$,

$$E_m = \frac{eN_d}{\varepsilon\varepsilon_0}x_m. \qquad (21.5)$$

Але значення E_m невідоме. Відомо, що між межею кристалу і точкою $x = x_m$, де поле відсутнє, існує потенційний бар'єр висотою W_s, що означає, що між цими точками існує різниця потенціалів $U_s = W_s / e$. Цього достатньо, щоб визначити і величину E_m, і ширину бар'єру x_m. Дійсно, так як

$$E = -dU_s / dx, \text{ то } U_s = -\int_0^{x_m} E(x)dx.$$

Із рис.21.4,*б* видно, що U_s формально дорівнює площі трикутника $OE_m x_m$, тобто

$$U_s = \frac{1}{2} E_m x_m \cdot \qquad (21.6)$$

Тоді із співвідношень (21.5) і (21.6) отримаємо, що

$$x_m = \left(\frac{2\varepsilon\varepsilon_0 U_s}{eN_d}\right)^{\frac{1}{2}};$$
$$E_m = \left(\frac{2eN_d U_s}{\varepsilon\varepsilon_0}\right)^{\frac{1}{2}}. \qquad (21.7)$$

Проведемо деякі числові оцінки. Приймемо, що $U_s = 0{,}2$В. Значення діелектричної проникності для найбільш широко розповсюджених напівпровідників Ge, Si і GaAs є величими: 16, 12,5 і 12 відповідно. Тоді, якщо $N_d = 10^{16}$м$^{-3}$, то за формулами (21.7) отримаємо: $E_m = 2{,}4 \cdot 10^3$ В/м, $x_m = 170$ мкм. Якщо ж при тій самій приповерхневій різниці потенціалів $U_s = 0{,}2$В рівень легування $N_d = 10^{26}$м$^{-3}$, то $E_m = 2{,}4 \cdot \cdot 10^8$В/м, $x_m = 0{,}0017$ мкм.

В першому випадку поле при поверхневого бар'єру проникає в глибину напівпровідника на товщину порядку 300 000 атомних шарів. У другому випадку – 3÷4 шарів.

Зрозуміло, якщо ми маємо справу з дірковим напівпровідником, то всі приведені судження і всі отримані формули є справедливими. Необхідно тільки замінити в них величину N_d на $(-N_a)$ – концентрацію акцепторної домішки.

§21.3.*p–n* перехід та його властивості

Якщо кристал напівпровідника легований так, що одна частина, має провідність p – типу (діркову), а інша – провідність n – типу (електронну), то на границі між цими частинами виникає шар з особливими властивостями, який називається $p-n$ переходом або, по–іншому, електронно–дірковим переходом.

Способи отримання $p-n$ переходів

Якщо взяти два куски напівпровідника (рис.21.5,*а*): один p –, інший n – типу, акуратно їх відшліфувати і міцно притиснути один до одного, то ми отримаємо об'єкт, властивості якого не мають нічого спільного із властивостями p–n переходу. Із §21.2 відомо, що на поверхні, як би її акуратно не обробляли, завжди будуть порушення кристалічної решітки. Утворюється спектр поверхневих станів, тобто виникають приповерхневі потенціальні бар'єри.

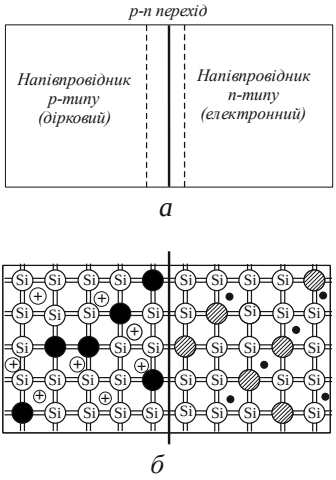


Рис.21.5.На межі p– і n– типів напівпровідника виникає p – n перехід (*а*). Показаний кристал кремнію, в лівій частині якого акцепторна домішка (темні кружечки), в правій – донорна (заштриховані кружечки). Позначення: ⊕ – дірки, ● – електрони (*б*).

Крім цього, на поверхні будуть окиси, адсорбовані сторонні атоми …. Як щільно не притиснуті один до одного куски p– і n– типу напівпровідників, «рани», які нанесені кристалічній решітці на поверхні, не залікуєш.

Яким же повинен бути контакт p– і n– напівпровідників? Звернувшись до рис.21.5,*б*, зазначимо важливу обставину. Кристалічна

решітка напівпровідника нічого про *p–n* перехід "не знає", вона просто його не помічає. На самій межі, яка розділяє $p-$ і $n-$ області, чи зліва, чи справа від неї, ми бачимо все ту ж кристалічну решітку кремнію, з тими ж зв'язками, які ніяк не порушуються тим, що в різні частини кристалу введені різні домішки. Тільки такий перехід має всі ті чудові властивості, про які будемо говорити в подальшому і саме такий перехід будемо вивчати.

Проблема створення якісних $p-n$ переходів вирішувалась і багато в чому вирішена на протязі більше півстоліття зусиллями тисяч спеціалістів. Розроблено досить багато способів отримання переходів і всі ці способи доведені до високої технічної досконалості. *Для всіх способів характерний спільний прийом*. Береться кристал напівпровідника $p-$ або $n-$ типу, а потім якимось способом на певну глибину кристалу вводяться домішки, які створюють в даному напівпровіднику провідність протилежного характеру. Якщо вихідний кристал був $n-$ типу, то вводиться акцепторна домішка, якщо $p-$ типу, то донорна. Домішка вводиться в кристал в більшій кількості, ніж вихідна. Тому в цій частині кристалу, куди попала домішка, тип провідності змінюється. Відбувається, як говорять, перекомпенсація вихідної домішки. На межі між областю, де тип провідності кристалу змінився, і областю, де він залишився вихідним, виникає $p-n$ перехід.

Сплавлення. Сплавлення один із самих старих способів створення *p–n* переходів. Перші *p–n* переходи були отримані цим методом в 1950р. В основі методу лежить утворення сплаву напівпровідникового матеріалу і металу, атоми якого є донорами або акцепторами в напівпровіднику.

На рис.21.6 показаний приклад утворення сплавного $p-n$ переходу в $n-$Ge при вплавленні в нього індія (In). Таблетка індію, яка має форму тонкого диску, клалась на очищену поверхню кристалу германію (рис.21.6,*а*). Потім кристал разом з таблеткою повільно нагрівають. Як тільки температура перевищить температуру плавлення індію (156℃) таблетка починає плавитись, перетворюючись в краплю рідини. При підвищенні температури форма краплі, як видно із рис.21.6,*а*, змінюється. Сили поверхневого натягу рідини стараються надати краплі форму кулі, а сили змочування – розмазати краплю по поверхні кристалу. Індій при $t \simeq 300℃$ має форму кульки, при більш високій температурі ($\approx 500℃$) розтікається по поверхні германію.

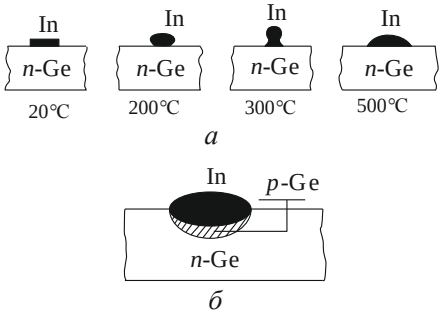


Рис.21.6.

Кристалічний германій при температурі $\sim 500℃$ дуже добре розчиняється в рідкому індії. Розчинення продовжується до тих пір, поки не утвориться насичений розчин германію і індію. Після цього пластинку германія починають повільно охолоджувати. Із зменшенням температури кількість германію, здатного розчинятись в індії,теж зменшується; надлишковий германій, який випадає із розчину, знову кристалізується на нерозчиненій частині пластинки (рекристалізується). Цей рекристалізований шар (рис.21.6,*б*) є збагачений атомами індію.

Атоми індію в германію є акцепторами. Концентрація індію в рекристалізованому шарі складає $\sim 10^{25}$м$^{-3}$ і значно перевищує концентрацію донорної домішки у вихідному кристалі. Тому на межі рекристалізованої області і нерозчиненої частини кристалу утворюється $p-n$ перехід.

В 50–ті роки 20 століття метод сплавлення відігравав важливу роль в розвитку напівпровідникової електроніки. Саме на його основі були створені перші промислові германієві діоди і транзистори.

Дифузія. Дифузійний метод виготовлення $p-n$ переходів з'явився незабаром після методу сплавлення як деяка його модифікація. Метод дифузії виявився більш ефективним і на довгі роки був основним методом створення $p-n$ переходів, а потім і інтегральних схем.

Розглянемо процес дифузії детальніше. Нехай напівпровідниковий кристал розміщений в атмосфері газу, який містить велику концентрацію атомів речовини, яка є в даному напівпровіднику донорною (або акцепторною) домішкою. На поверхні любого кристалу є безліч порушень кристалічної гратки. Тому атоми домішки із газу–дифузанта легко захоплюється поверхнею напівпровідника. Поверхня стає збагаченою домішками. Оскільки ідеальних кристалів не існує, то дифузія в глибину кристалу буде відбуватись, головним чином, по вакансіям, міжвузловинах та інших порушеннях в гратці кристалу при його нагріванні. Дифузійний потік домішкових атомів направлений від поверхні, де домішкових атомів багато, в глибину напівпровідника, де їх немає.

На рис.21.7 показана схема проведення процесу дифузії. Кварцева труба 1 розміщена в печі, яка створює всередині труби температуру $1000-1300$℃. В трубі знаходиться кварцева касета 3 із закріпленими на ній пластинками кремнію 4. Речовина домішки (так званий дифузант) поступає в трубу у вигляді газу, концентрація якого строго контролюється. На протязі всього часу, поки поверхня напівпровідника обдувається дифузантом, атоми домішки дифундують в глибину пластинки. Чим більший час дифузії, тим глибше проникають домішки в кристал. $p-n$ перехід при дифузії виникає внаслідок того, що кристал, в якому проводиться дифузія, має вихідну об'ємну домішку, яка протилежна по типу від домішки, яка вводиться при дифузії. Якщо вихідна домішка, наприклад, донорного типу, то для створення $p–n$ переходу використовують дифузію акцепторів.

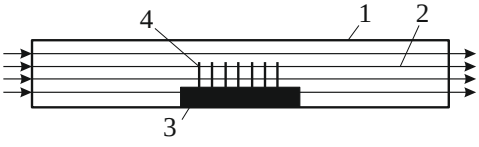


Рис.21.7.

Глибина залягання $p-n$ переходів, які виготовляються методом дифузії, залежить від об'ємної концентрації вихідної домішки в кристалі і лежить в межах від долей мікрометра до сотень мікрометрів.

Іонна імплантація. Це один із сучасних методів виготовлення $p-n$ переходів, який дозволяє отримувати $p–n$ переходи дуже маленьких розмірів з глибиною залягання, яка точно контролюється і є досить малою (до сотих долей мікрометра). Суть методу іонної імплантації, або, як його деколи називають методом іонного легування, в тому, що в напівпровідник проникають атоми домішки, які рухаються строго в одному напрямку з певною і досить великою швидкістю, на відміну від методу дифузії, в якому атоми домішки рухаються хаотично.

Установка для проведення іонної імплантації схематично показана на рис.21.8. В камері 1, де відбувається процес, створюється і підтримується глибокий вакуум. Іони домішки виникають в джерелі іонів 2, яке складається із матеріалу домішки, який нагрівається до температури, щоб почалось випаровування атомів з поверхні. Біля поверхні матеріалу домішки, горить електрична дуга, яка іонізує практично всі частинки, які в неї попадають.

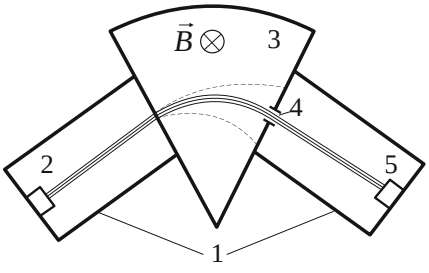


Рис.21.8.

Позитивні іони домішки, які утворились, витягуються через вихідний отвір джерела сильним електричним полем. Особлива система електродів (яку називають електростатичною лінзою) надає пучку іонів необхідну голкоподібну форму. Прискорені полем до великих швидкостей і сфокусовані іони попадають в магнітний аналізатор 3. Це найважливіший вузол, який забезпечує одну із найголовніших переваг методу іонного легування: практично повне очищення пучка іонів від сторонніх домішок. В аналізаторі пучок іонів попадає в магнітне поле (силові лінії магнітного поля $\vec{B}$ на рис.21.8 перпендикулярні площині рисунка). На іони діє сила Лоренца. Радіус траєкторії іонів в магнітному полі залежить від їх швидкості і від маси іонів. Тому, якщо в пучку є сторонні частинки крім іонів легуючої домішки, вони відхиляються в аналізаторі не на той кут, на який відхиляються іони домішки.

В вихідній стінці аналізатора є щілина 4, через яку проходять тільки ті іони, які відхилились строго на заданий кут. “Вибирати” тип іонів в пучку дуже легко: достатньо змінити величину індукції магнітного поля аналізатора.

“Відфільтровані” іони, які вилітають із аналізатора, підлітають до напівпровідникової пластинки 5. При зіткненнях з граткою напівпровідника іони поступово витрачають свою енергію, віддаючи її атомам, які попадаються на їхньому шляху, і в кінці кінців зупиняються. Шлях, який іон проходить до повної зупинки, або глибина проникнення, залежить від маси і енергії іона. Енергію іонів можна точно регулювати, змінюючи прискорюючий іони електричний потенціал. Таким чином, проникнення іонів домішок в гратку відбувається на строго задану глибину. Наприклад, іони бора B^+ при енергії 20 кеВ проникають в кремнієву пластинку на глибину 0,071 мкм, при енергії 100 кеВ – на 0,327 мкм, а при енергії 300 кеВ– на 0,559 мкм.

Концентрація введеної домішки визначається сумарною дозою опромінення. Ця величина з високою точністю контролюється часом іонного бомбування. Таким чином, іонна імплантація – один із найточніших методів виготовлення p–n переходів.

На шляху іонного пучка розміщують електростатичні пластини, такі, як в електронно-променевій трубці осцилографів. Подаючи на ці пластини напругу, яка змінюється за заданим законом, який виробляє і управляє ним ЕОМ, можна відхиляти іонну “голку” в потрібну точку пластини. Таким чином, іонний промінь “малює” необхідну конфігурацію $p-n$ переходів. Закінчивши весь “малюнок” для однією домішки і змінивши магнітне поле аналізатора і енергію пучка,

можна зразу ж виконати інший рисунок, який створить іншу систему $p-n$ переходів на іншій глибині, і т.д. Такі установки використовуються для виготовлення інтегральних схем, в яких на одній напівпровідниковій пластині розміщується сотні тисяч і більше $p-n$ переходів.

Потенціальний бар'єр в області контакту p – і n – напівпровідників

З'ясуємо, чому між двома областями одного і того самого напівпровідника, які відрізняються один від одного тільки характером введеної домішки, виникає потенційний бар'єр.

На рис.21.9,*а* зліва показаний напівпровідник $p-$типу, справа $n-$типу. Дірковий напівпровідник містить при звичайних умовах позитивно заряджені дірки – рухомі частинки, які здатні переносити заряд (створювати струм), і негативно заряджені акцептори, закріплені в кристалічній решітці. Електронний напівпровідник, навпаки, містить негативно заряджені електрони провідності і позитивно заряджені іонізовані донори, які закріплені в решітці.

На рис.21.9,*б* наведена енергетична діаграма напівпровідників $p-$ і $n-$типу. Це, як правило, сильно леговані напівпровідники, в яких концентрація основних носіїв заряду визначається рівнем легування (§19.10; формули (19.62); (19.63)). Концентрація неосновних носіїв заряду в напівпровідниках визначається формулами (19.64) і (19.65) із §19.10. Положення рівня Фермі в напівпровіднику $p-$типу визначається формулою (19.67), а в напівпровіднику $n-$типу

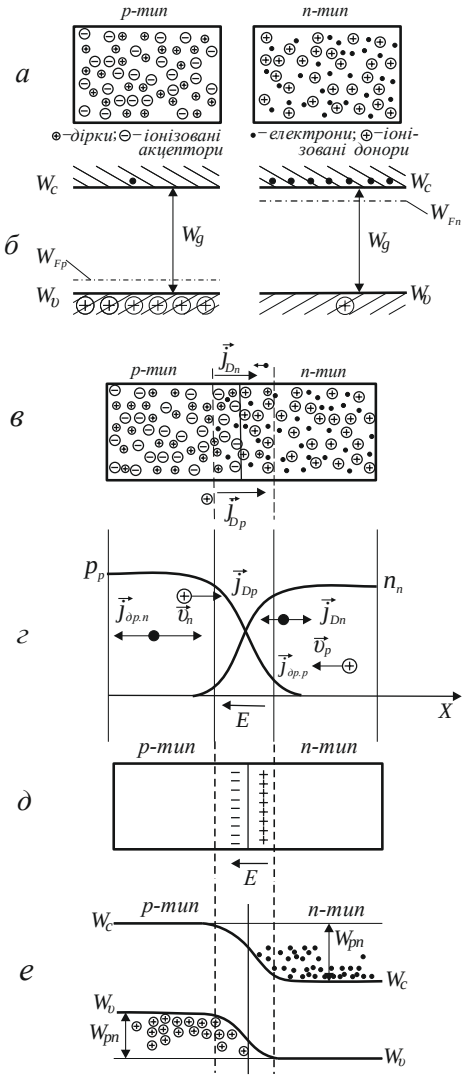


Рис.21.9.Утворення потенційного бар'єру на межі $p-n$ переходу: *а*– напівпровідники $p-$ і $n-$типу, в цілому електронейтральні; *б*– енергетична діаграма напівпровідників, W_{F_p} і W_{F_n} – енергії Фермі, W_g – ширина забороненої зони; *в*–дифузія електронів в напівпровідник $p-$типу і дірок в напівпровідник $n-$типу; *г*– напрямки дифузійних струмів основних і дрейфових струмів неосновних носіїв заряду; *д*– біля межі $p-n$ переходу виникає подвійний заряджений шар; *е*– наявність подвійного зарядженого шару на межі $p-n$ переходу еквівалентно існуванню на границі енергетичного бар'єру. Цей бар'єр перешкоджає проникненню електронів із напівпровідника $n-$типу в напівпровідник $p-$типу і проникненню дірок із напівпровідника $p-$типу в напівпровідник $n-$типу.

формулою (19.66). Рівні Фермі в таких напівпровідниках знаходяться близько біля "стелі" валентної зони W_v і "дна" зони провідності W_c. Таке розташування рівнів Фермі відображає той факт, що напівпровідники сильно леговані і концентрація основних носіїв заряду значно переважає концентрацію неосновних носіїв заряду. Напівпровідники однакові (наприклад, кремній), тому ширина забороненої зони W_g в них однакова.

На рис.21.9,*в* показано, що відбудеться, коли такі напівпровідники утворюють контакт. Повинен виникнути дифузійний потік електронів із напівпровідника *n*–типу в напівпровідник *p*−типу. І навпаки – дифузійний потік дірок із *p*− в *n*–напівпровідник. Густини дифузійних струмів (рис.21.9,*г*) визначаються формулами (19.99) і (19.100) із §19.12. При цьому частина напівпровідника *p*−типу, біля *p*−*n* переходу буде заряджатись негативно, а частина напівпровідника *n*–типу біля *p*−*n* переходу буде набувати позитивний заряд, так як в ній залишається нескомпенсований позитивний заряд нерухомих донорів. Дифузія дірок із *p*− в *n*–область сприяє заряджанню біля *p*–*n* переходу напівпровідника *p*−типу негативно, а *n*–типу – позитивно. Таким чином, біля *p*–*n* переходу утворюється подвійний заряджений шар, негативний зі сторони *p*−області і позитивний зі сторони *n*–області (рис.21.9,*д*), який створює контактне електричне поле $\vec{E}$. Це поле буде протидіяти процесу дифузії. Воно виштовхує дірки із цього шару в *p*−напівпровідник, а електрони в *n*–напівпровідник.

Контактне поле $\vec{E}$ сприяє дрейфу електронів із *p*–напівпровідника в *n*−область і дрейфу дірок із *n*−напівпровідника в *p*–область (рис.21.9,*г*). Густини дрейфових струмів визначається формулою (19.95) і в нашому випадку можуть бути записані так:

$$j_{др.n} = en_p v_n = e\frac{n_i^2}{N_a}v_n, \qquad (21.8)$$

$$j_{др.p} = ep_n v_p = e\frac{n_i^2}{N_d}v_p. \qquad (21.9)$$

При написанні цих виразів враховані формули (19.64) і (19.65). Розглянемо процеси детальніше. Неосновні носії в *p*−області з концентрацією n_p у своєму хаотичному русі попадають в область контактного поля, яке є для них прискорюючим, і переходять в *n*–область напівпровідника. Якщо в області переходу їхня кількість не змінюється в результаті, наприклад, процесів рекомбінації або генерації, то величина струму, який створюється електронами, визначається тільки кількістю електронів, що попали на межу прискорюючого поля і майже не залежить від напруженості електричного поля. Аналогічна ситуація і із дрейфом дірок із *n*–області в *p*–область. Тому дрейфовий струм через *p*−*n* перехід буде дорівнювати:

$$j_s = j_{др.n} + j_{др.p} = e\left(\frac{v_n}{Na} + \frac{v_p}{N_\alpha}\right)n_i^2. \quad (12.10)$$

Враховуючи формулу (19.44) можемо зробити висновок, що дрейфовий струм неосновних носіїв

$$j_s \sim \exp\left\{-\frac{W_g}{kT}\right\}, \qquad (21.11)$$

тобто він різко залежить від температури і величини ширини забороненої зони W_g.

Таким чином, в області *p*−*n* переходу встановиться рівновага між потоками електронів і дірок. Але, якщо контактне електричне поле $\vec{E}$ (рис.21.9,*д*) сприяє переходу неосновних носіїв заряду через перехід, то для основних носіїв заряду воно є гальмівним, тобто бар'єром (рис.21.9,*е*). Проте, внаслідок хаотичного теплового руху, серед основних

носіїв заряду, яких багато, знайдуться дуже енергійні, енергія яких буде більша висоти бар'єру W_{pn}. Концентрація їх визначається розподілом Максвелла–Больцмана, тобто

$$n \approx n_n \exp\left\{-\frac{W_{pn}}{kT}\right\}, \qquad (21.12)$$

$$p \approx p_p \exp\left\{-\frac{W_{pn}}{kT}\right\}. \qquad (21.13)$$

Концентрація цих "гарячих" електронів і дірок дуже мала, але і концентрація неосновних носіїв заряду, які утворюють струм (21.10), теж дуже мала. Тому в області $p-n$ переходу встановлюється *динамічна* рівновага між потоками основних і неосновних носіїв заряду, результатом чого є виникнення контактної різниці потенціалів

$$U_{pn} = W_{pn} / e.$$

Рис.21.9,*е* пояснює, що означає існування в місці контакту $p-$ і n–напівпровідників енергетичного бар'єру. Із рисунка видно, що електрону для того, щоб попасти із n–області в $p-$область, необхідно подолати енергетичний бар'єр, висота якого W_{pn}. Такий же бар'єр необхідно подолати і дірці для того, щоб потрапити з $p-$ в n–область $p-n$ переходу.

Висота бар'єру. Як уже обговорювалось в §19.9 рівень Фермі в напівпровіднику відіграє роль "центру мас" між електронами і дірками і його розташування в забороненій зоні напівпровідника завжди ближче до тих носіїв заряду, яких більше. Тому можна вважати, що енергії мілких акцепторних і донорних рівнів W_a і W_d майже співпадають з енергіями Фермі W_{F_p} і W_{F_n} (рис.21.10).

Робота виходу електронів із напівпровідника $p-$типу $\Phi_p = W_0 - W_a = W_0 - W_{F_p}$, а для напівпровідника $n-$типу $\Phi_n = W_0 - -W_d = W_0 - W_{F_n}$, де W_0–енергія вакууму (енергія вільного електрону). Електрону, щоб потрапити з n–області в p–область необхідно подолати енергетичний бар'єр, висота якого

$$W_{pn} = \Phi_p - \Phi_n = W_{F_n} - W_{F_p}.$$

Із урахуванням формул (19.66) і (10.67), отримаємо

$$W_{pn} = W_g - kT \ln \frac{N_c N_v}{N_a N_d}, \qquad (21.14)$$

де враховано, що $W_g = W_c - W_v$. Цій формулі можна надати і такий вигляд:

$$W_{pn} = kT \ln \frac{n_n p_p}{n_i^2} \cdot \qquad (21.15)$$

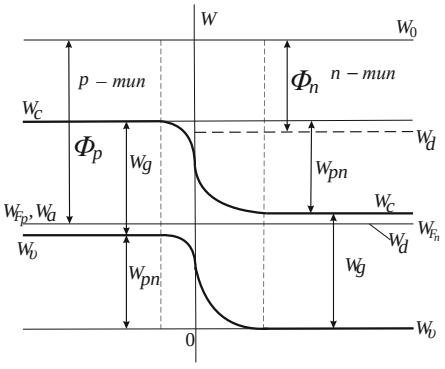


Рис.21.10.Енергетичний бар'єр на межі p–n переходу. $\Phi_p = W_0 - W_a$, $\Phi_n = W_0 - W_d$. W_0– енергія вакууму (енергія вільного електрона). W_a і W_d– енергії мілких акцепторних і донорних рівнів, які майже співпадають з енергіями Фермі (W_{F_p} і W_{F_n}) в $p-$ і n– напівпровідниках, відповідно.

При цьому враховано, що $n_n \approx N_d$, $p_p \approx N_a$, $n_i^2 = N_c N_v \exp\left\{-\frac{W_g}{kT}\right\}$.

Із (21.14) видно, що величина потенційного бар'єру *тим більша, чим сильніше леговані* p– і n–області напівпровідників. Якщо $N_d \approx N_c$, а $N_a \approx N_v$, то контактна різниця потенціалів $U_{pn} = W_{pn} / e$ матиме максимальне значення W_g/e. В §19.5 приводяться значення W_g при кімнатній температурі для різних напівпровідників. Із приведеної таблиці видно, що U_{pnmax} може приймати значення від 0,17 В у InSb до 3 В у SiC. Зрозуміло, що такі значення U_{pnmax} є чисто теоретичні, на практиці U_{pn} є меншим.

Збіднений шар. Між *p*– і *n*– областями в *p*–*n* переході існує енергетичний бар'єр, максимальна висота якого складає ~ 0,7 еВ для германієвого, ~ 1 еВ для кремнієвого, ~ 1,4 еВ для GaAs переходу. Більшість електронів в зоні провідності (і дірок в валентній зоні) мають енергію порядку *kT* (тобто всього 0,026 еВ при кімнатній температурі). Зрозуміло, що для більшості носіїв бар'єр на межі *p*– і *n*– областей є непроникний. Кількість електронів і дірок, які здатні подолати бар'єр на межі *p*– і *n*– областей дуже не значна і за допомогою формул (21.12) або (21.13) можемо оцінити їх долю для Ge при T=300 K:

$$\frac{n}{n_n} \sim \exp\left\{-\frac{0{,}70}{0{,}026}\right\} \sim 10^{-12}.$$

Розподіл електричного поля в бар'єрі. Ширина бар'єру

Оскільки область енергетичного бар'єру практично повністю збіднена вільними носіями, можна досить просто розрахувати всі основні характеристика бар'єру. Справа від межі *p*−*n* переходу, в *n*–області, в кожному кубічному метрі напівпровідника міститься N_d позитивно заряджених донорів. Оскільки концентрація електронів і дірок в області бар'єру дуже незначна, то густина об'ємного заряду ρ в бар'єрі буде визначатись концентрацією донорів $\rho = eN_d$. Точно так само в *p*–області (зліва від переходу) густина об'ємного заряду визначається концентрацією негативно заряджених акцепторів: $\rho = -eN_a$ (рис.21.11,*а,б*). У випадку, якщо густина об'ємного заряду ρ постійна, визначення розподілу поля в бар'єрі $E(x)$ і ширина бар'єру x є проста задача, враховуючи розв'язану задачу в §21.2. Залежність $E(x)$ в *p*-*n* переході показана на рис.21.11,*в*. Напруженість електричного поля максимальна в площині *p*-*n* переходу і лінійно спадає при віддаленні від переходу.

Скориставшись формулою (21.4) отримаємо, що глибина проникнення поля в *p*−область буде

$$x_p = \frac{\varepsilon\varepsilon_0 E_m}{eN_a}, \qquad (21.16)$$

а в *n*−область –

$$x_n = \frac{\varepsilon\varepsilon_0 E_m}{eN_d}. \qquad (21.17)$$

Враховуючи це, із трикутника на рис.21.11,*в* легко знайти, що

$$U_{pn} = \frac{1}{2} E_m \cdot x, \qquad (21.18)$$

де x – ширина *p*−*n* переходу (основа трикутника). Тому легко отримати, що

$$x = x_p + x_n = \sqrt{\frac{2\varepsilon\varepsilon_0 U_{pn}}{e}\left(\frac{1}{N_d}+\frac{1}{N_a}\right)}. \qquad (21.19)$$

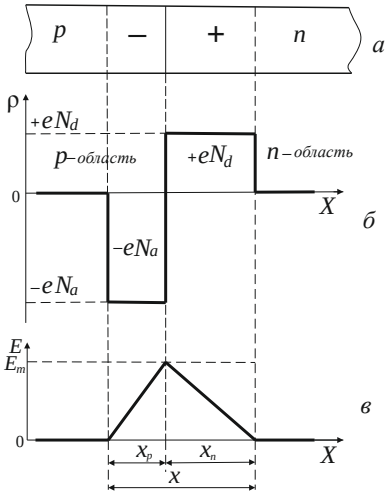


Рис.21.11.Розподіл об'ємного заряду і електричного поля в *p*–*n*− переході: *а* – якісний розподіл заряду; *б* – залежність густини об'ємного заряду $\rho(x)$; *в* – розподіл напруженості електричного поля $E(x)$; N_d − концентрація донорів в *n*− області; N_a − концентрація акцепторів в *p*− області; x_p − і x_n − глибина проникнення електричного поля в *p*−область і *n*−область відповідно; $x = x_p + x_n$ − ширина *p*−*n* переходу.

На практиці використовуються різко несиметричні $p-n$ структури. Так називаються $p-n$ переходи, в яких одна область легована сильніше, ніж інша. На рис.21.12 показаний

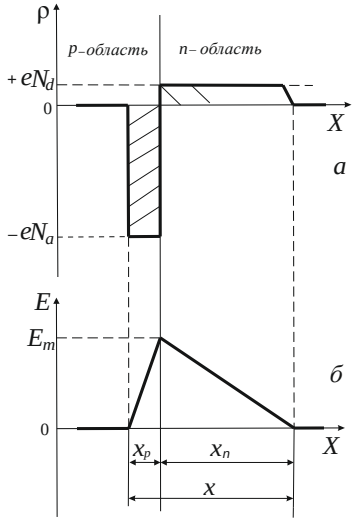


Рис.21.12.Розподіл об'ємного заряду $\rho(x)$ $(а)$ і поля $E(x)$ $(б)$ в різко несиметричному p–n переході $(N_a \gg N_d)$.

розподіл поля в $p-n$ переході, в якому $p-$область легована значно сильніше, ніж $n-$область, тобто $N_a \gg N_d$. Із рисунка видно, що в цьому випадку ширина області об'ємного заряду x_p в $p-$області значно менша, ніж x_n, і розподіл поля $E(x)$ практично має вид прямокутного трикутника. Практично вся напруга на бар'єрі $U_{pn} = \frac{1}{2}E_m x_n$ падає на $n-$область. В цьому випадку розрахунок основних параметрів бар'єру можна здійснити за допомогою формул (21.17) і (21.18):

$$x = x_n \cong \left(\frac{2\varepsilon\varepsilon_0 U_{pn}}{eN_d}\right)^{1/2}, \quad (21.20)$$

$$E_m \cong \left(\frac{2eN_d U_{pn}}{\varepsilon\varepsilon_0}\right)^{1/2}. \quad (21.21)$$

В кремнієвому потужному високовольтному випрямляючому діоді концентрація донорів в $n-$області $N_d = 10^{19}\text{м}^{-3}$. (Концентрація акцепторів в $p-$ області $N_a \approx 10^{24}\text{м}^{-3}$, тобто в 100.000 раз більша і перехід дійсно різко несиметричний). Підставляючи в формули (21.20) і (21.21) значення $U_{pn} = 1{,}1$ В, $\varepsilon = 12{,}5$, $N_d = 10^{19}\text{м}^{-3}$, отримаємо $x \approx x_n \approx$ 12мкм, а $E_m \approx 1{,}8 \cdot 10^5$ В/м.

В кремнієвих діодах, які призначенні для випрямлення низьких напруг, концентрація донорів в $n-$області може бути порядку 10^{23}м^{-3}. В такому діоді ширина області об'ємного заряду складає всього $\sim 0{,}12$ мкм, а максимальне поле в p – n переході $\sim 1{,}8 \cdot 10^7$ В/м.

Зворотнє зміщення. Порушимо рівновагу між дрейфовим і дифузійним струмами в області p–n переходу. Для цього

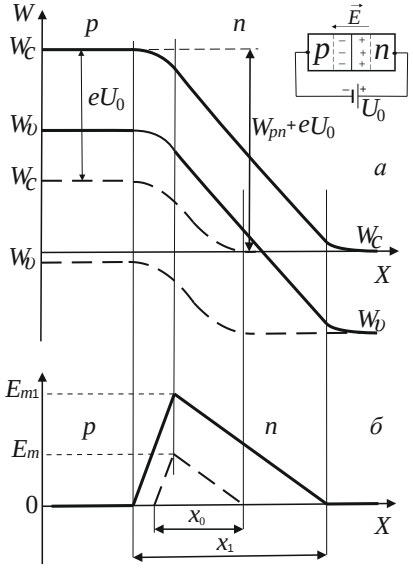


Рис.21.13.Енергетична діаграма $(а)$ і розподіл електричного поля $(б)$ в p–n переході. Штрихові лінії – випадок відсутності зовнішньої напруги. Суцільні лінії – до діода прикладена зворотня напруга U_0.

досить прикласти до діода, який містить p–n перехід, напругу, або як часто говорять, зміщення U_0 від зовнішнього джерела (рис. 21.13). Якщо "плюс" джерела під'єднаний до $n-$області діода, а "мінус" – до $p-$області, то до p–n переходу прикладена зворотня напруга (або

зворотнє зміщення). Саме такий випадок показаний на рис.21.13.

Діод має три області: $p-$область, $n-$ область і область об'ємного заряду (сам p–n перехід), який знаходиться між ними. По відношенню до зовнішнього джерела U_0 всі ці три області включені послідовно. Область об'ємного заряду збіднена вільними носіями заряду, електронами і дірками. Тому вона має найбільший опір. Значить, практично вся прикладена напруга U_0 падає на область об'ємного заряду.

Звертаємо увагу на те , що поле $\vec{E}$ від зовнішнього джерела U_0 при зворотному зміщенні на діоді співпадає за напрямком з полем всередині області об'ємного заряду (рис.21.13). Це означає, що спад напруги від зовнішнього джерела U_0 додається до напруги на переході U_{pn}, яка існувала і до прикладення зовнішнього зміщення. В результаті висота потенціального бар'єру між $p-$ і $n-$ областями стає $U_{pn}+U_0$, а висота енергетичного бар'єру $e\left(U_{pn}+U_0\right)$ (рис.21.13, *а*).

Для різко несиметричного переходу ширину бар'єру і максимальну напруженість електричного поля знайдемо із формул (21.20) і (21.21):

$$x_1 \cong x_{1n} \approx \left(\frac{2\varepsilon\varepsilon_0(U_{pn}+U_0)}{eN_d}\right)^{1/2}, \quad (21.22)$$

$$E_{m1} = \left(\frac{2eN_d(U_{pn}+U_0)}{\varepsilon\varepsilon_0}\right)^{1/2}. \quad (21.23)$$

Із формули (21.22) видно, що при прикладанні до діода зворотної напруги вільні носії заряду витісняються із областей, які прилягають до p–n переходу: область об'ємного заряду стає ширшою (рис.21.13,*б*). Для високовольтного кремнієвого випрямляючого діода при $U_0 = 1000$ В за формулами (21.22) і (21.23) знайдемо $x = x_{n1} = 370$ мкм, $E_{m1} = 5400$ кВ/м.

Зі збільшенням U_0 максимальна напруженість E_m зростає. Проте для кожного напівпровідникового матеріалу існує граничне значення напруженості електричного поля E_i, перевищувати яке не слід. В полі, напруженість якого E_i, для Si і GaAs вона досягає значень $\sim 3\cdot 10^7$ В / м, електрони і дірки набувають таку велику енергію, що здатні при зіткненні з атомами кристалічної решітки напівпровідника створювати нові електрони і дірки. Напруженість E_i отримала назву *напруженості пробою.* Напругу пробою p–n переходу можна визначити за допомогою формули (21.23), а саме:

$$U_i \approx \frac{\varepsilon\varepsilon_0 E_i^2}{2eN_d}\cdot \quad (21.24)$$

Якщо $N_d \approx 10^{23}$м$^{-3}$, $E_i = 3\cdot 10^7$ В/м , $\varepsilon = 12$, то $U_i = 3$В, а якщо $N_d \approx 10^{20}$м$^{-3}$, то $U_i = 3000$ В.

Зворотний струм. Коли до p–n переходу зовнішня напруга не прикладена, то струм неосновних носіїв заряду врівноважується зустрічним потоком "гарячих" основних носіїв заряду. Але коли до p–n переходу прикладене зворотнє зміщення і висота бар'єру W_{pn} збільшиться, то "гарячі" носії вже не в стані подолати бар'єр. На потік неосновних носіїв прикладена напруга ніяк не впливає. Тому зворотний струм визначається потоком неосновних носіїв заряду (21.10) і дуже різко залежить від температури. Вище сказане належить до германієвого діоду. Для кремнієвого і арсенід–галієвого переходів зворотний струм не насичується, а монотонно зростає пропорційно $U_0^{1/2}$. В цих діодах домінуючу роль при зворотньому зміщенні відіграє струм, який генерується в області об'ємного заряду. Деколи його називають

просто *генераційним струмом*. Носії, які створюють генераційний струм виникають в самому *p–n* переході, в області об'ємного заряду, де існує сильне електричне поле. Будь-який носій, електрон або дірка, які виникають в цій області, будуть негайно підхвачені електричним полем і викинуті: електрони в $n-$ область, дірки – в $p-$область. Генераційний струм пропорційний ширині бар'єру (області об'ємного заряду) x. Так як згідно формули (21.22) $x \sim U_0^{1/2}$, то зрозуміло, чому генераційний струм пропорційний $U_0^{1/2}$.

Відношення густини генераційного струму j_g до густини струму насичення j_s (формула(21.10)) дорівнює

$$\frac{j_g}{j_s} = \frac{N \cdot x}{n_i L}, \qquad (21.25)$$

де $N_d = N_a = N, D_n = D_p = D, L_n = L_p = L$ і час життя $\tau_p = \tau_n = \tau$. Вираз (21.25) дає можливість порівняти відносні величини j_g і j_s для будь-якого *p–n* переходу при будь-якій температурі і будь-якій прикладеній напрузі. Із (21.25) видно, що з ростом ширини забороненої зони напівпровідника W_g роль j_g різко, експоненціально зростає, оскільки згідно (19.44), концентрація n_i експоненційно зменшується. Це пояснює, чому вклад генераційного струму в зворотний струм в кремнієвих, а особливо в арсенід–галієвих переходах значно більший, ніж в германієвих.

Пряме зміщення. На рис.21.14 показаний *p–n* перехід, до якого прикладене пряме зміщення. "Плюс" джерела під'єднаний тепер до $p-$ області діода, "мінус" – до $n-$області. Поле від зовнішнього джерела U_0 направлене назустріч полю, яке існує всередині $p-n-$області. Значить, спад напруги від зовнішнього джерела U_0 вираховується із бар'єрної різниці потенціалів U_{pn}, яка існувала до прикладання зовнішнього зміщення. В результаті висота потенціального бар'єру між $p-$ і $n-$ областям стане дорівнювати $U_{pn} - U_0$, а висота енергетичного бар'єру відповідно $e(U_{pn} - U_0) = W_{pn} - eU_0$. При відсутності

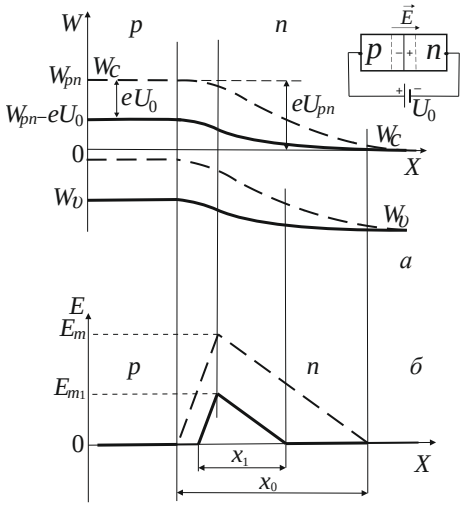


Рис.21.14.Енергетична діаграма (*а*) і розподіл напруженості електричного поля (*б*) в $p-n$ переході. Штрихові лінії – випадок відсутності зовнішньої напруги. Суцільні лінії – до діода прикладене пряме зміщення.

зміщення електрони із $n-$ області в $p-$ область і дірки із $p-$ області в $n-$ область "прориваються" тільки ті, енергія яких більша висоти бар'єру eU_{pn} (формули (21.12) і (21.13)). Струм, який вони створюють, компенсується дрейфовим струмом неосновних носіїв (формула (21.10)). Якщо висота бар'єру зменшується на величину eU_0, то тепер вже із $n-$ в $p-$область здатні "прориватись" всі електрони, енергія яких більша висоти бар'єру $W_{pn} - eU_0$. Концентрація таких електронів згідно (21.12) дорівнює:

$$n_1 \cong n_n \exp\left\{-\frac{W_{pn} - eU_0}{kT}\right\}.$$

Таким чином, число електронів, здатних подолати бар'єр, що понизився, зростає

в $\frac{n_1}{n_n} = \exp\left\{\frac{eU_0}{kT}\right\}$ раз.

Наприклад: при $U_0 = 0{,}1$ В, $T = 300\ K$, $\frac{n_1}{n_n} \cong 6 \cdot 10^{16}$ раз. Струм, який утворюють ці електрони, в стільки ж раз зростає і складає тепер величину

$$j_{n1} = j_{ns}\exp\left\{\frac{eU_0}{kT}\right\}, \qquad (21.26)$$

де враховано, що при відсутності зміщення дрейфовий струм неосновних електронів j_{ns} (формула (21.8)) утворюється струмом "гарячих" електронів (формула (21.12)). Таким чином , електронна складова струму через *p–n* перехід при його прямому включенні буде:

$$\begin{aligned} j_n(U_0) &= j_{n1} - j_{ns} = \\ &= j_{ns}\left(\exp\left\{\frac{eU_0}{kT}\right\} - 1\right). \end{aligned} \qquad (21.27)$$

Аналогічно, діркова складова прямого струму

$$j_p(U_0) = j_{ps}\left(\exp\left\{\frac{eU_0}{kT}\right\} - 1\right). \qquad (21.28)$$

Повний струм через *p–n* перехід дорівнює сумі електронної і діркової складових і дорівнює

$$\begin{aligned} j(U_0) &= j_n(U_0) + j_p(U_0) = \\ &= j_s\left(\exp\left\{\frac{eU_0}{kT}\right\} - 1\right), \end{aligned} \qquad (21.29)$$

де $j_s = j_{ns} + j_{ps}$.

Формула (21.19) описує залежність струму через *p–n* перехід від прикладеної напруги при будь-яких зміщеннях і прямих, і зворотніх. Тільки при прямій прикладеній напрузі величину U_0 треба вважати додатною, а при зворотній – від'ємною.

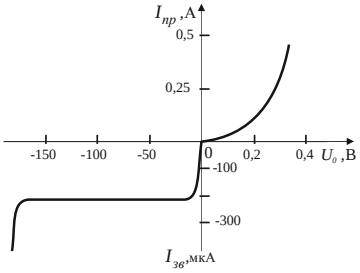


Рис.21.15. Вольт–амперна характеристика германієвого *p – n* переходу.

Оцінимо відношення струму при прямому зміщенні до струму при зворотному зміщенні за допомогою формули (21.29) при кімнатній температурі $T = 300\ K$ ($kT = 0{,}026$ еВ при $U0$=1В:

$$\frac{j_{пр.}}{|j_{зв.}|} = \frac{\exp\left\{\frac{eU_0}{kT}\right\} - 1}{\left|\exp\left\{-\frac{eU_0}{kT}\right\} - 1\right|} =$$

$$= \frac{\exp\{38\} - 1}{|\exp\{-38\} - 1|} \approx \exp\{38\} \simeq 10^{16}.$$

Таким чином, струм при прямому зміщенні значно перевищує струм при зворотньому зміщенні. Тобто *p–n* перехід має односторонню провідність. Цю властивість *p–n* переходу в основному і використовують на практиці. На рис.21.15 приведена вольт–амперна характеристика германієвого діода з *p–n* переходом. Зверніть увагу на різницю в масштабах по осях струму і напруги, які відповідають прямому і зворотньому зміщенню. При $U_0 \simeq -200$В, як видно із рис. 21.15, струм змінюється через *p–n* перехід, а напруга на діоді стала. В такому режимі пробою працює стабілітрон, який стабілізує напругу на *p–n* переході.

Примітка.

1. При отриманні формули (21.29) вважалось, що вся прикладена до діода напруга U_0 падає на *p–n* переході, тобто, що опір *p–n* переходу набагато більший, ніж всі інші послідовно ввімкнені з ним опори: опір *n*–області, *p*–області, опір контактів до діода і т.п. При зворотній напрузі таке припущення справедливе. Опір бар'єру (*p–n* переходу) є великий вже при нульовому зміщенні. При прямому зміщенні справа дещо інша. Із збільшенням прямої напруги опір бар'єру різко зменшується! Значить не вся прикладена напруга U_0 падає на *p–n* переході. Значна частина напруги падає на послідовно ввімкнених з переходом контактах, а також *n*–

і $p-$областях. Тому висота бар'єру p–n-переходу знизиться не на eU_o, а на дещо меншу величину. В результаті ситуація, коли

$$eU_0 > W_{pn}\left(n_1 \cong n_n \exp\left\{-\frac{W_{pn}-eU_0}{kT}\right\} > n_n\right)$$

стає неможливою. Найбільше, чого можна досягти, збільшуючи пряму напругу на діоді, – це дуже сильно зменшити висоту бар'єру (як деколи говорять – "спрямити" бар'єр). Але досягнутий ефект вимагає, щоб густина струму через діод складала десятки і навіть сотні тисяч $\text{А}/\text{см}^2$.

2.Увипадку різко несиметричного переходу (сильно легована $p-$область) при любих зміщеннях основний вклад в загальний струм в області бар'єру вносять дірки. Ця область діода називається емітером (від латинського emitto – випускаю). Вона здатна випускати (інжектувати) в слаболеговану $n-$область таке число дірок (неосновних носіїв для $n-$області), що їх концентрація може значно перевищувати вихідну концентрацію основних носіїв (електронів в $n-$області).

Слаболегована область діода називається базою (від грецького basis – основа). Саме в слаболегованій області переходу, базі, в основному розміщується область об'ємного заряду при зворотньому зміщенні (рис.21.13). Параметри бази – її ширина і рівень легування – визначають напругу пробою p–n структури (формули (21.22), (21.23), (21.24)). У відповідності з прийнятою термінологією на рис.21.15 аналізується робота діода з p – емітером і n – базою.

§ 21.4. Ємність електронно–діркового переходу

Якщо до p–n переходу прикладена зворотня напруга U_0, то по відношенню до змінної напруги зворотньозміщений p–n перехід поводить себе як конденсатор. Ємнісний опір p–n переходу визначається за формулою (9.77) $X_C = 1/(\omega C)$, де ω – циклічна частота напруги, C – ємність p–n переходу. Ємність p–n переходу можна визначити за формулою ємності плоского конденсатора

$$C = \varepsilon\varepsilon_0 S/x, \qquad (21.30)$$

де x – ширина області об'ємного заряду, S – площа переходу. Так як x визначається за формулою (21.22), то можна визначати ємність переходу при любій прикладеній напрузі і різних рівнях легування. Із розділу 8(§8.2) відомо, що конденсатор проводить змінний електричний струм тому, що лінії струму провідності в середині конденсатора переходять в лінії струму зміщення (змінне електричне поле) і таким чином, електричне коло стає замкненим. Пояснення проходження змінного струму через конденсатор може бути і таке. Коли до конденсатора прикладена змінна напруга, то він в певну частину періоду накопичує електричний заряд, а потім, в другу частину періоду, віддає його в зовнішнє коло. Періодичне заряджання і розряджання конденсатора підтримує в колі змінний струм.

Тепер розглянемо зворотнозміщений p–n перехід, до якого одночасно з постійною зворотною напругою U_0 прикладена ще і невелика змінна напруга, амплітуда якої ΔU_0 (рис.21.16). Суцільна крива на рис.21.16 відповідає розподілу об'ємного заряду ρ в p–n переході при зворотній напрузі U_0. Штриховою лінією показаний розподіл об'ємного заряду ρ, коли до діода прикладена напруга $U_0 + \Delta U_0$.

Напруга на діоді стала більша – область об'ємного заряду розширилась (формула (21.22)). Вільні носії – електрони в

$n-$області і дірки в $p-$області відштовхуються в зовнішнє коло із області об'ємного заряду, яка звільняється від носіїв. В $n-$області при збільшенні напруги від U_0 до $U_0+\Delta U_0$ в зовнішнє коло виштовхуються всі електрони, які знаходились при напрузі U_0 на ділянці між суцільною і штриховою лініями. В $p-$ області точно так же із ділянки між суцільною і штриховою кривими витісняються дірки.

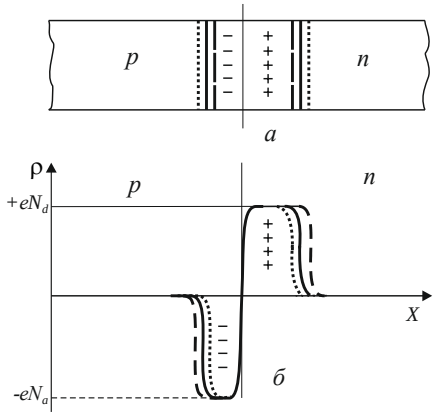


Рис.21.16.Розподіл об'ємного заряду в *p-n* переході: суцільна лінія – до діода прикладена зворотня напруга U_0; штрихова – зворотня напруга дорівнює $U_0+\Delta U_0$; точкова – $U_0-\Delta U_0$.

Точковою кривою показаний розподіл об'ємного заряду, коли до діода прикладена напруга $U_0-\Delta U_0$. Напруга на діоді стала менша – ширина області об'ємного заряду зменшилась. А це означає, що із зовнішнього кола забираються вільні носії, які заповнюють тепер ділянки діода, де їх не було при напрузі U_0. Якщо напруга зменшилась від U_0 до $U_0-\Delta U_0$, то вільні носії, які відбираються із зовнішнього кола, заповнюють області між суцільною і точковими кривими.

Таким чином, зрозуміло, що *p–n* перехід володіє здатністю в певну частину періоду накопичувати вільні носії заряду, а в іншу – віддавати їх в зовнішнє коло, тобто має ємність. Цей тип ємності носить назву бар'єрної ємності.

Бар'єрну ємність визначимо як $C=\Delta q\,/\,\Delta U_0$, де Δq величина заряду, яка накопичується або віддається *p–n* переходом при зміні напруги на величину ΔU_0. Для різко несиметричного переходу вся прикладена напруга падає на слаболеговану область і повна ширина об'ємного заряду практично визначається за формулою (21.22). При збільшенні напруги на величину ΔU_0 ширина області об'ємного заряду збільшується на величину Δx_n. Значення Δx_n можна визначити із формули (21.22). При цьому у зовнішнє коло буде витіснений заряд електронів на одиницю площі поперечного перерізу $\Delta q=eN_d\cdot\Delta x_n$ (точно такий же заряд дірок буде витіснений із $p-$області). Відношення $\Delta q/\Delta U_0$ визначить величину ємності *p–n* переходу на одиницю площі C_0. Тоді

$$C_0=\frac{dq_0}{dU_0}=eN_d\frac{dx_n}{dU_0}=eN_d\cdot\frac{1}{2}\left(\frac{2\varepsilon\varepsilon_0}{eN_d}\right)^{1/2}\cdot\frac{1}{(U_{pn}+U_0)^{1/2}}=\left(\frac{e\varepsilon\varepsilon_0N_d}{2(U_{pn}+U_0)}\right)^{1/2}=\frac{\varepsilon\varepsilon_0}{x},\qquad(21.31)$$

де U_0 – напруга на *p–n* переході, x– ширина області об'ємного заряду, яка визначається формулою (21.22). Із формули (21.31) випливають два важливих висновки. *По-перше*, бар'єрна ємність *p–n* переходу залежить від напруги. Чим більше прикладене до діоду зворотнє зміщення U_0, тим менша ємність. Конденсатори, ємність яких залежить від напруги, називаються нелінійними. Тому часто говорять, що *p–n* перехід є нелінійна ємність. *По-друге*, при однаковій зворотній напрузі на переході бар'єрна ємність тим більша, чим вищий рівень легування: $C\sim N_d^{1/2}$. Оцінимо величину бар'єрної ємності для високовольтного випрямляючого діода з кремнію, для якого $S\sim10-30$ см^2. При $S=10$ см^2, $N_d=10^{19}$м$^{-3}$ і відсутності

зовнішньої напруги на діоді ємність C буде

$$C = S\left(\frac{e\varepsilon\varepsilon_0 N_d}{2U_{pn}}\right)^{1/2} \approx 1\cdot 10^{-8}\Phi = 10^4 \text{пФ}.$$

Навіть при $U_0 = 1000$В ємність діода буде суттєва величина: C≈300 пФ.

Порівняємо ємнісний опір діода з його активним опором, який визначається зворотним струмом. Ці опори можна вважати з'єднаними паралельно. При кімнатній температурі і напрузі на діоді $U_0 = 100$В типове значення зворотнього струму $I_{зв}$ для кремнієвого діода такого типу складає ~0,5 мА. Таким чином, активний опір $R = U_0/I \simeq 2\cdot$ $\cdot 10^5$ Ом. Ємність при $U_0 = 100$В за формулою (21.31)дорівнює:

$$C = SC_0 = S\left(\frac{e\varepsilon\varepsilon_0 N_d}{2U_0}\right)^{1/2} \sim 10^{-9}\Phi.$$

При частоті $\nu \geq 800$ Гц, $X_C = 1/(C\cdot 2\pi\nu) = 2\cdot$ $\cdot 10^5$ Ом $>$ R. Тобто, при частоті напруги $\nu > 800$ Гц опір діода буде визначатись не тільки зворотним струмом, а і ємністю структури. При промисловій частоті $\nu = 50$ Гц $X_C \sim 3\cdot 10^6$Ом і опір R визначає в основному опір діода. Існують, однак, діоди, які призначені для роботи в колах високих частот. Для таких діодів наявність бар'єрної ємності відіграє дуже важливу роль. Деколи ємнісний струм, обумовлений бар'єрною ємністю заважає діоду виконувати ту роботу, для якої він призначений. В цьому випадку величину C стараються зробити по можливості меншою, щоб збільшити опір. Цього досягають за допомогою зменшення площі діодів, які призначені для роботи в колах надвисоких частот (НВЧ). Площа S для таких діодів складає $10^{-4}-10^{-6}$см^2 $(10^{-2}-10^{-4}$мм$^2)$. Величина бар'єрної ємності таких p–n структур дорівнює десятим, а деколи і сотим долям пікофаради. Але існують і такі діоди, дія яких прямо грунтується на властивостях бар'єрної ємності переходу. Вони використовуються як конденсатори, ємністю яких можна управляти прикладеною напругою, в колах високої і надвисокої частоти.

Варікапи. Назва приладу (від англійського variable capacitor–змінний конденсатор) дуже точно характеризує його призначення. Варікапи використовуються як конденсатори змінної ємності, величина якої змінюється в залежності від прикладеної напруги. Чим більша напруга, тим менша ємність. Кожний, хто налаштовує радіоприймач на потрібну станцію, з точки зору радіотехніки зайнятий тим, що, змінюючи ємність змінного конденсатора, налаштовує вхідний контур радіоприймача на потрібну частоту. Одне із перших застосувань варікапів була заміна механічного змінного конденсатора в радіоприймачі.

Напівпровідникові ємності, які управляються напругою, лежать в основі надчутливих перетворювачів постійної напруги в змінну. З їх допомогою здійснюється прийом і підсилення надслабих сигналів, в тому числі і сигналів із космосу. Вони використовуються для генерації сигналів з частотою в сотні мільярдів герц. В параметричних підсилювачах в LC – контурі, який налаштований на частоту сигналу, який підсилюється, як ємність C включається варікап. Параметр варікапа, ємність, два рази за період коливань сигналу, який підсилюється, в потрібній фазі змінюється зовнішнім джерелом коливань. В результаті досягається підсилення, яке називають параметричним. Основна перевага параметричних підсилювачів – їх рекордно висока чутливість. Головним параметром варікапа є коефіцієнт зміни ємності K_C, який визначається як відношення максимальної ємності варікапа C_{max} до мінімальної C_{min}:

$K_C = C_{max} / C_{min}$. Так як ємність варікапу залежить від прикладеної до нього напруги, то зрозуміло, що величина K_C залежить від діапазону, в якому можна змінювати зворотнє зміщення на варікапі. Мінімальна зворотня напруга U_0 приймається за нульову. При цьому напруга на *p–n* переході мінімальна і дорівнює U_{pn} і згідно формули (21.21)

$$C_{max} = S\left(\frac{e\varepsilon\varepsilon_0 N_d}{2U_{pn}}\right)^{1/2}. \qquad (21.32)$$

Максимальна напруга на *p–n* переході обмежується процесом ударної іонізації і визначається з формули (21.24). Враховуючи, що $U_i \gg U_{pn}$, із (21.31) отримаємо значення мінімальної ємності варікапу

$$C_{min} = S\left(\frac{e\varepsilon\varepsilon_0 N_d}{2U_i}\right)^{1/2} = S\frac{eN_d}{E_i}. \quad (21.33)$$

Тоді коефіцієнт зміни ємності K_C буде визначатись як:

$$K_C = \frac{C_{max}}{C_{min}} = \left(\frac{\varepsilon\varepsilon_0 E_i^2}{2eU_{pn}N_d}\right)^{1/2}. \quad (21.34)$$

Із (21.34) видно, що чим менша концентрація домішки в слабо легованій області, тим більша величина K_C. На практиці значення N_d вибирають порядку $10^{21} - 10^{22}$м$^{-3}$. Величина K_C в реальних варікапах лежить в межах від ~2 до 15.

§ 21.5. Фотодіоди і світлодіоди

В основі роботи будь-якого фотоелектричного пристрою лежить елемент, електричні параметри якого: опір, струм або напруга змінюються під дією освітленості. Одним із найбільш розповсюджених елементів такого типу є фотодіод – *p–n* перехід зміщений у зворотньому напрямку. Якщо на напівпровідник падає світло, енергія фотона якого $W_Ф$ більша, ніж ширина забороненої зони W_g, то кожний поглинутий фотон породжує в напівпровіднику електронно–діркову пару (рис.19.32,*а*). Якщо електрон і дірка виникли під дією світла в області об'ємного заряду зворотньо зміщеного *p–n* переходу, то любий носій, електрон чи дірка, які народжені в області об'ємного заряду будуть негайно підхоплені електричним полем і викинуті: електрон – в *n*− область, дірка – в *p*− область.

В пітьмі, без освітлення, генераційний струм *p–n* переходу (так званий темновий струм) визначається числом електронів і дірок, які є в області об'ємного заряду внаслідок теплової генерації. При освітленні зворотньозміщеного переходу світлом з енергією квантів $W_Ф > W_g$ струм через перехід збільшується. Фотострум, який виникає під дією світла буде в стільки раз більшим темнового струму, в скільки раз число носіїв, які створюються в області об'ємного заряду світлом, буде більшим, ніж число електронів і дірок, які виникають внаслідок теплової генерації. Фотострум, підсилений і перетворений відповідною електронною схемою, приводить в дію захисні, охоронні, сигнальні, блокуючі і тому подібні пристрої.

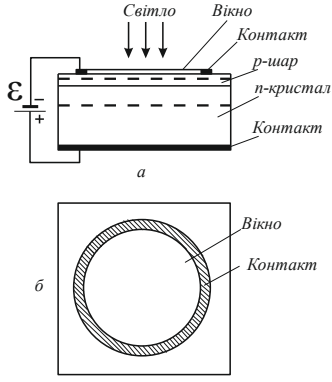


Рис.21.16. Конструкція типового фотодіода.

На рис.21.16,*а* показана типова конструкція фотодіода. Із рисунка видно, що *p–n* перехід зміщений у зворотньому напрямку. Верхній контакт до діоду виконаний у формі

кільця (рис.21.16,*б*). Область p–n переходу всередині кільця, так зване вікно, доступне для світла. В фотодіоді площина p–n переходу розташовується близько до поверхні кристалу: це необхідно для того, щоб світло, яке падає на кристал, не встигло поглинутись на шляху до області об'ємного заряду. Як правило, площина p–n переходу розташовується в фотодіодах на відстані в декілька десятих долей мікрометрів від поверхні. Такий p–n перехід можна виготовити за допомогою дифузії або іонної імплантації.

Величина зворотної напруги U_0, яка прикладається до p–n переходу, складає зазвичай 10–30В. Із збільшенням зворотної напруги, як відомо із формули (21.22) область об'ємного заряду розширюється. Величина U_0 і концентрація домішки в слаболегованій $n-$ області кристалу N_d вибирається таким чином, щоб практично весь світловий потік, який падає на поверхню світлодіода, поглинався в області об'ємного заряду і давав вклад в фотострум. Як правило, необхідна для цього товщина збідненої області x_n складає декілька мікрометрів.

Найбільше розповсюдження тримали фотодіоди на основі Ge і Si. Однак розроблені і використовуються також фотодіоди із GaAs, InAs, InSb, GaP, трикомпонентних напівпровідникових сполук GaInAs, GaAlAs, чотирикомпонентної сполуки GaInAsP.

Одна з найважливіших причин для розробки фотодіодів на основі різних напівпровідникових сполук в тому, що фотодіод, який виготовлений із певного матеріалу, здатний ефективно реагувати на світло з довжиною хвилі, що лежить тільки в певному діапазоні. Цю властивість фотодіодів називають селективність – від латинського selection (вибір,відбір), що показано на рис. 21.17. На рис.21.17 для фотодіода, який виготовлений з напівпровідника з шириною забороненої зони W_g, показана якісна залежність фотоструму від енергії падаючих фотонів $W_Ф = h\nu$. При $W_Ф < W_g$ енергія фотонів недостатня для створення в матеріалі електронно-діркових пар і фотострум тому не виникає. При $W_Ф > W_g$ фотострум спочатку швидко збільшується із зростанням $W_Ф$, а потім починає зменшуватись. Справа в тому, що зі збільшенням $W_Ф$ різко зростає коефіцієнт поглинання світла в матеріалі, тобто падаючий світловий потік із зростанням енергії фотонів поглинається все ближче і ближче до поверхні фотоприймального вікна (рис.21.16,*б*). Як відомо із §21.2 на поверхні завжди існує велика кількість дефектів і домішок, які створюють ефективні центри рекомбінації. Тому електрони і дірки, які виникають під дією світла біля поверхні, швидко рекомбінують і не дають вклад у фотострум.

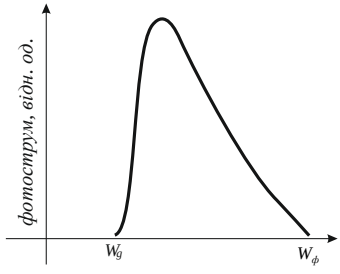


Рис.21.17. Якісна залежність фотоструму від енергії $W_Ф$ фотонів світла , яке падає на фотодіод.

Таким чином, область максимальної фоточутливості діода визначається величиною W_g і відповідною умовою $W_Ф \geq W_g$. Напівпровідники з невеликою шириною забороненої зони, так звані *вузькозонні напівпровідники*, використовуються для створення фотодіодів, які чутливі до світла з малою енергією квантів, тобто з великою довжиною хвилі. Напівпровідники з великим значенням W_g *(широкозонні напівпровідники)*

використовуються для створення приймачів ультрафіолетового світла.

Важливою властивістю фотодіодів є швидкодія. Швидкодію фотоелектричних приладів оцінюють по тому, як швидко спадає фотострум через прилад після виключення світла. Якщо різко перекрити світловий потік, який падає на приймальне вікно фотодіода, то фотострум зникне, як тільки електрони і дірки, які збуджені світлом, будуть винесені полем із області об'ємного заряду. Як відомо, типове значення напруженості електричного поля в зворотньозміщеному *p–n* переході велике. Воно складає $\sim 10^6 - 10^7 \text{В}/\text{м}$. В таких сильних полях дрейфова швидкість електронів і дірок $v_{др}$ не залежить від напруженості поля (§19.12) і дорівнює $\sim 10^5\,\text{м/с}$ (рис.19.40). Шлях, який повинні пройти електрон і дірка, перш ніж вони будуть винесені полем із *p–n* переходу дорівнює, очевидно, ширині області об'ємного заряду x. Для типових фотодіодів значення x складає декілька мікрометрів $((1 \div 10) \cdot \cdot 10^{-6}\text{м})$. Таким чином, час, який визначає швидкодію фотодіода τ, складає приблизно

$$\tau \sim \frac{x}{v_{др}} \sim 10^{-10} \div 10^{-11}\text{с}.$$

Фотодіод здатний реагувати на зміни інтенсивності світла за стомільярдні долі секунди. Висока швидкодія дозволяє використовувати фотодіоди в швидкісних системах передачі і обробки інформації волоконно–оптичного зв'язку. Приймачами швидко модульованих світлових сигналів в таких системах служать саме фотодіоди. Фотоелектричні пристрої дозволяють вводити в ЕВМ тисячі чисел і команд в секунду.

Світлодіод (світловипромінюючий діод), напівпровідниковий прилад, який перетворює електричну енергію в енергію оптичного випромінювання на основі явища інжекційної електролюмінесценції, що відбувається в напівпровідниковому кристалі з *p–n* переходом або гетеропереходом або контактом метал–напівпровідник.

Будемо розглядати напівпровідниковий *p–n* перехід при прямому зміщенні. Можна сказати, що любий прямо зміщений *p–n* перехід є світлодіодом. Основні носії електрони і дірки, які попадають із емітера (інжектуються) в базу, рекомбінують. Частина цих носіїв рекомбінують з випромінюванням фотонів. Якась частина фотонів, не зазнавши поглинання в самому діоді, вилітають назовні. Для виготовлення світлодіодів придатні матеріали з достатньо великою ймовірністю випромінювальної рекомбінації. Германій і кремній цим вимогами не відповідають. В них переважна частина електронів і дірок рекомбінує без випромінювання фотона. А в арсеніді галію (GaAs) і трикомпонентних напівпровідникових сполуках на його основі: GaAsP і GaAlAs ймовірність випромінювальної рекомбінації дуже висока (близька до одиниці). Тому більша частина світлодіодів випускається на основі цих матеріалів. Колір, тобто довжина хвилі світла, яке випромінюється світлодіодом, визначається енергією фотона $W_Ф$, який випромінюється при рекомбінації. В більшості випадків ця енергія близька ширині забороненої зони напівпровідника W_g. Так, довжина хвилі випромінювання арсенід–галієвого $(W_g \simeq 1{,}4\ \text{еВ})$ світлодіода дорівнює $\lambda = hc/W_g = 1{,}24/1{,}4\ \text{мкм} = 0{,}89\ \text{мкм}$. Таке випромінювання людське око не бачить. Воно відповідає інфрачервоній області спектру. Приймачами такого інфрачервоного випромінювання є кремнієві фотодіоди, для яких енергія фотона $W_Ф \approx 1{,}4\ \text{еВ}$ приблизно відповідає максимуму їх

спектральної чутливості (рис.21.17). Добавка в гратку GaAs атомів фосфору (P) або алюмінію (Al) призводить до збільшення ширини забороненої зони W_g в сполуці $GaAs_{1-x}P_x$. При збільшенні в напівпровідниковій сполуці долі відносного вмісту фосфору x, ширина забороненої зони напівпровідника збільшується від 1,4еВ (GaAs, x=0) до 2,26еВ (GaP, x=1). Виготовляючи твердий розчин GaAsP з різним вмістом P, отримують світло- діоди, які випромінюють червоне світло ($W_Ф = 1{,}6 \div 1{,}9$ еВ, $\lambda_ч = 0{,}653 \div 0{,}775$ мкм). Отримати таким способом жовті або зелені світлодіоди не вдається, хоч енергія квантів, що відповідає цим кольорам, менша, ніж ширина заборненої зони чистого GaP ($W_g = 2{,}26$ еВ, $\lambda = 0{,}549$ мкм – відповідає зеленій області спектру). Справа в тому, що при $x \geq 0{,}45$ ймовірність випромінювальної рекомбінації в GaAsP різко падає і світлодіоди стають малоефективними. Для збільшення ефективності свічення в твердий розчин GaAsP або в чистий GaP добавляють спеціальні домішки (азот N, цинк Zn, кисень O), які збільшують ймовірність випромінювальної рекомбінації. Сьогодні для створення "різнокольорових" світлодіодів використовується широкий вибір напівпровідникових матеріалів. З використанням GaAlAs отримують інфрачервоні і червоні джерела світла. На основі сполук GaAsP і GaP виготовляють світлодіоди із кольором свічення від червоного до зеленого. Чотирикомпонентна напівпровідникова сполука GaInAsP дозволяє створювати інфрачервоні світлодіоди з довжиною хвилі $\lambda \sim 1 \div 1{,}3$ мкм, які є найкращими джерелами для волоконно–оптичних ліній зв'язку. А на основі карбіду кремнію (SiC) отримують світлодіоди всіх основних кольорів: червоного, зеленого і синього.

На сьогодні головна сфера використання світлодіодів – це різні індикатори. В наш вік інформаційного вибуху поява зручних, універсальних, дешевих і ефективних приладів для індикації інформації, якими є світлодіоди, дозволяє вирішувати багато важливих проблем. Світлодіоди випромінюють некогерентне випромінювання з вузьким спектром. Довжина хвилі випромінювання залежить від напівпровідникового матеріалу і його легування. Яскравість випромінювання більшості світлодіодів знаходиться на рівні $10^{-3} - 10^{-5}$ кд / м^2 (кандела на квадратний метр – одиниця СІ яскравості плоскої поверхні, площа якої 1 м^2, що світиться в перпендикулярному до неї напрямку, при силі світла 1 кд. Кандела (від латинського candela – свічка) – одиниця сили світла в СІ). Коефіцієнт корисної дії (ККД) світлодіода видимого випромінювання складає від 0,01% до декількох відсотків. У світлодіодів інфрачервоного випромінювання з метою знижень витрат на повне внутрішнє відбивання і поглинання в тілі кристалу для останнього вибирають напівсферичну форму, а для покращення характеристик направленості випромінювання світлодіод розміщують в параболічний або конусний відбивач. ККД світлодіода з напівсферичною формою кристалу досягає 40%.

Сонячні батареї. За кожну секунду Земля отримує від Сонця у вигляді випромінювання $\sim 3 \cdot 10^{17}$ Дж енергії, що в 30000 раз більше енергії, яка виробляється за цей час всіма електростанціями світу. Способи перетворення сонячної енергії в електричну досліджуються дуже давно. Перша пропозиція по використанню для цієї мети напівпровідників була в 1884р., коли

спробували виготовити сонячний елемент на основі сполуки селена з золотом. Вже більше 100 років продовжується вперта робота по створенню і вдосконаленню напівпровідникових перетворювачів сонячної енергії.

Сонячний елемент – це все той же *p–n* перехід і робота його дуже подібна на роботу вже відомого фотодіода. Однак, на відміну від фотодіода зовнішнє зміщення до сонячного елемента не прикладається.

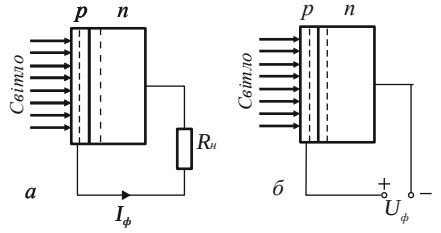


Рис.21.18. Виникнення фотоструму (*а*) і фото – е.р.с. (*б*) в *p–n* переході під дією світла.

Повернемось до рис.21.16 і уявимо собі, що ніякого зовнішнього зміщення до діода не прикладено, ніякого джерела у зовнішньому колі не має. А нижній (суцільний) і верхній (кільцевий) контакти просто з'єднані один з одним через опір навантаження $R_н$ (рис.21.18,*а*). Як відомо, і при відсутності зміщення між *p*– і *n*–областями існує потенціальний бар'єр. Біля *p–n* переходу є область об'ємного заряду (рис.21.12,*а*), в якому існує електричне поле (рис.21.12,*б*). Це означає, що і без зовнішнього зміщення електронно – діркові пари, які виникатимуть під дією світла в області об'ємного заряду, будуть розділятися полем: дірки будуть викидатись в *p*–область, електрони в *n*–область. Такий рух дірок і електронів викликає у зовнішньому колі струм, який буде направлений від *p*–області до *n*–області – фотострум. Фотострум $I_ф$, який протікає через опір навантаження $R_н$ (цим опором може бути нагрівальний елемент, двигун, акумулятор, який заряджається і т.п), здійснює корисну роботу. Енергія світла перетворюється в електричну енергію.

Якщо розімкнути провід, який з'єднує *p*– і *n*–області діода, то струм протікати не буде. При неперервному падінні світла на вікно сонячного елемента в матеріалі неперервно створюються електронно– діркові пари. Вони розділяються полем *p–n* переходу. Дірки викидаються в *p*– область і створюють там надлишковий позитивний заряд. Електрони переміщуються в *n*–область і створюють негативний заряд. Між *p*– і *n*–областями елемента виникає різниця потенціалів – фото-е.р.с. $U_ф$ (рис.21.18,*б*). Може здатись, що якщо *p–n* перехід освітлюється достатньо довго, то величина фото – е.р.с може стати досить великою. Дійсно, чим довше світло падає на перехід, тим більше дірок накопичується в *p–*, а електронів – в *n*–області, і, як наслідок, тим більша повинна бути різниця потенціалів між *p*– і *n*– парами переходу. Насправді так не відбувається. Полярність фото – е.р.с – "плюс" зі сторони *p*– і "мінус" зі сторони *n*–області – відповідають прямому зміщенню на *p–n* переході (рис.21.14). Пряме зміщення на переході знижує висоту бар'єру і полегшує відтік дірок із *p*–області в *n*, а електронів із *n*–області в *p* (рис.21.14,*а,б*). В результаті дії двох протилежних механізмів – накопичення носіїв під дією світла і їхній відтік через зниження висоти бар'єру – при кожному значенні інтенсивності світла, яке падає на елемент, в стаціонарному стані встановлюється певна величина фото – е.р.с. $U_ф$. Із збільшенням інтенсивності світла $U_ф$ росте дуже слабо, пропорційно логарифму інтенсивності. При дуже великій інтенсивності світла, коли під дією фото – е.р.с. бар'єр стає практично

“спрямленим”, величина U_ϕ дорівнює висоті бар'єру неосвітленого переходу:

$$U_{\phi max} \simeq U_{pn} \simeq W_g / e.$$

ККД сучасних сонячних елементів виріс до 10-20%. Головна складність на шляху до збільшення ефективності сонячних елементів в тому, що сонячне світло складається із фотонів різних енергій. До того ж, для кожного сонячного елемента, з якого би напівпровідника він не був виготовлений, існує крива спектральної чутливості (рис.21.17). Із кривої легко зрозуміти, що ефективно перетворюється в фотострум тільки кванти світла з енергією $W_\Phi \approx W_g$. Якщо вибрати для виготовлення сонячного елемента напівпровідник з великою шириною забороненої зони W_g, то практично весь спектр сонячного світла пройде через такий елемент, не створивши електронно – діркових пар ($W_\Phi < W_g$). Якщо, навпаки, виготовити сонячний елемент із матеріалу з малим значенням W_g, то практично всі фотони, що містяться в сонячному світлі, будуть мати енергію $W_\Phi \gg W_g$. Із рис.21.17 видно, що такі фотони дають дуже малий вклад у фотострум. Вони поглинаються біля самої поверхні елемента, далеко від *p–n* переходу, і генеровані ними електронно – діркові пари швидко гинуть із–за поверхневої рекомбінації.

Розрахунки показують, що максимальну ефективність перетворення енергії сонячного світла в електричну енергію повинні мати елементи, які виготовлені із напівпровідника з $W_g \approx 1{,}5$ еВ ($\lambda = 1{,}24/1{,}5$ мкм $\approx 0{,}83$ мкм). Біля цієї оптимальної величини W_g знаходяться значення забороненої зони добре вивчених і освоєних в промисловості матеріалів Si і GaAs. Із них, головним чином, і виготовляють сонячні елементи. Типові значення фото-е.р.с. сонячного елемента складає ~0,5 В, максимальна сила фотоструму I_ϕ~50мА при робочій площі елемента S~2см^2. Для отримання більших напруг (10-20 В) елементи з'єднують послідовно, а для отримання робочих струмів – паралельно. Так отримують сонячні батареї.

§21.6. Інші види контактів

Важливим елементом більшості напівпровідникових приладів є контакт метала і напівпровідника. Історично цей контакт був основою одного з перших напівпровідникових випрямляючих діодів, який на сьогодні не має широкого застосування. Однак, завдяки досягенням сучасних технологій, інтерес до такого типу контактів все більше зростає. Це стосується перш за все випрямляючого контакту метал – напівпровідник, який називається бар'єром Шотткі.

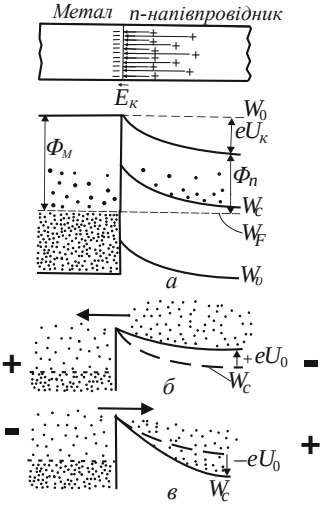


Рис. 21.19. Енергетичні діаграми біля контакту металу з напівпровідником *n*– типу ($\Phi_M > \Phi n$ при рівновазі (*а*); при прямому (*б*); зворотньому (*в*) зміщеннях. W_0 – енергія вакууму.

Бар'єр Шотткі утворюється біля межі розділу металу і напівпровідника (рис.21.19). Якщо робота виходу електронів із металу (Φ_M) більша, ніж із напівпровідника (Φ_n), то рівноважний стан встановлюється при контактній різниці потенціалів $U_к = (\Phi_M - \Phi_n)/e$, яка виникає в результаті переходу електронів із напівпровідника в метал. Контактне поле $E_к$, направлене так як показано на рис.21.19,*а* і практично не буде проникати в метал. Воно локалізується в деякому приповерхневому шарі напівпровідника, товщина якого може бути визначена, як і в електронно-дірковому переході, за формулою (21.20). Цей шар збіднюється основними носіями заряду і просторовий заряд в ньому створюється перш за все позитивно іонізованими донорними атомами. Збіднений шар, який називають деколи запірним, забезпечує контакту нелінійні властивості.

При підключенні зовнішньої напруги U_0 в прямому напрямку (рис.21.19,*б*) поле джерела послаблює контактне поле і зменшує висоту потенційного бар'єру на величину eU_0. Електрони, основні носії *n*– напівпровідника, переходячи в метал, створюють прямий струм тим більший, чим нижчий потенціальний бар'єр. Підвищення потенційного бар'єру при зворотній напрузі (рис.21.19,*в*) зменшує прямий струм і через контакт буде протікати зворотний струм, який обумовлений рухом термічно збуджених електронів металу в прискорюючому полі контактної області. Таким чином, в створенні струму приймають участь тільки електрони і перехід їх із напівпровідника в метал не супроводжується явищем дифузії і рекомбінації. Надлишковий заряд електронів практично миттєво розподіляється в об'ємі, створюючи дрейфовий струм. Це і забезпечує відсутність в контакті метал – напівпровідник бар'єрної ємності і, відповідно, його високу швидкодію. Діоди, створені на основі такого контакту (діоди Шотткі), можуть працювати на частотах порядку 100 ГГц. Вольт–амперна характеристика контакту метал–напівпровідник майже така ж, як і для електронно – діркового переходу. Однак на її вигляд суттєво впливає форма потенціального бар'єру біля поверхні металу. Сили дзеркального відображення згладжують бар'єр і зовнішня напруга викликає ефект аналогічний ефекту Шотткі при емісії електронів в вакуум (розділ 20). Зовсім іншими властивостями володіє контакт металу й електронного напівпровідника, якщо $\Phi_M < \Phi_n$ (рис.21.20). В цьому випадку контактна різниця потенціалів

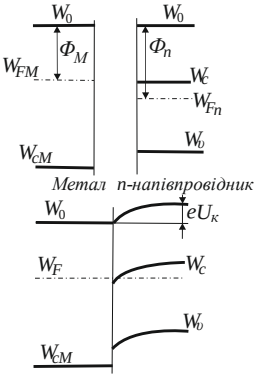


Рис.21.20.Енергетична діаграма контакту метал – напівпровідник при $\Phi_M < \Phi_n$.Омічний контакт.

має протилежний напрямок і електричне поле викликає збагачення контактного шару напівпровідника електронами. Тому такий контакт має малий опір, а його вольт – амперна характеристика майже лінійна. Контакти з лінійною характеристикою називають омічними і широко використовуються для створення виводів напівпровідникових приладів.

Аналогічні властивості проявляє і контакт типу n^+n (знак "+" означає високу ступінь легування напівпровідника домішкою).

Оскільки із збільшенням концентрації донорної домішки рівень Фермі розташовується ближче до зони провідності і зменшується робота виходу, а за концентрацією електронів n^+– напівпровідник наближається до металів, то енергетична діаграма (рис.21.21) і електричні властивості n^+n контакту мало відрізняються від омічного контакту метал – напівпровідник.

Досить різноманітними властивостями володіють гетеропереходи. Гетеропереходи – це контакт двох різних напівпровідників, які мають відповідно і різні ширини заборонених зон. Вид енергетичної діаграми і властивості гетеропереходу залежать від значень ширини забороненої зони W_{g1} і W_{g2}, енергії електронної спорідненості ($W_0 - W_c$, $W_0 - W_v$, де W_0 – енергія вакууму, W_c і W_v – енергія "дна" зони

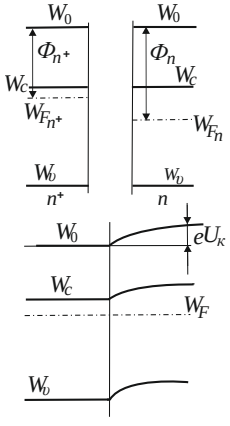


Рис.21.21.Енергетична діаграма контакту n^+n.

провідності напівпровідника n–типу і W_v енергія "стелі" валентної зони напівпровідника p–типу), від виду і концентрації домішки в областях, які контактують. Розглянемо контакт двох напівпровідників з різним характером провідності (неізотипний) або з однаковим типом провідності (ізотипний). Утворення контактної різниці потенціалів викликає, як і у звичайному електронно–дірковому переході, викривлення енергетичних зон (рис.21.22,*а*). Однак через різницю в ширині заборонених зон в місці контакту виникають розриви дозволених енергетичних рівнів ΔW_c і ΔW_v. Наявність таких розривів забезпечує різну висоту потенціального бар'єру для різнойменних носіїв: $eU_к + \Delta W_v$ – для дірок, які переходять із p–області в n– області, і $eU_к - \Delta W_c$ – для електронів, які рухаються в оберненому напрямку. Така різниця дозволяє забезпечити, по суті, односторонню інжекцію, в даному випадку електронів із n– в p–напівпровідник. Необхідно мати на увазі, що потенціальний пік, який утворюється на границі в зоні провідності може бути досить вузьким і тому прозорим для електронів із-за тунельного ефекту.

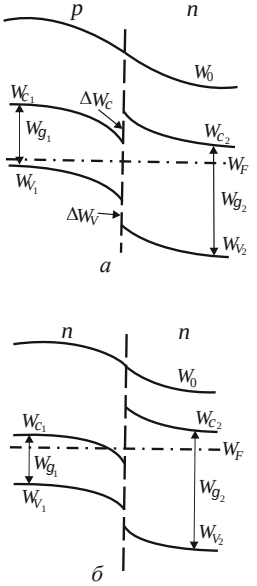


Рис.21.22.Енергетична діаграма гетеропереходу: *а*–неізотипний, *б*–ізотипний.

Властивості односторонньої провідності можуть виявлятись і в ізотипних *nn*

– або *pp* – переходах. На рис.21.22,*б* зображена енергетична діаграма електронно–електронного *nn*–гетеропереходу. Як і в звичайних електронно – діркових переходах, умова переходу носіїв в протилежних напрямках тут будуть різними, що викликано різницею в ширинах заборонених зон. Головною особливістю даного переходу є те, що прямий струм створюється практично тільки електронами і процес дифузії неосновних носіїв відсутній. Відсутність перехідних процесів, які пов'язані з кінцевим часом життя неосновних носіїв, дозволяє використовувати такі переходи, як і переходи з бар'єром Шотткі, для дуже високих частот.

Наведені приклади звичайно не вичерпують всі можливі типи гетеропереходів. Їхня різновидність пов'язана з відмінністю в енергії електронної спорідненості, в ширині забороненої зони, концентрації і типі домішки провідників, які контактують. При розгляді властивостей гетеропереходів необхідно мати на увазі одну особливість. Не зважаючи на те, що гетеропереходи виготовляються із напівпровідників, які мають однотипну будову кристалічної гратки, відмінність в їх параметрах неминуче викликає існування великої кількості поверхневих станів в місці контакту, вплив яких на властивості гетеропереходу може бути дуже значним. Наприклад, на поверхневих станах може накопичуватись заряд, поле якого буде суттєво визначати форму потенціальних бар'єрів в переході. Поверхневі стани можуть виявитись ефективними центрами рекомбінації електронів і дірок в місці контакту. Цілеспрямоване використання поверхневих станів дозволяє ще більше урізноманітнити властивості гетеропереходів, спектр використання яких сьогодні дуже широкий (наприклад, в напівпровідникових лазерах).

Контрольні питання

1.Чим визначається глибина проникнення електричного поля в напівпровідник ?

2.Як визначається контактна різниця потенціалів *p*- і *n*- областей напівпровідника?

3.Які $p - n$ переходи називають різко несиметричними?

4.Які знаєте основні параметри $p - n$ переходу і чим вони обумовлені?

5.Від чого суттєво залежить густина струму неосновних носіїв заряду через $p - n$- перехід?

6.Що означає зворотнє зміщення для $p - n$- переходу?

7.Ширина бар’єру і струм через $p - n$-перехід при зворотньому зміщенні.

8.Що означає пряме зміщення для $p - n$-переходу?

9.Висота бар’єру і струм через *p*–*n*-перехід при прямому зміщенні.

10.Чим визначається максимальне значення висоти потенційного бар’єру $p - n$-переходу?

11. Чим обумовлена бар’єрна ємність $p - n$-переходу?

12.Основні принципи роботи фотодіодів і світлодіодів

13.Контакт метал-напівпровідник. В якому випадку такий контакт є омічним?

14.В чому різниця між *p*-*n* переходом і гетеропереходом?

Розділ 22.Елементи квантової електроніки

Квантова електроніка – галузь фізики, яка вивчає методи підсилення і генерації електромагнітних коливань і хвиль, які базуються на використанні вимушеного випромінювання, а також властивості квантових підсилювачів і генераторів і їх застосування. Лазери – це генератори і підсилювачі когерентного випромінювання в оптичному діапазоні довжин хвиль, робота яких базується на індукованому (вимушеному) випромінюванні квантових систем – атомів, іонів, молекул, що знаходяться в станах, які суттєво відрізняються від стану термодинамічної рівноваги. Слово ЛАЗЕР і МАЗЕР (квантовий генератор або підсилювач діапазону надвисоких частот (НВЧ)) походять із початкових букв англійської фрази Light (Microwave) Amplification by Stimulated Emission of Radiation, що означає підсилення світла (НВЧ хвиль) індукованим (вимушеним) випромінюванням. Лазери, як і мазери, генератори і підсилювачі НВЧ діапазону, називають ще *квантовими* генераторами (підсилювачами), так як поведінка частинок, які приймають участь в їх роботі, описується законами квантової механіки. Принципова різниця лазерів від усіх інших джерел світла (теплових, газорозрядних та інших), які по суті справи є джерелами оптичного шуму, у високій ступені когерентності лазерного випромінювання. Зі створенням лазерів в оптичному діапазоні з'явились джерела випромінювання когерентних хвиль в радіодіапазоні, які здатні успішно використовуватись для цілей зв'язку та передачі інформації, а по багатьом властивостям – направленості випромінювання, смузі частот, які передаються, низькому рівні шумів, концентрації енергії в просторі і в часі і т.п. – переважають класичні пристрої радіодіапазону.

§22.1.Фізичні основи квантової електроніки

Процеси випромінювання і поглинання світла реальними атомами, молекулами та іншими квантовими системами описується квантовою електродинамікою. Проте ці процеси можна дослідити на основі простої феноменологічної теорії, яку розвинув Ейнштейн в 1916р. В цій теорії вперше використано уявлення про *вимушене випромінювання.*

Будемо вважати, що вільний атом може знаходитись тільки в стаціонарних станах з визначеними енергіями W_1, W_2, ... Перехід атома із одного стаціонарного стану в інший може відбуватись скачком в результаті *поглинання* або *випромінювання* електромагнітного випромінювання, причому для такого елементарного процесу виконується закон збереження енергії: $W_m - W_n = \hbar\omega$ – енергія фотона, який поглинається або випромінюється, дорівнює різниці енергії відповідних стаціонарних станів атома. Ці квантові уявлення про будову атома і характер його взаємодії з випромінюванням, які узагальнили гіпотезу Планка про гармонічний осцилятор, були введені Бором в 1913р. і повністю підтверджуються сучасною квантовою теорією.

Розглянемо систему N_0 частинок (атоми, молекули), які знаходяться при сталій температурі T (рис.22.1). Енергія електронів в атомах дискретна. Кожному значенні енергії відповідає енергетичний рівень. Є рівень з найменшою енергією. Такий рівень енергії

називають основним. Всі інші рівні енергії називаються збудженими. Якщо електрон в атомі знаходиться на основному рівні (або стані), то будемо вважати, що атом знаходиться

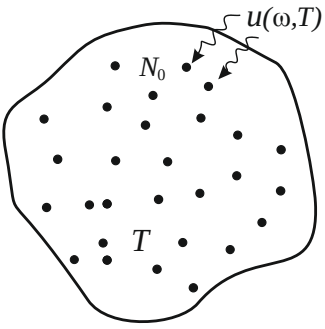


Рис.22.1.Точками позначені частинки (атоми, молекули і т.п.).

в основному стані. Якщо електрон знаходиться на збудженому рівні (стані), то будемо вважати, що атом знаходиться у збудженому стані. В рівноважному стані ($T = \text{const}$) розподіл атомів за енергетичними рівнями (станами) підпорядковується класичному розподілу Больцмана

$$N = N_0 \exp\left\{-\frac{W}{kT}\right\}, \qquad (22.1)$$

де N_0 – загальне число атомів, N – число атомів на енергетичному рівні, енергія якого W, k – стала Больцмана.

Обмежимося розглядом двох енергетичних рівнів: основного і першого збудженого (рис.22.2). Енергія основного рівня W_1, а збудженого W_2. За формулою (22.1) кількість атомів, що знаходяться в основному стані (на рівні W_1) буде –

$$N_1 = N_0 \exp\left\{-\frac{W_1}{kT}\right\}, \qquad (22.2)$$

а в збудженому (W_2) –

$$N_2 = N_0 \exp\left\{-\frac{W_2}{kT}\right\}. \qquad (22.3)$$

В рівноважному стані завжди $N_2 < N_1$, так як $W_2 > W_1$. Кількість атомів, що знаходяться наданому енергетичному рівні (в даному енергетичному стані), називається його *заселеністю*. Тобто, згідно (22.1), чим більша енергія рівня, тим менша його заселеність. Нехай вільний атом, який не піддається зовнішнім впливам, знаходиться в збудженому стані з енергією W_2. Згідно квантових уявлень, самочинний перехід атома за відсутності зовнішніх впливів зі збудженого стану W_2 в основний W_1 з випромінюванням фотона (*спонтанне випромінювання*) відбувається миттєво, скачком (рис.22.2, перехід 1). В який

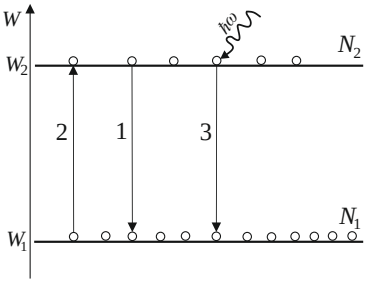


Рис.22.2.Кружечками позначені атоми на енергетичних рівнях.

саме момент відбувається цей перехід, передбачити неможливо. Момент випромінювання фотона є випадковою величиною, судження про яку можуть мати тільки статистичний характер . Позначимо через A_{21} ймовірність спонтанного переходу атома за одиницю часу зі збудженого стану в основний. Розглянемо сукупність дуже великого числа однакових атомів, взаємодією між якими можна знехтувати. Нехай в момент часу t в першому збудженому стані W_2 знаходиться N_2 атомів. На протязі проміжку часу від t до $t + dt$ частина зі них спонтанно перейде в основний стан. Неможливо вказати, які саме атоми здійснюють перехід, але, знаючи ймовірність A_{21}, можна вказати середнє число dN_{21} таких переходів:

$$dN_{21} = A_{21} N_2 dt. \qquad (22.4)$$

Якщо при цьому ніяких процесів збудження атомів не відбувається, то зміна числа N_2 збуджених атомів за проміжок часу від t до $t + dt$ згідно (22.4), дорівнює

$$dN_2 = -A_{21}N_2dt. \qquad (22.5)$$

Коефіцієнт Ейнштейна A_{21} характеризує атом, який розглядається, і не залежить від часу. Розв'язок рівняння (22.5) має вигляд:

$$N_2\,(t) = N_{20}\exp\{-t/\tau\}, \quad (22.6)$$

де N_{20} – число збуджених атомів при t = 0, $\tau = 1/A_{21}$.

Таким чином, за відсутності зовнішніх впливів в результаті спонтанного випромінювання число збуджених атомів зменшується за експоненційним законом. Проміжок часу $\tau = 1/A_{21}$, на протязі якого N_2 зменшується в e =2,72… раз, дорівнює середньому часу життя атомів у збудженому стані. За таким самим законом (22.6) повинно зменшуватись з часом світіння газу збуджених атомів. Згідно квантових уявлень, спонтанне випромінювання – це сукупність незалежних переходів: один зі збуджених атомів може повернутись в основний стан через короткий проміжок часу, інший може прожити в збудженому стані значно довше, але середній для великої сукупності атомів час життя τ має цілком певну величину. На квантовій мові це означає, що спонтанному випромінюванню атома при переході із збудженого стану в основний відповідає вузький, але кінцевий інтервал частот $\Delta\nu \sim 1/\tau$.

Оскільки частота випромінювання визначається умовою Бора $h\nu = W_2 - W_1$, то розклад частот свідчить, про кінцеву ширину ΔW енергетичного рівня збудженого стану W_2. Цей висновок знаходиться у повній відповідності зі співвідношенням невизначеності Гайзенберга між енергією і часом $\Delta W \cdot \Delta t \geq \hbar$ або $\tau \geq \hbar/\Delta W$, яке пов'язує невизначеність в зміні енергії системи з невизначеністю моменту часу, коли ця зміна відбудеться.

Експериментальна перевірка закону (22.6) дала для червоної лінії водню ($\lambda = 0{,}6562$ мкм) $\tau = 1{,}5 \cdot 10^{-8}$ с, а для лінії ртуті ($\lambda = 0{,}2537$ мкм) $\tau = 9{,}8 \cdot 10^{-8}$ с.

Статистичний, випадковий характер процесів спонтанного випромінювання приводить до того, що фази, напрямки розповсюдження і стан поляризації світлових хвиль, які випромінюються окремими атомами, не узгоджені одні з одними. Це значить, що спонтанне випромінювання некогерентне. В електромагнітному полі крім спонтанного випромінювання будуть відбуватись і процеси збудження атомів, тобто переходи із основного стану в збуджений з поглинанням фотонів з енергією $\hbar\omega = W_2 - W_1$. Ймовірність такого переходу в одиницю часу пропорційна густині енергії електромагнітного поля $u(\omega, T)$ на частоті переходу ω і деякому коефіцієнту B_{12}, який характеризує ймовірність збудження атома (рис.22.1,перехід 2). Середнє число dN_{12} переходів з основного стану в збуджений за проміжок часу від t до $t + dt$ пропорційне також числу N_1 атомів в основному стані:

$$dN_{12} = N_1B_{12}u(\omega, T)dt. \qquad (22.7)$$

Коефіцієнт Ейнштейна B_{12} залежить від вибраних станів W_1 і W_2 атома і може бути розрахований методами квантової механіки. Тут його розглядатимемо як феномелогічну постійну. B_{12} не залежить від спектральної густини енергії $u(\omega, T)$ поля. Пропорційність числа dN_{12} актів поглинання (22.7) величині $u(\omega, T)$ спостерігається тоді, коли спектральна густина $u(\omega, T)$ плавно залежить від частоти ω

біля частоти переходу. Допустимо тепер, що атоми знаходяться в термодинамічній рівновазі з полем випромінювання. Тоді на основі принципу детальної рівноваги число переходів з випромінювання фотонів і з поглинанням повинно бути однаковим. Якщо прирівняти праві частини (22.4) і (22.7), можна знайти вид функції $u(\omega,T)$, тобто спектральну густину рівноважного випромінювання. Однак при цьому для $u(\omega,T)$ отримується не формула Планка (16.21), а її граничний випадок при $\hbar\omega/(kT) >> 1$, тобто формула Віна (16.30). Щоб таким способом отримати формулу Планка (16.21), яка підтверджується на практиці, необхідно допустити, як вперше показав Ейнштейн, що електромагнітне поле викликає не тільки переходи 2 (рис.22.2) з основного стану в збуджений, але й обернені переходи 3 (рис.22.2) зі збудженого стану в основний, який супроводжується випромінюванням фотонів. Такі переходи під дією зовнішнього поля, на відміну від спонтанного, отримали назву індукованого, або вимушеного (стимульованого) випромінювання.

Число вимушених переходів $dN_{21}^{в}$ за проміжок часу від t до $t+dt$ пропорційне спектральній густині енергії $u(\omega,T)$ поля на частоті переходу, числу N_2 атомів в збудженому стані W_2 і деякому коефіцієнту B_{21}, який характеризує ймовірність такого переходу в атомі. Із врахуванням спонтанного випромінювання повне число переходів за проміжок часу від t до $t+dt$ із збудженого стану в основний дорівнює

$$dN_{21} = N_2(A_{21} + B_{21}\,(\omega,T))dt. \quad (22.8)$$

В стані термодинамічної рівноваги необхідно прирівняти праві частини виразів (22.7) і (22.8) . Тоді

$$u(\omega,T) = \frac{A_{21}/B_{12}}{\frac{N_1}{N_2}-\frac{B_{21}}{B_{12}}}\cdot \quad (22.9)$$

Легко переконатись в тому, що для даної пари рівнів енергії W_1 і W_2 коефіцієнти Ейнштейна B_{12} і B_{21} дорівнюють один одному. Дійсно, при дуже високій температурі густина енергії $u(\omega,T)$ є настільки великою, що у формулі (22.8) можна знехтувати першим доданком порівняно з другим. Це означає, що при високій температурі вимушене випромінювання переважає спонтанне. Прирівнявши для цих умов праві частини (22.7) і (22.8), отримаємо $N_1B_{12} = N_2B_{21}$. Але при рівновазі і при $kT/(\hbar\omega) \to \infty$ відношення населеності рівнів, як видно із (22.2) і (22.3, буде:

$$\frac{N_2}{N_1} = \exp\left\{-\frac{W_2-W_1}{kT}\right\} = \\ = \exp\{-\hbar\omega/(kT)\} \Rightarrow 1. \quad (22.10)$$

Це означає, що при цих умовах $N_2 = N_1$, тому B_{12}=B_{21}. Коефіцієнти B_{12} і B_{21} залежать тільки від властивостей атома і не залежать від зовнішніх умов, при яких відбуваються переходи. Тому рівність $B_{12} = B_{21}$, яка отримана для граничного випадку $T \to \infty$, справедлива завжди, в тому числі і при відсутності теплової рівноваги.

Тоді для довільної температури з урахування (22.10) і $B_{12} = B_{21}$ формула (22.9) перепишеться так:

$$u(\omega,T) = \frac{A_{21}/B_{21}}{\exp\{\hbar\omega/(kT)\}-1}\cdot \quad (22.11)$$

Цей вираз співпадає з формулою Планка (16.21) при $\frac{A_{21}}{B_{21}} = \frac{\hbar\omega^3}{\pi^2c^3}$. Таким чином, всі три коефіцієнти Ейнштейна зв'язані між собою.

Найбільш важливий результат, який витікає з виведення формули Планка, за Ейнштейном, – це висновок про існування вимушеного випромінювання. Тільки в цьому випадку теорія Ейнштейна не протирічить законам теплового випромінювання. При

$\hbar\omega >> kT$ (справедлива формула Віна (16.30)) для підтримки термодинамічної рівноваги практично достатньо спонтанного випромінювання. Вимушене випромінювання в таких умовах значно менше поглинання так як при $kT << \hbar\omega$ термічно збуджених атомів мало. Навпаки, при низьких частотах і високих температурах, коли $\hbar\omega/(kT) << 1$, при тепловій рівновазі переходи з вимушеним випромінюванням відбуваються майже так часто, як і спонтанні переходи з випромінюванням.

Вимушене випромінювання має чудові властивості, на що вперше звернув увагу П.Дірак в 1927 р., застосувавши квантову механіку до поля випромінювання. В кожному акті вимушеного випромінювання відбувається збільшення на одиницю числа фотонів в тій моді (частоті) випромінювання, під дією якої відбувався перехід. Всі фотони однієї моди тотожні. Це значить, що новий фотон не відрізняється від фотонів, які викликали це випромінювання. Частота, фаза, напрямок розповсюдження і поляризація хвиль, які випромінюються при вимушених переходах, точно такі самі, як і у випромінювання, яке викликає переходи.

Розглянемо зв'язок між коефіцієнтами Ейнштейна і коефіцієнтом поглинання α у формулі (15.6). Нехай в розрідженому середовищі розповсюджується паралельний пучок монохроматичного випромінювання, частота якого ω = $(W_2 - W_1)/\hbar$. Зміна числа dN фотонів в потоці при проходженні шару, товщина якого $dx = cdt$ (c – швидкість випромінювання в середовищі) внаслідок процесів вимушеного поглинання і вимушеного випромінювання у відповідності з (22.7) і (22.8) дорівнює

$$dN = (N_2B_{21} - N_1B_{12})u(\omega,T)dx/c. \quad (22.12)$$

Вкладом спонтанного випромінювання в зміну числа фотонів направленного пучка можна знехтувати, оскільки спонтанне випромінювання розповсюджується у всіх напрямах і у фіксований напрямок пучка, який розглядається, попадає незначна його частина. Формула (22.12) дозволяє з'ясувати, які умови необхідні для безпосереднього експериментального виявлення вимушеного випромінювання. Так як B_{21}=B_{12}, то

$$dN = (N_2 - N_1)B_{12}u(\omega,T)dx/c.$$

Звідки видно, що в розповсюдженні направленого пучка вклад поглинання переважає над вкладом вимушеного випромінювання і інтенсивність пучка зменшується у всіх випадках, коли $N_2 < N_1$, тобто число збуджених атомів менше числа атомів в основному стані (закон розподілу (22.1)). Так буває, якщо пучок розповсюджується в середовищі, яке знаходиться в стані термодинамічної рівноваги або близькому до нього. Щоб спостерігати наростання інтенсивності направленого пучка (негативне поглинання) і тим самим експериментально виявити вимушене випромінювання, необхідно створити в середовищі нерівноважний стан, при якому число атомів на більш високому енергетичному рівні було би більшим, ніж на низькому ($N_2 > N_1$). Перша спроба виявити вимушене випромінювання у видимій області спектру в дослідах з парами ртуті, атоми якої переводились в нерівноважний стан електричним розрядом, була зроблена В.А. Фабрикантом (1939 р.). Ним же вперше була висловлена ідея використання вимушеного випромінювання для підсилення світла.

Знайдемо закон зміни інтенсивності пучка по мірі його розповсюдженні в середовищі. Середній потік енергії

випромінювання (інтенсивність $\langle \Pi \rangle = I$) за одиницю часу дорівнює добутку $u(\omega, T)$ на швидкість світла c: $I = u(\omega, T) \cdot c$. Оскільки зміна *потоку* фотонів на інтервалі dx дорівнює cdN, то зміна потоку енергії dI на тому ж відрізку dx дорівнює $\hbar\omega c \cdot dN$. Помножимо обидві частини рівності (22.12) на $\hbar\omega c$ і знайдемо, що:

$$dI = (\hbar\omega/c)(N_2 - N_1)B_{12}u(\omega, T)cdx$$

або

$$dI = -\left(\frac{\hbar\omega}{c}\right)(N_1 - N_2)B_{12}\, Idx.$$

Звідки

$$dI/dx = -\alpha I(x), \qquad (22.13)$$

де

$$\alpha = \left(\frac{\hbar\omega}{c}\right)(N_1 - N_2)B_{12}. \quad (22.14)$$

Якщо населенності рівнів N_1 і N_2 не залежать від x (однорідне середовище), то розв'язок (22.13) має вид

$$I(x) = I_0 \exp\{-\alpha x\}, \qquad (22.15)$$

де I_0 – інтенсивність направленого пучка при $x = 0$. Якщо $\alpha > 0$, що буває при $N_2 < N_1$, то інтенсивність випромінювання експоненціально зменшується по мірі розпосюдження пучка і ми приходимо до закону Бугера (15.6), а формула (22.14) виражає коефіцієнт поглинання α через коефіцієнт Ейнштейна B_{12} і населенності рівнів N_1 і N_2.

Для тих випадків, коли в середовищі виконується умова $N_2 > N_1$, відома як *інверсна населеність*, в (22.14) $\alpha < 0$ (негативний коефіцієнт поглинання) і інтенсивність хвилі у відповідності з (22.15) наростає по мірі її розпосюдження. Підсилення падаючого пучка світла здійснюється внаслідок того, що при $N_2 > N_1$ переходи з вимушеним випромінюванням фотонів відбуваються частіше, ніж переходи з поглинанням. Так як фотони, що виникають при вимушеному випромінюванні, тотожні з фотонами, які викликають випромінювання, то когерентні властивості початкового пучка повністю зберігаються. Такий принцип дії *квантового підсилювача* випромінювання. Різні способи створення необхідного для роботи підсилювача середовища з інверсної населеністю (*активного середовища*) розглянуті в подальшому. Важливо зауважити, що для створення активного середовища завжди потрібно підводити ззовні додаткову енергію, яка потім при вимушеному випромінюванні частково перетворюється в енергію електромагнітного випромінювання, яке підсилюється.

Квантовий підсилювач можна перетворити в генератор випромінювання, якщо ввести в ньому *позитивний обернений зв'язок*. Для цього частина випромінювання, яке виходить із підсилювача, повинно повертатись в активне середовище. Тоді відпадає необхідність у вхідному сигналі, так як відбувається самозбудження системи. Обернений зв'язок можна здійснити, розмістивши активне середовище в резонатор, який утворений двома паралельними дзеркалами.

§22.2. Лазери

Принцип роботи лазера або оптичного квантового генератора (ОКГ) заснований на трьох фундаментальних ідеях, які народились в різний час в різних областях фізики.

Перша ідея пов'язана з використанням *вимушеного випромінювання* світла атомними системами, яке відкрив Ейнштейн в 1917р. теоретично вивчаючи некогерентне теплове випромінювання. Як показав Дірак, фотони

вимушеного випромінювання не відрізняються від фотонів, які викликали це випромінювання.

Друга ідея полягає в застосуванні *термодинамічно нерівноважних систем*, в яких можливе підсилення, а не поглинання світла. Вона була висловлена В.А.Фабрикантом в 1940р.

Третя ідея, яка бере початок в радіофізиці, полягає в використанні *позитивного оберненого зв'язку* для перетворення підсилювальної системи в генератор когерентного випромінювання.

Те, що перші квантові прилади з'явились в радіодіапазоні (НВЧ – надвисокі частоти), пояснюється тим, що класична радіофізика не могла вирішити традиційними для неї методами ряд дуже важливих з практичної точки зору проблем, таких, як освоєння більш коротких хвиль, створення в короткохвильовому діапазоні високостабільних, малошумливих і достатньо потужних приладів. Це наполегливо заставляло шукати інший підхід розв'язку цих проблем і призвело до створення в НВЧ-діапазоні принципіально нових приладів – квантових підсилювачів і генераторів.

Перший квантовий генератор (лазер), який працював на квантових переходах з довжиною хвилі 1,25 см, був створений у 1954 р. Н.Г.Басовим і А.М. Прохоровим (СРСР) і групою під керівництвом Ч.Таунса (США). Активним середовищем в них був пучок молекул аміаку: інверсна населеність на робочому переході досягалась просторовим розподілом молекул аміаку, які знаходились в різних квантових станах, при проходженні пучка через неоднорідне електричне поле (див.формулу (2.8)). Перший оптичний квантовий генератор, який працював в імпульсному режимі, був створений в 1960р. Мейманом. Активним середовищем в ньому служив стрижень з кристалу рубіна, який збуджувався світлом від лампи спалаху. Роком пізніше Джаваном, Беннетом і Ерріотом був створений перший лазер неперервної дії з газоподібним активним середовищем (суміш неону і гелію). Сьогодні існує багато типів лазерів, які відрізняються способами збудження активного середовища, робочою спектральною областю, потужністю, часовими і спектральними характеристиками випромінювання і т.п.

Існують різні способи отримання необхідного для лазера активного середовища, яке підсилює випромінювання. Згідно рис.22.2 для створення стаціонарної інверсії ($N_2 > N_1$), необхідно, щоб швидкість накачки (процеси 2) і час життя атомів в збудженому стані W_2 були більшими, ніж для нижнього рівня W_1. З цього витікає, що віддати перевагу необхідно селективній накачці і, що інверсія може бути досягнута не тільки внаслідок переважного заселення верхнього рівня W_2 (рис.22.2), але і внаслідок високої швидкості очистки нижнього рівня W_1.

Існують такі способи створення активного середовища (інверсна заселеність енергетичних рівнів W_1 і W_2 (рис.22.2)):

1.Внаслідок поглинання світла (оптична накачка). Підбираючи джерело світла із відповідним спектром випромінювання, можна забезпечити високу селективність (вибірковість) накачки. Найбільш успішно цей вид накачки використовується в твердотільних (на кристалах і склах) лазерах і в лазерах на барвниках.

2.Внаслідок непружних зіткнень атомів і молекул із вільними електронами, при яких частина енергії електрона йде на збудження атома або молекули. Вільні електрони можуть

створюватись або під час газового розряду, або вводитись в газ у вигляді пучка, який формується в прискорювачі.

3.Внаслідок непружних зіткнень атомів робочої речовини зі збудженими атомами або іонами допоміжного газу з передачею енергії збудження від них робочій речовині. В деяких типах зіткнень передача енергії носить резонансний характер і досягається висока ступінь селективності заселення рівнів.

4.В процесі спеціально підібраних хімічних реакцій (хімічна накачка); при цьому збуджуються коливальні рівні молекул, причому збудження може бути селективним.

5.Внаслідок нагріву (теплова накачка) . Цей метод використовується для накачки коливальних рівнів в молекулах, інверсія на переходах між якими здійснюється за рахунок різного часу життя (часу релаксації) для верхнього і нижнього лазерних рівнів при швидкому адіабатичному (без теплообміну) розширенні газу. На цьому принципі базується робота газодинамічних лазерів.

6. В напівпровідникових лазерах неперервної дії нерівноважний (інверсний) стан досягається при пропусканні електричного струму через $p-n$ перехід.

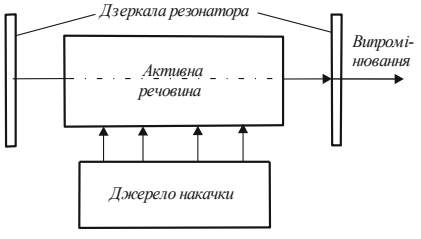


Рис.22.3.

Щоб перетворити підсилювач (активне середовище) в генератор, необхідно організувати обернений позитивний зв'язок. Для його здійснення частина випромінювання, яке генерується, повинна залишатись всередині активного середовища і викликати вимушене випромінювання все новими і новими збудженими атомами. З цією метою активне середовище розміщують в оптичний резонатор (рис.22.3), який утворений двома паралельними плоскими або сферичними дзеркалами, одне із яких напівпрозоре. Розглянемо як відбувається генерація ("запалювання") ОКГ.

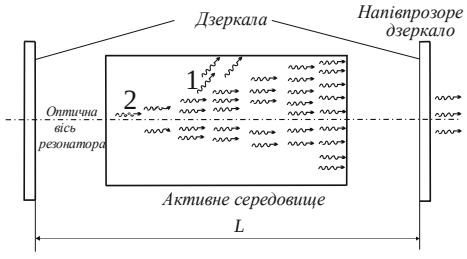


Рис.22.4.

Нехай є активне середовище (рис.22.4), тобто середовище з інверсною населеністю енергетичних рівнів ($N_2 > N_1$). Якщо якийсь атом спонтанно випромінив хвилю в напрямку під кутом до оптичної осі (рис.22.4, хвиля 1), то ця хвиля в цьому напрямку індукує вимушене випромінення, яке рано чи пізно покине активне середовище. Тобто, всяке спонтанне випромінення в напрямку під кутом до оптичної осі резонатора не має шансів бути підсиленим і вийти з ОКГ. Тільки хвиля, яка спонтанно випромінена атомом в напрямку оптичної осі резонатора (рис.22.4,хвиля 2) буде підсилюватись, зустрічаючи на своєму шляху розповсюдження збуджені атоми (атоми на енергетичному рівні W_2). Підсилення хвилі тим більше, чим більший шлях в активному середовищі вона пройде. Після відбивання від дзеркала її підсилення продовжуватиметься. Цей процес багаторазово повторюється, а енергія хвилі в резонаторі

стрімко наростає. В цьому і проявляється позитивний зворотний зв'язок, який здійснюється за допомогою дзеркал. Підсилення хвилі відбувається згідно формули (22.15) стільки раз, скільки раз хвиля проходить через активне середовище. Якщо активне середовище заповнює простір між дзеркалами, то відбившись від дзеркала хвиля знову на довжині резонаторами L підсилюється в $\exp\{|\alpha|L\}$ раз, де $|\alpha|$ – коефіцієнт підсилення (формула (22.14)). Необхідно зазначити, що система із двох паралельних дзеркал, володіє резонансними властивостями – резонує тільки на певних частотах і виконує в лазері ту роль, яку у звичайних низькочастотних генераторах відіграє коливальний контур.

Нехай на подвійній довжині резонатора вкладається ціле число довжин хвиль: $2L = \lambda q$, де q – ціле число. Цим забезпечується позитивний обернений звязок: при другому і всіх наступних проходженнях через резонатор хвиля узгоджена за фазою з вихідною, тобто фактично не відрізняється від неї. Всі елементарні хвилі, які виникають в різні моменти часу при вимушених переходах атомів, які розташовані в різних точках активного середовища, створюють одну когерентну хвилю. Таким чином, оптичний резонатор збільшує ефективну відстань, яку хвиля, що розповсюджується вздовж його осі, проходить в активному середовищі, і сприяє формування когерентного монохроматичного випромінювання, яке виходить через напівпрозоре дзеркало.

Якщо оптичний резонатор розглядати як коливальну систему, в якій власні нормальні коливання (*моди*) мають вид стоячих електромагнітних хвиль з вузлами на дзеркалах, то частоти ν таких мод визначаються із умови $2L = q\lambda_q$, де q – ціле число, тобто утворюють еквідистанційний спектр:

$$\nu_q = v/\lambda_q = c/(n\lambda_q) = q(c/(2nL)), \quad (22.16)$$

де n – показник заломлення активного середовища, c – швидкість світла в вакуумі. Тому лазер – це автоколивальна система, в якій можливі незгасаючі коливання на одній із власних частот резонатора. Частотна відстань між сусідніми модами (q відрізняється на одиницю) дорівнює

$$\Delta\nu = c/(2nL). \quad (22.17)$$

Однак існує низка причин, які будуть зменшувати амплітуду хвилі всередині резонатора. Це втрати при відбиванні від дзеркал, а також втрати на розсіювання в середовищі та деякі інші. Якщо ввести для описання сумарних втрат коефіцієнт втрат β, то формула (22.15) перепишеться так:

$$I = I_0\exp\{-(\alpha + \beta)x\}. \quad (22.18)$$

Для виникнення стійкої генерації необхідно, щоб $(\alpha + \beta < 0)$, тобто, втрати випромінювання повинні бути меншими, ніж підсилення.

Необхідно також підкреслити, що корисна дія дзеркал виявляється і в тому, що безперервно збільшується кількість атомів, які випромінюють вимушено, оскільки за кожний цикл відбивання зростає густина енергії електромагнітних хвиль $u(\omega, T)$. Зазначимо, що цей процес не буде продовжуватися нескінченно, оскільки починаючи із деякого значення $u(\omega, T)$, наступає явище насичення в ОКГ. Інверсія в заселеності енергетичних рівнів та кількість вимушених переходів з випромінюванням зменшується, отже, і підсилення уповільнюється. Ефект насичення – це одне із основних явищ нелінійної оптики, який відіграє важливу роль в розумінні механізму генерації випромінювання ОКГ та

визначає умови стандартності процесу підсилення в лазерах.

На рис.22.5 приведена частотна залежність коефіцієнта підсилення α в робочому середовищі (вона визначається шириною і формою спектральної лінії робочого середовища, наприклад, лінії при переході атома із рівня W_2 на рівень W_1 (рис.22.2), і набором власних частот відкритого резонатора. Для відкритих резонаторів з високою добротністю пропускання резонатора $\Delta\nu_р$, яка визначає ширину резонансних кривих окремих мод (криві 2−4 на рис.22.5), і навіть відстань між сусідніми модами $\Delta\nu$ є меншою від ширини лінії підсилення $\Delta\nu_п$ (рис.22.5). Тому в контур підсилення попадає декілька типів мод коливань резонатора.

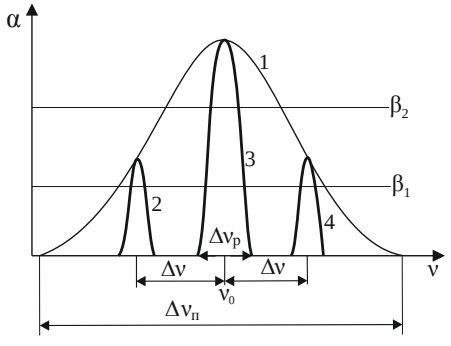


Рис.22.5.Частотна залежність коефіцієнта підсилення α в робочому середовищі (крива 1 і типи коливань резонатора (2−4), які потрапляють в контур підсилення. $\Delta\nu_п$ −ширина лінії підсилення. При рівні втрат β_1 можуть генеруватися три моди (2,3,4), а при β_2 – одна центральна (3) ($\beta_2 > \beta_1$).

Таким чином, лазер не обов'язково генерує на одній частоті, частіше навпаки, генерація відбувається одночасно на декількох типах коливань, для яких підсилення α більше втрат в резонаторі β (три моди на рис.22.5 при $\beta = \beta_1$). Для того, щоб лазер працював на одній частоті (в одночастотному режимі), необхідно, як правило, вживати спеціальних заходів (наприклад, збільшити втрати, як показано на рис.22.5 при $\beta = \beta_2$) або змінити відстань між дзеркалами так, щоб в контур підсилення попала тільки одна мода. Це забезпечить високу монохроматичність $(\Delta\lambda/\lambda)$ лазерного випромінювання.

Тобто, якщо коефіцієнт підсилення в робочому середовищі перекриває втрати в резонаторі для певних типів коливань, то на них виникає генерація. Затравкою для її виникнення є, як і в будь – якому генераторі, шум. В лазерах цим шумом є спонтанне випромінювання.

Необхідно зазначити ще одну властивість лазерного випромінювання – це його вузьку спрямованість (малу розбіжність). Вона є наслідком того, що хвиля, яка виходить з ОГК, є результатом складання величезної кількості висококогерентних хвиль від багатьох атомів.

Тут доречно згадати явище дифракції на решітці, де розподіл інтенсивності світла на екрані є результатом інтерференції когерентних хвиль, якій ідуть від великої кількості щілин. Причому, чим більше щілин решітки, тим більша інтенсивність максимуму ($\sim N^2$) і менша його ширина ($\sim\frac{1}{N}$), і тим швидше спадають інтенсивності інших максимумів. Випромінювання лазера з певним припущенням можна вважати граничним випадком дифракції. Випромінювання атомів активного середовища когерентне, кількість їх величезна, а тому випромінювання активного середовища зосереджене у вузькому дифракційному куті, а інтенсивність на осі променя у N^2 разів більша інтенсивності хвилі окремого атома, однак, цей кут через дифракцію на вихідній щілині отвору ОКГ не може бути зведений до нуля в принципі. Таким чином, значна когерентність лазера і

визначає незначну розбіжність його випромінювання.

§22.3.Властивості лазерного випромінювання

1.*Потужність* лазерного випромінювання в неперервному режимі може бути порядку $10^5 - 10^6$ Вт, в імпульсному – до $10^{12} - 10^{16}$ Вт, при цьому вдається досягнути інтенсивність порядку $10^{12} - 10^{16}$ Вт/см^2. Суттєво, що ці потужності можуть бути сконцентровані в надзвичайно вузьких спектральних ($\Delta\lambda -$ мале) і часових інтервалах.

2.*Тривалість* імпульсу випромінювання в лазерах може складати $10^{-12} - 10^{-13}$ с і спеціальними методами доводиться до 10^{-15} с (за цей час світло проходить всього $3 \cdot 10^{-5}$ см), тобто лазер володіє дивовижно високою ступінню концентрації енергії по часу.

3.*Монохроматичність* лазерного випромінювання визначається як $\Delta\nu_p/\nu_0$ (або $\Delta\lambda_p/\lambda_0$), де $\Delta\nu_p -$ ширина лінії генерації (рис.22.5), яка в принципі обумовлена шумами, $\nu_0 -$ її центральна частота, при роботі лазера на одній частоті і в неперервному режимі. Використовуючи спеціальні методи стабілізації, вдається отримати відносну стабільність частоти $\Delta\nu_p/\nu_0 = 10^{-12} - 10^{-13}$ ($\Delta\nu_p = 50 - 500$ Гц).

Висока ступінь монохроматичності лазерного випромінювання визначає високу спектральну густину енергії – високу ступінь концентрації світлової енергії в дуже малому спектральному інтервалі. Висока монохроматичність полегшує фокусування лазерного випромінювання, так як при цьому хроматична аберація лінзи (залежність показника заломлення матеріалу лінзи n від довжини хвилі λ) стає несуттєвою.

4.*Когерентність*. Лазери володіють надзвичайно високою, порівняно з іншими джерелами світла, ступінню когерентності: часовою і просторовою (див. §12.5 і §12.6). При роботі лазера в одномодовому режимі досягається повна просторова когерентність, що визначає високу направленість лазерного випромінювання і робить можливим його фокусування в пляму надзвичайно малих розмірів (порядка довжини хвилі λ).

Часова когерентність (формули (12.37) і (12.38)) може бути записана так:

$$t_{когер} = \lambda^2/(c \cdot \Delta\lambda) = 1/\Delta\nu_p, \quad (22.19)$$

тобто, визначається монохроматичністю і виявляється дуже високою. Так для лазера на $He - Ne$, який працює неперервно в одномодовому режимі $\Delta\nu_p \cong 1$ кГц і $t_{когер} = 10^{-3}$с. Тоді довжина когерентності $l_{когер} = t_{когер} \cdot c$ ($c -$ швидкість світла) складає $3 \cdot 10^5$ м (300 км), в той час як для нелазерних джерел світла (наприклад, натрієва лампа) $t_{когер} = 10^{-10}$ с ($l_{\text{когер}} = 0{,}03$ м). Таким чином, з використанням лазерів можна спостерігати інтерференційну картину навіть при різниці ходу променів в декілька кілометрів.

5.*Направленість* лазерного випромінювання в основному визначається тим, що в відкритому резонаторі можуть збуджуватись тільки такі хвилі, які напрямлені по осі резонатора (рис.22.4) або під дуже малими кутами до неї. При високій ступені просторової когерентності кут розбіжності лазерного променя θ може бути доведений до граничного, який визначається дифракцією на вихідному отворі лазера. Типові значення θ складають: для газових лазерів $(0{,}5 \div 5) \cdot 10^{-3}$ радіан, у твердотільних $(2 \div 20) \cdot 10^{-3}$

радіан, у напівпровідникових $(5\div50)\cdot10^{-3}$ радіан.

6.*Яскравість*. Через високу направленість випромінювання лазерні джерела світла мають дуже високу яскравість, завдяки чому на мішені можна отримати дуже велику інтенсивність світла. Так, гелій - неоновий лазер з потужністю всього 10 мВт і розбіжністю випромінювання 10^{-4} радіан при поперечній площі пучка 0,1 $см^2$ має яскравість 10^6 $Вт/(см^2\cdot стерадіан)$, що набагато перевищує яскравість Сонця $(130\ Вт/(см^2\cdot стерадіан))$.

Примітка. Яскравість (L), поверхнево-просторова густина світлового потоку, який йде від поверхні, дорівнює відношенню світлового потоку $d\Phi$ до геометричного фактору $d\Omega dS\cos\theta$:

$$L = d\Phi/(d\Omega dS\cos\theta)\,.$$

Тут $d\Omega$ – заповнений випромінюванням тілесний кут, dS – площадка, яка випромінює або приймає випромінювання, θ – кут між перпендикуляром до цієї площини і напрямком випромінювання. В нашому випадку $d\Omega = \pi tg^2\,\theta \cong 3\cdot10^{-7}$ стерадіан, $\cos\theta \approx 1$, тому

$$L = \frac{10\cdot10^{-3}}{3\cdot10^{-7}\cdot10^{-1}}\ \frac{Вт}{см^2\cdot стерадіан} \sim$$
$$\sim 0{,}3\cdot10^{6}\ \frac{Вт}{см^2\cdot стерадіан}\cdot$$

§22.4.Особливості конструкції і роботи деяких типів лазерів

1. *Лазер на кристалі рубіну*. Рубін – це кристал окису алюмінію Al_2O_3 (корунд), в якому невелика частинка іонів алюмінію ($\approx 0{,}05\%$) при вирощуванні кристалу заміщена іонами хрому Cr^{3+}. Сам корунд в видимій області прозорий і основну роль в роботі лазера грають іони хрому. Спрощена схема енергетичних рівнів хрому показана на рис.22.6,*а*. Інверсія населеності створюється між основним станом (рівнем) W_1 і збудженим станом (рівнем) W_2. Цьому переходу відповідає довжина хвилі $\lambda = 0{,}6943$ мкм в червоній області спектру. Вище рівня W_2 лежать широкі смуги енергетичних рівнів W_3 і W_3'. Переходам на них із основного стану відповідають дві широкі смуги поглинання в зеленій і синій областях спектру. З цим поглинанням зв'язане рожеве забарвлення рубіну.

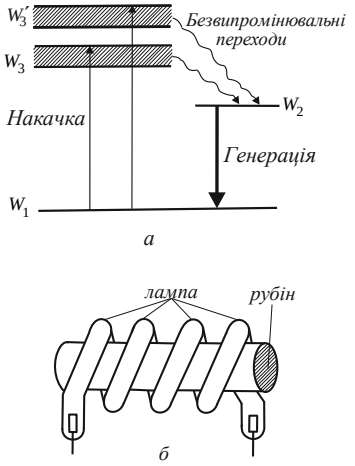


Рис.22.6.

Циліндричний стержень рубіну, довжина якого декілька сантиметрів і діаметр біля 1 см, освітлюється потужним імпульсом білого світла від лампи спалаху, яка обвиває його у вигляді спіралі (рис.22.6,*б*). Тривалість імпульсу біля 1 мс. При достатній енергії спалаху більша частина іонів хрому, поглинаючи світло, переходить в стани W_3 і W_3'. Потім іони хрому за час порядку 10^{-8} с безвипромінювально переходять на рівень W_2, передаючи надлишок енергії коливанням кристалічної гратки. Час життя (час релаксації) іонів хрому в збудженому стані W_2 складає декілька мілісекунд, що на декілька порядків перевищує типові величини часу життя

збуджених станів ($10^{-7} \div 10^{-8}$с). Збуджені рівні із таким великим часом життя називається *метастабільними*. При недостатній енергії спалаху на рівні W_2 виявиться менше половини всіх іонів хрому. Тоді обернений їхній перехід в основний стан відбувається спонтанно (за час ≈ 1 мс) і супроводжується люмінесценцією (свіченням) на довжині хвилі 0,6943 мкм.

Якщо ж в результаті поглинання світла спалаху і наступних безвипромінювальних переходів на рівні W_2 накопичиться більше половини всіх іонів хрому, то між рівнями W_1 і W_2 виникає інверсія населеностей $(N_2 > N_1)$. Такий механізм утворення інверсії населеності називається *оптичною накачкою.* Якщо рубіновий стрижень розміщений в зовнішньому оптичному резонаторі або має посріблені плоскопаралельні торці, в ньому виникає короткий імпульс лазерної генерації на довжині хвилі 0,6943 мкм. Лазерний імпульс має складну часову структуру і складається із нерегулярної послідовності окремих імпульсів тривалістю біля 1 мкс. Через малу тривалість імпульсу (порядка 1 мс) потужність рубінового лазера в імпульсі досягає декількох кіловат при порівняно невеликій енергії (декілька джоулів).

2. *Лазер на неодимовому склі*. Поряд з рубіновим лазером, який працює по *трьохрівневій схемі*, широке розповсюдження отримали *чотирьохрівневі лазери* на іонах рідкоземельних елементів (неодим, самарій), які впроваджені в кристалічну (флюорит кальцію) або скляну матрицю. Ці іони мають структуру енергетичних рівнів (рис.22.7), яка надзвичайно добре підходить для лазерів. При кімнатній температурі різниця енергій $W_2 - W_1$ рівнів 2 і 1 в декілька разів перевищує kT, тому рівень W_2 практично пустий (слабо заселений). За допомогою сильного світла лампи накачки іони неодиму збуджується в широку смугу рівнів 4 (W_4). Час життя іона в стані 4 дуже малий ($\tau \geq \hbar/\Delta W$, а ΔW порівняно значна величина), тому з великою ймовірністю відбуваються безвипромінювальні переходи в метастабільний стан 3(W_3). Інверсія населеності

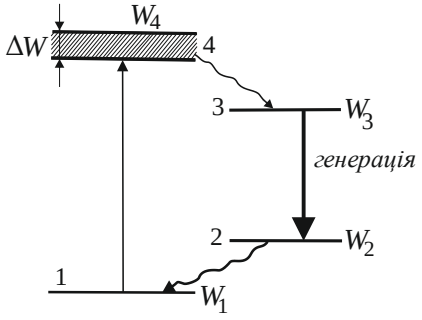


Рис.22.7.

між рівнями 2 і 3 лазерного переходу досягається дуже легко, так як нижній рівень 2 практично незаселений. Лазер на неодимовому склі генерує випромінювання на довжині хвилі 1,06 мкм у вигляді імпульсів з дуже великими енергіями (біля 1 кДж). За чотирьохрівневою схемою працює лазер неперервної дії на кристалах іттрій-алюмінієвого граната (YAG) з домішкою неодима і з вихідною потужністю до 1 кВт.

3.*Газові лазери* за багатьма характеристиками переважають лазери інших типів. Вони перекривають широкий спектральний діапазон (від субміліметрового до ультрафіолетового). Серед газових лазерів завжди можна знайти лазер, який має принаймні одну з наступних властивостей: високу ступінь монохроматичності випромінювання, малу розбіжність випромінювання (аж до дифракційної границі), граничне значення потужності в неперервному режимі, високий ККД і т.п.

Будова найбільш поширеного гелій-неонового лазера схематично показана на

рис.22.8. Газорозрядна трубка з внутрішнім діаметром 1−10 мм і довжиною від декількох десятків сантиметрів до 1,5−3 м має торцеві

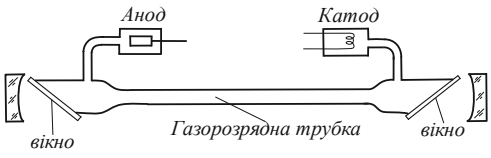


Рис.22.8.Розрядна трубка газового лазера з вікнами, які встановлені під кутом Брюстера.

плоскопаралельні скляні або кварцеві вікна, які встановлені під кутом Брюстера до її осі. Для лінійно поляризованого світла з електричним вектором в площині падіння коефіцієнт відбивання дорівнює нулю (§14.2, рис.14.8;14.10). Тому брюстеровські вікна забезпечують лінійну поляризацію випромінювання лазера і виключають втрати енергії при розповсюдженні світла із активного середовища до дзеркал і назад. Трубка розміщується в резонатор, який утворений двома дзеркалами. Пропускання дзеркала, через яке виводиться випромінювання, вибирається біля 1−2%, іншого − менше 1%.

Розрядна трубка заповнюється сумішшю гелію і неона в молярному відношенні 5:1 при тиску $10^2 \div 10^3$ Па. До електродів вмикається джерело з напругою в декілька кіловольт. Типова сила струму при розряді – десятки міліампер. Тліючий розряд створює умови для виникнення інверсії населеності рівнів в неоні. Гелій використовується для резонансного збудження неона.

На рис.22.9 наведена спрощена схема енергетичних рівнів неону (Ne). Випромінювання з довжинами хвиль 0,6328 і 1,150 мкм відповідають переходам $W_4 \to W_2$ і $W_3 \to W_2$. Атоми неону збуджуються (накачуються) в результаті їхнього зіткнення з електронами в газовому розряді, при цьому вони переходять зі стану W_1 в стани W_3 і W_4.

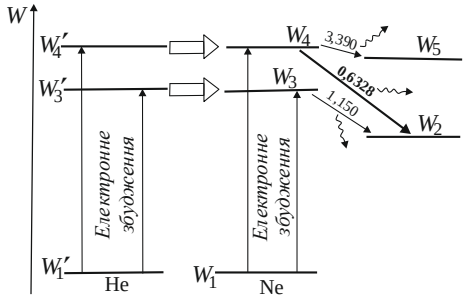


Рис.22.9.Схема енергетичних рівнів He та Ne. Числа на рисунку означають довжини хвиль випромінювання в мікрометрах.

Енергії електронів газового розряду вистачає для створення інверсного заселення рівнів W_3 і W_2. Випромінювання, яке виникає при цьому, припадає на інфрачервону область спектру ($\lambda = 1{,}15$ мкм). Рівні ($W_4 \to W_2$) і ($W_4 \to W_5$) залишаються заселеними неінверсно.

Інтенсивне заселення цих рівнів можна забезпечити, якщо додати до неону атоми гелію (He). Гелій має два метастабільних (наче стабільних) стани W_3' і W_4', які «заселяються» при зіткненнях атомів з електронами і за величиною енергії близькі до рівнів W_3 і W_4 неону. Час життя атомів у цих станах набагато більший часу життя в звичайних станах (10^{-8}с). Пов'язано це з тим, що для гелію переходи $W_3' \to W_1'$ $W_4' \to W_1'$ є забороненими, тобто не можуть бути здійсненими шляхом випромінювання, а будь-яким іншим, наприклад, в процесі зіткнення з атомами неону (горизонтальні стрілки на діаграмі). У результаті таких зіткнень атоми неону набувають енергії і їх кількість на рівнях W_3 і W_4 значно зростає. Виникає інверсне заселення рівнів W_4 і W_2, а, отже, можливе підсилення випромінювання ($\lambda = 0{,}6328$ мкм). Висока оптична однорідність газового активного

середовища дозволяє отримувати випромінювання з дуже високим ступенем часової і просторової когерентності.

4. *Напівпровідниковий лазер*, лазер на основі напівпровідникового кристалу. На відміну від лазерів інших типів в напівпровідниковому лазері використовуються випромінювальні квантові переходи між дозволеними енергетичними зонами, а не дискретними рівнями енергії. В напівпровідниковому *активному середовищі* може досягатись дуже великий показник оптичного підсилення α до 10^6 м$^{-1}$ (формула (22.14), завдяки чому розміри активного елемента напівпровідникового лазера виключно малі (довжина резонатора $L{\sim}50$ мкм $\div$ 1 мм). Крім компактності, особливостями напівпровідникового лазера є мала інерційність ($\sim 10^{-9}$с), високий коефіцієнт корисної дії (до 50%), можливість спектрального перетворювання і великий вибір речовин для генерації в широкому спектральному діапазоні від $\lambda = 0{,}3$ мкм до 30 мкм. Активними частинками в напівпровідниковому лазері є надлишкові (нерівноважні) електрони провідності і дірки, тобто вільні носії заряду, які можуть інжектуватись, дифундувати і дрейфувати в активному середовищі. Найважливішим способом накачки в напівпровідниковому лазері є інжекція носіїв через $p-n$ перехід або гетероперехід, яка дозволяє здійснювати безпосереднє перетворення електричної енергії в когерентне випромінювання (інжекційний лазер). Іншим способом накачки служить електричний пробій (наприклад, в так званих стримерних лазерах), бомбування електронами (напівпровідниковий лазер з електричною накачкою) й освітлення (напівпровідниковий лазер з оптичною накачкою). Напівпровідникові лазери вперше створені на $p-n$ переході в кристалі арсеніда галію (GaAs) з електричною накачкою.

Оптичне підсилення в напівпровідниках виникає під дією інтенсивної накачки при виконанні умов інверсії населеності рівнів біля дна W_c в зоні провідності і стелі W_v в валентній зоні (рис.22.10). При цьому ймовірність заповнення

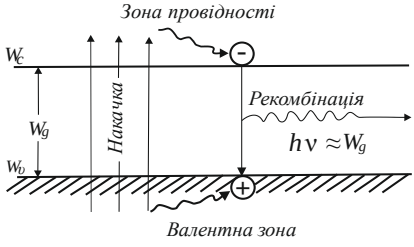


Рис.22.10.

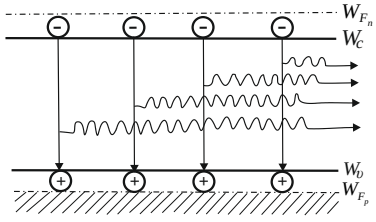


Рис.22.11.

електронами верхніх робочих рівнів в дозволеній зоні (зоні провідності) більша, ніж нижніх рівнів (валентної зони). В цьому випадку вимушені випромінювальні переходи переважають над поглинальними переходами (рис.22.11). Величина оптичного підсилення залежить не тільки від інтенсивності накачки, але й від інших факторів: ймовірності випромінювальної рекомбінації, внутрішнього квантового виходу випромінювання, температури. Як лазерні матеріали використовуються прямозонні напівпровідники (наприклад, GaAs, CdS, PbS), в яких квантовий вихід випромінювання може досягти 100%. На непрямозонних напівпровідниках (Ge, Si) створити лазер не вдається.

Різноманітність напівпровідникових лазерних матеріалів дозволяє перекрити широкий спектральний діапазон з допомогою напівпровідникових лазерів. Інжекційний напівпровідниковий лазер – це напівпровідниковий діод, дві плоскопаралельні грані якого перпендикулярні площині $p-n$ переходу (гетеропереходу), служать дзеркалами оптичного резонатора (коефіцієнт відбивання $\sim 30\%$ рис.22.12.). Деколи застосовуються зовнішні резонатори. Інверсія заселеності досягається при великому прямому струмі через діод внаслідок інжекції надлишкових носіїв в шар, який прилягає до переходу. Генерація когерентного випромінювання виникає в смузі краєвої люмінесценції (випромінювання), якщо оптичне підсилення здатне перевищити втрати енергії, які пов'язані з виходом випромінювання назовні, поглинанням і розсіянням всередині резонатора. Струм, який відповідає початку генерації, називається пороговим. Густина порогового струму в інжекційних напівпровідникових лазерах, як правило, $\sim 1 кА/см^2$.

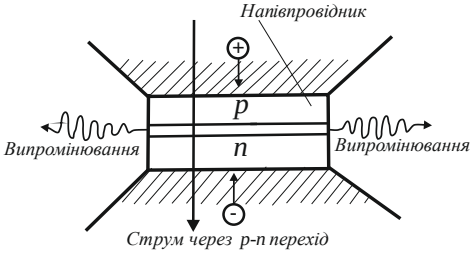


Рис.22.12.

Найбільше розповсюдження отримали напівпровідникові лазери на снові гетероструктур (гетеролазери), які мають найбільш низькі порогові густини струму при температурах 300 К. Гетеролазер містить два гетеропереходи, один типу $p-n$, що інжектує електрони (емітер), і другий, типу $p-p$, що обмежує дифузійне розтікання носіїв заряду з активного шару: активна область знаходиться між ними. В так званих смугових лазерах активна область в формі вузької смужки, ширина якої $1-20$ мкм, протягується вздовж осі резонатора від одного дзеркала до другого.

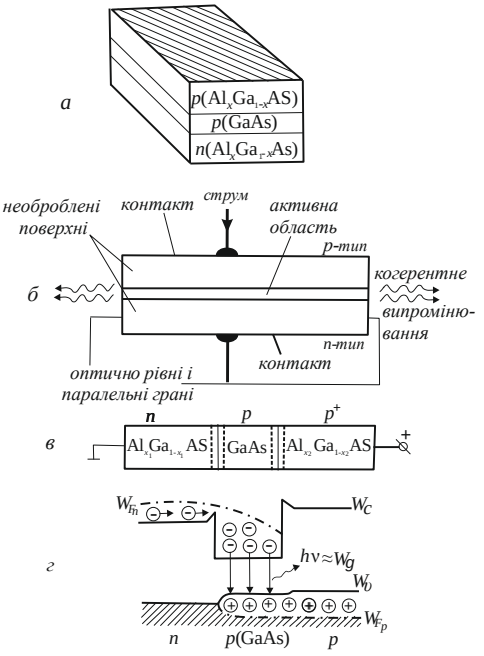


Рис.22.13. Схема гетеролазера з двохсторонньою гетероструктурою на основі AlGaAs (*а,б,в*) і його енергетичною діаграмою (*г*). W_c і W_v – краї зони провідності і валентної зони. W_{F_n} і W_{F_p} – енергії Фермі для електронів і дірок.

Через малі розміри активної області пороговий струм смугових гетеролазерів досить малий ($5-100$ мА) для отримання неперервної генерації при $T = 300\,K$. Потужність випромінювання таких напівпровідникових лазерів (~ 100 мВт) обмежена перегріванням активної області. В коротких імпульсах напівпровідникові лазери випромінюють більшу потужність (до 100 Вт), яка обмежена оптичним руйнуванням торцевих граней. Багатоелементні інжекційні напівпровідникові

лазери дають в імпульсі потужність до 10 кВт. Напівпровідники, з яких можуть бути виготовлені гетеролазери, при різному хімічному складі повинні мати однаковий період кристалічної решітки. Використовуються багатокомпонентні тверді розчини, серед яких можна знайти неперервні ряди речовин з постійним періодом решітки (ізоперіодичні системи). Наприклад, в гетеролазері на основі твердих розчинів $Al_xGa_{1-x}As$ гетероструктура складається з шарів (рис.22.13); $p(Al_xGa_{1-x}As)$; $p(GaAs)$ і $(Al_xGa_{1-x}As)$. В структурі з двома гетеропереходами носії зосереджені всередині активної області ($p-GaAs$ на рис.22.13,*г*), яка обмежена з обох сторін потенціальними бар'єрами; випромінювання теж обмежене цією областю внаслідок скачкоподібного зменшення показника заломлення за її межами. Ці обмеження сприяють підсиленню стимульованого випромінювання і відповідно, зменшенню порогової густини струму.

§22.5.Фізичні основи волоконно–оптичних ліній зв'язку

Телефонні розмови, телевізійні програми, комп'ютерна інформація – все це, закодовані в певній послідовності ультракороткі лазерні спалахи, які передаються на великі відстані по дуже тонкому кварцевому волокну. Хоч суто зовнішньо волоконно-оптична лінія зв'язку (скорочено ВОЛС) нічим не відрізняється від традиційної провідної: кабель, приймальний і передаючий модулі по її кінцям, ретранслятори (регенератори), які вбудовуються в кабель через певні відстані для відновлення сигналу, який послабшав, комутатори там, де лінія розгалужується, стики, які з'єднують все в єдине ціле (рис.22.14). Звичайно, "начинка" всіх цих блоків – волокна, лазери, фотодіоди, оптико-електронні перемикачі – надають цим лініям зв'язку нові якості.

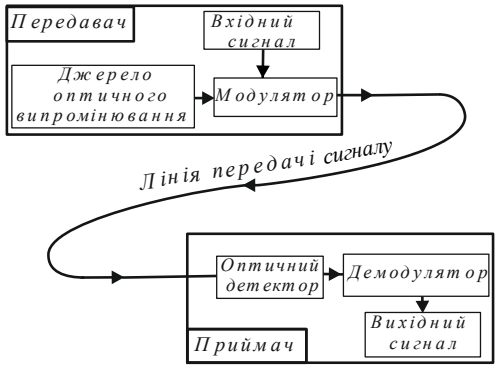


Рис.22.14.Структурна схема оптичної системи зв'язку.

Волоконна оптика як оригінальний науково-технічний напрямок зародився в кінці п'ятдесятих років 20 століття, коли навчились виготовляти спеціальні скляні волокна – високопрозорі, тонкі, гнучкі і досить міцні. Їх створення є технологічною вершиною майже вікових пошуків оптиками світловодів, які придатні для передачі світла криволінійними траєкторіями. У другій половині 19 століття була проблема освітлювати декілька приміщень однією електричною дугою (її зручно було "запалити" десь на стороні), подати світло в цехи, де виготовляли порох, але так, щоб не сталося вибуху. Пробували порожнисті дзеркальні трубки, суцільні скляні і кварцеві волокна і навіть потік рідини, але все це було не те. Накінець знайшли: двохшарове скляне волокно – серцевина зі скла з одним показником заломлення, оболонка з іншим – чудово передає світловий потік від вхідного торця до вихідного, затримуючи його всередині серцевини незалежно від того, чи витягнути волокно в струну, чи згорнуте котушкою.

З'явився ідеальний засіб для підсвітки важкодоступних місць для ендоскопії, для

передачі зображень, для перетворення світлових полів. Волоконна оптика кардинально перетворила багато напрямів приладобудування, медицини, вимірювальної техніки, нею зацікавились спеціалісти з засобів відображення інформації, логіко-математичного аналізу. Революційну роль волоконної оптики порівнювали з напівпровідниками, які в той період уже зробили крок від транзисторів до мікросхем. Але про оптичний зв'язок чути було мало. Збільшення довжини волокна від одного метра до десяти зменшувала світлопередачу до десятих і сотих долей відсотка. В 1970 р. в одному із випусків Праць американського інституту радіоінженерів навіть було висловлено резюме: оптичний зв'язок не відбувся… Але саме в цьому році з'явилось повідомлення американської фірми "Corning Glass" про виготовлення кварцевого волокна із такою значною прозорістю, що лазерний сигнал міг пробігти по ньому більше кілометра з невеликим ослабленням. Світловод для оптичного зв'язку був створений.

Оптичне волокно

Досвід свідчить, що кожний новий технічний прорив, як правило, базується на "трьох китах": фізика–матеріал–технологія. Фізика світлопередачі оптичним волокном базується на законах відбивання і заломлення світла. Розглянемо проходження світлового випромінювання через границю розділу двох середовищ. В геометричній оптиці світлові хвилі зображуються променями, які зазнають змін на межі розділу середовищ з різними оптичними властивостями, які характеризуються показником заломлення $n = c/v$, який в загальному випадку показує, у скільки разів швидкість c розповсюдження світла в вакуумі більша швидкості v розповсюдження світла в середовищі, яке розглядається (§11.3).

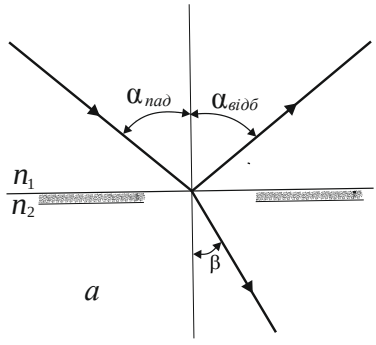


Рис.22.15,*а*.Відбивання і заломлення світлового променя на межі двох середовищ $n_1 < n_2$.

Як відомо, середовище, у якого показник заломлення більший, називається оптично більш густим, у противному випадку – менш густим, тому при падінні променя світла (світлової хвилі) на границю розділу таких середовищ в загальному випадку виникає відбита і заломлена хвилі (рис.22.15,*а*). У

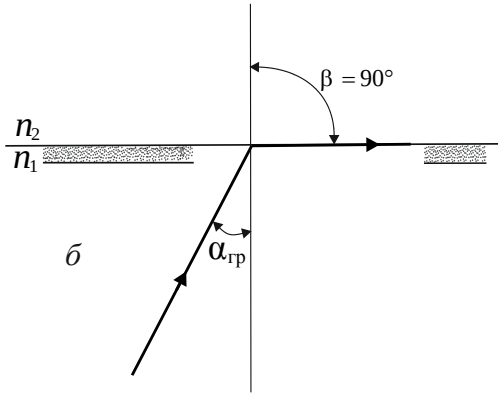


Рис.22.15,*б*. Повне внутрішнє відбивання світлового променя на межі двох середовищ $n_1 > n_2$.

відповідності з законом Снелліуса, кути падіння $\alpha_{пад}$, відбивання $\alpha_{відб}$ і заломлення β зв'язані наступними співвідношеннями:

$$\alpha_{пад} = \alpha_{відб} \text{ і } n_1 \sin\alpha_{пад} = n_2 \sin\beta\,. \quad (22.20)$$

Якщо промінь переходить з оптично більш густого середовища в менш густе $n_1 > n_2$ (рис.22.15,*б*), то $\beta > \alpha_{пад}$. Шляхом збільшення

кута падіння можна досягнути випадку, при якому заломлений промінь буде ковзати по межі розділу середовищ, не переходячи в друге середовище (рис.22.15,*б*). Кут падіння, при якому має місце даний ефект, називається граничним кутом $\alpha_{гр}$ повного внутрішнього відбивання. Очевидно, що для всіх кутів падіння, більших критичного $(\alpha_{пад} > \alpha_{гр})$, буде мати місце тільки відбивання, а заломлення буде відсутнім. Це явище називаться *повним внутрішнім відбиттям* і, оскільки вся потужність світлового променя практично повністю повертається в область більш густого середовища, на цьому ефекті базується принцип передачі оптичного випромінювання по оптичному волокну.

Розповсюдження світлових променів в оптичних волокнах

В загальному випадку оптичне волокно (рис.22.16) складається з серцевини,

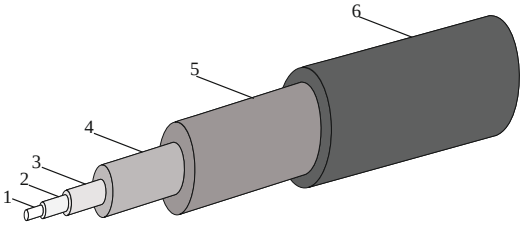


Рис.22.16 Загальний вид типового оптичного волокна: 1– серцевина; 2– оболонка; 3–силіконове покриття; 4–буферна облонка; 5– зміцнююча оболонка; 6– зовнішня оболонка.

по якій відбувається розповсюдження світлових хвиль, і оболонки, яка призначена, з одного боку, для створення кращих умов відбивання на межі розділу "серцевина – оболонка", а з іншого – для зниження випромінювання енергії в оточуючий простір. З метою підвищення міцності і тим самим надійності волокна поверх оболонки, як правило, накладають первинне захисне зміцнююче покриття. Така конструкція оптичного волокна є типовою і використовується в багатьох оптичних волокнах як базова. Розглянемо тепер структуру світловоду, по якому передається інформація в сучасних оптичних лініях зв'язку. Оптичне волокно складається з серцевини (1 на рис.22.16) й оболонки (2 на рис.22.16), які є осісиметричним діелектричним хвилеводом. Хвилеводна структура утворюється внаслідок зменшення показника заломлення від осі хвилевода до його оболонки для виконання умови повного внутрішнього відбиття (рис.22.15,*б*). Якщо зміна показника заломлення відбувається скачком (рис.22.17,*а*, $n_2 < n_1$.), то волокна називаються ступеневими. Волокна з плавною зміною показника заломлення називаються градієнтними. Для таких волокон показник заломлення n залежить від відстані r від осі волокна до його периферії, причому $dn/dr < 0$, тобто, зі збільшенням r, n зменшується (рис.22.17,*б*).

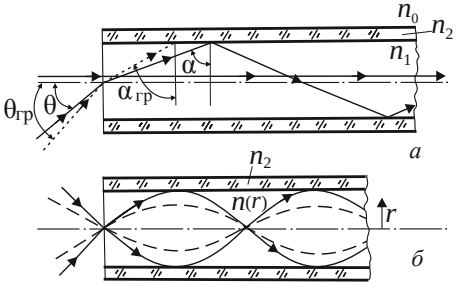


Рис.22.17.Розповсюдження світла в ступеневому (*а*) і градієнтному (*б*) волоконному світловоді.

В ступеневому волокні промені зазнають багатократного відбивання від межі між серцевиною і оболонкою. В градієнтному хвилеводі відбивання не має, а відбувається викривлення траєкторії променя внаслідок оптичної неоднорідності середовища (зміна показника заломлення в напрямку, який перпендикулярний осі). В результаті промінь, який перетинає вісь, знову повертається до неї,

причому повертання це плавне. Траєкторія променя в такому хвилеводі задовольняє принципу мінімуму оптичної довжини (принцип Ферма). Із рис.22.17,*а* видно, що промені, які розповсюджуються по серцевині ступеневого світловода під кутом $\alpha > \alpha_{гр}$ до його осі, зазнають повного внутрішнього відбиття на межі розділу серцевина-оболонка, і тому розповсюджуються вздовж світловода. Інші промені проникають в оболонку. Граничний кут можна знайти за формулою, що визначається зі співвідношення (22.20):

$$\sin\alpha_{гр} = n_2/n_1.$$

Кожному значенню кута α відповідає значення $\theta -$ кута падіння на торець світловода (рис. 22.17,*а*). Кут $\theta_{гр}$, при якому промінь буде розповсюджуватись по світловоду під кутом $\alpha_{гр}$, називається вхідним кутом світловоду. Як правило, його характеризують апертурою $A = n_0 \sin\theta_{гр}$, де $n_0 -$ показник заломлення зовнішнього середовища. Для ступеневого світловода *апертура* $A = \sqrt{n_1^2 - n_2^2}$, а для градієнтного справедливі співвідношення $\sin\theta_{кр} = \sqrt{n^2 - n_2^2}$ і $A = \sqrt{n_{max}^2 - n_2^2}$, де $n_{max} -$ найбільше значення показника заломлення серцевини світловода. У ступеневих світловодах, які використовуються в системах зв'язку, апертура дорівнює 0,18…0,23, у градієнтних – 0,13…0,18.

Променеве наближення справедливе тільки для описання розповсюдження випромінювання в волокнах з поперечним перерізом, набагато більшим довжини хвилі світла, оскільки в таких світловодах світло розповсюджується у відповідності з законами геометричної оптики.

Якщо ж розміри поперечного перерізу волокна можна порівняти з довжиною хвилі випромінювання, то для аналізу розповсюдження випромінювання через такі світловоди необхідно використати електромагнітну теорію світла і рівняння Максвелла. Строгий розв'язок цих рівнянь показує, що по світловоду малого перерізу одночасно може розповсюджуватись певний дискретний набір електромагнітних хвиль, яких називають модами або власними хвилями. Хвилі в серцевині, які відповідають променям, що утворюють малі і великі кути з віссю світловода, називають *модами нижчих і вищих порядків* відповідно. На рис.22.18 показано, що частина променів, які ввійшли в світловод, розповсюджується тільки в його серцевині. Це направлені моди a і b. Частина променів розповсюджується в оболонці, це моди оболонки c. Накінець частина променів виходить назовні, це витікаючі моди d.

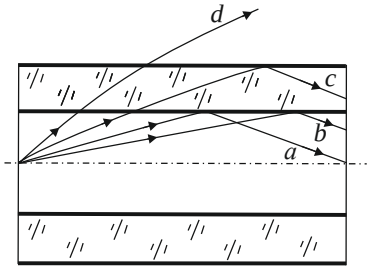


Рис.22.18.Моди в ступеневому волоконному світловоді.

Таким чином, кожна мода є коливання, яке характеризується певною просторовою структурою електричного і магнітного поля і відповідною фазовою швидкістю. Будь-яке оптичне випромінювання, яке розповсюджується по світловоду, можна представити як суперпозицію мод, тобто як результат додавання коливань, що розповсюджуються в ньому. Число цих мод можна розрахувати, якщо відомі довжина хвилі, радіус серцевини світловода і різниця показників заломлення серцевини й оболонки.

В волоконних світловодах, товщина яких від 1 до 5 мкм, розповсюджується одна мода – одне коливання. В більш товстих (від 10 до 100 мкм) світловодах може одночасно розповсюджуватись декілька коливань (мод). На практиці використовуються як одномодові, так і багатомодові світловоди.

Технологія виготовлення світловодів

Революційність технологічного прориву в створенні надпрозорих волокон пов'язана з вибором кварцу як основного матеріалу. Хімічно кварц – це діоксид кремнію SiO_2, в надчистому виді його можна синтезувати методом парафазного осаджування, який розроблений в 1970 році (рис.22.19). Через

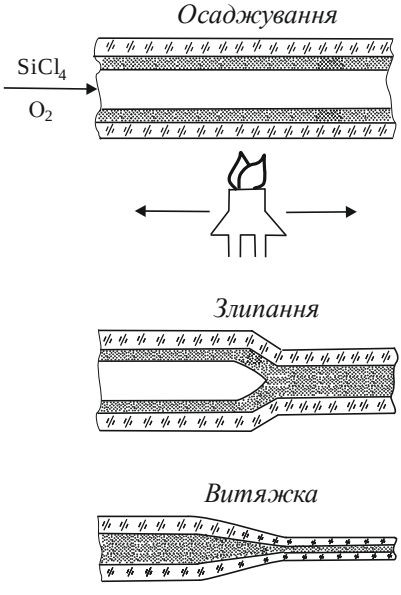


Рис.22.19.Основні етапи виготовлення кварцевого двохшарового світловода.

розпечену кварцеву вихідну трубу пропускають суміш чотирихлористого кремнію ($SiCl_4$) і кисню. Під дією високої температури (~1400℃) в цій суміші відбувається реакція з утворенням кварцу, який осідає на внутрішній поверхні труби. В процесі осаджування в потік добавляють газоподібні реактиви, які містять бор, фосфор, германій – ці присадки трохи збільшують або зменшують показник заломлення кварцу. Після отримання шарів заданих складів і товщин трубу нагрівають ще більше до розм'ягчення і злипання. Отриману таким чином вихідну заготовку потім перетягують через серію *фільєр* (французьке filiere, від fil – волокно, нитка, деталь машин у вигляді ковпачка або пластини з отворами, діаметр яких 0,8 ÷ 0,06 мм, для формування волокон), які послідовно звужуються. Установку для витяжки волокна за розмірами можна порівняти з триповерховим будинком.

Отримати волокна з високою прозорістю – це для цілей оптичного зв'язку умова "необхідна", але "недостатня". Адже інформація кодується послідовністю імпульсів, значить, необхідно мати можливість передавати по волокну саме світлові імпульси, і чим коротші вони будуть, тим вища буде пропускна здатність каналу зв'язку. Важливим параметром оптичного волокна є дисперсія, яка визначає його пропускну здатність для передачі інформації.

Дисперсія оптичних волокон

Дисперсія – це розсіювання з часом спектральних або модових складових оптичного сигналу, яке призводить до збільшення тривалості імпульса оптичного випромінювання при розповсюдженні його по оптичному волокну (рис.22.20) і визначається різницею квадратів тривалостей імпульсів на виході і вході оптичного волокна:

$$\tau = \sqrt{\tau_{вих}^2 - \tau_{вх}^2}\,, \qquad (22.21)$$

де значення $\tau_{вих}$ і $\tau_{вх}$ визначаються на рівні половини амплітуди імпульсів.

Як видно із рис.22.18 і 22.20,*а,б* різні промені з числа апертурних проходять на шляху

розповсюдження до вихідного торця різні відстані: досить порівняти промінь, який падає вздовж осі, з променем, який падає під кутом $\theta_{гр}$ (рис.22.17,*а*). Це призводить до розмиття імпульсу – із прямокутного на вході він стає дзвоноподібним, розмазаним на виході (рис.22.20,*а*), Сусідні імпульси набігають один на одного, виникають помилки, збої, а то і повна втрата інформації. До цієї *хвилеводної* (*міжмодової*) *дисперсії* додається ще і звичайна "*хроматична" дисперсія* – розмиття імпульсу світла, що обумовлене залежністю фазової $v_ф = c/(n(\lambda))$ і групової $v_{гр} = v_ф - -\lambda\frac{dv_ф}{d\lambda} = v_ф\left(1 + \frac{\lambda}{n}\frac{dn}{d\lambda}\right)$ (див. §15.2) швидкостей світла в кварці від довжини хвилі (рис.22.20,*б*).

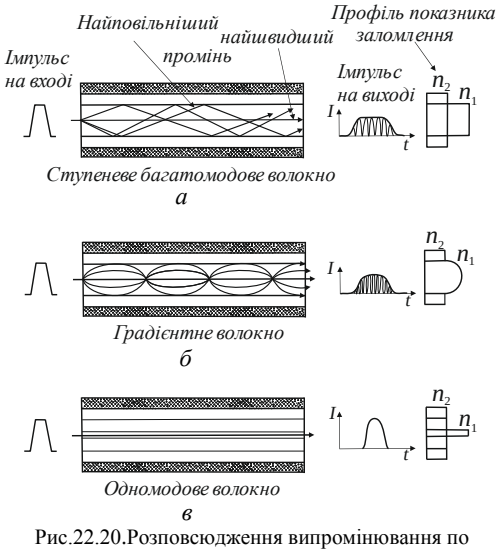


Рис.22.20.Розповсюдження випромінювання по ступеневому (*а*), градієнтному багатомодовому (*б*) і одномодовому оптичному волокну (*в*).

Причиною міжмодової дисперсії є некогерентність джерел випромінювання, які реально працюють в спектрі довжин хвиль від λ до $\lambda + \Delta\lambda$. Міжмодова дисперсія переважає в багатомодових оптичних волокнах. Все це в сукупності призвело до того, що в перших волокнах 70-х років 20 століття при використанні тодішніх не зовсім досконалих лазерів розмиття імпульсу характеризувалось дисперсією $\tau{\sim}40$ нс/км. Це означає, що 10-кілометрова лінія з такого волокна придатна для одночасної передачі тільки декількох телефонних розмов. Міжмодова дисперсія в градієнтних оптичних волокнах, як правило, на порядок нижча, ніж в ступеневих волокнах. Це обумовлене тим, що внаслідок зменшення показника заломлення від осі оптичного волокна (рис.22.20,*б*) до оболонки швидкість $(v = c/n)$ розповсюдження променів вздовж їхньої траєкторії змінюється. Так, на траєкторіях, які близькі до осі, швидкість менша (більший n), а віддалених – більша (менший n). Значить, промені, які розповсюджуються найкоротшими траєкторіями, що близькі до осі, мають меншу швидкість, а промені, які розповсюджуються по більш протяжних траєкторіях, мають більшу швидкість. В результаті час розповсюдження променів вирівнюється і збільшення тривалості імпульсу стає меншим. В 1973 році було встановлено, що випромінювання з довжинами хвиль $\lambda = 1{,}3 - 1{,}5$ мкм майже не зазнають дисперсії в кварці. Світлопропускання в цій області спектру набагато вище, ніж при $\lambda = 0{,}85$ мкм, а саме на такій довжині хвилі працювали тогочасні лазери. Тому і постало питання про перехід в нову спектральну область, що вимагало створення нових джерел і приймачів світла. Для боротьби з хвилевидною (модовою) дисперсією довелось звернутись до одномодових волокон (рис.22.20,*в*) з діаметром серцевини порядку 5 мкм. Саме вони і складають основу сучасного високошвидкісного волоконно-оптичного зв'язку. Для створення джерел світла для ВОЛЗ був синтезований новий клас

напівпровідникових сполук (§22.2), які включають чотири компоненти: індій (In) – галій (Ga) – миш'як (As) – фосфор (P). Варіюючи склад, отримують необхідну довжину хвилі випромінювання лазера.

Приймачі оптичного випромінювання

Одним із базових компонентів ВОЛЗ, за допомогою яких здійснюється реєстрація оптичного випромінювання, є приймачі оптичного випромінювання. Останні виготовлені на основі фоточутливих напівпровідникових елементів, що використовують явище фотопровідності (внутрішнього фотоефекту §19.11). Принцип роботи фотодіодів розглянутий в §21.4.

На сьогоднішній день найбільш розповсюдженим типом детектора оптичного випромінювання є $p-i-n$ структура фотодіода – **PIN** фотодіод (рис.22.21). Основною

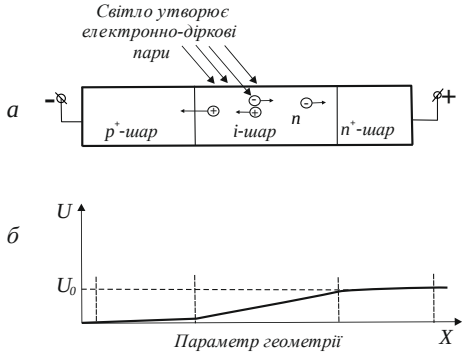


Рис.22.21.Структура включення і розподіл потенціалу $p-i-n$ фотодіоду.

відмінністю даного типу діодів від звичайних діодів з p – n переходами є наявність між сильно легованими p^+ і n^+ шарами напівпровідника так званого i – шару, який є слабо легованим напівпровідником n – типу, товщина якого декілька десятків мікрометрів (рис.22.21,*а*). В i – шарі вільні носії заряду відсутні, то його деколи називають збідненим шаром. Оскільки сильне легування збільшує провідність p^+ і n^+ шарів, то зовнішня різниця потенціалів U, яка прикладена до них, створює в i – шарі градієнт електричного поля (рис.22.21,*б*). При подачі на $p-i-n$ структуру зворотної напруги, то через те, що в i – шарі не має вільних носіїв, даний шар поляризується і через опір навантаження протікає малий постійний струм I_m, який носить назву темнового струму. Значення цього струму визначається параметрами матеріалу напівпровідника, топологією $p-i-n$ структури і температурою оточуючого середовища. При дії на i–шар зовнішнім випромінюванням в ньому утворюються вільні електрон – діркові пари, які швидко розділяються і, прискорюючись прикладеним електричним полем, рухаються до електродів в протилежних напрямках, створюючи у зовнішньому колі діода електричний струм I_ϕ (фотострум). PIN діоди на велику довжину хвилі (1,3–1,5 мкм) виготовляють із Ge, InGaAs, InGaAsP.

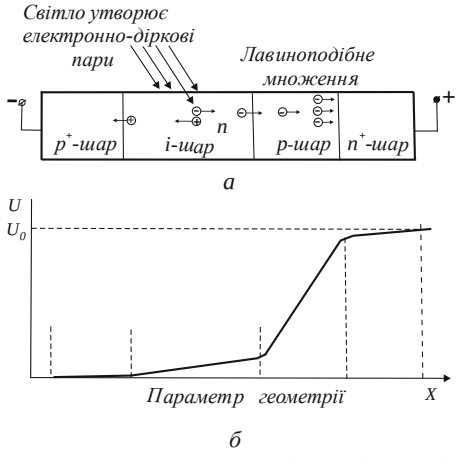


Рис.22.22.Структура включення і розподіл потенціалу APD фотодіода.

Більш високу чутливість, ніж PIN діоди, мають лавинні фотодіодні детектори APD (avalanche photon detector). Основна відмінність APD фотодіодів від PIN діодів, які

мають $p^+ - i - n^+$ структуру в тому, що вони містять додатковий *p*–шар (рис.22.22,*а*). При цьому створюється $p^+ - i - p - n^+$ структура з таким профілем розподілу легуючих домішок, при якому найбільша напруженість електричного поля має місце в *p*–шарі (рис.22.22,*б*), що забезпечує тим самим найменшу провідність цього шару. З цієї причини зворотнє зміщення в APD діодах настільке велике, що збіднений шар збільшується до розмірів *i* – шару не знижуючи, а навіть збільшуючи напруженість електричного поля. У зв'язку з цим електронно – діркові пари, які створюються світлом, розділяються і прискорюються цим полем в збідненому шарі, отримуючи енергію, достатню для іонізації атомів кристалу. Зіштовхуючись з нейтральними атомами, носії викликають зростаючу в геометричній прогресії концентрацію електронів і дірок, так званий лавинний ефект (рис.22.22,*а*). Цей процес носить назву помноження первинного фотоструму.

Використовуються короткохвильові Si APD діоди і довгохвильові InGaAs, Ge APD, InGaAsP /InP та інші.

Введення і виведення світла у світловодах

Для введення випромінювання в світловод і виведення світла на фотодіод використовуються лінзи. Найбільш ефективні лінзи з малою фокусною відстанню, однак у таких лінз великі спотворення. У зв'язку з цим замість звичайних лінз застосовуються так звані градієнтні лінзи, в яких фокусна відстань зменшується не внаслідок геометрії, а внаслідок використання матеріалів, в яких показник заломлення плавно змінюється. В градієнтних стержневих лінзах, яких називають граданами або селфоками, промінь, який ввійшов в лінзу, розповсюджується криволінійною траєкторією.

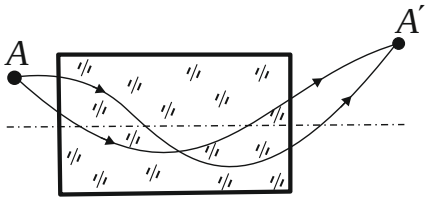


Рис.22.23. Розповсюдження світла в градані.

На рис.22.23 показано розповсюдження променів в градієнтній стержневій лінзі, в якій показник заломлення змінюється за параболічним законом. В залежності від довжини градани можуть бути багато −, одно −, напів − і чверть крокові. Останні два зображені на рис.22.24. Градієнтні стержневі лінзи використовуються для з'єднання окремих волокон, джерела або приймача випромінювання з волокном, в різних розгалужувачах світлових потоків, оптичних перемикачах і т.д.

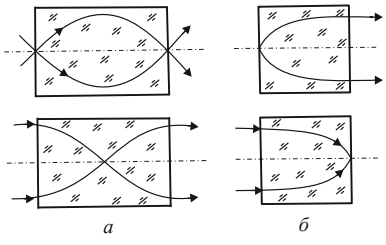


Рис.22.24. Напівкрові (*а*) і чверть крокові (*б*) градани.

Найбільш часто застосовуються чверть крокові градієнтні стержневі лінзи, які як видно із рис.22.24, дозволяють перетворити розбіжний пучок в паралельний, або навпаки – сфокусувати паралельний пучок.

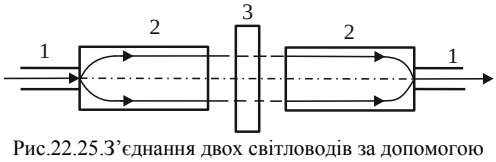


Рис.22.25.З'єднання двох світловодів за допомогою граданів.

На рис.22.25 показано з'єднання двох світловодів 1 за допомогою чвертькрокової градієнтної стержневої лінзи 2.

При співвісному їх розташуванні світло із одного волокна попадає в друге. Розмістивши між градієнтними лінзами світлофільтр або послаблювач потужності 3, можна отримати елемент, який називається аттенюатором.

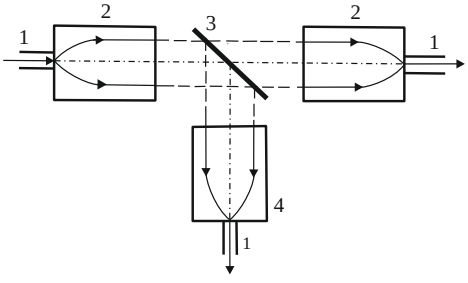


Рис.22.26. Спрямований відгалужувач на граданах.

На рис.22.26 показана схема спрямованого відгалужувача. В ньому між градієнтними лінзами 2 розміщене напівпрозоре дзеркало 3, яке частину випромінювання пропускає, а частину направляє в лінзу 4, яка зв'язана ще з одним світловодом. Втрати світла в таких розгалужувачах мінімальні.

Переваги ВОЛЗ

Перше – дуже мале згасання сигналу у волокні:

$$B = \frac{10}{L}\lg(P_{вх}/P_{вих}),$$

де B – величина згасання в децибелах на кілометр, $P_{вх}$ і $P_{вих}$ – вхідна і вихідна потужність сигналу, L – довжина світловода.

На сьогоднішній день створені світловоди з втратами 0,5 ÷ 1дБ/км. Це означає, що потужність сигналу зменшується в два рази на відстані $L\sim 6$ км. Тобто відстань між ретрансляторами може досягати 100 – 200 км, що набагато перевищує той же показник у провідних лініях. Але і це не межа – теоретичні оцінки показують, що в спектральній області $\lambda \simeq 2 \div 2{,}5$ мкм можна досягнути передачі сигнала без ретранслятора на 1000 і більше кілометрів. Звичайно, для цього доведеться замінити і кварц, і лазери, і фотодіоди.

Друге – це надвисока пропускна здатність: швидкість передачі десятки Гбіт/с вже давно перестала бути рекордною, а типові значення складають сотні і тисячі Мбіт/с. Виявилось, що якщо одночасно пропускати по волокну потоки світлових імпульсів з дещо відмінними довжинами хвиль, то вони не змішуються один з одним і на прийомному кінці можуть бути розділені. Завдяки цьому можна збільшити в десятки, а то і в сотні раз пропускну здатність вже прокладених комунікацій. Але й це ще не все. Якщо вводити в волокно лазерні імпульси підвищеної потужності, то утворюються так звані солітони (поодинокі хвилі), які при розповсюдженні не тільки не "розлазяться", але можуть навіть звужуватись. Швидкість передачі інформації може досягати до 1 Тбіт/с. Сьогодні лабораторні досліди дають швидкість передачі інформації по оптоволокну 1,125 Тбіт/с. Це у 50 тисяч разів перевищує середню швидкість підключення приватних користувачів до інтернету. За таких умов передачі інформації, щоб отримати всі дані, зібрані космічним телескопом Хаббл за 20 років спостереження, потрібно трохи більше, ніж півхвилини.

Третє – переносником інформації у ВОЛЗ є фотон, який не має електричного заряду, тому світловий сигнал, який розповсюджується волокном, зовсім не чутливий до зовнішніх електромагнітних впливів: розряди блискавок, багаточисленні радіопередавачі, електродвигуни, силові ядерні установки і т.п. Волоконним лініям не потрібне

важке і дороге екранування, вони захищені від завад самим принципом своєї роботи. Традиційна радіоапаратура не тільки піддається впливу електромагнітних хвиль, вона і сама їх випромінює (носієм інформації є заряджена частинка – електрон, нерівномірний рух якого супроводжується випромінюванням електромагнітних хвиль). На цій основі і виник радіошпіонаж. І тільки ВОЛЗ, у яких потайність передачі повідомлень закладена в самій фізиці, дозволяє гарантовано убезпечитись від несанкціонованого доступу до інформації.

Четверте – волоконні кабелі легкі і компактні, в десятки і сотні разів легші і дешевші, ніж металічні. Основу металічних кабелів складають мідь і свинець – не тільки дуже дорогі, але і досить рідкісні в земній корі. Волокно виготовляється з кварцу, а він зі звичайного піску, запаси якого величезні.

Один із апологетів волоконного зв'язку співставляє його по важливості з винайденням парової машини, електричної лампочки, транзистора… .

Контрольні питання

1.Назвіть фізичні принципи, які лежать в основі роботи будь – якого квантового генератора.

2.Як можна пояснити необхідність існування квантового переходу (вимушеного випромінювання), поняття про який вперше ввів А.Ейнштейн?

3.У чому полягає суть ідеї В.Фабриканта щодо можливості підсилення випромінювання, яке проходить через певне середовище?

4.Чи можна здійснити підсилення випромінювання в середовищі, яке знаходитися в рівноважному стані? Відповідь пояснити.

5.Розкажіть, як конкретно реалізуються принципи, які лежать в основі роботи будь – якого квантового генератора, в лазері (на прикладі гелій – неонового лазера).

6.Пояснити трирівневу енергетичну схему створення інверсної населеності, яка реалізована в гелій – неоновому лазері, рубіновому та інших. Чим вона ефективніша від дворівневої схеми?

7.Як підсилюється випромінювання в активному середовищі?

8.Яке призначення мають дзеркала в оптичному генераторі?

9.Які властивості притаманні лазерному випромінюванню? Чому саме тільки лазерному?

10. Застосування ОКГ.

11.Які фізичні основи волоконного зв'язку?

12. Які існують типи оптичних волокон?

13.Які є типи дисперсії в оптичних волокнах і як дисперсія впливає на передачу інформації?

14.Які оптичні волокна є одномодовими, а які багатомодовими?

15.Джерела світла і фотоприймачі в оптоволоконних лініях зв'язку.

16.Який тип лінз використовується в оптоволоконних лініях зв'язку?

Розділ 23.Атомні ядра

§23.1.Атомні ядра та їхні характеристики

Атомне ядро, центральна масивна частина атома, яка складається із протонів і нейтронів, яких називають нуклонами (нуклон—від латинського nucleus—ядро). Маса ядра атома приблизно в $4 \cdot 10^3$ раз більша маси всіх електронів, що входить до складу атома. Розміри атомного ядра становлять $\sim 10^{-14} \div$

$\div 10^{-15}$ м. Електричний заряд ядра позитивний і за абсолютною величиною дорівнює сумі зарядів атомних електронів нейтрального атома. Атомне ядро було відкрите Е. Резерфордом (див.§17.1) в дослідах по розсіюванню α – частинок при проходженні їх через речовину. В кінці 1913 року учень Резерфорда Г.Мозлі показав експериментально, що електричний заряд ядра атома (в одиницях абсолютної величини заряду електрона e) дорівнює порядковому номеру Z елемента в періодичній системі елементів. Після Мослі факт існування атомного ядра утвердився в фізиці остаточно.

В ядерній фізиці прийнято виражати маси в атомних одиницях маси (а.о.м.) Одна а.о.м. дорівнює $1{,}66\cdot10^{-27}$кг. Частіше массу виражають в одиницях енергії, для цього одну а.о.м. помножують на c^2 і виражають в мільйонах електрон-вольт (Мев), де c –швидкість світла в вакуумі. Одній а.о.м відповідає 931,2 МеВ. Маса спокою електрона (в Мев) є:

$$m_e = 9{,}1\cdot10^{-31}\cdot(3\cdot10^{8})^2/\ /(1{,}6\cdot10^{-19})\text{еВ} = 0{,}51\ \text{МеВ}. \qquad (23.1)$$

Протон. Протон (p) має позитивний заряд, що дорівнює заряду електрона, масу спокою $m_p = 1{,}672\cdot10^{-27}\text{кг} = 1836m_e =$ 938,26МеВ. Протон має спін, що дорівнює половині (S=1/2) і власний магнітний момент $\mu_p = +2{,}79\mu_я$, де

$$\mu_я = e\hbar/(2m_p) = 5{,}05\cdot10^{-27}\text{Дж/Тл} \qquad (23.2)$$

–одиниця магнітного момента, яка називається *ядерним магнетоном*, який в 1836 раз менший магнетона Бора $\mu_Б$ (§17.5). Тобто, власний магнітний момент протона приблизно в 660 разів менший, ніж магнітний момент електрона.

Нейтрон. Нейтрон (n) був відкритий у 1932р. англійським фізиком Д.Чедвиком. Електричний заряд цієї частинки дорівнює нулю, а маса

$$m_n = 1{,}675\cdot10^{-27}\text{кг} = \\ = 1839m_e = 939{,}55\ \text{МеВ} \qquad (23.3)$$

дуже близька до маси протона. Різниця мас нейтрона і протона $m_n - m_p$ складає 1,3 МеВ, тобто $2{,}5m_e$. Нейтрон має спін, який дорівнює половині (S=1/2), і, не дивлячись на відсутність електричного заряду, власний магнітний моментом

$$\mu_n = -1{,}91\mu_я,$$

де знак мінус вказує на те, що напрямки власних механічного і магнітного моментів нейтрона і протона протилежні.

У вільному стані нейтрон нестабільний (радіоактивний) – він самовільно розпадається, перетворюючись в протон з випусканням електрона (e^-) і ще однієї частинки, яку називають антинейтрино ($\tilde{\nu}$).

Характеристики атомного ядра

1.Зарядове число Z – дорівнює кількості протонів, які входять в склад ядра і визначає його заряд, який дорівнює $+Ze$. Число Z визначає порядковий номер хімічного елемента в періодичній таблиці. Тому його також називають *атомним номером ядра.*

2. Масове число A – число нуклонів (тобто сумарна кількість протонів і нейтронів) в ядрі. Число нейтронів в ядрі дорівнює $N=A-Z$.

3.Для позначення ядер застосовується символ

$$_ZX^A \quad \text{або} \quad {}^A_ZX,$$

де під X розуміють хімічний символ даного елемента.

Ядра утворюють такі групи:

–ізотопи. Ядра з однаковими Z, але різними A. Більшість хімічних елементів має по

декілька стабільних ізотопів. Так, наприклад, у кисню є три стабільних ізотопи $^{16}_{8}$O, $^{17}_{8}$O, $^{18}_{8}$O, у олова– десять і т.д. Водень має три ізотопи:

$^{1}_{1}$H – *звичайний водень*, або протій (Z=1,N=0),

$^{2}_{1}$H – важкий водень, або дейтерій (Z=1, N=1),

$^{3}_{1}$H –тритій (Z=1, N=2).

Протій і дейтерій стабільні, тритій радіоактивний.

–ізобари. Ядра з однаковим масовим числом A, наприклад, $^{40}_{18}$Ar і $^{40}_{20}$Ca.

–ізотони. Ядра з однаковим числом нейтронів $N=A-Z$, наприклад, $^{13}_{6}$C, $^{14}_{7}$N.

–ізомери. Радіоактивні ядра з однаковими Z і A, але з різним періодом напіврозпаду. Наприклад: існують два ізомери ядра $^{80}_{35}$Ba, у одного із них період напіврозпаду дорівнює 18 хвилин, в іншого – 4,4 години.

Відомо біля 1500 ядер, які відрізняються або Z, або A, або і тим, і іншим. Приблизно 1/5 частина цих ядер стійкі, інші радіоактивні. Багато ядер були отримані штучним шляхом за допомогою ядерних реакцій.

В природі зустрічаються елементи з атомними номерами Z від 1 до 92, крім технеція (Tc, Z=43) і прометія (Pm, Z=61).

Розміри ядра. Експерименти з використанням заряджених частинок, які розсіювались на ядрах, привели до висновку, що ядро можна вважати кулею, радіус якої може бути виражений формулою

$$R = r_0 A^{1/3}, \qquad (23.4)$$

де A–масове число данного ядра, r_0 = (1,3÷1,7)·10^{-15}м = (1,3÷1,7)Ф. (Ф=10^{-15}м – фермі – назва в ядерній фізиці одиниці довжини). В задачах і прикладах беруть компромісне значення $r_0 = 1{,}3$ Ф.

Густину маси ядра можна розрахувати за формулою

$$\rho_{\text{ядра}} = \frac{\text{маса ядра}}{\text{об'єм ядра}},$$

де масса ядра приблизно дорівнює A(а. о. м) , а об'єм дорівнює $\frac{4\pi R^3}{3} = \frac{4\pi\left(r_0 \mathrm{A}^{1/3}\right)^3}{3} = 4\pi r_0^3 A/3$. Тоді густина ядерної речовини $\rho_{\text{ядра}} = 3\text{а. о. м.}/(4\pi r_0^3) \approx 2 \cdot 10^{17}$кг/м^3.

Це неймовірно велика густина ядерної речовини, яка приблизно однакова для всіх ядер, так як не залежить від масового числа A.

§23.2. Ядерні сили

Згідно закону Кулона позитивно заряджені протони, які знаходяться на дуже близьких відстанях в атомному ядрі, сильно відштовхуються один від одного так, що вони повинні були б миттєво розлетітись. Тому не можна пояснити стабільне існування атомного ядра, якщо не зробити припущення, що нуклони в ньому піддаються впливу потужних ядерних сил притягання. Ці ядерні сили є проявом так званної сильної взаємодії.

Основні властивості ядерних сил:

1.Ядерні сили є силами притягання.

2.Ядерні сили досить значні. Тому таку взаємодію називають сильною. Енергія зв'язку, що припадає на один нуклон у ядрі досягає 7–8,5 МеВ. Сили взаємодії ядра з електронами атома забезпечують енергію зв'язку від десятків до тисяч електрон-вольт, а сили зв'язку між атомами в молекулах забезпечують енергію зв'язку в кілька електрон-вольт.

3.Ядерні сили є короткодіючими, на відміну від електромагнітних і гравітаційних сил. Радіус r дії ядерних сил приблизно дорівнює розміру нуклона (10^{-15}м). При $r > 10^{-14}$м ядерні сили практично дорівнюють нулю.

4.Ядерні сили мають властивість зарядової незалежності, тобто величина ядерних сил не залежить від електричного заряду взаємодіючих нуклонів. Ядерні взаємодії між двома протонами або між двома нейтронами, або між протоном і нейтроном однакові.

5.Ядерні сили мають властивість насичення, тобто кожен нуклон в ядрі взаємодіє лише з обмеженою кількістю найближчих до нього нуклонів. Насичення появляється в тому, що питома енергія зв'язку нуклонів у ядрі при збільшенні числа нуклонів не зростає, а залишається приблизно сталою (якщо не враховувати легкі ядра).

Властивості насичення ядерних сил приводять до такого важливого висновку: взаємодія між нуклонами не зводиться тільки до сил притягання. На відстанях між нуклонами $r<0{,}5\cdot 10^{-15}$м притягання між ними переходить у відштовхування. За відсутності у ядрі сил відштовхування між нуклонами ядра повинні б колапсувати, тобто стискатись у точку.

6.Ядерні сили залежать від *орієнтації спінів нуклонів*, які взаємодіють. Система з протона і нейтрона утворює ядро−дейтрон лише у випадку, коли спіни протона і нейтрона паралельні. Якщо ж спіни протилежно напрямлені, то нейтрон і протон ядра не утворюють.

7.Величина ядерних сил також *залежить від взаємної орієнтації спіна та орбітального моменту кожного нуклона.*

8.Ядерні сили *не є центральними силами*, тобто силами, які можна уявити такими, що діють вздовж лінії, яка з'єднує центри взаємодіючих нуклонів.

Для пояснення властивостей ядерних сил – насичення і короткодії – В.Гайзенберг висунув гіпотезу про те, що ядерні сили є «обмінні сили», тобто що ядерні сили між двома нуклонами забезпечуються третьою частинкою.

Довжина хвилі $\lambda_{\text{о.ч.}}$ обмінної частинки повинна відповідати радіусу $r_я$ дії ядерних сил. Виходячи із формули де Бройля (18.5),

$$\lambda_{\text{о.ч.}} \approx \frac{h}{m_{о.ч.}c} = r_я \leq 10^{-14}\text{м}$$

можна оцінити масу обмінної частинки:

$$m_{о.ч.} \approx \frac{h}{r_я c} = 243m_e.$$

Ця маса повинна у 300 разів перевищувати масу електрона.

У 1935р. японський фізик Х.Юкава висунув гіпотезу про те, що ядерні сили зумовлені невідомою на той час частинкою, маса якої становить 200-300 мас електрона. Оскільки ці гіпотетичні частинки за величиною маси займали проміжне місце між масою електрона і протона, то їх назвали мезонами (від грецького mezos – середній, проміжний). Такі частинки дійсно було виявлено у 1947 році в космічному випромінюванні. Виявилось, що існує три типи π – мезонів: π^+, π^-, π^0. Заряди π^+ і π^- за абсолютною величиною дорівнюють заряду електрона. Маса зарядженого мезона $m_{\pi^\pm}=273m_e$, маса π^0−мезона $m_{\pi^0}=264m_e$. Спін π−мезонів дорівнює нулю.

Ядерну взаємодію двох нуклонів, що знаходяться на відстані радіуса дії ядерних сил, можна представити у такий спосіб. Один нуклон випускає π−мезон, а другий нуклон його поглинає упродовж 10^{-23}с (час прольоту частинки з швидкістю світла відстані між нуклонами – ядерний час). Частинки, які існують лише в області дії ядерних сил протягом ядерного часу, називаються *віртуальними*.

Розглянемо обмінну взаємодію між нуклонами. В результаті віртуальних процесів

$$p \Leftrightarrow n + \pi^+, \quad n \Leftrightarrow p + \pi^-,$$

$$p \Leftrightarrow p + \pi^0, \quad n \Leftrightarrow n + \pi^0$$

нуклон виявляється оточеним хмарою віртуальних $\pi-$мезонів, які утворюють поле ядерних сил. Поглинання цих мезонів іншим нуклоном приводить до сильної взаємодії між нуклонами, яка здійснюється за однією з таких схем:

$$1)\ p + n \Leftrightarrow n + \pi^+ + n \Leftrightarrow n + p.$$

Протон випускає віртуальний π^+-мезон і перетворюється в нейтрон. Мезон поглинається нейтроном, який внаслідок цього перетворюється у протон. Потім такий процес відбувається у зворотному напрямку. Кожний із взаємодіючих нуклонів частину часу проводить в зарядженому стані, а частину в нейтральному.

$$2)\ n + p \Leftrightarrow p + \pi^- + p \Leftrightarrow p + n.$$

Нейтрон і протон обмінюються π^- мезонами.

$$3)\ p + p \Leftrightarrow p + \pi^0 + p \Leftrightarrow p + p,$$

$$4)\ n + n \Leftrightarrow n + \pi^0 + n \Leftrightarrow n + n$$

Два протони або два нейтрони обмінюються нейтральними π^0-мезонами. Перша з цих трьох схем обміну π-мезонами нуклонів експериментально підтверджується при вивченні розсіяння нейтронів на протонах. Безперечно, процес взаємодії між нуклонами значно складніший і наведені схеми не що інше, як намагання унаочнити дії обмінних сил. Обмінні сили – це суто квантово-механічне поняття.

§23.3.Енергія зв'язку ядра

Енергія зв'язку ядра $W_{зв}$ – це енергія, яку необхідно затратити, щоб розщепити ядро на окремі нуклони. Маса ядра $m_я$ завжди менша суми мас частинок, які входять в його склад. Це обумовлено тим, що при об'єднанні нуклонів в ядро виділяється енергія зв'язку нуклонів один з одним.

Енергія спокою частинки зв'язана з її масою співвідношенням $W_0 = mc^2$. Значить, енергія ядра, яке знаходиться в стані спокою, менше сумарної енергії нуклонів, які знаходяться в стані спокою і не взаємодіють між собою на величину

$$W_{зв} = \left(Zm_p + (A - Z)m_n - m_я\right)c^2. \quad (23.5)$$

Ця величина і є енергія зв'язку нуклонів в ядрі.

Співвідношення (23.5) практично не порушується, якщо вважати, що $m_p = m_н - m_e$, а $m_я = m_a - Zm_e$, де $m_н$ – маса атома водню, m_a – маса атома, який відповідає даному ядру, m_e – маса електрона. Тому формулу (23.5) можна переписати так:

$$W_{зв} = (Zm_н + (A - Z)m_n - m_a)c^2. \quad (23.6)$$

Формула (23.6) зручна тим, що в таблицях як правило приводяться не маси ядер, а маси атомів.

Величина

$$\Delta m = Zm_н + (A - Z)m_n - m_a \quad (23.7)$$

називається дефектом маси ядра. Якщо дефект маси виразити в а.о.м., то енергія зв'язку знаходиться за формулою:

$$W_{зв} = \Delta m(\text{а. о. м.}) \cdot 931{,}2\ \text{МеВ}. \quad (23.8)$$

Енергія зв'язку, яка приходиться на один нуклон, тобто

$$W_{пит} = W_{зв}/A \quad (23.9)$$

називається *питомою енергією* зв'язку нуклонів в ядрі.

Якщо $W_{зв} > 0$, то ядро стабільне (щоб розщепити його на складові частинки, йому необхідно надати енергію ззовні). Якщо $W_{зв} < 0$, то ядро нестабільне (буде розпадатись само по собі, або, як говорять, спонтанно).

Вирахуємо питому енергію зв'язку ядра $^{16}_{8}O$.

Розв'язок:

$$Zm_н = 8 \cdot 1{,}007825\ \text{а. о. м.} = 8{,}062600\ \text{а. о. м.}$$

$Nm_n = 8 \cdot 1{,}008665$ а. о. м. $= 8{,}069320$ а. о. м.

Маса атома $^{16}_{8}O$ дорівнює 16,000000 а.о.м. і тому дефект маси Δ = +0,131920 а.о.м. Тоді енергія зв'язку $W_{зв}$ = 0,131920 · 931,2 МеВ = =122,84 МеВ, а питома енергія зв'язку $W_{пит}$ = =122,84/18 = 7,68 МеВ/нуклон.

Для порівняння зауважимо, що енергія зв'язку валентних електронів в атомах має величину порядку 10 еВ, що в 10^6 раз менше питомої енергії зв'язку ядер. На рис.23.1 зображений графік, який показує залежність питомої енергії зв'язку $W_{зв}/A$ від масового числа

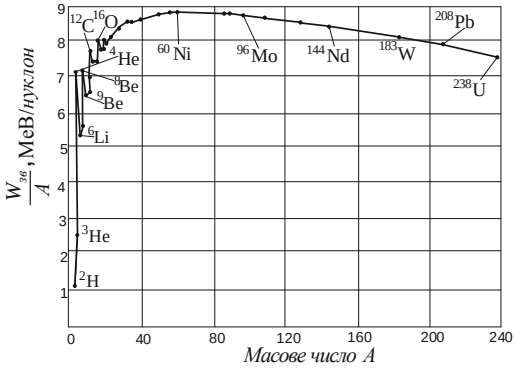


Рис.23.1.

A. Сильніше за все зв'язані нуклони в ядрах з масовими числами порядку 50-60 (тобто для елементів від Cr до Zn). Питома енергія зв'язку для цих ядер досягає 8,7 МеВ/нуклон. За винятком декількох нерегулярностей, які спостерігаються для ядер $^{4}_{2}He$, $^{12}_{6}C$, $^{16}_{8}O$, крива на рис.23.1. порівняно плавна. Вона різко зростає при малих значеннях A і вирівнюється при $A \geq 30$. З ростом A $W_{пит}$ поступово зменшується; для найважчого природного елемента урану вона складає 7,5 МеВ/нуклон. Така залежність $W_{пит}$ від A пояснюється таким чином. При малих A нуклони на поверхні ядра не з усіх сторін оточені іншими нуклонами, що зменшує ефективність взаємодії з сусідніми нуклонами. Виникає так званий поверхневий ефект, подібний до поверхневого натягу в рідинах.

Ще один ефект, який зменшує енергію зв'язку ядер, – це кулонівське електростатичне відштовхування між протонами. Цей ефект особливо проявляється для ядер з великим масовим числом A. Зв'язок між нуклонами стає слабшим, а ядра менш міцними.

Така залежність питомої енергії зв'язку ядра від масового числа робить енергетично можливим два процеси: 1) ділення важких ядер на декілька більш легких ядер і 2) злиття (синтез) легких ядер в одне ядро. Обидва ці процеси повинні супроводжуватись виділенням великої кількості енергії. Так, наприклад, поділ одного ядра з масовим числом A = 240 (питома енергія зв'язку дорівнює 7,5 МеВ/нуклон) на два ядра з масовими числами A = 120 (питома енергія зв'язку дорівнює 8,5МеВ/нуклон) призвів би до виділення енергії в 240 МеВ. Злиття двох ядер важкого водню $^{2}_{1}H$ в ядро гелію $^{4}_{2}He$ призвело би до виділення енергії 24 МеВ. Для порівняння скажемо, що при з'єднанні одного атома вуглецю з двома атомами кисню (згорання вугілля до CO_2) виділяється енергія порядку 5 еВ.

Ядра із значеннями масового числа A від 50 до 60 є енергетично найбільш вигідними. У зв'язку з цим виникає питання: чому ядра з іншими значеннями A є стабільними? Відповідь в наступному. Для того, щоб розділитись на декілька частин, важке ядро повинно пройти через ряд проміжних станів, енергія яких перевищує енергію основного стану ядра. Отже, для процесу ділення ядру необхідна додаткова енергія (*енергія активації*), яка потім повертається назад, додаючись до енергії, яка виділяється при поділі ядра внаслідок зміни енергії зв'язку. У звичайних умовах ядру не має

звідки взяти енергію активації, внаслідок чого важкі ядра не зазнають спонтанного поділу. Енергія активації може бути надана важкому ядру додатковим нейтроном, який захоплюється ним. Процес поділу ядер урану або плутонію під дією нейтронів, які захоплюються ядрами, лежить в основі роботи ядерних реакторів і звичайної атомної бомби.

Щодо легких ядер, то злиття їх в одне ядро можливе, якщо вони підійдуть одне до одного на досить близьку відстань ($\sim 10^{-15}$м = 1Ф). Такому зближенню ядер заважає кулонівське відштовхування між ними. Для того, щоб подолати це відштовхування, ядра повинні рухатись з великими швидкостями, які відповідають температурам порядку декількох сотень мільйонів кельвін. Із-за цієї причини процес синтезу легких ядер називають *термоядерною реакцією*. Термоядерні реакції протікають в надрах Сонця і зірок. В земних умовах поки що були здійснені некеровані термоядерні реакції при вибухах водневих бомб.

§23.4.Моделі атомного ядра

Спроби створення теорії ядра наштовхуються на серйозні труднощі: 1) недостатні знання про сили, які діють між нуклонами, 2) надзвичайна громіздкість квантової задачі багатьох тіл (ядро з масовим числом A – це система із A тіл). Ці труднощі вимушують іти шляхом створення ядерних моделей, які дозволяють описувати за допомогою порівняно простих математичних засобів певну сукупність властивостей ядра. Ні одна із подібних моделей не може дати вичерпне описання ядра, тому приходиться використовувати моделі, кожна із яких описує певні властивості ядра і своє коло явищ. В кожній моделі містяться довільні параметри, значення яких підбираються так, щоб отримати узгодження з експериментом. З багатьох моделей найбільш використовувані дві: краплинна та оболонкова.

Краплинна модель. Ця модель була запропонована Я.І.Френкелем в 1939 р. і потім розвивалась Н.Бором та іншими вченими. Вона ґрунтується на аналогії між поведінкою нуклонів у ядрі та молекул в краплині рідини. В обох випадках сили, що діють між складовими частинами короткодіючі та мають властивість насичення. Рідина має сталу густину, ядерна речовина також характеризується сталою густиною, незалежно від числа нуклонів у ядрі. Об’єм краплини, як і об’єм ядра, пропорційний числу частинок, що в них знаходяться.

При отриманні певної енергії краплина ядерної рідини переходить у збуджений стан. Отримана енергія в результаті зіткнень нуклонів швидко перерозподіляється між ними. Проте можуть виникнути такі умови, коли отримана енергія концентрується на поверхневому нуклоні або групі нуклонів. Якщо ця енергія більша від енергії зв’язку частинки в ядрі, частинка може подолати поверхневі сили ядерного притягання і вийти з ядра. Такий процес аналогічний випаровуванню молекули з поверхні рідини краплини.

Істотна відмінність ядра від краплини рідини в цій моделі в тому, що вона трактує ядро як краплину електрично зарядженої нестискуваної рідини (з такою самою густиною, як ядерна), яка підлягає законам квантової механіки. Краплинна модель ядра дозволила отримати напівемпіричну формулу для енергії зв’язку нуклонів у ядрі, пояснила механізм реакцій поділу ядер. Однак не змогла пояснити, наприклад, підвищену стійкість ядер, які містять магічні числа протонів і нейтронів.

Оболонкова модель ядра (1949-50 р.,

німецький фізик Х.Йенсен, американський фізик М.Гепперт-Майєр) базується на припущенні, що нуклони в ядрі розташовуються на дискретних рівнях (оболонках), заповнюючи їх згідно з принципом Паулі. Стійкість ядер залежить від ступеня заповнення таких оболонок. Вважається, що ядра з повністю заповненими оболонками найбільш стійкі. У відповідності з дослідом особливо стійкими є ядра, у яких число протонів, або число нейтронів (або обидва ці числа) дорівнюють 2, 8, 20, 28, 50, 82, 126. Ці числа отримали назву ***магічних.*** Ядра, у яких число протонів Z або число нейтронів N є магічним (тобто особливо стійкі ядра), теж називаються магічними. Ядра, у яких магічними є і Z, і N, називаються *два рази магічними*.

Два рази магічних ядер відомо всього п'ять:

$$^{4}_{2}\text{He}\ (Z=2, N=2), \qquad ^{16}_{8}\text{O}\ (Z=8, N=8),$$
$$^{40}_{20}\text{Ca}\ (Z=20, N=20),\ ^{48}_{20}\text{Ca}\ (Z=20, N=28),$$
$$^{208}_{82}\text{Pb}\ (Z=82, N=126).$$

Ці ядра особливо стійкі. Зокрема, особлива стійкість ядра гелію $^{4}_{2}\text{He}$ проявляється в тому, що це єдина складна частинка, яка випромінюється важкими ядрами при радіоактивному розпаді (вона називається $\alpha-$частинкою).

Нові експериментальні дані про властивості атомних ядер, які появились в результаті подальших досліджень, виходили за рамки обох теорій. Так виникла узагальнена модель (синтез краплинної і оболонкової) – *оптична модель ядра* (пояснює взаємодію ядер з елементарними частинками та ін.).

§23.5.Радіоактивність. Основний закон радіоактивного перетворення атомних ядер

Радіоактивністю (від лат. radio – випромінюю, radius – промінь і activus – діючий) *називають самовільне і спонтанне перетворення ядер нестійких ізотопів одних елементів у ядра ізотопів інших елементів, що зумовлено внутрішніми причинами та супроводжується* $\alpha-$, $\beta-$, $\gamma-$*випромінюванням, а також інших частинок (нейтронів, протонів).*

Нижче наведено процеси, які характеризують різні типи реакцій ядерних розпадів природно радіоактивних ядер:

$\alpha-$розпад: $^{A}_{Z}\text{X} \rightarrow {}^{A-4}_{Z-2}\text{X} + {}^{4}_{2}\text{He};$

$\beta^{-}-$розпад: $^{A}_{Z}\text{X} \rightarrow {}_{Z+1}{}^{A}\text{X} + {}_{-1}{}^{0}\beta + \bar{\nu};$

$\beta^{+}-$розпад: $^{A}_{Z}\text{X} \rightarrow {}_{Z-1}{}^{A}\text{X} \rightarrow {}^{0}_{1}\beta + \nu;$

$\beta^{-}-$захоплення: $^{A}_{Z}\text{X} + {}_{-1}{}^{0}\beta \rightarrow {}_{Z-1}{}^{A}\text{X} + \nu;$

$\gamma-$розпад: $(^{A}_{Z}\text{X})^{*} \rightarrow {}^{A}_{Z}\text{X} + \gamma.$

Тут ν і $\bar{\nu}$ позначають нейтрино і антинейтрино.

Символом $(^{A}_{Z}\text{X})^{*}$ позначено ядро, яке знаходиться у збудженому стані. Коли воно повертається в свій основний стан $^{A}_{Z}\text{X}$, то випромінює γ – квант.

Радіоактивність, яка спостерігається в ядрах, що існують у природних умовах, називається *природною*. Радіоактивність ядер, які отримані за допомогою ядерних реакцій, називається *штучною*. Між природною та штучною радіоактивністю немає принципової різниці.

Природні радіоактивні перетворення ядер, які відбуваються самочинно, називаються *радіоактивним розпадом*. Ядро, що виникло внаслідок розпаду, називають *дочірним* ядром, а ядро, яке розпалось – *материнським*.

Теорія, яка пояснює це явище, ґрунтується на припущенні, що радіоактивний розпад є спонтанним процесом.

Численні досліди показали, що на швидкість радіоактивного розпаду не впливають ніякі зовнішні зміни температури, тиск, наявність електричних і магнітних полів, вид хімічної сполуки, її агрегатний стан. З цього випливає, що радіоактивний процес відбувається в глибинних частинах атома, тобто в його ядрі. Радіоактивний розпад – це властивість самого атомного ядра, залежить вона тільки від його внутрішнього стану.

Внаслідок динамічності цього процесу природно припустити, що число ядер $|dN|$, які розпадаються за інтервал часу від t до $t+dt$ пропорційне до проміжку часу dt і кількості ядер, які ще не розпались на момент часу t:

$$-dN = \lambda N dt. \qquad (23.10)$$

Тут λ – стала величина, яку називають *сталою розпаду, або радіоактивною сталою*. Знак "–" вказує на те, що загальна кількість радіоактивних ядер під час розпаду зменшується $(dN < 0)$. Стала розпаду

$$\lambda = \frac{\frac{|dN|}{N}}{dt} \qquad (23.11)$$

дорівнює відносному зменшенню кількості ядер, які зазнають розпаду за одиницю часу. Стала λ визначає швидкість радіоактивного розпаду.

Розділивши в рівнянні (23.10) змінні, та, проінтегрувавши, отримаємо:

$$\int_{N_0}^{N} \frac{dN}{N} = -\lambda \int_0^t dt\,; \ln\frac{N}{N_0} = -\lambda t.$$

Звідки:

$$N = N_0 e^{-\lambda t}. \qquad (23.12)$$

Цей вираз виражає *закон радіоактивного розпаду*, згідно з яким кількість ядер, які не розпались, зменшується експоненційно. У цій формулі N_0 – початкова кількість ядер, які не розпались на момент часу $t = 0$, N – кількість ядер, які не розпались на момент часу t.

Кількість ядер, які розпались за час t, визначається виразом

$$N_0 - N = N_0\left(1 - e^{-\lambda t}\right). \qquad (23.13)$$

На рис.23.2 зображено залежність $\ln\frac{N}{N_0}$

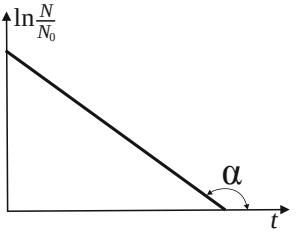


Рис.23.2.

від часу t. Вона дає змогу експериментально за нахилом прямої визначити сталу розпаду λ, тому що tgα = λ.

Середній час життя. Середнім часом життя радіоактивного ядра називають середній час життя ядер в якому-небудь зразку, який містить ці радіоактивні ядра.

Кількість ядер $|dN(t)|$, що розпалась за проміжок часу від t до $t+dt$ визначаються модулем виразу (23.10)

$$|dN(t)| = \lambda N(t) dt.$$

Час життя кожного з цих ядер дорівнює t. Отже, сумарний час життя всіх N_0 ядер, що були спочатку, одержують інтегруванням виразу $tdN(t)$. Поділивши цю суму на число ядер N_0, одержимо середній час життя τ радіоактивного ядра:

$$\tau = \frac{1}{N_0}\int_0^\infty t|dN(t)| = \frac{1}{N_0}\int_0^\infty t\lambda N(t)dt =$$

$$= \frac{1}{N_0}\int_0^\infty t\lambda N_0 e^{-\lambda t}dt = \int_0^\infty t\lambda e^{-\lambda t}dt = \frac{1}{\lambda}\cdot$$

$$\tau = 1/\lambda\,. \qquad (23.14)$$

Стала розпаду λ є оберненою величиною до середнього часу життя τ певного радіоактивного елементу. Тому (23.12) перепишеться так:

$$N = N_0 e^{-\frac{t}{\tau}}, \qquad (23.15)$$

де τ – це час, за який кількість ядер при радіоактивному розпаді зменшиться в e = =2,71828... раз.

Для характеристики стійкості ядер відносно розпаду, для оцінки тривалості життя певного радіоактивного ізотопу вводять поняття періоду піврозпаду $T_{1/2}$.

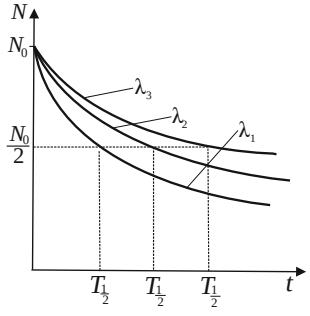


Рис.23.3.

Періодом піврозпаду називається час, протягом якого початкова кількість ядер певної речовини розпадається наполовину.

З визначення $T_{1/2}$ випливає, що

$$\frac{N_0}{2} = N_0 e^{-\lambda T_{1/2}} \quad , \quad e^{-\lambda T_{1/2}} = 2.$$

Звідси отримуємо:

$$T_{1/2} = \frac{\ln 2}{\lambda} = \frac{0{,}693}{\lambda} = 0{,}693\tau. \qquad (23.16)$$

Періоди піврозпаду різних природних ізотопів досить відрізняються один від одного. Так $T_{1/2} = 4{,}51 \cdot 10^9$років для $^{238}_{92}U$ і $T_{1/2} = 3 \cdot 10^{-7}$ с для $^{212}_{84}$Po.

На рис.23.3 наведені графіки кривих розпаду при різних значеннях λ. *Кількість атомів, що розпадається за одну секунду називається активністю елементу.*

Активність A визначається такими формулами:

$$A = \left|\frac{dN}{dt}\right| = \lambda N = \lambda N_0 e^{-\lambda t} = \frac{N \ln 2}{T} \cdot \qquad (23.17)$$

Отже, активність обернено пропорційна до періоду піврозпаду і зменшується з часом за експоненційним законом.

За одиницю активності препарату беруть один розпад за секунду – беккерель (Бк). Часто користуються позасистемною одиницею, яку називають Кюрі (Кі):

$$1\text{Кі} = 3{,}7 \cdot 10^{10}\text{Бк}.$$

($3{,}7 \cdot 10^{10}$ розпадів за 1 секунду відбувається в 1 грамі чистого радію ^{226}Ra).

Якщо відбувається ланцюжок радіоактивних перетворень, то закон розпаду дає змогу робити висновок лише про зменшення з часом материнської речовини. Якщо ж дочірня речовина також радіоактивна, то одночасно відбувається і нагромадження цієї речовини і її зменшення внаслідок розпаду. За час dt розпадеться $\lambda_м N_м dt$ ядер атомів материнської речовини із загальної кількості $N_м$ цих ядер, що були на момент часу t. Якщо з кожного ядра материнської речовини утворюється одне ядро дочірньої, то за проміжок часу dt кількість ядер дочірньої речовини збільшиться на $\lambda_м N_м dt$. Водночас з кількості ядер $N_д$ дочірньої речовини, що були на момент часу t, за проміжок часу dt розпадеться $\lambda_д N_д dt$.

Загальна зміна $dN_д$ кількості ядер дочірньої речовини за одиницю часу дорівнюватиме:

$$\frac{dN_д}{dt} = \lambda_м N_м - \lambda_д N_д.$$

У випадку рухомої рівноваги між материнською і дочірньою речовинами $\frac{dN_д}{dt} = 0$ і $\lambda_м N_м = \lambda_д N_д$ – умова радіоактивної рівноваги. Звідси

$$\frac{N_м}{N_д} = \frac{\lambda_д}{\lambda_м} = \frac{T_{1/2м}}{T_{1/2д}}.$$

Числа атомів обох речовин прямо пропорційні до періодів їх піврозпадів.

Радіоактивний розпад записують у вигляді рівняння

$$^{A_1}_{Z_1}\mathrm{X} \rightarrow {}^{A_2}_{Z_2}\mathrm{Y} + a,$$

де $^{A_1}_{Z_1}\mathrm{X}$ – материнське радіоактивне ядро, його маса m_x, $^{A_2}_{Z_2}\mathrm{Y}$ – дочірнє ядро (продукт розпаду), його маса m_y, a – частинка, яка випускається, її маса m_a.

При радіоактивному розпаді справджується *закон збереження енергії*, згідно з яким повна енергія материнського ядра дорівнює повній енергії продуктів розпаду

$$m_м c^2 = (m_д + m_a)c^2 + W_p,$$

де W_p – повна енергія розпаду, яка дорівнює зміні енергії спокою материнського ядра $m_м$, що перетворюється в кінетичну енергію дочірнього ядра $m_д$, яке вилітає, і енергію $(m_a\ c^2)$ γ –квантів.

При радіоактивному розпаді ядер виконується *закон збереження зарядових чисел*:

$$Z_я = \sum_i Z_i, \qquad (23.19)$$

де $Z_я$ – зарядове число ядра, яке зазнало розпаду, Z_i – зарядові числа ядер і частинок, що виникли внаслідок розпаду.

Крім того, виконується *правило збереження масових чисел*:

$$A_я = \sum_i A_i. \qquad (23.20)$$

Ці співвідношення при радіоактивному розпаді формують у вигляді правил зміщення, які дають змогу встановити, яке ядро виникає внаслідок розпаду певного материнського ядра. При цьому розрізняють правила зміщення для випадків можливого α– або β–розпаду.

Ядра, які виникають внаслідок радіоактивного розпаду, також можуть бути радіоактивними. Послідовність такого процесу приводить до виникнення ланцюжка радіоактивних перетворень, який закінчується стабільним елементом. Сукупність елементів, яка утворює такий ланцюжок, називається радіоактивним сімейством.

Сімейства називають за найбільш довгоживучим «родоначальником»: сімейство торію $^{232}_{90}$Th, нептунію $^{237}_{93}$Np, урану $^{238}_{92}$U і актинідію $^{235}_{89}$Ac. Кінцевими продуктами такого розпаду відповідно є: $^{208}_{82}$Pb, $^{209}_{83}$Bi, $^{206}_{82}$Pb, $^{207}_{82}$Pb, тобто єдине сімейство нептунію (штучно радіоактивного ядра) закінчується нуклідом Bi. Всі інші (природно–радіоактивні ядра) – нуклідами Pb.

Дія γ–випромінювання (а також інших видів іонізуючого випромінювання) на речовину характеризується дозою іонізуючого випромінювання. Розрізняють такі одиниці випромінювання:

Поглинута доза випромінювання (D) – фізична величина, що дорівнює відношенню енергії поглинутого випромінювання до маси опромінюваної речовини. Одиниця поглинутої дози випромінювання – грей (Гр): 1Гр=1Дж/кг – доза випромінювання, при якій опромінюваній речовині масою 1кг передається енергія довільного іонізуючого випромінювання в 1Дж.

Експозиційна доза випромінювання (X) – фізична величина, що дорівнює відношенню суми електричних зарядів всіх іонів одного знаку, створених електронами, звільненими в опромінюваному повітрі (при повному використанні іонізуючої здатності електронів), до маси цього повітря.

Одиниця експозиційної дози випромінювання – кулон, поділений на кілограм Кл/кг, часто користуються позасистемною одиницею – рентген (Р):

$$1\mathrm{P} = 2{,}58 \cdot 10^{-4}\ \text{Кл/кг}.$$

При експозиційній дозі, яка дорівнює одному рентгену, в 1м^3 сухого повітря при нормальному атмосферному тиску виникає сумарний заряд іонів одного знаку, величина яких $0{,}33 \cdot 10^{-3}$Кл.

Еквівалентна доза іонізуючого випромінювання H−це добуток поглинутої дози на середній коефіцієнт якості k іонізуючого випромінювання в даному елементі об'єму біологічної тканини стандартного складу (О−76,2%, С−11,1%, Н−10,1%, N−2,6%). Так як коефіцієнт k безрозмірна величина, то розмірність H співпадає з розмірністю D. $[H]$ 1Дж/кг = 1Зв (зіверт).

Зіверт дорівнює еквівалентній дозі, при якій добуток поглинутої дози в біологічній тканині стандартного складу на середній коефіцієнт якості дорівнює 1Дж/кг.

Еквівалентна доза іонізуючого випромінювання є основною величиною, яка визначає рівень радіаційної небезпеки при хронічному опроміненні людини в малих дозах. Більш широко використовуються такі одиниці вимірювання H як мілізіверт (мЗв), мікрозіверт (мкЗв). При опроміненні всього тіла людини на протязі року допускається еквівалентна доза іонізуючого випромінювання не більше 250 Зв.

Потужність еквівалентної дози іонізуючого випромінювання (потужність еквівалентної дози) $dH/dt=\dot{H}$. $[\dot{H}]$=1Зв/с. Зіверт за секунду дорівнює потужності еквівалентної дози, при якій за 1сек створюється еквівалентна доза в 1 Зв.

Час перебування в полі випромінювання при низьких рівнях іонізуючого випромінювання вимірюється, як правило, годинами (шестигодинний робочий день, 36−годинний робочий тиждень). Тому надають перевагу такій одиниці потужності еквівалентної дози як мікрозіверт за годину (мкЗв/год) незалежно від розміру величини.

Допустима середньорічна потужність еквівалентної дози при опроміненні всього тіла людей, які працюють, дорівнює 28 мкЗв/год при 36−годинному робочому тижні.

Природний фон створює потужність еквівалентної дози в межах 0,05−0,2 мкЗв/год.

Біологічна доза – величина, яка вказує вплив випромінювання на організм. Одиниця біологічної дози – біологічний еквівалент рентгена (**бер**): 1 бер – доза довільного виду іонізуючого випромінювання, яка здійснює таку саму біологічну дію, яку здійснює доза рентгенівського або $\gamma-$ випромінювання в 1Р: (1бер=10^{-2} Дж/кг = 10мЗв).

Потужність дози випромінювання – величина, яка дорівнює відношенню дози випромінювання до часу опромінювання.

§23.6.Ядерні реакції

Ядерною реакцією називається процес сильної взаємодії атомного ядра з елементарною частинкою або з другим ядром, що призводить до перетворення ядра (ядер). Частинки починають взаємодіяти при зближенні їх до відстані 10^{-15}м, при якій починають діяти ядерні сили.

Характер ядерної реакції визначається видом частинки, її енергією, властивостями речовини і умовами взаємодії. Зокрема, зіткнення частинок з ядром викликає різні ядерні реакції, в результаті яких відбувається зміна заряду чи маси частинки або випускаються γ – кванти.

Типовим прикладом ядерної реакції є процес взаємодії швидких α-частинок з ядрами азоту, при якій спостерігається виліт протонів.

$$^{4}_{2}\mathrm{He}+^{14}_{7}\mathrm{N} \rightarrow {}^{18}_{9}\mathrm{F} \rightarrow {}^{1}_{1}\mathrm{H}+^{17}_{8}\mathrm{O}.$$

Вперше ядерну реакцію такого типу здійснив Резерфорд в 1919 р. Кількість відомих нині ядерних реакцій сягає уже тисячі.

За механізмом взаємодії ядерні реакції можна поділити на два види:

–прямі ядерні реакції;

–реакції з утворенням проміжного ядра.

Прямі ядерні реакції відбуваються при дуже високих енергіях частинок і за дуже короткий час, необхідний для того, щоб частинка пролетіла через ядро ($10^{-23} \div 10^{-21}$с). Наприклад, швидкий протон може вибити з поверхні ядра один з нуклонів і полетіти разом з ним.

Більшість же ядерних реакцій з кінетичною енергією частинок меншою 10МеВ проходять з утворенням проміжного ядра. Така ядерна взаємодія відбувається у два етапи. На першому етапі відбувається захоплення частинки ядром і виникнення проміжного ядра, яке знаходиться у збудженому стані. Енергія збудження W^* складається з кінетичної енергії частинки W_k і енергії зв'язку нуклона $W_{зв}$, який приєднався:

$$W^* = W_k + W_{зв}.$$

Енергія зв'язку нуклона в середньому дорівнює 8 МеВ. Тому складове ядро отримує досить велику енергію збудження.

В ядерній взаємодії беруть участь лише ближні до частинки нуклони, тому до них переходить основна частина енергії збудження. Енергія, отримана такими нуклонами, передається потім за участю ядерних сил сусіднім і поступово статистично розподіляється між всіма нуклонами ядра. При цьому відбувається значний розігрів ядра. Енергія збудження, розподілена рівномірно між всіма нуклонами, недостатня для подолання енергії зв'язку нуклона. Однак з часом флуктуації збудженого ядра приводять до того, що в одному нуклоні зосереджується енергія, достатня для відриву його від ядра ($W_{нук} > W_{зв}$). Тоді настає другий етап ядерної реакції – викидання нуклона з ядра.

При захопленні нейтрона з утворенням складового ядра швидкість нейтрона, який вилетів, звичайно менша, ніж швидкість захопленого первинного нейтрона. Оскільки первинні і вторинні нейтрони не можна розрізнити, таке явище сприймається як дещо сповільнене розсіяння нейтрона з втратою частини його енергії. Такий процес має назву *непружного (резонансного) розсіяння частинок*.

Проміжок часу, який потрібний нуклону з енергією порядку 1 МеВ ($v \sim 10^7$ м/с) для того, щоб пройти відстань, яка дорівнює діаметру ядра ($\sim 10^{-14}$м), називається ядерним часом (або ядерним часом прольоту $\tau_я$):

$$\tau_я \cong \frac{10^{-14}м}{10^7 м/с} = 10^{-21}\text{с}.$$

Середній час життя складового ядра ($10^{-1} \div 10^{-2}$с) на багато порядків перевищує ядерний час $\tau_я$. Отже, розпад проміжного ядра є процес, який не залежить від першого етапу реакції, що полягає в захопленні частинки (проміжне ядро якби забуває спосіб свого утворення). Одне і те саме ядро може розпадатися різними шляхами, причому характер цих шляхів і їх відносна ймовірність не залежить від способу утворення проміжного ядра.

Під час ядерної реакції зберігається кількість нуклонів і сумарний заряд. Має місце лише перерозподіл нуклонів і заряду між ядрами та частинками. Збереження заряду і кількості нуклонів дає змогу якісно визначити можливі напрямки перебігу ядерної реакції.

Ядерні реакції супроводжуються зміною кінетичної енергії взаємодіючих частинок. *Для ядерних реакцій виконуються закони збереження енергії і імпульсу*.

§ 23.7.Поділ ядер

В експериментах, які були проведені Енріко Фермі, Еміліо Сегре та їхніми співробітниками в 1934 р., ядра урану бомбувались нейтронами. Опромінений нейтронами уран проявляв при цьому штучну $\beta-$радіоактивність з декількома невеликими періодами піврозпаду. Оскільки на той час було відомо, що ядра урану розпадаються з випромінюванням $\alpha-$частинок з дуже великим періодом піврозпаду, то Фермі і його співробітники запропонували, що в них в урані утворились трансуранові елементи (з Z > 92). Це припущення, однак, виявилось неправильним. Отто Ган і Фріц Штрассман незабаром показали за допомогою дуже ретельного хімічного аналізу, що одним із елементів, які утворюються після бомбування урану нейтронами, був ізотоп барію $^{141}_{56}$Ba. І тоді в1939 р. Отто Фріш і Ліза Мейтнер висунули важливу гіпотезу, що ядра урану після бомбування повільними (тепловими) нейтронами зазнають реакції поділу, в результаті якої одне ядро урану розщеплюється на два ядра приблизно однакової маси, але менших розмірів.

Ці експерименти стали можливими внаслідок відкриття Фермі і співробітників, суть якого в тому, що повільні (теплові) нейтрони отримуються шляхом гальмування нейтронних пучків за допомогою пропускання їх через матеріали багаті воднем, наприклад, через воду або парафін.

Реакція поділу ядер урану схематично може бути представлена рівнянням:

$$^{235}_{92}\mathrm{U} + ^{1}_{0}n \rightarrow ^{236}_{92}\mathrm{U}^{*} \rightarrow \mathrm{X} + \mathrm{Y} + \text{нейтрони}, \quad (3.21)$$

де $^{1}_{0}n$ – повільний (тепловий) нейтрон, $^{236}_{92}\mathrm{U}^{*}$ – надзвичайно нестабільний ізотоп урану, X і У – ядра-фрагменти, які отримуються при поділі ядра урану. Ядра-фрагменти в реакції поділу можуть бути, звичайно, різні, як це в дійсності і проявляється в експерименті. Типовими реакціями ядерного ділення є такі реакції:

$$^{235}_{92}\mathrm{U}+^{1}_{0}n\rightarrow ^{236}_{92}\mathrm{U}^{*}\rightarrow \rightarrow ^{144}_{56}\mathrm{Ba}+^{89}_{36}\mathrm{Kr}+3^{1}_{0}n+Q, \quad (23.22)$$

$$^{235}_{92}\mathrm{U}+^{1}_{0}n\rightarrow ^{236}_{92}\mathrm{U}^{*}\rightarrow \rightarrow ^{140}_{54}\mathrm{Xe}+^{94}_{38}\mathrm{Sr}+2^{1}_{0}n+Q. \quad (23.23)$$

(тут Q – енергія, яка виділяється при реакції).

$\beta-$активність, яку спостерігали Фермі і його співробітники, відбувалась від радіоактивних розпадів ядер-фрагментів, наприклад, від ядерних реакцій:

$$^{140}_{54}\mathrm{Xe}\xrightarrow{\beta^{-}}{}^{140}_{55}\mathrm{Cs}\xrightarrow{\beta^{-}}{}^{140}_{56}\mathrm{Ba}\xrightarrow{\beta^{-}}{}^{140}_{57}\mathrm{La}\xrightarrow{\beta^{-}}{}^{140}_{58}\mathrm{Ce},$$

$$^{94}_{38}\mathrm{Sr}\xrightarrow{\beta^{-}}{}^{94}_{39}\mathrm{Y}\xrightarrow{\beta^{-}}{}^{94}_{40}\mathrm{Zr} \quad (23.24)$$

для реакції (23.23).

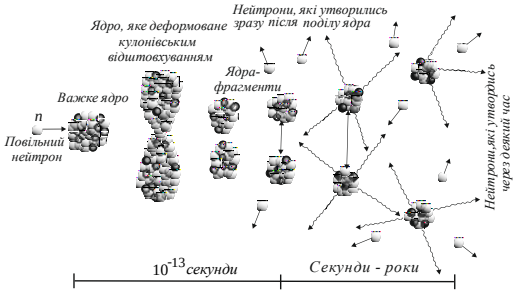


Рис.23.4.

Коли ядро $^{235}_{92}$U захоплює тепловий нейтрон, енергія захопленого нейтрона розподіляється між нуклонами, які входять в склад ядра. Потім, подібно рідкій краплині, ядро починає сильно деформуватись до тих пір, поки кулонівське відштовхування його протонів не розірве це ядро на два ядра-фрагменти, як це показано на рис.23.4. Розподіл

ядер−фрагментів, які отримуються при поділі ядра урану, що захопило нейтрон, за масовими числами показано на рис.23.5.

Всі ядра−фрагменти, які утворюються при поділі ядра урану, мають надлишкове число нейтронів. Із-за цього надлишку нейтронів вони є сильно нестабільними. Збуджені ядра--фрагменти майже зразу випромінюють нейтрони на протязі дуже короткого проміжку часу порядку 10^{-13}с. Крім цих зразу випущених нейтронів, є однак, інша група *затриманих за часом* нейтронів, які випромінюються ядрами-фрагментами з періодом піврозпаду порядку декількох секунд (рис.23.4). Ці затримані за часом нейтрони дуже важливі для здійснення контролю за ядерними реакціями, які відбуваються в атомному реакторі.

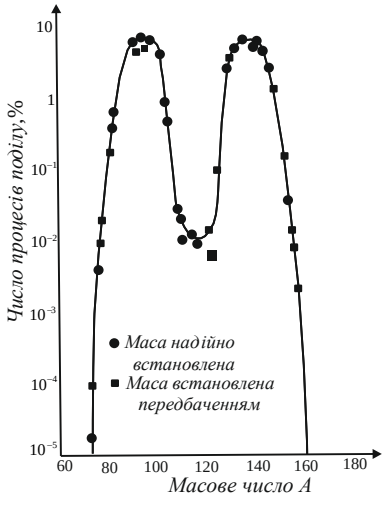


Рис.23.5.Крива виходу продуктів на один поділ. При поділі на теплових нейтронах утворюються переважно ядра-фрагменти зі співвідношенням мас 2:3. Найімовірнішими продуктами поділу з виходом ≈ 6,5 % є ядра-фрагменти з масовими числами 95 і 139. Поділ на однакові ядра-фрагменти є малоймовірними (≈ 0,01 %).

При ядерній реакції поділу виділяється досить велика кількість енергії. Ще трохи енергії виділяється при радіоактивному розпаді ядер-фрагментів.

Енергія ядерного поділу вивільняється у формі кінетичної енергії ядер-фрагментів і у вигляді кінетичної енергії електронів, фотонів, нейтрино і нейтронів, які випускаються ядрами-фрагментами.

Оцінимо порядок величини енергії, яка виділяється при реакції ядерного поділу (23.22) і (23.23). В цих реакціях нестабільне ядро $^{235}_{92}$U розщеплюється в кінці кінців на ядра-фрагменти $^{140}_{58}$Ce, $^{94}_{40}$Zr, 6β−частинок і 2 нейтрони. Маси до і після поділу наступні.

До поділу:	Після поділу:
235,0439а.о.м.($^{235}_{92}$U)	139,9054а.о.м.($^{140}_{58}$Ce)
1,0087 а.о.м.($^{1}_{0}n$)	93,9036а.о.м.($^{94}_{40}$Zr)
$\sum$ 236,0526 а. о. м.	2,0173 а.о.м.($2^{1}_{0}n$)
	0,0033а.о.м.($6\beta^{-}$)
	$\sum$ 235,8296 а. о. м.

Дефект маси реакції $\Delta m = (236{,}0526 - 235{,}8296$ а.о.м. $= 0{,}223$а.о.м. Енергетичний ефект реакції знайдемо за формулою (23.8):

$$Q = 0{,}223 \cdot 931{,}3 \text{ МеВ} = 208 \text{Мев}.$$

На один нуклон, який приймає участь в реакції приходиться $\frac{208}{236} = 0{,}88$ Мев/нуклон.

Для порівняння зазначимо, що енергія, яка виділяється при α−розпаді одного ядра, порядку 5МеВ, а енергія, яка виділяється в одному акті в хімічній реакції згорання, порядку 4 еВ.

Нейтрони, які отримуються в реакції поділу, можна знову використовувати для здійснення нових актів реакції поділу ядер і при цьому буде звільнятись все більше і більше

енергії. В «атомному котлі», який побудував Фермі в Чикагському університеті, була здійснена вперше самопідтримуюча ланцюгова ядерна реакція. Вона почалася 2 грудня 1942 р.

Тепловими нейтронами діляться ще ядра плутонія $^{239}_{94}\mathrm{Pu}$, урану $^{233}_{92}\mathrm{U}$, технецію $^{230}_{90}\mathrm{Th}$. Два останні ізотопи в природі не зустрічаються, їх отримують штучно. Одна із схем реакції поділу ядер плутонію наступна:

$$^{239}_{94}\mathrm{Pu}+^{1}_{0}n\rightarrow \rightarrow ^{140}_{54}\mathrm{Xe}+^{96}_{40}\mathrm{Zr}+4^{1}_{0}n+188\text{МеВ}. \quad (23.25)$$

Зазначимо, що в цій реакції в одному акті поділу виділяється чотири нейтрони, що приблизно в два рази більше, ніж в реакції поділу (23.23).

Реакцію поділу атомних ядер можна викликати, бомбуючи ядра не тільки нейтронами, але й іншими частинками, наприклад, протонами, α-частинками і γ−квантами. В реакції поділу ядер можуть приймати участь, крім урану, плутонію, також і більш легкі елементи. Типовий приклад такої реакції поділу – реакція поділу ядер міді, яка викликається протонами з енергією 60 МеВ, прискорених на циклотроні:

$$^{63}_{29}\mathrm{Cu}+^{1}_{1}p\rightarrow ^{38}_{17}\mathrm{Cl}+^{25}_{13}\mathrm{Al}+^{1}_{0}n. \quad (23.26)$$

§23.8.Ланцюгова реакція поділу. Ядерний реактор

Для практичного застосування поділу важких ядер найважливіше значення має виділення великої енергії при кожному акті поділу і поява при цьому 2−3 нейтронів. Кожен з цих нейтронів взаємодіє з сусідніми ядрами речовини, що, у свою чергу, спричиняє в них реакцію поділу, тобто відбувається лавиноподібне зростання кількості актів поділу. *Така реакція поділу називається ланцюговою.*

Схематично ланцюгова реакція поділу ядер зображена на рис.23.6. Ланцюгова реакція поділу характеризується *коефіцієнтом розмноження k нейтронів, який дорівнює відношенню кількості нейтронів у даному поколінні до їх кількості в попередньому поколінні.*

Виявляється, що не всі вторинні нейтрони, які утворюються, викликають наступний поділ ядер, що призводить до зменшення *k*. По-перше, внаслідок скінчених розмірів активної зони (простір, де відбувається ланцюгова реакція) і великої проникної здатності нейтронів, частина з них покине активну зону раніше, ніж буде захоплена яким-небудь ядром. По-друге, частина нейтронів захоплюється ядрами домішок, що не діляться, і які завжди присутні в активній зоні.

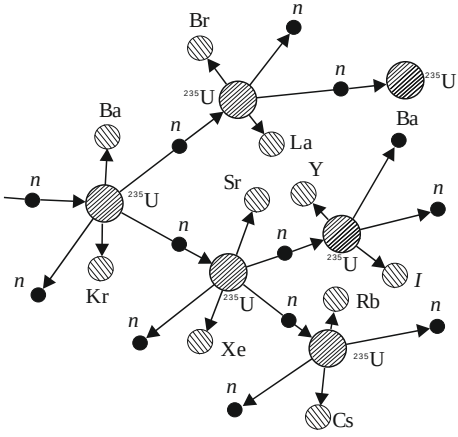


Рис.23.6.

Коефіцієнт розмноження залежить від природи речовини, що ділиться, а для даного ізотопу – від його кількості, а також від розмірів і форми активної зони.

Мінімальні розміри активної зони, при яких можливе здійснення ланцюгової реакції, називають *критичними розмірами*. Мінімальна маса речовини, що ділиться, яка необхідна для здійснення ланцюгової реакції, називається

критичною масою. Для зменшення втрат нейтронів і критичних параметрів речовини її оточують відбивачем – шаром неподільної речовини, яка має малий ефективний поперечний переріз для захоплення нейтронів і великий переріз для їх розсіяння.

Обчислимо швидкість розвитку ланцюгової реакції. Нехай τ – середній час життя одного покоління нейтронів , а N – число нейтронів у попередньому поколінні. В даному поколінні їх кількість дорівнює kN, тобто приріст кількості нейтронів за одне покоління $dN = kN - N = N(k-1)$. Приріст кількості нейтронів за одиницю часу, тобто швидкість наростання ланцюгової реакції

$$\frac{dN}{dt} = \frac{N(k-1)}{\tau}, \quad \frac{dN}{N} = \frac{k-1}{\tau}dt.$$

Інтегруючи цей вираз, отримуємо

$$N = N_0 e^{(k-1)\frac{t}{\tau}},$$

де N_0 – кількість нейтронів в початковий момент часу, а N– їх кількість в момент часу t. При $k>1$ йде наростаюча реакція, кількість поділів безперервно росте, і реакція може стати вибуховою.

При $k=1$ йде самопідтримуюча реакція, при якій кількість нейтронів з часом не змінюється. При $k<1$ йде згасаюча реакція.

Ланцюгові реакції діляться на *керовані* і *некеровані*. Вибух атомної бомби є некерованою реакцією. При масі, близькій до критичної, нейтрони швидко розмножуються і реакція набуває вибухового характеру. На цьому ґрунтується дія атомної бомби. Ядерним зарядом такої бомби є два або більше шматків майже чистого ^{235}U або ^{239}Pu. Маса кожного шматка менша критичної внаслідок чого ланцюгова реакція не виникає.

В земній атмосфері завжди є деяка кількість нейтронів, народжених космічними променями. Тому, щоб викликати вибух, досить з'єднати шматки ядерного заряду в один шматок з масою, яка більша критичної. Це має виконуватись дуже швидко і з'єднання шматків повинно бути дуже щільним. В противному випадку ядерний заряд розлетиться на частини перш, ніж встигне прореагувати значна частина речовини, яка ділиться. Для з'єднання шматків використовують звичайну вибухову речовину. Ланцюгова реакція в атомній бомбі відбувається на швидких нейтронах. При вибуху встигає прореагувати тільки частина ядерного заряду.

Керовані ланцюгові реакції здійснюються в ядерних реакторах. Як сировинні і подільні речовини в реакторах використовують $^{235}_{92}$U, $^{239}_{94}$Pu, $^{238}_{92}$U а також $^{232}_{90}$Th. У природній суміші ізотопів урану ізотопу $^{238}_{92}$U у 140 разів більше, ніж ізотопу $^{235}_{92}$U.

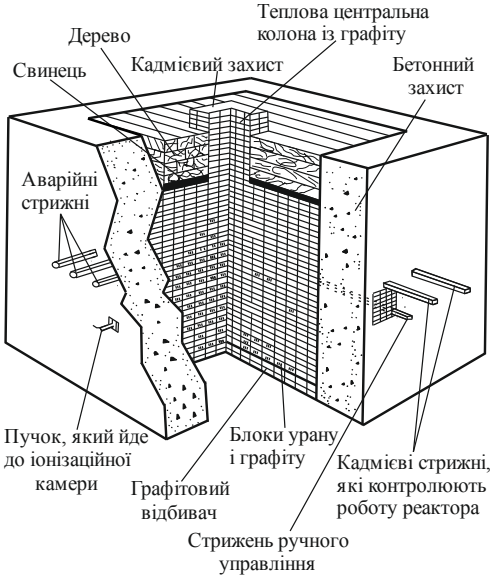


Рис.23.7.

Велике значення в ядерній енергетиці має не лише здійснення ланцюгової реакції поділу, але і керування нею. *Пристрій, в якому*

здійснюється і підтримується керована ланцюгова реакція поділу, називається ядерним реактором.

В першому ядерному реакторі, який схематично зображений на рис.23.7, містив 52 тонни урану у вигляді блоків, які були перекладені приблизно 472 тоннами графіту, який виконував роль сповільнювача нейтронів. В реакторі було 5 контролюючих стрижнів (покритих кадмієм−дуже в'язким матеріалом). Для забезпечення безпеки було ще три стрижні з постійно прикладеним до них механічним навантаженням, яке заштовхувало їх в реактор (у випадку, якщо під час аварії відкаже електромережа).

Речовиною, яка ділиться в реакторі, є природний (або дещо збагачений ізотопом ^{235}U) уран. Щоб перешкодити радіаційному захопленню нейтронів ядрами ^{238}U (яке стає особливо інтенсивним при енергії нейтронів ~ 7МеВ), порівняно невеликі блоки урану розміщують на деякій відстані один від одного, а проміжки між блоками заповнюють сповільнювачем, тобто речовиною, в якій нейтрони сповільнюються до теплових швидкостей. Такі теплові нейтрони ефективно взаємодіють з ядрами ^{235}U.

Сповільнення нейтронів здійснюється внаслідок пружного розсіювання. В цьому випадку енергія, яка втрачається частинкою, що сповільнюється, залежить від співвідношення мас частинок, які стикаються. Максимальна кількість енергії втрачається у випадку, якщо обидві частинки мають однакову масу. З цієї точки зору ідеальним сповільнювачем повинна бути речовина, яка містить звичайний водень, наприклад, вода (маси протона і нейтрона приблизно однакові). Ядра сповільнювача повинні сильно пружно розсіювати нейтрони. Цій умові задовільняють дейтрон (ядро важкого водню – дейтерію D), а також ядра графіту (C) і берилію (Be). Для зменшення енергії нейтрона від 2 МеВ до теплових енергій (~0,03 Мев) у важкій воді (D_2O) достатньо біля 25 зіткнень, в C і Be – приблизно 100 зіткнень.

Тепловиділяючі елементи (твели) – це блоки з радіоактивного матеріалу, що знаходиться в герметичній оболонці, яка слабо поглинає нейтрони. За рахунок енергії, що виділяється при поділі ядер, твели розігріваються і їх поміщають в потік теплоносія. Проходячи через активну зону, теплоносій у вигляді газу, води або розплавленого металу нагрівається і передає теплоту через спеціальний пристрій робочому тілу, наприклад, воді в паротурбогенераторі, а потім знову поступає в активну зону реактора. Активна зона оточується тепловідбивачем, що зменшує витік нейтронів.

Керування ланцюговою реакцією здійснюється спеціальними керуючими стрижнями з матеріалів, що сильно поглинають нейтрони (наприклад, кадмій). При повністю вставлених стрижнях реакція не йде. При поступовому вийманні стрижнів k росте і при певному положенні доходить до одиниці. В цей момент реактор починає працювати. В міру його роботи кількість матеріалу, який ділиться в активній зоні, зменшується і відбувається її забруднення фрагментами поділу. Щоб реакція не припинилась, з активної зони за допомогою автоматичного пристрою поступово виймаються керуючі стрижні. В реакторі є аварійні стрижні, введення яких при збільшенні інтенсивності реакції зразу її припиняє. Кожний реактор має біологічний захист – систему екранів із захисних матеріалів.

§23.9.Реакції синтезу легких ядер

Сонце втрачає величезну кількість енергії, але з часом помітно не охолоджується. Цей факт був загадковим для фізиків до тих пір, поки вони не відкрили існування ядерної реакції. В 1938р. Ханс Бете висунув гіпотезу, що за величезну кількість енергії, яка вивільняється на Сонці і в інших зорях, відповідальні ядерні реакції, в яких два легких ядра об'єднуються або зливаються і утворюють одне більш важке ядро. Такі реакції називаються ядерними реакціями синтезу.

Як видно з рис.23.1, об'єднання двох легких ядер, які належать до швидко зростаючої ділянки цієї кривої, в одне більш важке ядро дає ядро з більшою енергією зв'язку в розрахунку на один нуклон, ніж в ядер, які об'єднуються. Оскільки отримується більша питома енергія зв'язку, сумарна маса вихідних ядер, які зливаються, повинна зменшуватись і тому енергетичний ефект реакції повинен бути позитивним і в реакціях синтезу енергія дійсно буде вивільнятися.

Оскільки для синтезу ядер необхідні дуже високі температури (§23.3), цей *процес називається термоядерною реакцією*.

Для прикладу наведемо деякі реакції синтезу легких ядер та значення теплових ефектів Q:

$$^2_1\mathrm{H}+^2_1\mathrm{H}=^3_1\mathrm{H}+^1_1\mathrm{H}, \quad (Q=4{,}0\text{Мев}),$$

$$^2_1\mathrm{H}+^2_1\mathrm{H}=^3_2\mathrm{H}+^1_0n, \quad (Q=3{,}3\text{Мев}),$$

$$^2_1\mathrm{H}+^3_1\mathrm{H}=^4_2\mathrm{He}+^1_0n, \quad (Q=17{,}6\text{Мев}),$$

$$^6_3\mathrm{Li}+^2_1\mathrm{H}=^4_2\mathrm{He}+^4_2\mathrm{He}. (Q=22{,}4\text{Мев}).$$

Виділення великих кількостей енергії пояснюється тим, що питома енергія зв'язку ядер різко збільшується при переході від ядер водню ($^2_1\mathrm{H}, ^3_1\mathrm{H}$) до літію $^6_3\mathrm{Li}$ і особливо до гелію $^4_2\mathrm{He}$.

Реакції синтезу атомних ядер мають ту особливість, що в них енергія, яка виділяється на один нуклон, значно більша, ніж в реакціях поділу важких ядер. Справді, якщо під час поділу ядра $^{235}_{92}\mathrm{U}$ виділяється енергія приблизно 0,84 МеВ на один кулон, то в реакції синтезу приблизно 3,5 МеВ на один нуклон, тобто в 4 рази більше.

Оцінимо температуру перебігу термоядерної реакції на прикладі синтезу ядра дейтерію $^2_1\mathrm{H}$. Для об'єднання ядер дейтерію їх треба наблизити на відстань, що дорівнює радіусу дії ядерних сил $r = 2\cdot10^{-15}$м, долаючи при цьому кулонівський бар'єр $e^2/(4\pi\varepsilon_0 r) \approx 0{,}7$Мев. На кожне ядро, що стикається, припадає 0,35 МеВ. Середній енергії теплового руху $3kT/2$, яка дорівнює 0,35МеВ, відповідає температура порядку $2\cdot10^9$ К. Отже, реакція синтезу ядер дейтерію може відбуватися лише при температурі, що перевищує на два порядки температуру центральних ділянок Сонця ($1{,}3\cdot10^7$ К).

Однак для перебігу реакцій синтезу атомних ядер достатньо температур порядку 10^7 К. Це пов'язано з двома факторами: 1) при температурах, характерних для реакцій синтезу атомних ядер, речовина знаходиться в стані плазми, розподіл частинок за швидкостями якої описується законом Максвелла, тому завжди є деяка кількість ядер, енергія яких набагато перевищує середнє значення; 2) синтез ядер може відбуватися внаслідок тунельного ефекту.

При високих температурах і тисках, які існують всередині Сонця і зірок, молекули речовини дисоційовані на атоми, а високо іонізовані атоми утворюють гарячу й активну плазму. В цих умовах кулонівське відштовхування ядер може бути легко подолане

і ядра можуть зливатись одне з одним. В результаті вивільняється величезна енергія.

У зв'язку з тим, що водень – найбільш поширений елемент у Всесвіті, Бете запропонував, що в зірках здійснюється так званий «вуглецевий цикл». Вважається, що це найбільш важливий ланцюжок ядерних реакцій, при якому енергія виділяється внаслідок реакції злиття ядер:

$$
\begin{aligned}
&а)\ {}^{1}_{1}\mathrm{H}+{}^{12}_{6}\mathrm{C} \rightarrow {}^{13}_{7}\mathrm{N}+\gamma;\\
&б)\ {}^{13}_{7}\mathrm{N} \rightarrow {}^{13}_{6}\mathrm{C}+{}_{+1}^{0}e+\nu;\\
&в)\ {}^{1}_{1}\mathrm{H}+{}^{13}_{6}\mathrm{C} \rightarrow {}^{14}_{7}\mathrm{N}+\gamma;\\
&г)\ {}^{1}_{1}\mathrm{H}+{}^{14}_{7}\mathrm{N} \rightarrow {}^{15}_{8}\mathrm{O}+\gamma; \qquad (23.27)\\
&д)\ {}^{15}_{8}\mathrm{O} \rightarrow {}^{15}_{7}\mathrm{N}+{}_{+1}^{0}e+\nu;\\
&е)\ {}^{1}_{1}\mathrm{H}+{}^{15}_{7}\mathrm{N} \rightarrow {}^{12}_{6}\mathrm{C}+{}^{4}_{2}\mathrm{He}.
\end{aligned}
$$

Розглядаючи уважно цей ланцюжок реакцій, переконуємось, що ядро вуглецю виступає як каталізатор; з нього ланцюжок реакцій починається, ним же він і закінчується. В приведеному ланцюжку реакцій витрачається чотири ядра водню (протона – ${}^{1}_{1}\mathrm{H}$) і утворюється 3γ–кванта, 2 нейтрино (ν), 2 позитрона (${}_{+1}^{0}e$) і одне ядро гелію (${}^{4}_{2}\mathrm{He}$). Так що по суті у вуглецевому циклі йде реакція

$$4\,{}^{1}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He}+2\,{}_{+1}^{0}e+2\nu+Q.$$

Вирахуємо різницю мас Δm для цієї реакції:

До реакції:	Після реакції:
4,031300 а.о.м.($4\,{}^{1}_{1}\mathrm{H}$)	4,002603 а.о.м.(${}^{4}_{2}\mathrm{He}$)
	0,001098а.о.м($2\,{}_{+1}^{0}e$)
	$\sum$ 4,003701а.о.м.

Тоді Δm = (4,031300- 4,003701)а. о. м. = = 0,0276 а. о. м., що дає тепловий ефект реакції

Q = 0,0276а.о.м·931,2Мев/а.о.м.=25,7Мев.

Цю енергію виносять із собою γ–кванти, нейтрони і ядро гелію, тобто α–частинки.

Ще в одному циклі ядерних реакцій, яка ймовірно існує, приймають участь в основному протони, і тому він називається «протонно-протонним циклом». В ньому відбувається наступний ланцюжок ядерних реакцій:

$$
\begin{aligned}
&а)\ {}^{1}_{1}\mathrm{H}+{}^{1}_{1}\mathrm{H} \rightarrow {}^{2}_{1}\mathrm{H}+{}_{+1}^{0}e+\nu;\\
&б)\ {}^{2}_{1}\mathrm{H}+{}^{1}_{1}\mathrm{H} \rightarrow {}^{3}_{2}\mathrm{He}+\gamma; \qquad (23.28)\\
&в)\ {}^{3}_{2}\mathrm{He}+{}^{3}_{2}\mathrm{He} \rightarrow {}^{4}_{2}\mathrm{He}+2\,{}^{1}_{1}\mathrm{H}.
\end{aligned}
$$

В цьому ланцюжку реакції *а*) і *б*) повинні відбуватись по два рази, щоб отримались два ядра ${}^{3}_{2}\mathrm{He}$, які потрібні для реакції *в*). В результаті всіх цих реакцій із чотирьох ядер водню (протонів) утворюється одна α–частинка, 2 позитрони і 2γ–кванти. Енергія, яка виділяється в цьому циклі, дорівнює 26,2 МеВ.

Вперше штучна термоядерна реакція була здійснена в СРСР (1953), а потім в США у вигляді вибуху водневої (термоядерної бомби), що є некерованою реакцією. Вибуховою речовиною, в якій проходила реакція ${}^{3}_{1}\mathrm{H}+{}^{2}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He}+{}^{1}_{0}n$, є суміш дейтерію і тритію, а запалом – звичайна атомна бомба, під час вибуху якої виникає необхідна для перебігу термоядерної реакції температура.

Теоретичною основою для перебігу штучних керованих термоядерних реакцій є реакції типу ${}^{2}_{1}\mathrm{H}+{}^{3}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He}+{}^{1}_{0}n$, що відбуваються у високотемпературній плазмі. Однак завдання полягає не тільки в створенні умов, потрібних для інтенсивного виділення енергії в термоядерних процесах, а здебільшого

в підтриманні цих умов. Для здійснення термоядерної реакції, яка сама себе підтримує, потрібно, щоб швидкість виділення енергії в системі, де відбувається реакція, була не менша за швидкість відведення енергії від системи.

Розрахунки показують, що для забезпечення самопідтримної керованої термоядерної реакції температуру дейтерієвої плазми треба довести до кількох сотень мільйонів градусів. При температурах ~ 10^8К реакція має повну інтенсивність і супроводжується виділенням великої енергії. Так, при температурі $1{,}17\cdot10^8$К потужність, яка виділяється в одиниці об'єму плазми при злитті дейтерієвих ядер, становить 3кВт/м^3, у той час як при температурі $1{,}17\cdot10^6$ К вона дорівнює лише 10^{-17}кВт/м^3.

Основною причиною втрат енергії високотемпературною плазмою є її велика теплопровідність, яка швидко зростає ($\sim T^{5/2}$) при високих температурах. Енергія з плазми може відводитися завдяки дифузії гарячих частинок з області, де відбувається реакція, на стінки апарата, в якому міститься плазма. Якщо плазму не теплоізолювати від контакту з будь-якими навколишніми речовинами, то її можна нагрівати лише до кількох сотень тисяч градусів, тому що енергія, яка виділяється внаслідок реакції синтезу, виходитиме на стінки. Інакше кажучи, треба втримати плазму в заданому об'ємі, не допускаючи її розширення. Ідею ефективної магнітної термоізоляції плазми стосовно проблеми керованого термоядерного синтезу запропонували А.Д.Сахаров і І.Є.Тамм у 1950 р. Якщо пропустити через плазму у формі стовпа вздовж його осі великий електричний струм, то магнітне поле цього струму створює електродинамічні сили, які прагнутимуть стиснути плазмовий стовп. Отже, стовп плазми буде відірваний від стінок і стягнутий у плазмовий шнур. Однак утримувати плазмовий шнур в такому стані не вдається: відбуваються швидкі радіальні коливання – він то розширюється, то знову стискається. Внаслідок нестабільності, нестійкості плазми у плазмовому шнурі виникають деформації, які змінюють геометричну форму шнура. Результатом цього є порушення термоізоляції, інтенсивна взаємодія плазми зі стінками, що приводить до забруднення дейтерію речовиною стінок і швидкого охолодження плазми.

Основним питанням, розв'язання якого дозволить здійснити керовані термоядерні реакції, є з'ясування умов, за яких високотемпературна плазма в магнітному полі належної конфігурації може зберегти стійкість. Розв'язання цього питання поряд з пошуками шляхів підвищення температури плазми є головним напрямом, в якому розвиваються дослідження керованих термоядерних реакцій.

Можливість реалізації термоядерної реакції зводиться до необхідності виконання двох вимог: наявності деякої мінімальної температури і певного обмеження для добутку концентрації n частинок у плазмі на час τ їх утримання в плазмі. Ці умови називаються *критерієм Лоусона:*

$$\begin{cases} n\tau > 10^{14}\,\dfrac{\text{с}}{\text{см}^3}, T > 10^8 K, \\ {}^2_1\text{H} + {}^3_1\text{H} \rightarrow {}^4_2\text{He} + {}^1_0 n; \end{cases}$$

$$\begin{cases} n\tau > 10^{15}\,\dfrac{\text{с}}{\text{см}^3}, \quad T > 10^9 K, \\ {}^2_1\text{H} + {}^2_1\text{H} \rightarrow {}^3_1\text{H} + {}^1_1\text{H}. \end{cases}$$

Контрольні питання

1. Сучасні уявлення про будову атомного ядра.

2.Що таке дефект маси ядра?

3.Що таке енергія зв’язку ядра? Як вона практично визначається?

4.Накресліть приблизний графік залежності питомої енергії зв’язку від масового числа ядра. Який висновок можна зробити з такої залежності питомої енергії зв’язку від масового числа?

5.Які існують види радіоактивності?

6.Закон радіоактивного розпаду. Які є характеристики радіоактивного розпаду?

7.Які існують дози радіоактивного випромінювання?

8.Дати визначення активності радіоактивної речовини. В яких фізичних одиницях вимірюється активність радіоактивної речовини?

9.Отримання енергії при поділі важких ядер. Принцип дії атомної бомби.

10.Отримання енергії при синтезі легких ядер. Принцип дії водневої (термоядерної) бомби.

11.Отримання енергії в ядерних реакторах на атомних електростанціях.

Предметний показчик

А
Абсолютне чорне тіло 284
Активність елемента 490
Автоколивання 174
Акцептори 396
Анізотропія штучна 264
Аналізатори 253
Антиферомагнетик 121
Аперіодичний процес 164
Апертура:
-інтерференції 222
-світловода 475
Атом водню 311
Атомна модель 305
Атомна одиниця маси 482

Б
База 444
Бар'єр:
-висота 438
-потенціальний 353,436
- Шотткі 453
Батарея сонячна 451
Бер 492
Биття 158
Білінза Бі'є 216
Біпризма Френеля 215
Боровський радіус 315

В
Варікап 446
Вектор:
-електричного зміщення (індукції) 38
-намагнічування 100
-Пойтинга 201
-поляризації 36
-Умова 182
Вестона елемент 69,70
Взаємна енергія струмів 140
Видність 214
Випромінювальна здатність 283
Випромінювання
- вимушене 459
-гальмівне 331
-елементарного диполя 204
-рівноважне 282
-спонтанне 457
-теплове 281
-характеристичне 332
Висота звуку 190
Вітер електричний 50
Вісь симетрії 259
Власна енергія струму 138
Власна частота 155
Власний напівпровідник 394
Внутрішній фотоефект 399
Вольт-амперна характеристика 412,421,443
Вузол електричного кола 71

Г
Гальванічний елемент 68
Генератор квантовий 456
Генератор оптичний 461,463
Герца досліди 201
Гіпотеза Планка 290
Гіромагнітне відношення 102
Гістерезис:
- діелектричний 44
- магнітний 114
Головна площина 253,260
Головний максимум 245
Градани 479
Група хвиль 183
Групова швидкість 184,273
Густина енергії випромінювання 283
Густина енергії:
-об'ємна 283
-спектральна 283
Густина зарядів:
-зв'язаних 37
-лінійна 21
-об'ємна 22
-поверхнева 20
Густина потоку енергії 182
Густина струму 61
Гучність звуку 191

Д
Декремент згасання логарифмічний 163
Дефект маси 485
Джерело струму 66
Диполь:
-електричний 31
-магнітний 91
Дисперсійна область 247
Дисперсія:
-аномальна 272
-відносна 272

-кутова 246
-лінійна 246
-міжмодова 477
-нормальна 272
-речовини 272
-світла 270
-середня 272
-хвиль 183
-хроматична 477
Дифракційна решітка 244
Дифракція 227
Дифракція:
- Френеля 228, 230
- Фраунгофера 228, 240
Дифузія 407
Дифузія носіїв заряду 407
Дифузійна:
-довжина 411
-швидкість 411
Дихроїзм 253,264
Діаграма випромінювання диполя 207
Діамагнетизм 107
Діелектрики 29,30, 376
Діелектрики:
-іонні 33,35
-неполярні 34
-полярні 34
Дзеркала Френеля 215
Дірка (носій заряду) 336
Ділянка кола:
-неоднорідна 67
-однорідна 64
Добротність коливальної системи 163
Довгохвильова межа фотоефекту 298, 401
Довжина вільного пробігу 408
Довжина когерентності 220
Додавання взаємно перпендикулярних коливань 158
Додавання однаково напрямлених коливань 157
Доза іонізуючого випромінювання:
-біологічна 492
-випромінювання 491
-еквівалентна 491
-поглинута 491
- експозиційна 490
Домен 34
Домени:
- електричні 41, 414
-магнітні 119,120
Домішковий напівпровідник 395
Досліди:
-Герца 202
-Девіссона і Джермера 337
-Квінке 357
-Резерфорда 305,306,307
-Столєтова 112, 295
-Фарадея та Верде 270
-Франка і Герца 313
-Штерна і Герлаха 323
Дрейф носіїв заряду 405
Дуалізм корпускулярно-хвильовий 335

Е

Еквіпотенційні поверхні 26
Екстраструм 135
Електрична стала 13
Електрети 46
Електричний:
-вітер 50
-диполь 31
Електричний заряд 11
Електризація через вплив 48
Електроємність відокремленого провідника 52
Електрон 12
Електрона заряд питомий 12
Електрони провідності 375,381,392,394
Електронний газ:
-вироджений 381
-не вироджений 381
Електрорушійна сила 66
Елемент струму 75
Елементарна комірка 362
Електропровідність:
-металів 384-388
-питома 65
-темнова 401
Емісія:
-автоелектронна 360
-термоелектронна 418
- холодна 360
Емітер 444
Енергетична світність 283
Енергетичні зони 373
Енергія:
-активації 395,486
-взаємна 140
-власна струму 138
-електричного поля 58
-електромагнітної хвилі 200

-зарядженого провідника 58
-зв'язку електрона 309
-зв’язку ядра 485
-конденсатора 58
-магнітного поля 139
-"нульова" 383
-пружної хвилі 181
-струму 137
-Фермі 360,379,381,382
Ефект:
- Ганна 411- 416
- Доплера 193,278-280
- Зеємана 322
- Керра 265
- Комптона 303
- Коттона-Мутона 265, 267
-обернений п'єзоелектричний 46
-п'єзоелектричний 45
-піроелектричний 41
-Поккельса 267
- тунельний 359
- Шотткі 423, 453
- Фарадея 269
-Фредерікса 370
- Холла 97
Ефективна:
-густина станів 391
- маса 379

Є

Ємність: -бар'єрна 445
- електрична 52
-конденсаторів: плоского 54, сферичного 56, циліндричного 55
- кулі 53
Ємність *p-n* переходу 445

З

З’єднання конденсаторів:
-паралельне 56
- послідовне 57
Закон:
- Ампера 76
- Біо-Савара-Лапласа 78
- Богуславського-Ленгмюра 422
- Брюстера 256
-Бугера-Ламберта-Бера 275
-взаємозв’язку маси і енергії 485
- Джоуля-Ленца 65
- діючих мас 393
-замкнутого кола 68
- збереження електричних зарядів 12
- зміщення Віна 288
- Кірхгофа 284-285
- Кулона 13
-Кюрі 111
-Ленца 127
- Малюса 254
- Мозлі 333
- Ома 62 - 63
- Ома для густини струму 64
- Ома в диференціальній формі для неоднорідної ділянки кола 67
- Ома в диференціальній формі для однорідної ділянки кола 64
- Ома для однорідної ділянки кола 63
- Ома для неоднорідної ділянки кола 68
-Ома для замкненого кола 68
-повного струму 86,147
-радіоактивного розпаду 489
-Релея 277
-Стефана-Больцмана 286
-Фарадея 129
-фотоефекту 296
Заряд:
-зв’язаний 37
-питомий 98
-просторовий 421
Заселеність:
-енергетичних рівнів 457
-інверсна 461
Захист електростатичний 50
Звук 189:
-гучність 191
-інтенсивність 190
-тиск 190
-тональний 190
-швидкість 192
Зв’язок:
-гомеополярний (ковалентний) 365
-іонний 364
-металічний 366
Здатність інтегральна випромінююча 283, 286
Здатність поглинальна 284
Зіверт 492
Зміщення:
- електричне 38

- зворотне 441
- пряме 442
Зона:
- валентна 375
- заборонена 375
- провідності 375
Зони:
- енергетичні 373,377
- Френеля 230
Зонна пластинка 233

І
Ізобари 483
Ізомери 483
Ізотони 483
Ізотопи 483
Імпеданс 172
Інверсна населеність 461
Інверсний шар 430
Індуктивність 134
Індуктивність взаємна 136
Індукція:
-електростатична 48
-магнітна 207
-магнітного поля 78,80
Інтенсивність хвилі 189,201
Інтерференційний максимум, мінімум 185,212
Інтерференція:
- Жамена 225
- Майкельсона 226
- світла 210
- хвиль 185

К
Кандела 450
Квазімонохроматичне світло 219
Квазістаціонарні процеси 154
Квантування Бора-Вільсона-Зоммерфельда 317
Квантові числа 327
Квантовий:
- генератор 461
- підсилювач 461
Кільця Ньютона 218
Кірхофа правила 71
Класична теорія електропровідності металів 384
Когерентність 186
Коефіцієнт:
-дисперсії 272
-екстинкції 277
-дифузії 408
-згасання 160
-поглинання 461
-питомого обертання 268
-потужності 173
Коливання 148:
-гармонічні 152
-вимушені 161
-електромагнітні 205
-згасаючі 157
Кольори тонких плівок 217
Конденсатори 53:
-з'єднання паралельне 56
-з'єднання послідовне 57
-плоскиі 54
-сферичниі 55
-циліндричні 55
Контакт двох:
- металів 424
- напівпровідників 432
Контактна різниця потенціалів 425
Контур:
-зі струмом в магнітному полі 90
-ідеалізований 155
-коливальний 151
Крива намагнічування 112
Кристали 362
Критична маса, розміри 496
Куперовська пара 390
Кут Брюстера 256
Кут повороту площини поляризації 268
Кюрі точка 43,115

Л
Лазер 461
Лазери:
- газові 469
- на неодимовому склі 468
- напівпровідникові 471
- рубіновий 467
Ларморова прецесія 108
Лінії:
-індукції магнітного поля 80
-напруженості електричного поля 15
Лінійна густина заряду 21
Лінійчастий спектр поглинання 276

Логарифмічний декремент
згасання 163

М
Магічні числа 488
Магнетон Бора 103, 321
Магнетрон 99
Магнітострикція 120
Магнітна:
-напруженість поля 104
-стала 73
Мазер 456
Маса фотона 298
Масове число 482
Маятник пружний 153
Мезони 484
Метали 371,376
Метод зон Френеля
Моди в світловоді 475
Модель ядра:
-краплинна 487
-оболонкова 487
Момент диполя 31
Момент механічний власний 324-325
Момент орбітальний магнітний 102,321
Монохроматична хвиля 199

Н
Надпровідність 388
Намагніченість 100
Намагнічування:
-самодовільне 119
-спонтанне 119
Напруга 67
Напруженість:
-електричного поля 14
-магнітного поля 104
-поля як градієнт потенціалу 27
-поля поблизу зарядженого провідника 49
-поля точкового заряду 14
Напруженість пробою 441
Напівпровідники 372, 376:
-акцепторні 395-396
-донорні 395
Напруга діюча 173
Напрям електричного струму 59
Населеність інверсна 460
Насичення магнітне 112
Нейтрон 482
Носії заряду:
-неосновні 397
-нерівноважні 400
-основні 397
Нуклон 481

О
Об'ємна густина енергії:
-електричного поля 58
-магнітного поля 139
-теплового випромінювання 282
Опір:
-випромінювання диполя 208
-від'ємний 175
-ємнісний 172
-індуктивний 172
-критичний 165
-питомий 63- 64
-провідника 63
Оптична:
-активність 267
-вісь 259
-густина 209
-довжина шляху 212
-накачка 467
-різниця ходу 212

П
Парамагнетизм 110
Перехід електронно-дірковий 432 - 447
Період:
-дифракційної решітки 244
-коливань гармоніяних 156
-коливань згасаючих 162
-півроспаду 490
Петля гістерезису 114
Питома:
-енергії зв'язку 485
-теплова потужність 65
Площина коливань 251
Площина поляризації хвилі 200
Повне внутрішнє відбиття 474
Поверхні еквіпотенціальні 26
Поглинальна здатність 284
Поглинання світла 277
Подвійне променезаломлення 258
Поле електричне в діелектрику 37
Поле електричне в речовині 36, 427
Поле електричне:
-вихрове 132,133
-зарядженого циліндра 21

-зарядженої кулі 22
-двох площин заряджених різнойменно 20,21
-однорідне 16
-потенціальне 24
-сфери 22
Поле магнітне 73,77
Поляризатори 253,263
Поляризаційні призми 262
Поляризація:
-діелектриків 33
-електронна 33
-залишкова 44
-іонна 33,35
-орієнтаційна 33,35
-самодовільна 40, 41
-світла 250,252
-спонтанна 40, 43
Поляриметрія 268
Полярoїди 264, 268
Порядок інтерференції 212
Постулати Бора 313
Постулати Ейнштейна 297, 456
Потенціал:
-поля системи зарядів 25
-точкового заряду 25
Потік вектора:
-електричного зміщення 38
-магнітної індукції 87, 148
-напруженості електричного поля 16,17
Потік енергії хвилі 182
Потокозчеплення 129, 135
Потужність дози випромінювання 492
Правила Кірхгофа 71
Правило:
-"лівої руки" 84
-"правої руки" 85
Принцип:
-відповідності 299
-Гюйгенса-Френеля 228
-доповнюваності 339
-невизначеності 341
-Паулі 328
-причинності 349
Принцип суперпозиції:
-електричних полів 15
-магнітних полів 79
-хвиль 183, 185
Провідник відокремлений 52
Провідники в електричному полі 47
Провідність:
- власна 395
-діркова 397
-домішкова 395
-питома 64
-електронна 397
-напівпровідників 393
Промені рентгенівські 330
Промінь:
-звичайний 259
-незвичайний 259
Проникність відносна:
- діелектрична 38
- магнітна 105
Просвітлення оптики 224
Просторове квантування 321
Просторова когерентність 221
Протон 482
Пуассона рівняння 28
Пьєзоелектрики 45

Р

Радіоактивність 488
Радіус Бора 315
Реактанс 173
Реактор ядерний 497
Реакції:
-поділу ланцюгові 496
-термоядерна 487,499
-ядерні 494
-ядерного поділу 495
Резонанс 167
Резонанс :
-напруги 168
-параметричний 170
-струму 168
Резонатор оптичний 463
Рентген (Р) 491
Рівень:
- гучності 191
-Фермі в напівпровідниках 393
Рівні:
-акцепторні 395
-донорні 396
Рівняння:
- Ейнштейна для фотоефекту 297
-Максвелла 148,149
- Пуасонна 28
- хвильове 180
-Шредінгера 348,349

Рідкі кристали 367-369
Різниця:
-потенціалів 25
-ходу хвиль геометрична 186,212
Робота:
-виходу електронів 360,419, 420
-сил електростатичного поля 24
-сторонніх сил над зарядами 66-67
Роздільна здатність 247
Розміри критичні 497
Розпад:
-альфа 488
- бета 488
-радіоактивний 490
Розподіл Больцмана 380
Розсіяння:
-комптонівське 303
-молекулярне 278
-світла 277
Рухливіть носіїв заряду 64,406

С

Самоіндукція 133
Світло 208:
-квазімонохроматичне 219
Світловолокно:
-градієнтне 474
-ступеневе 474
Світлодіоди 447-450
Світло поляризоване:
-еліптично 250-252
-лінійно 250-252
-по колу 250-252
-плоско (лінійно) 250-252
-частково 250-252
Сегнетоелектрики 41
Селективність 450
Середній час життя 489
Середовище:
-активне 463
-пружне 176
Серія: -Бальмера 311
-Брекета 311
-Лаймана 311
-Пашена 311
-Пфунта 311
-Хемфрі 311
Селфоки 479
Сила:
-Ван дер Ваальса 369
-затримуюча 114
-коерцитивна 44,114
-консервативна 24
-Лоренца 95
-струму 61
Сили:
- сторонні 66
-ядерні 483
Скін-ефект 142
Смуги однакового:
-нахилу 217
-товщини 218
Соленоїд 83,88
Сонячні батареї 451,452
Спектр:
-рентгенівський 330,331
-характеристичний 332
Співвідношення:
-Ейнштейна 409
-невизначеностей Гайзенберга 339
Спін електрона 102,324
Спіраль Корню 239
Сприйнятливість:
-діелектрична 38
-магнітна 105
Стала:
-Керра 266
-Рідберга 310
-радіоактивного розпаду 489
-Холла 97
Стан системи:
-нерівноважний 460,461
-рівноважний 457
Статистика:
- класична 380
-Фермі-Дірака 379
Струм:
-вихровий 140
-генераційний 442
-діючий 173
-дифузійний 60
-електричний 59
-ефективний 173
-зворотний 441
-змінний 172
-зміщення 146
-індукційний 141
-конвекційний 60
-насичення 423
-повний 147
-пороговий 471

-постійний 61
-провідності 59
-термоелектронний 422,423
Ступінь поляризації 252

Т
Тверде тіло 361, 362
Температура:
-Дебая 388
-кольорова 294
-критична 389
-радіаційна 293
-яскравісна 293
Температурний коефіцієнт опору 63, 288
Теорема:
-єдиності в електростатиці 51
-Лармора 108
-Остроградського-Гаусса для електростатичного поля 19
- Остроградського-Гаусса для магнітного поля 87
-про циркуляцію вектора напруженості магнітного поля 104
Тиск світла 301
Тіло сіре 284
Томсона формула 156
Тороїда магнітне поле 89
Тунельний ефект 359

У
Умова вектор 182
Умови максимумів та мінімумів:
-дифракційної решітки 245
-інтерференції 212,213

Ф
Фарада 53
Ферити 117,121
Феромагнетизм 112
Фон 191
Фонони 388
Формула:
-Бальмера 310
-Больцмана 380
-Вульфа-Брега 250
-де-Бройля 336
-Лоренца 96
-Коші 271
-Планка 290
-Релея-Джінса 289
-Річардсона-Дешмана 423
-Томсона 156
Фотодіоди: 447
-*APD* 478
-*p-i-n* 478
Фотоефект:
-вентильний 399
-внутрішній 399
-зовнішній 294
Фотон 298-301
Фронт хвилі 177
Функції власні 349
Функція:
-видності 213
-густини станів 379
-Фермі-Дірака 379
-хвильова 347

Х
Характеристика спектральна 403
Хвилі:
-біжучі 179
-де Бройля 336
-електромагнітні 195
-довжина 178
- інтенсивність 201
-звукові 189
-когерентні 186
-некогерентні 186
-плоскі 179
-поздовжні 177
-поперечні 177
-стоячі 187
-сферичні 180
Хвильова поверхня 179
Хвильове рівняння 180
Хвильовий опір 156
Хвильовий пакет 183
Холодна емісія 360

Ц
Циркулярція вектора:
-індукції магнітного поля 85
-напруженості електростатичного поля 26
-напруженості магнітного поля 104
Цуг хвильовий 203

Ч
Час:
-життя нерівноважних

неосновних носіїв заряду 404
-життя середній 489
-когерентності 220
-релаксації 163
-ядерного прольоту 493
Частинки віртуальні 484
Частота:
-власна 156
-основна 189
-резонансна 166,167
-хвилі 178
Числа квантові 327
Число:
-зарядове 482
-масове 482
-хвильове 178
Червона межа 298, 401

Ш

Шар:
-запірний 440
-збіднений 439
-інверсний 430
Швидкість:
- групова 184,273
-дифузійна 411
-дрейфова 61,406
- звуку 192
-середня квадратична 405
-світла 198
-фазова 180
-хвилі 180
Ширина:
-дифракційного максимуму 243
-інтерференційної смуги 213
-*p-n* переходу 439
Шкала електромагнітних хвиль 195,196

Я

Явище:
-електромагнітної індукції 124
-взаємної індукції 136
Ядерна модель атома 308
Ядерні сили 483
Ядро атома 482
Яма потенційна 351
Яскравість 467

Список використаної літератури

1. Калашніков С.Г. Електрика. Переклад з другого російського видання. - Київ: «Радянська школа», 1964. - 630 с.
2. Савельев И.В. Курс общей физики. Т.2. Электричество и магнетизм. Волны. Оптика - Москва: Наука, 1978. - 480 с.
3. Савельев И.В. Курс общей физики. Т.3. Квантовая оптика. Атомная физика. Физика твердого тела. Физика атомного ядра и элементарных частиц. - Москва: Наука, 1979. - 304 с.
4. Лопатинський І.Є., Зачек І.Р., Ільчук Г.А., Романишин Б.М. Фізика. Підручник. - Львів: Афіша, 2009. - 386 с.
5. Парселл Э. Берклеевский курс физики. Электричество и магнетизм: Учебное руководство: Пер. с англ./Под ред.А.Н.Шальникова и А.О.Вайсенберга – 3-е изд., испр. - Москва: Наука. Гл.ред. физ. -мат. лит., 1983.- 416 с.
6. Кучерук І.М., Горбачук І.Т. Загальна фізика. Електрика і магнетизм: Навч. посібник/ за заг. ред.В.Й.Сугакова - Київ: Вища школа, 1990. - 367 с.
7. Кучерук І.М., Дущенко В.П. Загальна фізика. Оптика. Квантова фізика: Навч. посібник. - Київ: вища школа, 1991.- 463 с.
8. Бутиков Е.Н. Оптика: Учеб. пособие для вузов/Под ред, Н.Н.Калитеевского - Москва: Высш. шк., 1986. - 512 с.
9. Сивухин Д.В. Электричество. - Москва: Наука. Главн. редак. физ.-мат.лит. 1983.- 687 с.
10. Сивухин Д.В. Оптика. - Москва: Наука. Главн. редак. физ.- мат. лит.1980. - 751 с.
11. Акоста В., Кован К., Грем Б. Основы современной физики. Пер. с англ. В.В.Толмачева, В.Ф.Трифонова; под ред. А.Н.Матвеева, - Москва: Просвещение, 1981. - 495 с.
12. Левинштейн М.Е., Симин Г.С. Барьеры. - Москва: Наука. Гл. ред. физ. - мат.лит., 1987. - 320 с.
13. Шпольский Э.В. Атомная физика. Т.1, Введение в атомную физику. - Москва: Наука. Гл.ред. физ.- мат.лит.,1974. - 575 с.
14. Блохинцев Д.И. Основы квантовой механики. - Москва, Ленинград.: Государственное издательство технико-теоретической литературы, 1949. - 588 с.
15. Герасимов С.М., М.В.Белоус С.М., Москалюк В.А. Физические основы электронной техники: учебное пособие для вузов. - Киев: Вища школа. Головное издательство. 1981. - 386 с.
16. Шалимова К.В. Физика полупроводников. Учебник. Изд. 2-е перераб. и доп. - Москва: «Энергия», 1976. - 416 с.

17. Василевский А.М., Кропоткин М.А., Тихонов В.В. Оптическая электроника. - Ленинград: Энерго-атом - издат. Ленинградское отделение, 1990. - 176 с.
18. Ландсберг Г.С. Оптика: Учеб. пособие для вузов. - Москва: Наука. Гл.ред. физ.- мат.лит.1976. - 926 с.
19. Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. Электродинамика. Перевод с англ. А.В.Ефремова, Г.Н.Копьелова, Ю.А.Симонова. Под редак. Я.А.Смородинского – 2- ое издание. -Москва.: Мир. 1977. – 347 с.
20. Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. Электричество и магнетизм. Перевод с англ. А.В.Ефремова, Г.Н.Копьелова, Ю.А.Симонова. Под редак. Я.А.Смородинского – 2 - ое издание.-Москва: Мир, 1977.– 300 с.
21. Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. Физика сплошных сред. Перевод с англ. А.В.Ефремова, Г.Н.Копелова, Ю.А.Симонова. Под редак. Я.А.Смородинского – 2-ое издание.- Москва: Мир, 1977. – 288 с.
22. Воловик П. М. Фізика: для університетів. – К.; Ірпінь: Перун, 2005. - 864 с.
23. Гауэр Дж. Оптические системы связи: Пер.с англ. – Москва: Радио и связь. 1989. – 504 с.
24. Иванов А.Б. Волоконная оптика: компоненты, системы передачи, измерения. – Москва: Компания САЙРУС СИСТЕМ. 1999. - 671 с.
25. Лукіянець Б.А., Понеділок Г.В., Рудавський Ю.К. Планетарна модель атома. Квантування Бора - Вільсона - Зоммерфельда. Світ фізики (наук.популяр.журнал, №4. 2009. 3 – 14 с.).
26. Филонович С.Р. Кавендиш, Кулон и электростатика. – Москва: Знание. 1988. – 64 с. - (Новое в жизни, науке, технике. сер.«Физика»; №8).
27. Смородинский Я.А. Родственники фотона. – Москва: Знание, 1986. – 64 с. – (Новое в жизни, науке, технике. сер. «Физика»; №11).
28. Мигдаль А.Б. Квантовая физика и Нильс Бор. - Москва: Знание. 1987. – 64с. – (Новое в жизни, науке, технике. сер.«Физика»; №3).
29. Франк – Каменецкий Д.А. Электрическое сопротивление – квантовое явление. – Москва: Наука. Глав.редак. физ.-мат. лит. 1984. (Научн.попул. физ.-мат. журнал "Квант", №12, 1984. 2 - 9 с.).
30. Эдельман В. Металлы – Москва: Наука. Глав.редак. физ.-мат. лит. 1981. (Научн. попул.физ.-мат.журнал "Квант", №5, 1981. 5 - 13 с.).