

WOJSKOWA AKADEMIA TECHNICZNA

Antoni Rogalski

**PODSTAWY FIZYKI
DLA ELEKTRONIKÓW**

WARSZAWA 2002

SPIS TREŚCI

PRZEDMOWA	9
Rozdział 1. WPROWADZENIE	13
1.1. Czym jest fizyka?	13
1.2. Wstęp matematyczny	14
1.2.1. Pochodna funkcji	14
1.2.2. Rachunek całkowy	16
1.2.3. Liczby zespolone	18
1.2.4. Działania na wektorach	20
1.2.5. Analiza wektorowa	25
1.2.6. Prawdopodobieństwa. Wartości średnie	28
Rozdział 2. PRAWA ZACHOWANIA	30
2.1. Zachowanie pędu	30
2.2. Zachowanie momentu pędu	33
2.3. Zachowanie energii	35
Rozdział 3. PODSTAWY MECHANIKI RELATYWISTYCZNEJ	39
3.1. Niezmienniczość Galileusza	39
3.2. Postulaty teorii względności	42
3.3. Transformacje Lorentza	44
3.4. Konsekwencje transformacji Lorentza	46
3.4.1. Dodawanie prędkości	46
3.4.2. Skrócenie długości	47
3.4.3. Wydłużenie przedziałów czasowych	48
3.5. Mechanika relatywistyczna	49
3.5.1. Masa i pęd	49
3.5.2. Definicja siły	52
3.5.3. Relatywistyczna energia kinetyczna	52
3.5.4. Energia całkowita	54
3.6. Grawitacja a ogólna teoria względności	56
Rozdział 4. ELEKTROSTATYKA	58
4.1. Ładunek elektryczny	58
4.2. Prawo Coulomba	59
4.3. Pole elektryczne	60
4.4. Strumień pola	61
4.5. Prawo Gaussa	62
4.6. Niektóre zastosowania twierdzenia Gaussa	63
4.6.1. Równomiernie naładowana kula	63
4.6.2. Powierzchniowy rozkład ładunku	64
4.6.3. Liniowy rozkład ładunku	66
4.6.4. Powierzchnia przewodnika	66
4.7. Potencjał elektryczny	67
4.8. Pojemność elektryczna	70
4.9. Gęstość energii pola elektrycznego	71
4.10. Dielektryki	72
4.11. Twierdzenie Gaussa w przypadku obecności dielektryków. Wektor indukcji elektrycznej	74

Rozdział 5. PRĄD ELEKTRYCZNY I SIŁA MAGNETYCZNA	78
5.1. Prąd elektryczny	78
5.2. Prawo Ohma	79
5.3. Ciepło Joule'a	81
5.4. Obwody prądu stałego	81
5.4.1. Prawa Kirchhoffa	82
5.5. Siła magnetyczna	84
5.6. Indukcja magnetyczna	88
Rozdział 6. POLE MAGNETYCZNE	91
6.1. Prawo Ampere'a	91
6.2. Strumień magnetyczny	93
6.3. Prawo Biot-Savarta-Laplace'a	94
6.4. Pole magnetyczne niektórych konfiguracji prądów	95
6.4.1. Prostoliniowy przewodnik z prądem	95
6.4.2. Solenoid	96
6.4.3. Prąd kołowy	97
6.5. Oddziaływanie przewodników z prądem	98
6.6. Efekt Halla	100
6.7. Magnetyzm	101
Rozdział 7. INDUKCJA ELEKTROMAGNETYCZNA	105
7.1. Prawo Faradaya	105
7.2. Reguła Lenza	108
7.3. Indukcyjność. Samoindukcja	110
7.3.1. Indukcyjność wzajemna	111
7.4. Transformator	111
7.5. Energia pola magnetycznego	112
7.6. Równania Maxwella	114
7.6.1. Prąd przesunięcia	114
7.6.2. Równania Maxwella w postaci całkowitej	116
7.6.3. Równania Maxwella w postaci różniczkowej	118
Rozdział 8. DRGANIA I FALE	119
8.1. Drganie harmoniczne	119
8.1.1. Mechaniczne drgania harmoniczne	121
8.1.2. Elektryczne drgania harmoniczne	123
8.2. Składanie drgań harmonicznych równoległych o jednakowej częstotliwości. Dudnienie	125
8.3. Składanie drgań wzajemnie prostopadłych	127
8.4. Drgania swobodne tłumione	128
8.5. Drgania wymuszone	131
8.6. Amplituda i faza drgań wymuszonych. Rezonans	134
8.7. Prąd zmienny	135
8.7.1. Obwód zawierający rezystancję	135
8.7.2. Obwód zawierający indukcyjność	136
8.7.3. Obwód zawierający pojemność	137
8.7.4. Obwód RLC	138
8.7.5. Moc wydzielana w obwodzie prądu zmiennego	139
8.8. Procesy falowe	140
8.9. Fale biegnące	141
8.10. Przenoszenie energii przez fale	144

8.11. Paczka falowa. Prędkość grupowa	145
8.12. Interferencja fal	149
8.13. Fale stojące	151
8.14. Fale dźwiękowe	152
8.15. Zjawisko Dopplera	154
Rozdział 9. FALE ELEKTROMAGNETYCZNE	158
9.1. Równanie różniczkowe fali elektromagnetycznej	159
9.2. Promieniowanie płaskiego prądu	162
9.3. Oddziaływanie promieniowania z materią	165
9.3.1. Energia promieniowania	165
9.3.2. Pęd promieniowania	167
9.3.3. Odbicie promieniowania od przewodnika	168
9.3.4. Oddziaływanie promieniowania z dielektrykiem	169
9.3.5. Współczynnik załamania. Dyspersja	171
9.3.6. Pole promieniowania ładunków punktowych	174
9.4. Interferencja fal elektromagnetycznych	175
9.4.1. Interferencja fal promieniowanych przez dwa źródła punktowe	175
9.4.2. Interferencja fal od większej liczby źródeł	177
9.4.3. Siatka dyfrakcyjna	178
9.5. Dyfrakcja światła	180
9.5.1. Zasada Huygensa	180
9.5.2. Dyfrakcja na pojedynczej szczelinie	181
9.6. Koherentność i niekoherentność	183
9.7. Polaryzacja światła	185
9.7.1. Polaryzacja kołowa	185
9.7.2. Polaryzatory	187
9.7.3. Polaryzacja przez odbicie	189
9.8. Holografia	190
9.9. Optyka geometryczna	192
9.9.1. Prawo odbicia	193
9.9.2. Prawo załamania	195
9.9.3. Soczewki	195
Rozdział 10. FALOWA NATURA MATERII	197
10.1. Promieniowanie termiczne	197
10.2. Fotoefekt	202
10.3. Efekt Comptona	205
10.4. Dualizm korpuskularno-falowy	207
10.5. Funkcja falowa	208
10.6. Dyfrakcja elektronów	210
Rozdział 11. MECHANIKA KWANTOWA	213
11.1. Paczki falowe	213
11.2. Zasada nieoznaczoności	215
11.3. Właściwości paczek falowych	217
11.4. Częstka w studni potencjału	219
11.5. Równanie Schrödingera	221
11.6. Bariera potencjału. Efekt tunelowy	223
11.7. Oscylator harmoniczny	228

Rozdział 12. ATOM WODORU	232
12.1. Postulaty Bohra	232
12.2. Orbitalny moment pędu	234
12.3. Równanie Schrödingera dla atomu wodoru	236
12.4. Orbitalny moment magnetyczny	244
12.5. Spin elektronu	246
12.5.1. Spinowy moment pędu i spinowy moment magnetyczny	246
12.5.2. Całkowity moment pędu i całkowity moment magnetyczny	247
12.5.3. Oddziaływanie spin-orbita	250
Rozdział 13. FIZYKA ATOMOWA	252
13.1. Zakaz Pauliego	252
13.2. Atomy wieloelektronowe	253
13.3. Układ okresowy pierwiastków	255
13.4. Promieniowanie atomów wzbudzonych	258
13.4.1. Widma optyczne	258
13.4.2. Promieniowanie spontaniczne i wymuszone	259
13.4.3. Laser	263
13.5. Promieniowanie rentgenowskie	265
Rozdział 14. SIEĆ KRystaliczna CIAŁ STAŁYCH	269
14.1. Sieć krystaliczna i układy krystalograficzne	270
14.2. Oznaczenie węzłów, kierunków i płaszczyzn w kryształach	273
14.3. Proste struktury krystaliczne	274
14.4. Defekty sieci krystalicznej	275
14.5. Drgania sieci krystalicznej	278
14.6. Energia wewnętrzna i ciepło właściwe kryształu	284
14.7. Koncentracja fononów	286
14.8. Wiązania atomów w kryształach	287
14.8.1. Wiązania jonowe	289
14.8.2. Wiązania kowalencyjne	291
14.8.3. Wiązania metaliczne	294
14.8.4. Wiązania molekularne	294
Rozdział 15. ELEMENTY FIZYKI STATYSTYCZNEJ	296
15.1. Termodynamiczny opis układu	296
15.1.1. Pierwsza zasada termodynamiki	296
15.1.2. Druga zasada termodynamiki	297
15.1.3. Warunki równowagi termodynamicznej	298
15.1.4. Potencjał chemiczny	299
15.2. Statystyczny opis układu	300
15.3. Układy niezwyrodniałe i zwyrodniałe	301
15.4. Statystyki: klasyczna i kwantowa	302
15.5. Gaz doskonały	305
15.6. Gaz elektronów swobodnych	307
15.6.1. Przestrzeń fazowa. Funkcja gęstości stanów	307
15.6.2. Gaz elektronowy w temperaturze 0 K	309
15.6.3. Wpływ temperatury na gaz elektronowy	311

Rozdział 16. STRUKTURA PASMOWA CIAŁ STAŁYCH	316
16.1. Funkcja Blocha	318
16.2. Model Kröniga-Penneya	320
16.3. Przykłady pasm energetycznych w kryształach	326
16.4. Przewodniki, półprzewodniki i izolatory	328
16.5. Masa efektywna elektronu	330
16.6. Pojęcie dziury	332
Rozdział 17. PODSTAWY FIZYKI PÓLPRZEWODNIKÓW	337
17.1. Półprzewodniki samoistne	338
17.2. Półprzewodniki domieszkowe	343
17.3. Koncentracja nośników w półprzewodniku domieszkowym	345
17.4. Ruchliwość nośników w półprzewodnikach	351
17.5. Przewodnictwo elektryczne półprzewodników	356
17.5.1. Półprzewodniki samoistne	356
17.5.2. Półprzewodniki domieszkowe	357
Rozdział 18. ZJAWISKA NIERÓWNOWAGOWE W PÓLPRZEWODNIKACH	360
18.1. Relaksacja dielektryczna	360
18.2. Procesy generacji i rekombinacji	361
18.2.1. Rekombinacja bezpośrednia	363
18.2.2. Rekombinacja przez centra. Pułapkowanie	364
18.2.3. Rekombinacja powierzchniowa	365
18.3. Quasi-poziomy Fermiego	366
18.4. Transport nośników w półprzewodnikach	367
18.4.1. Prądy unoszenia i dyfuzji	367
18.4.2. Równanie ciągłości	371
18.5. Dyskusja ambipolarnego równania transportu	374
18.6. Dyfuzja nadmiarowych nośników prądu	375
18.7. Półprzewodnik niejednorodny	376
Rozdział 19. FIZYKA JĄDROWA	378
19.1. Pojęcia wstępne	378
19.2. Rozmiar jądra	379
19.3. Energia wiązania jądra. Defekt masy	380
19.4. Modele struktury jądra atomowego	381
19.4.1. Model kroplowy	381
19.4.2. Model powłokowy	384
19.4.3. Model kolektywny	385
19.5. Siły jądrowe	385
19.6. Przemiany jądrowe	387
19.6.1. Promieniotwórczość naturalna	387
19.6.2. Rozpad α	389
19.6.3. Rozpad β	392
19.6.4. Promieniowanie γ	394
19.7. Reakcje jądrowe	396
19.7.1. Charakterystyka reakcji jądrowych	396
19.7.2. Reakcje rozszczepienia	398
19.7.3. Reakcje łańcuchowe. Reaktor atomowy	400
19.7.4. Reakcje termojądrowe	402

Rozdział 20. ASTROFIZYKA I KOSMOLOGIA	405
20.1. Pojęcia wstępne	405
20.2. Gwiazdy	406
20.2. Czarne dziury	407
20.3. Ciśnienie kwantowo-mechaniczne	408
20.4. Ewolucja gwiazd	409
20.5. Kosmologia	414
20.5.1. Prawo Hubble'a	414
20.5.2. Promieniowanie cieplne Wszechświata	416
20.5.3. Model Wszechświata	417

PRZEDMOWA

Podręczników z fizyki dla studentów pierwszych lat studiów politechnicznych wydano już w Polsce wiele. Szczególnie cenione są zwłaszcza tłumaczenia wielotomowych kursów Feynmana, popularnego kursu "berkeleyowskiego", dwutomowa *Fizyka* Resnicka i Hallidaya, czy też książka Oreara. Jest również sporo książek i skryptów napisanych przez autorów polskich.

Mała ilość godzin przeznaczona na fizykę w programach studiów zmusza do selekcji wykładanych zagadnień przy konieczności sformułowania wykładu spójnego i obrazującego ważne zagadnienia fizyki. Wymogi te trudno jest spełnić w oparciu o jeden podręcznik. Dają się też zauważyć trudności sporządzania notatek z wykładów przez studentów. Wynikają one z dużego zagęszczenia wykładanych treści w czasie i niewłaściwego przygotowania studentów z fizyki na poziomie kursu średniego. Wymienione powody są głównymi, dla których podjąłem próbę napisania niniejszego skryptu, mającego na celu ułatwienie studentom nauki fizyki. W żadnym jednak przypadku skrypt nie może zastąpić podręczników z fizyki, w których omawia się wiele innych ważnych zagadnień. Należy zwrócić uwagę, że w nauczaniu fizyki na studiach technicznych, pracownie studenckie i ćwiczenia audytoryjne są równie ważną formą nauczania tego przedmiotu. Poznaje się na nich eksperymentalne i praktyczne zagadnienia fizyki, na które na wykładach nie ma czasu.

W powstaniu pierwszej wersji tego skryptu pomogli mi dwaj studenci Wydziału Elektroniki – Paweł Lewczyński i Jarosław Góra, którzy w roku akademickim 1991/1992 studiowali na I roku Wydziału Elektroniki. Wyżej wymienieni studenci w czerwcu 1992 roku zaproponowali mi, że napiszą skrypt na podstawie moich dobrze zredagowanych notatek. W krótkim czasie opanowali edytory tekstu i grafiki (*Word 2 for Windows* i *Corel Draw*) i po pół roku pracy przepisałi moje notatki. Druga wersja tego skryptu została poprawiona i uzupełniona przeze mnie. Uzupełniłem ją o dwa rozdziały poświęcone fizyce jądrowej i astrofizyce z kosmologią.

W tym miejscu należałoby sobie odpowiedzieć na pytanie: jaką rolę powinien spełniać kurs fizyki ogólnej (podstaw fizyki) na uczelniach technicznych? Na pewno nie jest on w stanie nauczyć fizyki. Sądzę, iż powinien dać jak gdyby "przedsmak" prawdziwej fizyki i zawierać elementy myślenia fizycznego, jakim posługują się naukowcy. Dzięki temu inżynier nabyłby pewnej kultury fizycznej dnia dzisiejszego, co powinno być chyba głównym celem nauczania fizyki w politechnikach.

W skrypcie unikałem dublowania kursu fizyki dla szkół średnich. Nie eksponuję klasycznej mechaniki, która jest stosunkowo dobrze znana przez studentów z nauczania fizyki w szkole średniej. Ponieważ jednak mechanika zawiera podstawowy pakiet pojęć i praw używanych w innych gałęziach fizyki, więc w skrypcie w sposób skrótowy zwracam uwagę na ten aspekt mechaniki (rozdział 2). Starłem się nie dzielić fizyki na "klasyczną" i "współczesną", ale raczej ująć ją w jedną całość. W tym zakresie niedoścignionym wzorem dla mnie jest książka Jay'a

Oreara *Fizyka*. Podział na fizykę "klasyczną" i "współczesną" wprowadzony na początku XX w. ma już znaczenie historyczne i obecnie nie ma głębszego uzasadnienia.

Większość zagadnień przedstawiono w sposób tradycyjny, gdyż sama formuła skryptu z jego ograniczoną objętością nie pozwala na eksperymentowanie w tym zakresie. Ponadto bardzo trudno jest zaproponować oryginalny i przystępny do samodzielnego studiowania układ treści biorąc pod uwagę różnorodność znakomitych podręczników fizyki dostępnych na polskim rynku wydawniczym. Szereg rysunków i sposobów ujęcia prezentowanych treści zaczerpnąłem z tych podręczników, a w szczególności z *Fizyki* J. Oreara (WNT, Warszawa) i z *Fizyki Kwantowej, Atomowej i Ciała Stałego* Z. Kleszczewskiego (Wydawnictwo Politechniki Śląskiej, Gliwice).

Skrypt ten adresowany jest dla studentów wydziałów politechnicznych o specjalności elektronicznej i elektrycznej, dlatego też najwięcej uwagi poświęcam oddziaływaniu elektromagnetycznemu i niektórym zagadnieniom fizyki ciała stałego, a w szczególności fizyki półprzewodników. Nie sposób przedstawić fizyczne podstawy teorii elektromagnetyzmu bez znajomości szczególnej teorii względności, dlatego też mechanikę relatywistyczną przedstawiłem w rozdziale 3 przed elektromagnetyzmem (rozdziały 4–7, częściowo 8 i rozdział 9). Ponieważ prawdziwe zrozumienie struktury materii i wiele innych zjawisk fizycznych opiera się na mechanice kwantowej, więc jej podstawy opisałem w rozdziałach 10 i 11. Dualizm korpuskularno-falowy cząstek i teorię kwantów można głębiej zrozumieć w następnych rozdziałach dzięki ich zastosowaniu w fizyce atomowej (rozdział 12 i 13) i fizyce ciała stałego (rozdział 14). Teoria gazu elektronów swobodnych przedstawiona jest w rozdziale 15. W kolejnych dwóch rozdziałach zwróciłem szczególną uwagę na zrozumienie fizycznych podstaw teorii struktury pasmowej (rozdział 16) i zjawisk elektronowych i optycznych w półprzewodnikach (rozdział 17). Rozdział 18 wychodzi poza ramy kursu fizyki i poświęcony jest niektórym elementarnym zagadnieniom fizyki półprzewodników, których znajomość jest niezbędna do zrozumienia działania przyrządów półprzewodnikowych. Tak pomyślany skrypt stanowić może pomoc do zrozumienia podstaw fizycznych elektromagnetyzmu i zjawisk transportu elektronowego w przyrządach półprzewodnikowych, które to zagadnienia są podstawą dalszych studiów politechnicznych na specjalnościach elektrycznych i elektronicznych. Skrypt kończą rozdziały poświęcone fizyce jądrowej i astrofizyce z kosmologią. Konieczność uzupełnienia skryptu o te dwa rozdziały wynika z zapotrzebowania społecznego współczesnej cywilizacji, dla której walka o ochronę środowiska naturalnego stała się priorytetowym zadaniem. Interesujący ostatni rozdział porządkuje wiedzę o Wszechświecie, którego obraz często deformowany jest fabułą filmów *science fiction*.

Studiowanie skryptu wymaga pewnych wiadomości z matematyki wyższej. Dla ułatwienia informacje te zawarte są we "Wstępie matematycznym", który zaleca się przestudiować na samym początku lektury skryptu.

Zdaję sobie sprawę, że skrypt nie jest doskonały, że zawiera w wielu miejscach nieprecyzyjne sformułowania. Jednakże mając na uwadze wyżej wymienione uwarunkowania

sądzę, że skrypt w obecnej postaci będzie przydatny szczególnie dla studentów Wydziału Elektroniki.

Chciałbym podziękować dr Jolancie Raczyńskiej za cenne uwagi dotyczące redakcji skryptu. Będę wdzięczny czytelnikom za uwagi, które mogłyby przyczynić się do ulepszenia skryptu.

Antoni Rogalski

ROZDZIAŁ 1

WPROWADZENIE

1.1. Czym jest fizyka?

Fizyka jest nauką przyrodniczą najbardziej podstawową i wszechogarniającą, wpływ zaś jej na rozwój nauk przyrodniczych był i jest ogromny. Właściwie fizyka odgrywa dziś rolę tego, co dawniej nazywano filozofią przyrody i z czego zrodziły się współczesne nauki przyrodnicze. Można powiedzieć, że fizyka stanowi system podstawowych idei uogólniających dane eksperymentalne i odzwierciedlających obiektywne prawa przyrody. Zajmuje się badaniem najprostszych a zarazem najbardziej ogólnych form ruchu materii. Materię uważamy za realność obiektywną, istniejącą niezależnie od świadomości ludzkiej i uzmysłowaną sobie przez człowieka. Cała historia fizyki i innych nauk przyrodniczych niezbicie dowodzi materialności świata, obiektywności i poznawalności praw jego rozwoju.

Świat fizyczny jest ogromny i złożony, jest widownią zjawisk i wydarzeń zdumiewających różnorodnością. To stwierdzenie oparte jest na ocenie rzędu wielkości pewnych interesujących nas liczb. Z obserwacji astronomicznych wynika, że rozmiar Wszechświata określony jego promieniem wynosi 10^{26} m, czyli 10^{10} lat świetlnych. Wartość tę znamy z dokładnością do czynnika około 3. Dla porównania podajmy, że odległość Ziemi do Słońca wynosi $1,5 \times 10^{11}$ m, a promień Ziemi jest równy $6,4 \times 10^6$ m. Liczba atomów we Wszechświecie jest bardzo duża. Przyjmuje się, że całkowita liczba protonów i neutronów we Wszechświecie jest rzędu 10^{80} , przy czym liczbę tę znamy z dokładnością do czynnika 100. Na Słońcu jest 10^{57} atomów, a na Ziemi 4×10^{51} . Najbardziej złożonym zjawiskiem we Wszechświecie jest życie. Człowiek, jedna z najbardziej złożonych form życia, składa się z około 10^{16} komórek. Komórka jest elementarną jednostką fizjologiczną zawierającą około 10^{12} – 10^{14} atomów.

W XX w. utrwalił się podział fizyki na fizykę klasyczną zajmującą się opisem makroświata, i fizykę współczesną zajmującą się opisem mikroświata (opisem budowy atomu, jądra atomowego, itp.). Obecnie bardziej celowe jest stosowanie nazw: fizyka makroświata i fizyka mikroświata. Każdy opis fizyczny jest przybliżonym opisem otaczającego nas świata i praw nim rządzących, przy czym fizyka makroświata daje nam jego pierwsze przybliżenie. Słupami granicznymi w tym podziale na fizykę makroświata i mikroświata są teoria względności i mechanika kwantowa.

Pomimo ogromnej różnorodności zjawisk fizycznych we Wszechświecie, do chwili obecnej odkryto zaledwie cztery typy oddziaływań. Ich porównanie zamieszczone jest w poniższej tabeli. Z tych oddziaływań wynikają wszystkie siły obserwowane we Wszechświecie.

Tabela 1.1. Oddziaływania fundamentalne

Oddziaływanie	Źródło	Intensywność względna	Promień działania
Grawitacyjne	Masa	10^{-39}	Dalekozasięgowe
Słabe	Wszystkie cząstki elementarne	10^{-15}	Krótkozasięgowe (10^{-15} m)
Elektromagnetyczne	Ładunki elektryczne	10^{-2}	Dalekozasięgowe
Jądrowe (silne)	Hadrony (protony, neutrony, mezony)	1	Krótkozasięgowe (10^{-15} m)

Studiowanie skryptu wymaga pewnych wiadomości z matematyki wyższej. Ponieważ zarówno wykłady z matematyki jak i z fizyki rozpoczynają się na pierwszym semestrze studiów, stąd też nie można zrealizować zgodności programowej obu przedmiotów. Z tego powodu, dla ułatwienia, informacje te zawarte są we "Wstępie matematycznym", który zaleca się przestudiować na samym początku lektury skryptu.

1.2. Wstęp matematyczny

Materiał przedstawiony w tym skrypcie wymaga znajomości jedynie dość elementarnej matematyki: różniczkowania i całkowania, działań na liczbach zespolonych, działań na wektorach, czy elementów rachunku prawdopodobieństwa. Są to zagadnienia dobrze znane z kursów matematyki, niemniej jednak celem jest przypomnieć je w specjalnym wstępie wiążąc je od zaraz z fizyką, tym bardziej, że program matematyki podejmuje te zagadnienia w późniejszych semestrach.

1.2.1. Pochodna funkcji

Ze względu na ograniczoną dokładność wszelkich przyrządów pomiarowych, sensowne wydaje się posługiwanie tylko skończonymi przyrostami wielkości fizycznych oznaczonych tradycyjnie symbolami Δ , np. Δx , Δt , ΔV , itd. Często zdarza się, że jedna wielkość fizyczna wyraża się przez stosunek przyrostów dwóch innych wielkości, jak np. prędkość i przyspieszenie. Jednakże zapis typu $\Delta y/\Delta x$ byłby nieodpowiedni, gdyż występujące w nim

przyrosty są niejednoznacznie określone; wartość tego stosunku zależy na ogół od wartości granicznej przy Δx dążącym do zera, o ile tylko Δx jest dostatecznie małe.

Wartość graniczna to pochodna funkcji $y(x)$ względem x

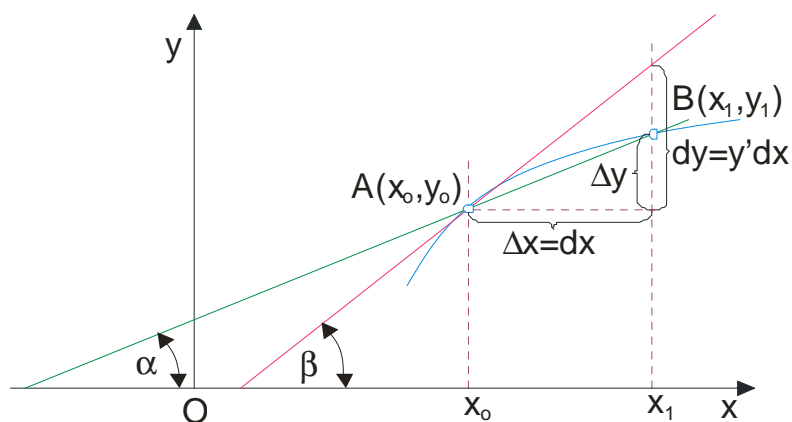
$$y' = \frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}. \quad (1.1)$$

Wyrażenie $dy = y'dx$ nazywa się różniczką funkcji $y(x)$, zaś dx różniczką argumentu x . Obliczanie pochodnej nazywamy różniczkowaniem.

Pochodna funkcji ma prostą interpretację geometryczną (rys. 1.1). Punkty A i B tego wykresu mają współrzędne x_0, y_0 oraz x_1, y_1 . Gdy punkt B zbliża się do punktu A (tzn. gdy $\Delta x \rightarrow 0$ i w granicznym przypadku pokrywa się z punktem A), prosta AB przechodzi w styczną do krzywej w punkcie A, a kąt β jest równy kątowi α jaki tworzy ta styczna z osią x . Zatem

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx} = \operatorname{tg} \alpha. \quad (1.2)$$

Możemy więc powiedzieć, że pochodna funkcji w danym punkcie jest równa tangensowi kąta nachylenia stycznej do wykresu funkcji w tym punkcie do osi x . Na rys. 1.1 pokazano również przyrost argumentu Δx i przyrost funkcji Δy . Nie ma istotnej różnicy między interpretacją geometryczną przyrostu argumentu Δx , a jego różniczką dx , jest natomiast zasadnicza różnica między przyrostem funkcji Δy a jej różniczką $dy = y'dx$.



Rys. 1.1. Interpretacja geometryczna pochodnej funkcji $y(x)$.

Pochodna sumy dwóch funkcji

Jeżeli $y = u + v$, przy czym u i v są funkcjami tego samego argumentu x , wówczas

$$\frac{dy}{dx} = \frac{d(u + v)}{dx} = \frac{du}{dx} + \frac{dv}{dx}. \quad (1.3)$$

Jeżeli $y = uv$, wówczas

$$\frac{dy}{dx} = \frac{d(uv)}{dx} = \frac{du}{dx}v + u\frac{dv}{dx}. \quad (1.4)$$

Pochodna ilorazu dwóch funkcji

Jeżeli $y = u/v$, mamy

$$\frac{dy}{dx} = \frac{\frac{du}{dx}v - u\frac{dv}{dx}}{v^2}. \quad (1.5)$$

Pochodna funkcji złożonej

Niech z będzie funkcją zmiennej y , zaś y funkcją zmiennej x ; np. $z = \cos y$, $y = 3x^2$, czyli $z = \cos 3x^2$. Wówczas

$$\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}. \quad (1.6)$$

Jeżeli funkcja f zależy od kilku zmiennych niezależnych, np. $f(x,y,z,t)$, wówczas pochodne po każdej z nich nazywa się pochodnymi cząstkowymi i oznacza nieco innym symbolem, np. $\partial f / \partial x$, $\partial f / \partial y$, itd. Pochodne te oblicza się identycznie, jak zwykle pochodne, traktując zmienne po których nie wykonuje się różniczkowania, jako stałe.

1.2.2. Rachunek całkowy

Operacją odwrotną do różniczkowania jest całkowanie (nieoznaczone). Całką nieoznaczoną lub funkcją pierwotną funkcji $y = f(x)$ nazywamy taką funkcję $F(x)$, której pochodna jest równa danej funkcji $f(x)$, czyli $dF(x)/dx = f(x)$. Całkę nieoznaczoną zapisujemy symbolicznie

$$F(x) = \int f(x) dx = \int dF(x). \quad (1.7)$$

Całką funkcji $f(x)$ jest każda funkcja będąca sumą funkcji $F(x)$ i dowolnej stałej C , ponieważ zawsze $d[F(x) \pm C]/dx = dF(x)/dx = f(x)$. Całka nieoznaczona funkcji $f(x)$

$$F(x) = \int f(x) dx + C. \quad (1.8)$$

Całkowanie przez zmianę zmiennej (metoda podstawienia)

Jeżeli w funkcji $f(x)$ za zmienną x podstawimy funkcję $x = \varphi(t)$, to

$$\int f(x) dx = \int f[\varphi(t)]\varphi'(t) dt. \quad (1.9)$$

Całkowanie przez części

Jeżeli u oraz v są funkcjami tej samej zmiennej x , to

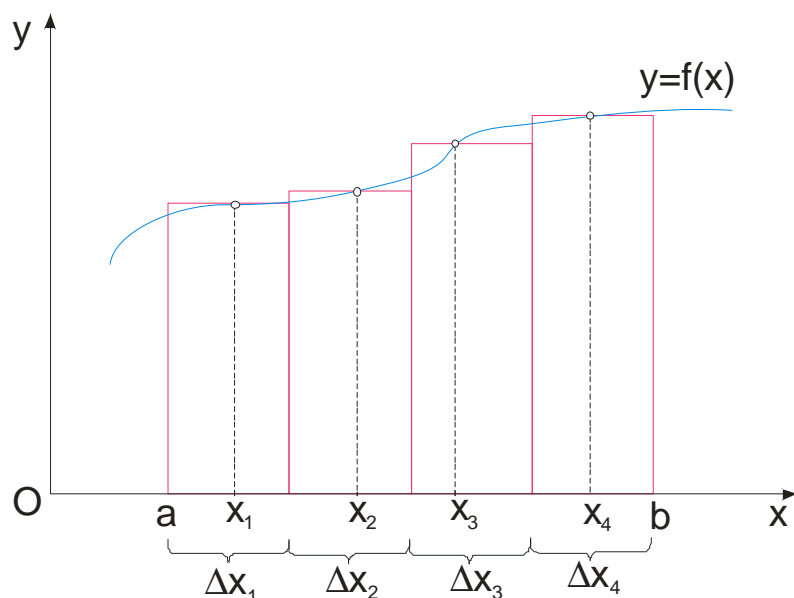
$$\int uv' dx = uv - \int u'v dx. \quad (1.10)$$

Całka oznaczona funkcji $f(x)$ w granicach od a do b jest definiowana jako

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a). \quad (1.11)$$

Z całką oznaczoną mamy do czynienia przy rozpatrywaniu wielkości globalnych, zależnych od wartości innej wielkości w pewnym skończonym przedziale argumentu. Klasycznym przykładem jest praca wzdłuż pewnej drogi. Jest ona równa sumie prac na dostatecznie małych odcinkach drogi, na jakie dzieli się ją w przypadku siły zależnej od położenia (rys. 1.2). Przy dostatecznie drobnym podziale można przyjąć, że siła na każdym z odcinków jest stała. Suma $\sum f_i(x_i)\Delta x_i$ równa jest polu figury ograniczonej krzywą schodkową. Wartość tej sumy nie jest jednoznacznie określona, gdyż zależy od sposobu podziału drogi (a,b) . W analizie matematycznej dowodzi się, że suma ta niewiele się różni od swej wartości granicznej przy wszystkich Δx_i dążących do zera. Ta granica to właśnie całka oznaczona. Możemy więc zapisać

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i) \Delta x_i. \quad (1.12)$$



Rys. 1.2. Interpretacja geometryczna całki oznaczonej

W fizyce mamy często do czynienia z całkami po krzywych, powierzchniach (strumienie), bądź też obszarach trójwymiarowych. Wszystkie takie całki rozumiemy w podobnym sensie, jak to opisywaliśmy powyżej. Obszar całkowania dzielimy myślowo na małe fragmenty; na każdym z nich funkcję całkowaną uważamy za stałą, a następnie tworzymy sumę iloczynów tych wartości i miar odpowiadających im fragmentów. W przypadku całkowania po krzywej, rolę Δx_i odgrywa długość Δs_i i -tego łuku krzywej; przy całkowaniu po powierzchni Δx_i należy zastąpić przez pole ΔS_i i -tego wycinka powierzchni; zaś w całkach objętościowych używamy elementów objętości ΔV_i . Graniczne wartości tak utworzonych sum nazywają się odpowiednio całkami: krzywoliniowymi, powierzchniowymi i objętościowymi. Możemy zatem zapisać

- całka krzywoliniowa

$$\int_C f(x, y, z) ds = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i, y_i, z_i) \Delta s_i, \quad (1.13)$$

- całka powierzchniowa

$$\int_S f(x, y, z) dS = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i, y_i, z_i) \Delta S_i, \quad (1.14)$$

- całka objętościowa

$$\int_V f(x, y, z) dV = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(x_i, y_i, z_i) \Delta V_i. \quad (1.15)$$

Jeżeli krzywa C lub powierzchnia S , na które rozciąga się całkowanie, jest zamknięta, to na symbolu całki zwykło się dopisywać kółko: \oint_C lub \oint_S .

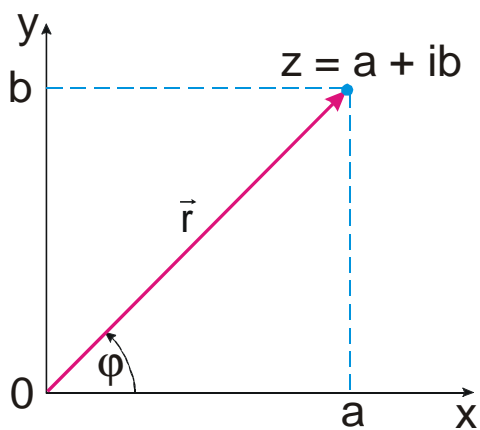
1.2.3. Liczby zespolone

Liczbą zespoloną z nazywamy liczbę

$$z = a + ib, \quad (1.16)$$

gdzie a i b są dowolnymi liczbami rzeczywistymi, zaś i – jednostką urojoną spełniającą związek $i^2 = -1$. Liczbę a nazywamy częścią rzeczywistą liczby zespolonej z , a liczbę b – częścią urojoną liczby z , co zapisujemy $a = \operatorname{Re}z$, $b = \operatorname{Im}z$. Zapis (1.16) nazywamy postacią algebraiczną liczby zespolonej.

Dwie liczby zespolone $z_1 = a_1 + ib_1$ oraz $z_2 = a_2 + ib_2$ są równe, gdy równe są ich części rzeczywiste i urojone, tzn. $a_1 = a_2$, $b_1 = b_2$. Nie istnieje natomiast pojęcie większej lub mniejszej liczby zespolonej.



Rys. 1.3. Interpretacja geometryczna liczby zespolonej.

Liczbę zespoloną można przedstawić jako punkt na płaszczyźnie zespolonej (rys. 1.3). Na osiach układu współrzędnych płaszczyzny zespolonej odkładamy współrzędne punktu będącego obrazem geometrycznym liczby z ; na osi rzeczywistej x liczbę a , zaś na osi urojonej y liczbę b . Korzystając z powyższej interpretacji geometrycznej, liczbę zespoloną można przedstawić w postaci trygonometrycznej

$$z = a + ib = r \cos \varphi + ir \sin \varphi = r(\cos \varphi + i \sin \varphi) = re^{i\varphi}. \quad (1.17)$$

Kąt φ nazywa się argumentem liczby zespolonej. Długość wektora wodzącego r nazywamy modułem lub wartością bezwzględną liczby zespolonej

$$|z| = r = \sqrt{a^2 + b^2}. \quad (1.18)$$

Liczbę zespoloną

$$z^* = a - ib = r(\cos \varphi - i \sin \varphi) = re^{-i\varphi}, \quad (1.19)$$

nazywamy liczbą zespoloną sprzężoną z liczbą $z = a + ib$. Zauważmy, że moduły liczb zespolonych sprzężonych są równe

$$|z^*| = \sqrt{a^2 + b^2} = |z|, \quad (1.20)$$

oraz, że

$$zz^* = a^2 + b^2 = r^2. \quad (1.21)$$

Łatwo sprawdzić, że

$$z_1 z_2 = r_1 e^{i\varphi_1} r_2 e^{i\varphi_2} = r_1 r_2 e^{i(\varphi_1 + \varphi_2)} = r_1 r_2 [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)], \quad (1.22)$$

$$|z_1 z_2| = |z_1| |z_2|, \quad (1.23)$$

$$\frac{z_1}{z_2} = \frac{r_1 e^{i\varphi_1}}{r_2 e^{i\varphi_2}} = \frac{r_1}{r_2} e^{i(\varphi_1 - \varphi_2)} = \frac{r_1}{r_2} [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)], \quad (1.24)$$

$$\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}. \quad (1.25)$$

$$\frac{z}{z^*} = \frac{re^{i\varphi}}{re^{-i\varphi}} = e^{2i\varphi}, \quad (1.26)$$

$$z^n = (re^{i\varphi})^n = r^n e^{in\varphi} = r^n [\cos(n\varphi) + i \sin(n\varphi)]. \quad (1.27)$$

1.2.4. Działania na wektorach

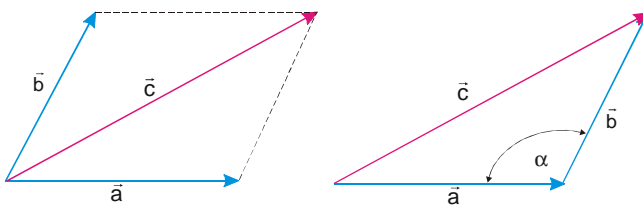
Większość podstawowych wielkości fizycznych ma charakter kierunkowy, w związku z czym reprezentowane są przez wektory. Początek wektora można w zasadzie umieszczać w dowolnym miejscu, chociaż w niektórych przypadkach jest on narzucony z góry (np. przy wektorze położenia lub siły).

Każdy wektor ma określoną wartość równą długości odcinka łączącego początek i koniec wektora. Długość wektora \vec{a} oznaczamy zwykle symbolem $|\vec{a}|$ lub po prostu a .

W rezultacie mnożenia wektora \vec{a} przez liczbę rzeczywistą λ otrzymuje się nowy wektor oznaczony jako $\lambda\vec{a}$, o tym samym ($\lambda > 0$) lub przeciwnym ($\lambda < 0$) zwrocie co wektor \vec{a} . Długość wektora $\lambda\vec{a}$ wynosi $|\lambda\vec{a}| = |\lambda|a$.

Dodawanie wektorów

Dwa wektory tego samego rodzaju dodaje się metodą równoległoboku, przy czym możliwe są dwa sposoby dodawania: albo przez sprowadzenie ich do wspólnego początku, albo przez umieszczenie początku drugiego wektora w końcu pierwszego (rys. 1.4).



Rys. 1.4. Dodawanie dwóch wektorów

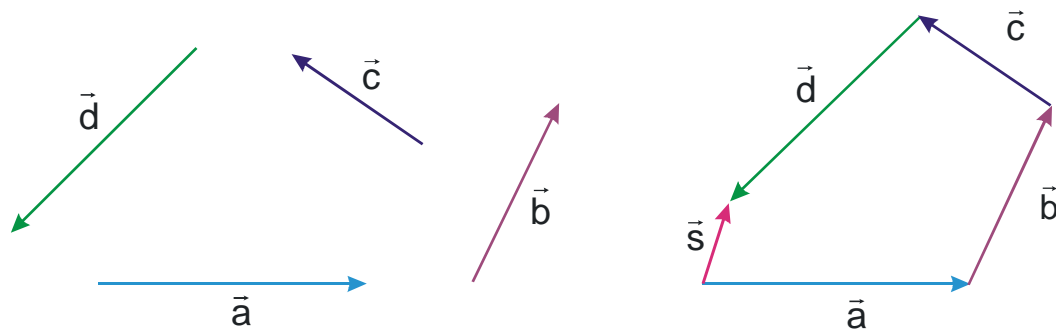
W pierwszym sposobie wypadkowa wektorów jest przekątną równoległoboku zbudowanego na wektorach \vec{a} i \vec{b} ; w drugim – wektor wypadkowy to odcinek o początku pokrywającym się z początkiem pierwszej składowej i końcu pokrywającym się z końcem drugiej składowej.

Wektor wypadkowy $\vec{c} = \vec{a} + \vec{b}$, a jego długość można wyznaczyć z twierdzenia cosinusów

$$c = \sqrt{a^2 + b^2 - 2ab \cos \alpha}, \quad (1.28)$$

gdzie α jest kątem między wektorami \vec{a} i \vec{b} .

Sumę większej liczby wektorów najlepiej jest tworzyć sposobem drugim, jak pokazano na rys. 1.5. Wynik dodawania nie zależy od kolejności poszczególnych składników. Łącząc początek wektora pierwszego z końcem wektora ostatniego otrzymujemy wektor wypadkowy $\vec{s} = \vec{a} + \vec{b} + \vec{c} + \vec{d}$.



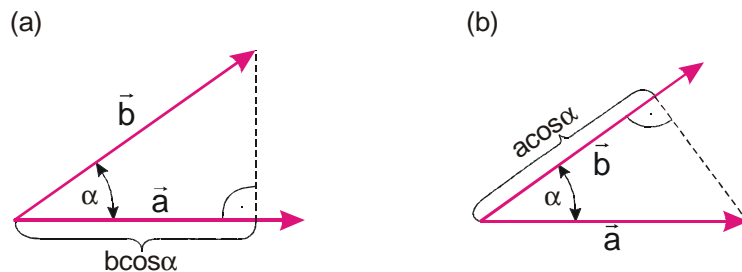
Rys. 1.5. Dodawanie geometryczne dowolnej liczby wektorów.

Iloczyn skalarny dwóch wektorów

Często w fizyce dwie wielkości wektorowe występują łącznie dając w rezultacie wielkość skalarną. Zazwyczaj jest to iloczyn długości jednego wektora przez rzut drugiego na pierwszy. Iloczyn taki nazywamy iloczynem skalarnym dwóch wektorów \vec{a} i \vec{b} , i oznaczamy symbolem $\vec{a} \cdot \vec{b}$

$$\vec{a} \cdot \vec{b} = ab \cos \alpha, \tag{1.29}$$

gdzie α jest kątem między wektorami \vec{a} i \vec{b} (rys. 1.6). Iloczyn skalarny dwóch wektorów jest więc skalarem. Z definicji iloczynu skalarnego widać, że jego wartość nie zależy od kolejności czynników, tzn. $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$.



Rys. 1.6. Interpretacja geometryczna iloczynu skalarnego wektorów.

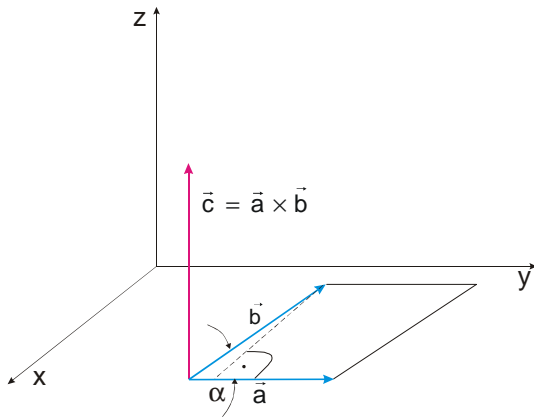
Iloczyn skalarny dwóch wektorów prostopadłych jest równy zero, gdyż $\cos 90^\circ = 0$. Klasyczny przykład iloczynu skalarnego to praca siły \vec{F} na odcinku $\Delta \vec{s}$ równa iloczynowi skalarnemu tych wektorów $\vec{F} \cdot \Delta \vec{s}$.

Iloczyn wektorowy dwóch wektorów

Sporo wielkości fizycznych o charakterze wektorowym wyraża się poprzez inne wielkości wektorowe przy pomocy tzw. iloczynu wektorowego. Iloczyn wektorowy \vec{a} i \vec{b} jest wektorem \vec{c} o kierunku prostopadłym do obu wektorów i zwrocie zgodnym z kierunkiem ruchu śruby prawoskrętnej wkręcanej tak, by pierwszy wektor \vec{a} nałożyć na drugi \vec{b} po mniejszym kącie

(rys. 1.7). Wartość iloczynu wektorowego równa jest - na mocy definicji - polu równoległoboku utworzonego przez wektory \vec{a} i \vec{b} , czyli

$$|\vec{c}| = |\vec{a} \times \vec{b}| = ab \sin \alpha. \quad (1.30)$$



Rys. 1.7. Konstrukcja geometryczna iloczynu wektorowego

Jest to więc iloczyn długości jednego z tych wektorów (obojętnie którego) przez składową drugiego wektora prostopadłą do niego.

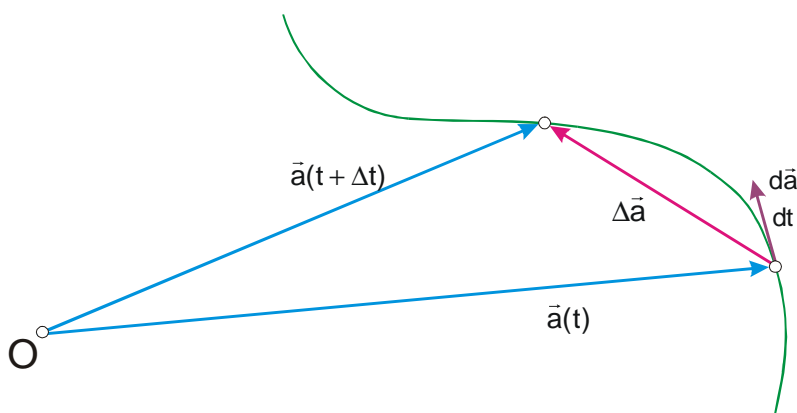
Przy pomocy pojęcia iloczynu wektorowego definiuje się różne wielkości fizyczne, jak np. moment pędu, moment siły, a ponadto zapisuje się szereg praw z mechaniki i elektrodynamiki.

Różniczkowanie wektorów

Jeżeli wielkość wektorowa jest funkcją pewnej zmiennej (np. czasu t), wówczas często zachodzi potrzeba obliczenia jej pochodnej po t . Pochodna wektora jest również wektorem, który otrzymuje się przez przejście do granicy z przyrostami skończonymi:

$$\frac{d\vec{a}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \vec{a}}{\Delta t} \quad ; \quad \Delta \vec{a} = \vec{a}(t + \Delta t) - \vec{a}(t). \quad (1.31)$$

Pochodna wektora ma na ogół inny kierunek niż wektor różniczkowany (rys. 1.8). Zgodność kierunków ma miejsce tylko wtedy, gdy wszystkie wektory $\vec{a}(t)$ mają ten sam kierunek. Jeżeli natomiast wszystkie wektory $\vec{a}(t)$ mają tę samą długość, to $d\vec{a}/dt$ jest prostopadły do \vec{a} . Pochodna ta jest różna od zera, gdyż zmienia się kierunek wektora \vec{a} .

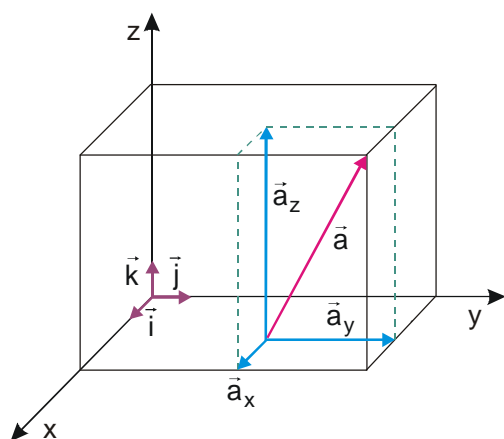


Rys. 1.8. Interpretacja różniczkowania wektora po czasie.

Współrzędne wektora

Wektory opisuje się przez podanie trzech liczb zwanych współrzędnymi wektora. W najprostszym przypadku są to trzy rzuty na trzy wzajemnie prostopadłe osie, mające wspólny początek umieszczony w początku wektora (rys. 1.9). Osie te nazywa się najczęściej osiami Ox , Oy , Oz . Zespół tych trzech współrzędnych wektora często utożsamia się z samym wektorem pisząc np. $\vec{a} = (a_x, a_y, a_z)$ choć poprawny jest tylko taki zapis, w którym wektor \vec{a} przedstawia się w postaci sumy trzech jego składowych w kierunkach osi układu współrzędnych:

$$\vec{a} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}, \quad (1.32)$$



Rys. 1.9. Składowe wektora w układzie prostokątnym.

gdzie wektory \vec{i} , \vec{j} i \vec{k} są wektorami jednostkowymi w kierunkach trzech osi współrzędnych.

Długość wektora wyraża się przez jego współrzędne w następujący sposób

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2}. \quad (1.33)$$

Podstawowe operacje na wektorach, zapisane przy użyciu współrzędnych, mają postać

$$\vec{a} \cdot \vec{b} = a_x b_x + a_y b_y + a_z b_z, \quad (1.34)$$

$$\begin{aligned} \vec{a} \times \vec{b} &= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = (a_y b_z - a_z b_y) \vec{i} + (a_z b_x - a_x b_z) \vec{j} + (a_x b_y - a_y b_x) \vec{k} = \\ &= (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x), \end{aligned} \quad (1.35)$$

$$\frac{d\vec{a}}{dt} = \left(\frac{da_x}{dt}, \frac{da_y}{dt}, \frac{da_z}{dt} \right), \quad (1.36)$$

Łatwo sprawdzić, że definicje te są zgodne z definicjami podanymi poprzednio bez stosowania współrzędnych.

Określone powyżej współrzędne wektora nazywane są jego współrzędnymi kartezjańskimi i są najbardziej naturalnymi współrzędnymi wektorowymi. Oprócz nich stosuje się również inne trójki liczb do scharakteryzowania wektora. Najczęściej stosowanymi współrzędnymi są: współrzędne biegunowe na płaszczyźnie, współrzędne walcowe w przestrzeni trójwymiarowej oraz współrzędne sferyczne (także przestrzenne). Współrzędnymi biegunowymi są: długość wektora a i kąt φ jaki on tworzy z dodatnim kierunkiem osi Ox (rys. 1.10). Związek między współrzędnymi kartezjańskimi (a_x, a_y) i biegunowymi (a, φ) jest następujący:

$$a_x = a \cos \varphi \quad ; \quad a_y = a \sin \varphi. \quad (1.37)$$

Współzrędnymi tymi posługujemy się często przy opisie ruchu odbywającego się w jednej płaszczyźnie.

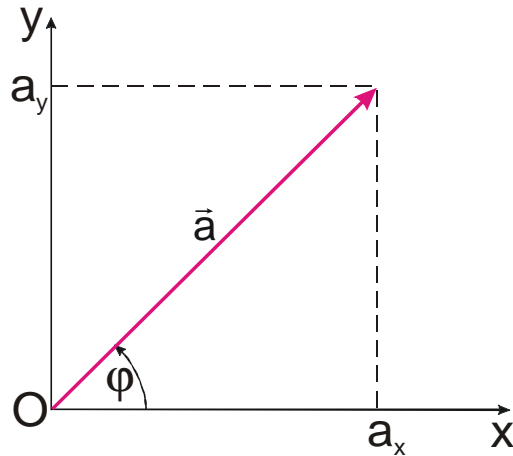
Współzrędnne walcowe (cylindryczne) to: długość rzutu wektora \vec{a} na płaszczyznę Oxy , kąt azymutalny φ w płaszczyźnie Oxy oraz współzrędnna kartezjańska a_z (rys. 1.11)

$$a_x = \sqrt{a_x^2 + a_y^2} \cos \varphi = a' \cos \varphi,$$

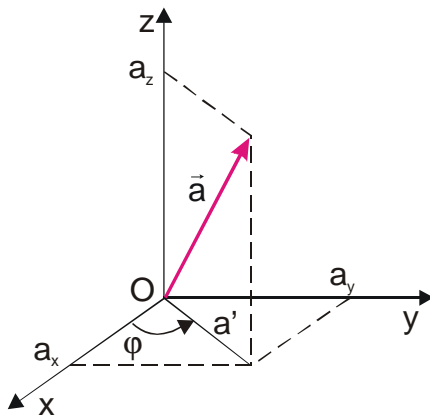
$$a_y = \sqrt{a_x^2 + a_y^2} \sin \varphi = a' \sin \varphi, \quad (1.38)$$

$$a_z = a_z.$$

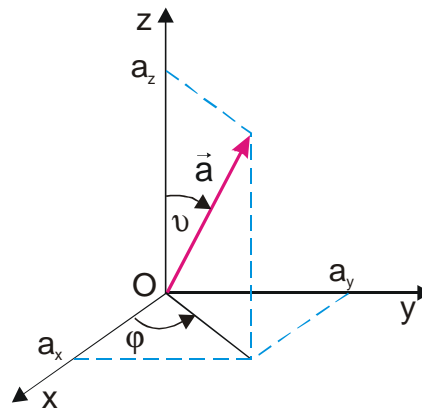
Współzrędnne te są stosowane w zagadnieniach wykazujących symetrię obrotową wokół osi Oz .



Rys. 1.10. Współzrędnne biegunowe w płaszczyźnie Oxy .



Rys. 1.11. Współzrędnne walcowe.



Rys. 1.12. Współzrędnne sferyczne.

Współzrędnymi sferycznymi są: długość wektora a , kąt biegunowy ϑ jaki tworzy wektor \vec{a} z dodatnią półosią Oz oraz kąt azymutalny φ (rys. 1.12). Związek ze współzrędnymi kartezjańskimi jest następujący:

$$a_x = a \sin \vartheta \cos \varphi,$$

$$a_y = a \sin \vartheta \sin \varphi, \quad (1.39)$$

$$a_z = a \cos \vartheta.$$

Współzrędnne te są wygodne w rozwiązywaniu zagadnień o symetrii sferycznej.

Wszystkie wymienione wyżej współzrędnne mogą być funkcjami czasu, a ich pochodne czasowe służą jako definicje odpowiednich prędkości.

1.2.5. Analiza wektorowa

Jeżeli funkcja $V(x,y,z)$ jest określona w każdym punkcie przestrzeni to mówimy, że funkcja $V(x,y,z)$ określa pewne pole skalarne. Funkcja $V(x,y,z)$, która przyporządkowuje każdemu punktowi pola pewną wielkość skalarną, nazywa się funkcją pola. Typowym przykładem takiej funkcji jest potencjał pola elektrostatycznego $V(x,y,z)$. W podobny sposób można zdefiniować temperaturę jako funkcję współrzędnych $T(x,y,z)$.

Jeżeli w każdym punkcie przestrzeni są określone trzy funkcje $A_1(x,y,z)$, $A_2(x,y,z)$ i $A_3(x,y,z)$, to można je traktować jako współrzędne wektora:

$$\vec{A}[A_1(x,y,z), A_2(x,y,z), A_3(x,y,z)].$$

Można zatem uważać, że każdemu punktowi przestrzeni został przyporządkowany pewien wektor $\vec{A} = \vec{A}(x,y,z)$. Przestrzeń, gdzie w każdym punkcie został zdefiniowany wektor według określonego prawa, nazywamy polem wektorowym. Tak więc każdemu punktowi pola elektrostatycznego można przyporządkować wektor natężenia pola \vec{E} , a każdemu punktowi pola magnetycznego wektor indukcji magnetycznej \vec{B} .

Gradient pola skalarnego

Niech funkcja $V(x,y,z)$ określa pewne pole skalarne. Punkty dla których funkcja ta ma stałą wartość [$V(x,y,z) = const$] leżą na pewnej powierzchni. Zmieniając wartość $const$ otrzymujemy rodzinę powierzchni, które nazywamy powierzchniami ekwipotencjalnymi.

Obliczmy różniczkę dV funkcji $V(x,y,z)$ przy przejściu od punktu (x,y,z) określonego wektorem wodzącym $\vec{r} = x\vec{i} + y\vec{j} + z\vec{k}$, do punktu $(x+dx, y+dy, z+dz)$ leżącego na bliskiej, sąsiedniej powierzchni ekwipotencjalnej i określonego wektorem wodzącym $\vec{r}' = (x+dx)\vec{i} + (y+dy)\vec{j} + (z+dz)\vec{k}$. Różniczka ta jest równa

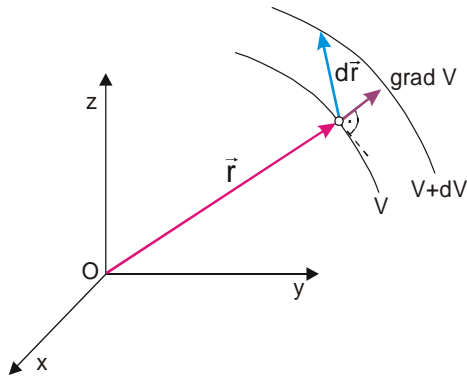
$$dV = \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz.$$

Różniczkę dV można przedstawić w postaci iloczynu skalarnego wektora

$$gradV = \frac{\partial V}{\partial x} \vec{i} + \frac{\partial V}{\partial y} \vec{j} + \frac{\partial V}{\partial z} \vec{k}, \quad (1.40)$$

nazwanego gradientem funkcji skalarnej $V(x,y,z)$ i wektora $d\vec{r} = \vec{i} dx + \vec{j} dy + \vec{k} dz$ gdyż

$$dV = \left(\vec{i} \frac{\partial V}{\partial x} + \vec{j} \frac{\partial V}{\partial y} + \vec{k} \frac{\partial V}{\partial z} \right) (\vec{i} dx + \vec{j} dy + \vec{k} dz) = gradV \cdot d\vec{r}. \quad (1.41)$$



Rys. 1.13. Gradient pola skalarnego.

$GradV$ jest wektorem prostopadłym do powierzchni ekwipotencjalnej i jest skierowany od powierzchni o potencjale niższym do powierzchni o potencjale wyższym (rys. 1.13). Długość wektora $gradV$ wynosi:

$$|gradV| = \sqrt{\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial y}\right)^2 + \left(\frac{\partial V}{\partial z}\right)^2} = \frac{dV}{dr}.$$

Operator nabla. Dywergencja i rotacja pola wektorowego

Operatorem nazywamy symbol określający przepis działania matematycznego na jakiejś wielkości. Np. symbol d/dx jest operatorem różniczkowania po zmiennej x . Oznaczony symbolem ∇ operator

$$\nabla = \frac{\partial}{\partial x} \vec{i} + \frac{\partial}{\partial y} \vec{j} + \frac{\partial}{\partial z} \vec{k}, \quad (1.42)$$

nazywany jest operatorem nabla lub operatorem Hamiltona. Sam operator nie oznacza żadnej wielkości, lecz działając na jakąś wielkość (skalar lub wektor) nabiera sensu wielkości. Operator nabla ma charakter wektora, działa więc na inne wielkości tak, jak gdyby był wektorem.

1. Iloczyn operatora nabla i skalara.

$$\nabla \lambda = \left(\frac{\partial}{\partial x} \vec{i} + \frac{\partial}{\partial y} \vec{j} + \frac{\partial}{\partial z} \vec{k} \right) \lambda = \frac{\partial \lambda}{\partial x} \vec{i} + \frac{\partial \lambda}{\partial y} \vec{j} + \frac{\partial \lambda}{\partial z} \vec{k} = grad \lambda. \quad (1.43)$$

2. Iloczyn skalarny operatora nabla i wektora.

$$\nabla \vec{a} = \left(\frac{\partial}{\partial x} \vec{i} + \frac{\partial}{\partial y} \vec{j} + \frac{\partial}{\partial z} \vec{k} \right) (a_x \vec{i} + a_y \vec{j} + a_z \vec{k}) = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}.$$

Sumę pochodnych cząstkowych kolejnych współrzędnych wektora \vec{a} , względem kolejnych zmiennych x, y, z , nazywamy dywergencją wektora \vec{a} i oznaczamy symbolem $div \vec{a}$. Zatem

$$div \vec{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}, \quad (1.44)$$

oraz

$$\nabla \vec{a} = div \vec{a}. \quad (1.45)$$

3. Iloczyn wektorowy operatora nabla i wektora.

$$\nabla \times \vec{a} = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) \vec{i} + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) \vec{j} + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \vec{k}.$$

Wektor występujący po prawej stronie tej równości nazywa się rotacją wektora \vec{a} .

$$\text{rot} \vec{a} = \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z} \right) \vec{i} + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x} \right) \vec{j} + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y} \right) \vec{k}. \quad (1.46)$$

Mnożąc wektorowo operator ∇ przez wektor otrzymujemy jego rotację

$$\nabla \times \vec{a} = \text{rot} \vec{a}. \quad (1.47)$$

4. Iloczyn skalarny dwóch operatorów nabla.

$$\nabla \nabla = \nabla^2 = \left(\frac{\partial}{\partial x} \vec{i} + \frac{\partial}{\partial y} \vec{j} + \frac{\partial}{\partial z} \vec{k} \right)^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Otrzymujemy w ten sposób nowy operator, zwany operatorem Laplace'a lub laplasjanem i oznaczamy symbolem

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (1.48)$$

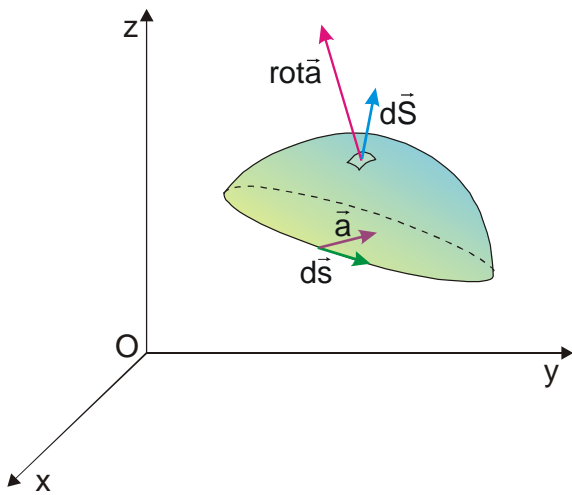
Laplasjan ma charakter skalara, a nie wektora jak operator nabla.

Twierdzenie Stokesa i twierdzenie Gaussa-Ostrogradzkiego

Podstawowe twierdzenia analizy wektorowej, twierdzenie Stokesa mówi, że dla pola wektorowego $\vec{a}(x, y, z)$ całka krzywoliniowa wektora \vec{a} po obwodzie zamkniętym C jest równa całce wektora $\text{rot} \vec{a}$ po powierzchni ograniczonej przez ten obwód (rys. 1.14).

$$\oint_C \vec{a} \cdot d\vec{s} = \int_S \text{rot} \vec{a} \cdot d\vec{S}. \quad (1.49)$$

Wektor $d\vec{s}$ o długości równej długości elementu ds jest wektorem stycznym do obwodu i wskazuje kierunek całkowania po obwodzie. Wektor $d\vec{S}$ jest prostopadły do powierzchni ograniczonej obwodem i ma długość równą polu elementu dS . Zwrot wektora $d\vec{S}$ wskazuje przesuw śruby prawoskrętnej obracającej się zgodnie ze zwrotem wektora $d\vec{s}$.



Rys. 1.14. Ilustracja do twierdzenia Stokesa.

Twierdzenie Gaussa-Ostrogradzkiego mówi, że dla pola wektorowego $\vec{a}(x, y, z)$ całka wektora \vec{a} po powierzchni zamkniętej S jest równa całce $\text{div} \vec{a}$ po objętości V ograniczonej powierzchnią S .

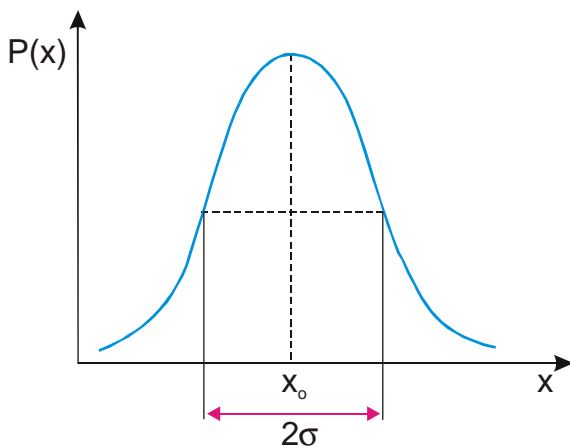
$$\oint_S \vec{a} \cdot d\vec{S} = \int_V \text{div} \vec{a} \cdot dV. \quad (1.50)$$

Wektor $d\vec{S}$ jest prostopadły do powierzchni i skierowany na zewnątrz powierzchni, element objętości $dV = dx dy dz$.

1.2.6. Prawdopodobieństwa. Wartości średnie

Rachunek prawdopodobieństwa i oparta na nim statystyka matematyczna należą do podstawowych narzędzi współczesnej fizyki. Prawie wszystkie prawa opisujące zachowanie się mikrocząstek formułowane są w kategoriach prawdopodobieństwa, a nie pewności - jak w fizyce klasycznej. Rachunkiem tym posługujemy się także w badaniu właściwości układów złożonych z bardzo dużej liczby cząstek. Jest on także podstawą rachunku błędów przy opracowywaniu danych pomiarowych.

W fizyce wystarcza elementarna definicja prawdopodobieństwa $P(x)$ jako graniczna wartość stosunku liczby zdarzeń (sytuacji) odpowiadających danej wartości x , do ogólnej liczby zdarzeń (sytuacji), możliwych do zaistnienia w określonych warunkach. Zmienna losowa x może być dyskretna lub ciągła; odpowiednio do tego mamy dwa rodzaje funkcji $P(x)$. Dodajmy, że funkcję określającą rozkład prawdopodobieństw nazywa się zwykle funkcją rozkładu (prawdopodobieństwa) lub krótko – rozkładem.



Rys. 1.15. Przykład funkcji rozkładu dla zmiennej losowej ciągłej.

Jawna postać funkcji rozkładu zależy oczywiście od konkretnej sytuacji i jej określenie jest często głównym celem rozwiązań problemów fizycznych.

Typowa funkcja rozkładu ma kształt dzwonu (rys. 1.15), z wyraźnie zaznaczonym maksimum dla pewnej wartości x_0 zmiennej x . Bardzo ważną charakterystyką takiej krzywej jest szerokość 2σ tego maksimum. Większej szerokości odpowiada większy "rozrzut" wartości zmiennej x .

Jeżeli wielkość x może przyjmować różne wartości z prawdopodobieństwem $P(x)$, to należy ją uśrednić. Sposób uśredniania zależy od charakteru tej wielkości. Przy dyskretnych wartościach tej zmiennej, równych x_1, x_2, x_3, \dots , wartość średnia $\langle x \rangle$ zmiennej x oblicza się według reguły:

$$\langle x \rangle = \sum_j x_j P(x_j). \quad (1.51)$$

Nietrudno zauważyć, że jest to zwykła średnia arytmetyczna. Dla ciągłej zmiennej losowej mamy analogicznie

$$\langle x \rangle = \int x P(x) dx, \quad (1.52)$$

przy czym całkowanie rozciąga się na cały przedział zmienności x .

Warto zwrócić uwagę na pewną subtelną różnicę między rozkładami $P(x)$ występującymi w dwóch powyższych definicjach. W przypadku dyskretnej zmiennej losowej prawdopodobieństwa $P(x_j)$ są liczbami bezwymiarowymi, natomiast we wzorze (1.52) wielkością bezwymiarową jest iloczyn $P(x)dx$. Iloczyn ten ma znaczenie prawdopodobieństwa wystąpienia wartości zmiennej losowej na odcinku dx wokół "bieżącej" wartości x . Samo $P(x)$ ma więc znaczenie gęstości prawdopodobieństwa, czyli prawdopodobieństwa odniesionego do jednostkowego przedziału wokół x .

Średnia wartość $\langle x \rangle$ jest zwykle zbliżona do x_0 , choć na ogół różna od niej.

Szczególnie ważną średnią jest tzw. odchylenie kwadratowe zdefiniowane jako

$$\sigma = \sqrt{\langle (x - x_0)^2 \rangle}. \quad (1.53)$$

Liczba ta określa "szerokość" rozkładu (rozmycie centralnego maksimum); w ten sposób definiuje się właśnie wprowadzoną wcześniej liczbę σ . W rachunku błędów σ nazywa się błędem średnim kwadratowym; liczby x_j to wyniki kolejnych pomiarów.

ROZDZIAŁ 2

PRAWA ZACHOWANIA

Początek XX w. rozpoczął erę niespotykanego postępu w rozwoju fizyki. Chociaż mechanika klasyczna rozwijała się prawie 400 lat, to jednak jej znajomość ma zasadnicze znaczenie do jasnego zrozumienia podstawowych praw współczesnej fizyki – np. teorii względności i mechaniki kwantowej.

Kinematyka, czyli badanie ruchu, została rozwinięta głównie przez Galileusza w XVII w., znakomitego włoskiego astronoma i matematyka. W najbardziej podstawowym znaczeniu kinematyka jest właściwie badaniem geometrii z dodatkiem nowego parametru – czasu. Badania przyczyn ruchu (dynamiki) rozwinął Newton w XVIII w., wielki angielski astronom, fizyk i matematyk.

Mechanika klasyczna z powodzeniem rozwiązała szeroki zakres problemów inżynierii, astronomii i fizyki. Rozwój fizyki pokazał jednak, że mechanika klasyczna nie może być stosowana uniwersalnie. Przejście do badań mikroskopowego świata atomów, elektronów, protonów, itp. pomogło w rozwoju nowych gałęzi fizyki, zwłaszcza teorii względności i mechaniki kwantowej. Teoria w fizyce nie jest traktowana jako prawda ostateczna, lecz jedynie jako model stosowany do rozwiązywania zagadnień i prowadzący do rozwiązań ściśle zgodnych z danymi eksperymentalnymi.

2.1. Zachowanie pędu

Wśród praw i modeli fizycznych najbardziej fundamentalnymi są prawa zachowania. Dzielą się one na dwie grupy: "zewnętrzne", elementarne prawa, do których należy prawo zachowania pędu, prawo zachowania momentu pędu, prawo zachowania energii; oraz "wewnętrzne" prawa zachowania, np. całkowitej liczby nukleonów w reakcji jądrowej, zachowanie liczby leptonowej, czy też barionowej.

Obecnie rozpatrzmy elementarne prawa zachowania w celu ustalenia podstaw do zagadnień omawianych w dalszych rozdziałach.

Zagadnienia mechaniki klasycznej można opisać traktując jako jej podstawę prawa Newtona, albo wychodząc z zasady zachowania pędu. Zastosujmy drugie podejście, gdyż zasada zachowania pędu jest prostsza a jej zastosowania są bardziej ogólne.

Jako układ odniesienia przyjmijmy układ Ox,y,z wzajemnie prostopadłych współrzędnych liniowych. Układ ten nazwiemy inercjalnym układem odniesienia; rozumiemy przez to, że

obowiązuje w nim mechanika klasyczna. Jeżeli możemy określić taki inercjalny układ odniesienia, to wszystkie inne układy odniesienia, które poruszają się względem niego ruchem jednostajnym prostoliniowym, są również układami inercjalnymi. *Istnienie "podstawowego układu odniesienia", jako takiego układu w którym spełnione są prawa Newtona, jest postulatem mechaniki newtonowskiej i teorii grawitacji, zwanym zasadą Macha.*

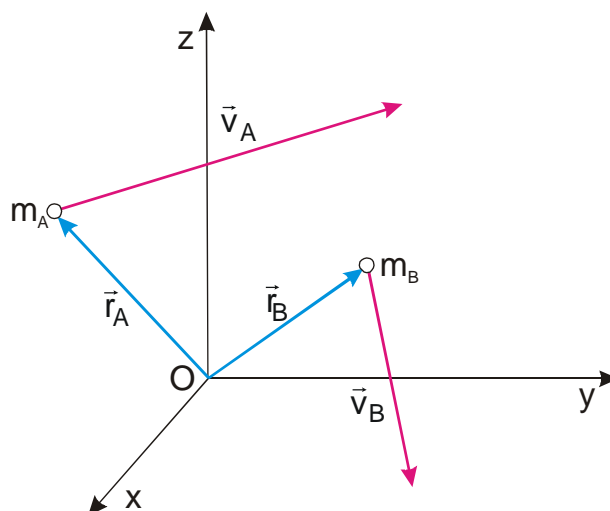
Pojęciem nierozłącznie związanym z pojęciem siły jest pojęcie masy bezwładnej (inercjalnej). *Masa bezwładna jest miarą oporu jaki stawia przyspieszane ciało.* Wiemy, że dla danej siły, im większa jest masa ciała na które ta siła działa, tym mniejszego przyspieszenia doznaje ciało. Klasycznie masę bezwładną traktujemy jako stałą uniwersalną, niezależną od wszelkich wpływów zewnętrznych, takich jak temperatura, ciśnienie lub prędkość.

Pęd cząstki o masie bezwładnej m poruszającej się z prędkością \vec{v} jest wektorem zdefiniowanym wzorem:

$$\vec{p} = m\vec{v} . \quad (2.1)$$

Jeżeli $\vec{i}, \vec{j}, \vec{k}$ oznaczać będą wektory jednostkowe równoległe odpowiednio do osi Ox, Oy, Oz układu współrzędnych oraz v_x, v_y, v_z oznaczać będą odpowiednie składowe wektora prędkości \vec{v} w tym układzie, to równanie (2.1) można będzie zapisać w następujący sposób:

$$\vec{p} = \vec{i}mv_x + \vec{j}mv_y + \vec{k}mv_z .$$



Rys. 2.1. Zasada zachowania pędu dla dwóch izolowanych cząstek $m_A\vec{v}_A + m_B\vec{v}_B = const$.

Zasada zachowania pędu brzmi: *całkowity pęd izolowanego układu cząstek pozostaje stały.* Przez układ izolowany rozumiemy układ wolny od jakiegokolwiek wpływu zewnętrznego. Dla układu izolowanego przedstawionego na rys. 2.1

$$m_A\vec{v}_A + m_B\vec{v}_B = const . \quad (2.2)$$

Dla układu złożonego z wielu cząstek mamy:

$$m_A \vec{v}_A + m_B \vec{v}_B + \dots + m_N \vec{v}_N = \sum_{i=1}^N m_i \vec{v}_i = \text{const} . \quad (2.3)$$

Z zasady zachowania pędu wyprowadzimy teraz trzy prawa ruchu Newtona.

W przypadku dwóch izolowanych cząstek zróżniczkowanie równania (2.2) względem czasu daje

$$m_A \frac{d\vec{v}_A}{dt} = -m_B \frac{d\vec{v}_B}{dt} .$$

Ponieważ $\vec{a} = d\vec{v} / dt$, mamy

$$m_A \vec{a}_A = -m_B \vec{a}_B . \quad (2.4)$$

Przyspieszenia są zatem odwrotnie proporcjonalne do mas bezwładnych, t.j. $a = F(1/m)$, gdzie F jest stałą proporcjonalności. Dochodzimy w ten sposób do definicji siły

$$\vec{F} = m\vec{a} . \quad (2.5)$$

Jest to drugie prawo Newtona: *siła działająca na ciało jest równa iloczynowi przyspieszenia i masy tego ciała*. W przypadku dwóch izolowanych cząstek oddziaływujących jedynie ze sobą (np. siłą elektryczną lub grawitacyjną), \vec{F}_A jest siłą jaką cząstka B wywiera na cząstkę A , a \vec{F}_B jest siłą jaką cząstka A wywiera na cząstkę B , czyli

$$\vec{F}_A = -\vec{F}_B .$$

Jest to zasada akcji i reakcji zwana trzecim prawem Newtona. Ponieważ dla pojedynczej swobodnej cząstki zarówno $\vec{F} = 0$, jak i $\vec{a} = 0$ oraz wiadomo że $\vec{a} = d\vec{v} / dt$, możemy więc wyciągnąć wniosek, że

$$\vec{v} = \text{const} .$$

Jest to sformułowanie prawa bezwładności, czyli pierwszego prawa Newtona: *ciało nie poddane oddziaływaniu żadnych innych ciał pozostaje w spoczynku, albo porusza się ruchem jednostajnym prostoliniowym*.

Drugie prawo Newtona można zapisać w postaci:

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

stąd otrzymujemy

$$\vec{F} dt = d(m\vec{v}) . \quad (2.6)$$

Jeżeli siła działa w ciągu skończonego czasu t , to mamy

$$\int_0^t \vec{F} dt = m\vec{v} - m\vec{v}_0 . \quad (2.7)$$

Całka ta zwana jest *popędem siły* \vec{F} . Widzimy, że jest ona równa zmianie pędu wywołanej działaniem siły w ciągu czasu t .

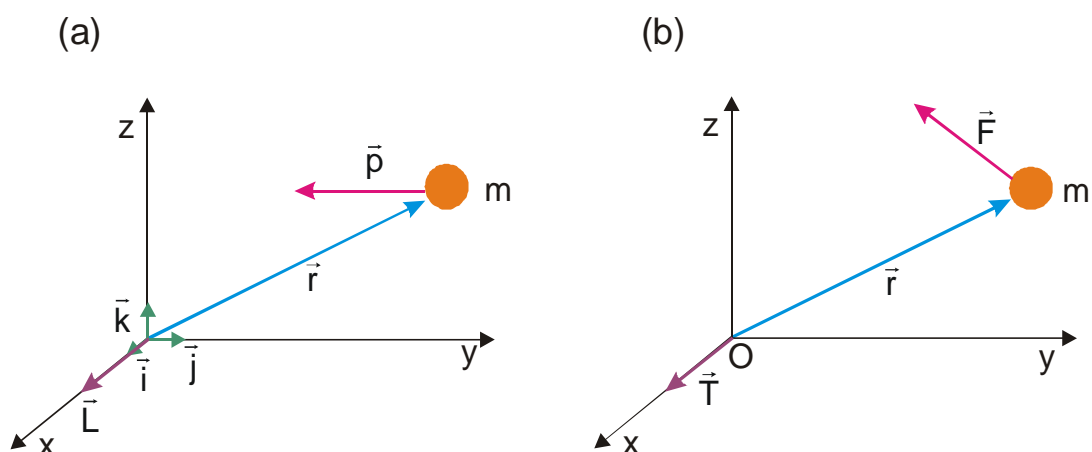
2.2. Zachowanie momentu pędu

Moment pędu (zwany też *krętem*) cząstki o pędzie \vec{p} i znajdującej się względem początku układu odniesienia O w punkcie określonym wektorem wodzącym \vec{r} (zwanym również wektorem położenia) jest wektorem zdefiniowanym wzorem:

$$\vec{L} = \vec{r} \times m \vec{v} = \vec{r} \times \vec{p}, \quad (2.8)$$

Definicję wektora momentu pędu wyjaśnimy za pomocą rys. 2.2a. Należy zauważyć, że moment pędu zależy od wyboru położenia punktu odniesienia. Wektor momentu pędu możemy wyrazić za pomocą wektorów jednostkowych i składowych pędu, jako

$$\vec{L} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x & y & z \\ p_x & p_y & p_z \end{vmatrix} = \vec{i}(yp_z - zp_y) + \vec{j}(zp_x - xp_z) + \vec{k}(xp_y - yp_x). \quad (2.8a)$$



Rys. 2.2. (a) Cząstka o masie m i pędzie \vec{p} w kierunku $-y$ będzie miała moment pędu $\vec{L} = \vec{r} \times \vec{p}$. (b) Cząstka o masie m , na którą działa siła \vec{F} (w płaszczyźnie yz) ma moment obrotowy względem początku układu równy $\vec{T} = \vec{r} \times \vec{F}$.

Przypomnijmy, że siła jest przyczyną ruchu postępowego. W podobny sposób moment siły (inaczej moment obrotowy), zwykle oznaczany symbolem \vec{T} , może być uważany za przyczynę ruchu obrotowego. Na rys. 2.2b, z siłą \vec{F} działającą na cząstkę, znajdującą się w położeniu określonym względem początku układu za pomocą wektora wodzącego \vec{r} , związany jest moment siły

$$\vec{T} = \vec{r} \times \vec{F}. \quad (2.9)$$

Aby otrzymać związek pomiędzy momentem pędu i momentem siły, różniczkujemy wyrażenie (2.8) względem czasu

$$\frac{d\vec{L}}{dt} = \frac{d\vec{r}}{dt} \times m\vec{v} + \vec{r} \times \frac{d}{dt}(m\vec{v}).$$

Ponieważ $d\vec{r}/dt = \vec{v}$, $(d\vec{r}/dt) \times m\vec{v} = 0$ oraz $\vec{F} = (d/dt)(m\vec{v})$, więc powyższe wyrażenie można sprowadzić do postaci

$$\frac{d\vec{L}}{dt} = \vec{r} \times \vec{F} = \vec{T}, \quad (2.10)$$

co oznacza, że *pochodna momentu pędu względem czasu t jest równa momentowi siły \vec{T} działającemu na tę cząstkę.*

Rozważając ruch planet stwierdzamy, że ciało stałe podlega działaniu siły przyciągania grawitacyjnego. Jest to siła skierowana zawsze wzdłuż promienia toru ciała pod warunkiem, że początek układu współrzędnych znajduje się w środku ciała przyciągającego. Ponieważ wektor wodzący \vec{r} i siła \vec{F} są wtedy zgodnie skierowane, więc $\vec{T} = \vec{r} \times \vec{F} = 0$ i z wyrażenia (2.10) otrzymujemy, że moment pędu \vec{L} takiego układu musi być wielkością stałą.

Dla układu wielu ciał i sił, wypadkowy moment siły jest równy

$$\vec{T} = \sum_{i=1}^n \vec{T}_i = \frac{d}{dt} \left(\sum_{i=1}^n \vec{L}_i \right). \quad (2.11)$$

Rozważmy układ na który nie działają zewnętrzne momenty sił. Nasza poprzednia analiza pokazała, że zgodnie z trzecim prawem Newtona momenty sił pochodzące od sił wewnętrznych działających pomiędzy dowolną parą cząstek, znoszą się wzajemnie. A więc wypadkowy moment siły jest zerowy:

$$\frac{d}{dt} \left(\sum \vec{L} \right) = 0$$

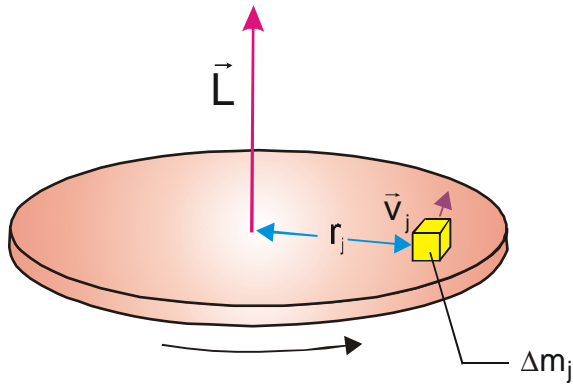
i dlatego

$$\sum \vec{L} = \text{const}. \quad (2.12)$$

Jest to prawo zachowania momentu pędu. Mówi ono, że *jeżeli wypadkowy moment sił zewnętrznych działających na układ jest równy zeru, to całkowity moment pędu tego układu jest stały.*

Rozważmy ciało stałe obracające się z prędkością kątową ω wokół przytwierdzonej osi przechodzącej przez środek masy ciała. Jeżeli element masy Δm_j położony jest w odległości r_j od osi obrotu, to jego prędkość $v_j = r_j \omega$ a moment pędu ciała jest sumą

$$L = \sum r_j \Delta m_j v_j = \sum r_j \Delta m_j (r_j \omega) = \left(\sum r_j^2 \Delta m_j \right) \omega.$$



Rys. 2.3. Obracający się dysk.

Wielkość w nawiasie nazywamy momentem bezwładności

$$I = \sum r_j^2 \Delta m_j.$$

W przypadku ciągłego rozkładu masy

$$I = \int r^2 dm. \quad (2.13)$$

Zatem:

$$L = I\omega. \quad (2.14)$$

Ponieważ moment siły określamy zależnością $T = dL/dt$, możemy więc napisać:

$$T = I \frac{d\omega}{dt} = I\alpha,$$

gdzie α oznacza przyspieszenie kątowe.

W układzie środka masy energia kinetyczna

$$K = \frac{1}{2} \sum \Delta m_j v_j^2 = \frac{1}{2} \sum \Delta m_j (r_j \omega)^2 = \frac{1}{2} \left(\sum \Delta m_j r_j^2 \right) \omega^2.$$

Tak więc

$$K = \frac{1}{2} I \omega^2, \quad (2.15)$$

czyli

$$K = \frac{1}{2} \frac{(I\omega)^2}{I} = \frac{1}{2} \frac{L^2}{I}$$

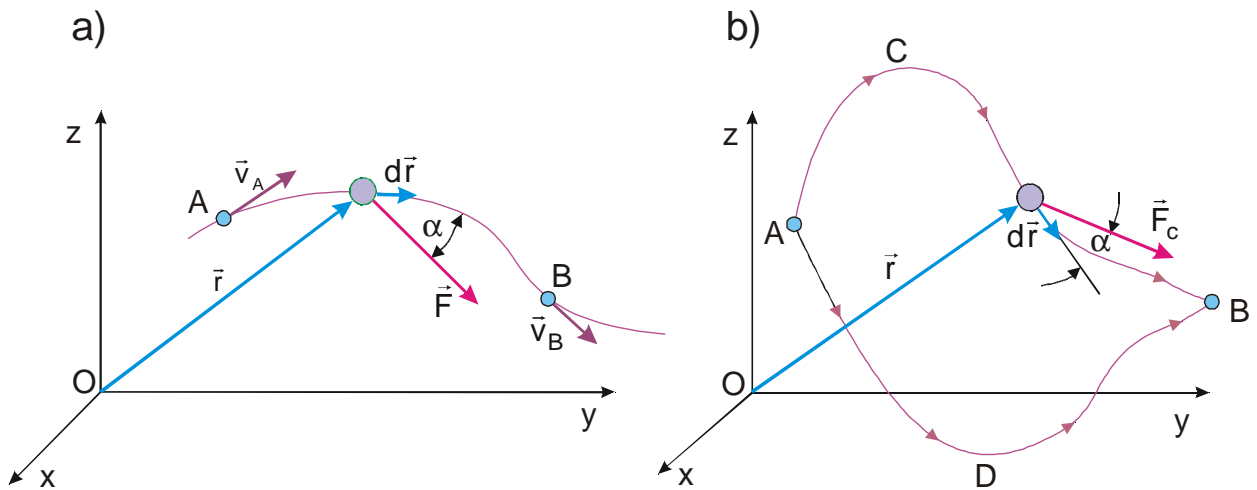
2.3. Zachowanie energii

Na rys. 2.4 przedstawiono cząstkę poruszającą się wzdłuż toru krzywoliniowego AB podlegającą działaniu siły \vec{F} , pod wpływem której przemieszcza się o odcinek \vec{r} . Różniczkowa praca siły F jest zdefiniowana jako:

$$dW = \vec{F} \cdot d\vec{r}. \quad (2.16)$$

Jeżeli siła F działa wzdłuż toru AB , to całkowita wykonana praca wynosi

$$W_{AB} = \int_A^B \vec{F} \cdot d\vec{r} = \int_A^B F \cos \alpha dr. \quad (2.17)$$



Rys. 2.4 (a) Praca wykonana przez siłę \vec{F} przy przesunięciu cząstki na odległość $d\vec{r}$ jest równa $dW = \vec{F} \cdot d\vec{r}$. (b) W przypadku siły zachowawczej \vec{F}_c praca $W_{AB} = \int_A^B \vec{F}_c \cdot d\vec{r}$ jest niezależna od toru łączącego punkty A i B.

Założmy, że \vec{F} jest wypadkową wszystkich sił działających na cząstkę. W tym wypadku

$$W_{AB} = \int_A^B \vec{F} \cdot d\vec{r} = \int_A^B m \frac{d\vec{v}}{dt} \cdot d\vec{r} = \int_A^B m\vec{v} d\vec{v},$$

ponieważ $d\vec{r} / dt = \vec{v}$. Po scałkowaniu otrzymujemy

$$W_{AB} = \int_{v_A}^{v_B} mvdv = \frac{1}{2}mv_B^2 - \frac{1}{2}mv_A^2 = K_B - K_A. \quad (2.18)$$

Wielkość $K = (1/2)mv^2$ nazywamy energią kinetyczną. Zasada równoważności pracy i energii mówi, że wypadkowa praca wykonana przez wszystkie siły działające na cząstkę równa jest odpowiedniej zmianie energii kinetycznej cząstki.

O sile \vec{F}_c na rys. 2.4b mówimy że jest siłą zachowawczą, jeżeli

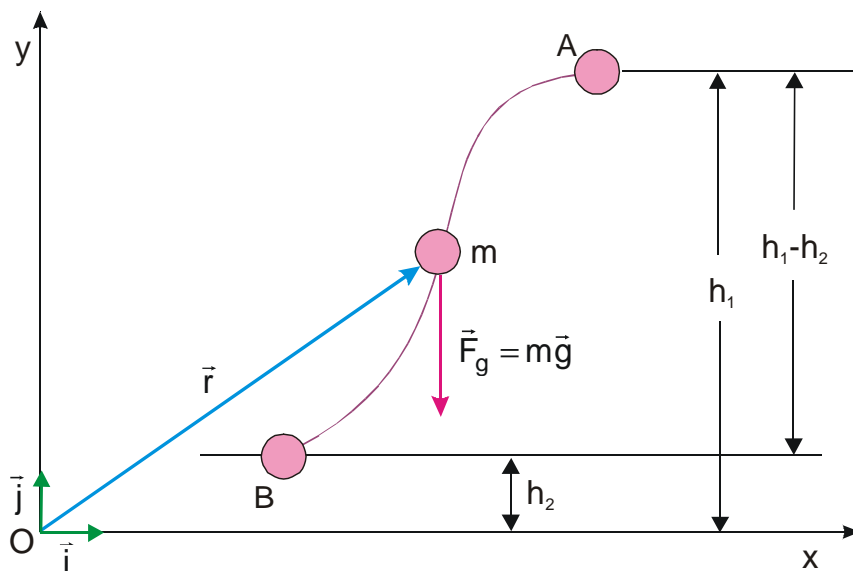
$$W_{AB} = \int_{ACB} \vec{F}_c \cdot d\vec{r} = \int_{ADB} \vec{F}_c \cdot d\vec{r} = const.$$

Powyższy wynik można wyrazić w następujący sposób: jeżeli praca wykonana przez siłę \vec{F}_c przemieszczającą cząstkę z punktu A do B jest niezależna od toru łączącego punkty A i B, to siła \vec{F}_c jest siłą zachowawczą. Dla przykładu, rozpatrzmy pracę wykonaną przez siłę grawitacyjną.

Na rys. 2.5 przedstawiono cząstkę o masie m poruszającą się od punktu A do B pod wpływem siły grawitacyjnej \vec{F}_g . Ponieważ $\vec{F}_g = -\vec{j}mg$, więc praca wykonana przez tę siłę jest równa

$$W_{AB} = \int_{h_1}^{h_2} (-\vec{j}mg) \cdot (\vec{i}dx + \vec{j}dy) = \int_{h_1}^{h_2} -mgdy = mg(h_1 - h_2) = mgh.$$

Ponieważ praca wykonana przez siłę grawitacyjną jest niezależna od tego po jakim torze porusza się cząstka między punktami A i B, więc jest to siła zachowawcza.



Rys. 2.5. Praca wykonana przez zachowawczą siłę grawitacyjną jest niezależna od drogi między punktami A i B.

Energię potencjalną definiujemy jako pracę wykonaną przez siłę zachowawczą

$$U_{AB} = \int_A^B \vec{F}_c \cdot d\vec{r} = U_A - U_B. \quad (2.19)$$

Skalarna funkcja położenia $U(x,y,z)$ jest funkcją energii potencjalnej związaną z siłą zachowawczą \vec{F}_c . Wielkości U_A i U_B są wartościami funkcji $U(x,y,z)$ wyznaczonymi w punktach końcowych toru. Energia potencjalna w każdym punkcie jest zdefiniowana za pomocą wyrażenia (2.19) w którym położenie B może być dowolnie wybrane. Zwykle B wybiera się w nieskończoności i przyjmuje, że $U_B = 0$. Wtedy energia potencjalna w dowolnym punkcie A wynosi

$$U_A = U_{AB} = \int_A^B \vec{F}_c \cdot d\vec{r} = - \int_B^A \vec{F}_c \cdot d\vec{r} = - \int_{\infty}^A \vec{F}_c \cdot d\vec{r}.$$

Energia potencjalna w dowolnym punkcie jest więc zdefiniowana jako praca wykonana przez równą, lecz przeciwnie skierowaną siłę, potrzebną do przemieszczenia cząstki z nieskończoności do danego punktu położenia.

Wróćmy do zasady równoważności pracy i energii wyrażonej wzorem (2.18): $W_{AB} = K_B - K_A$. Można ją teraz zapisać w taki sposób, aby włączyć zarówno siły zachowawcze jak i niezachowawcze.

$$W_{AB}(\text{zachowawcze}) + W_{AB}(\text{niezachowawcze}) = K_B - K_A. \quad (2.20)$$

Ponieważ z poprzednich rozważań wiemy, że $W_{AB}(\text{zachowawcze}) = U_A - U_B$, więc przegrupowując wyrazy we wzorze (2.19), otrzymamy:

$$W_{AB}(\text{niezachowawcze}) = (K_B - K_A) - (U_A - U_B),$$

lub

$$W_{AB}(\text{niezachowawcze}) = (K_B + U_B) - (K_A + U_A). \quad (2.21)$$

Jeżeli wszystkie siły są zachowawcze, czyli $W_{AB}(\text{niezachowawcze}) = 0$, to otrzymujemy

$$K_A + U_A = K_B + U_B = \text{const}. \quad (2.22)$$

Jest to prawo zachowania energii mechanicznej. Można je wyrazić w następujący sposób: *jeżeli wszystkie siły działające na cząstkę są zachowawcze, to całkowita energia cząstki w każdym jej położeniu jest wielkością stałą zwaną całkowitą energią mechaniczną.*

Jeżeli uwzględnimy wszystkie siły, zarówno zachowawcze jak i niezachowawcze, to praca wykonana przez siły niezachowawcze w wyrażeniu (2.21) pojawi się zawsze w postaci jakiejś formy energii. Jeżeli np. siła niezachowawcza jest siłą tarcia, to energia powstająca w wyniku jej działania ma postać energii wewnętrznej. Zasada zachowania energii w ujęciu uogólnionym, wynikającym z doświadczenia, brzmi: *energia układu izolowanego może przekształcać się z jednej postaci w inną, jednak energia całkowita w jej różnorodnych formach nie może być ani stworzona z niczego, ani też unicestwiona.*

ROZDZIAŁ 3

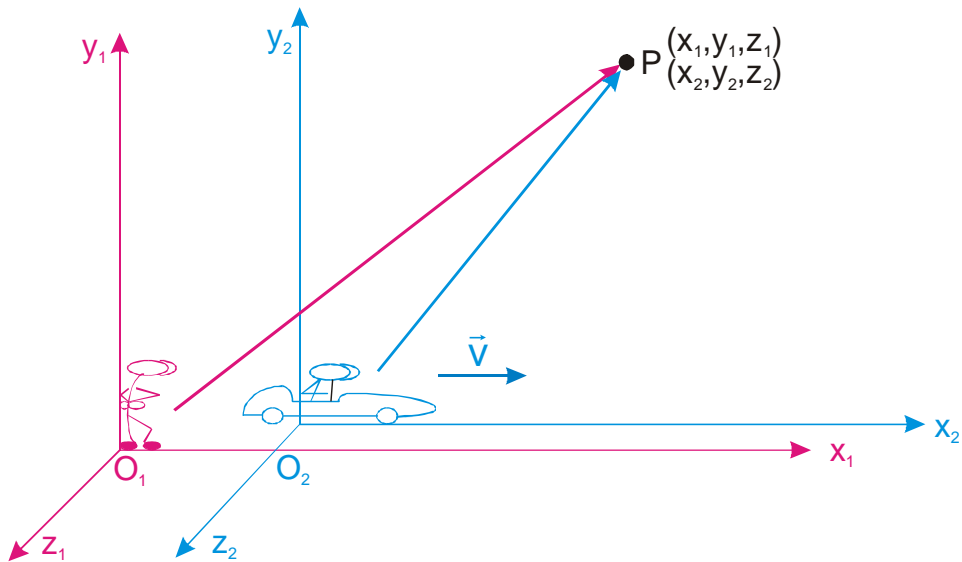
PODSTAWY MECHANIKI RELATYWISTYCZNEJ

W rozdziale tym omówimy podstawowe prawa szczególnej teorii względności Einsteina, które spowodowały rewolucyjny przewrót w poglądach na przestrzeń i czas. Teoria ta ma podstawowe znaczenie dla fizyki. Mechanika Newtona (zwana również mechaniką klasyczną) jest słuszna gdy prędkość ciała v jest znacznie mniejsza od prędkości światła c . W życiu codziennym spotykamy się z prędkościami znacznie mniejszymi od prędkości światła. Jednakże w wielu przypadkach $v \approx c$. I tak teoria światła wynika z teorii elektromagnetyzmu, a elektromagnetyzm to teoria relatywistyczna. W fizyce jądrowej i w fizyce cząstek elementarnych spotykamy się z cząstkami, które poruszają się z prędkością światła (np. foton, neutron). We wszechświecie galaktyki oddalają się z prędkościami bliskimi prędkości światła a natura gwiazd neutronowych, pulsarów i czarnych dziur związana jest z efektami relatywistycznymi. Również dla zrozumienia mechaniki kwantowej należy znać związki relatywistyczne pomiędzy masą, energią i pędem.

3.1. Niezmienniczość Galileusza

Różne zjawiska fizyczne opisujemy względem określonych układów odniesienia. Układy odniesienia możemy podzielić na dwie klasy – inercjalne i nieinercjalne. Układ inercjalny musi spełniać następujący warunek: *ciała lub układ ciał, na które nie działają żadne siły, musi być w spoczynku lub poruszać się ruchem jednostajnym prostoliniowym*. Możemy mieć bardzo dużo układów inercjalnych, które będą się poruszać względem siebie ruchem jednostajnym prostoliniowym. Wszystkie inne układy poruszające się z przyspieszeniem są układami nieinercjalnymi. Pierwsza zasada dynamiki Newtona nie jest prawem przyrody, lecz postulatem układu inercjalnego w przyrodzie. Fundamentalną trudność polegającą na tym, że do sformułowania praw mechaniki klasycznej koniecznym było postulowanie układu odniesienia, którego nie sposób zrealizować w praktyce, przezwyciężyła dopiero ogólna teoria względności Einsteina.

Dla wielu zjawisk, układ związany na stałe z Ziemią jest wystarczającym przybliżeniem układu inercjalnego. Przyspieszenie takiego układu odniesienia jest związane przede wszystkim z ruchem obrotowym Ziemi, jest to bardzo małe przyspieszenie.



Rys. 3.1. Punkt P nieruchomy w stacjonarnym układzie O_1 obserwowany jest z układu O_2 poruszającego się z prędkością \vec{v} względem układu O_1 .

Rozważmy teraz dwa układy inercjalne O_1 i O_2 poruszające się względem siebie wzdłuż osi x ze stałą prędkością \vec{v} (rys. 3.1). Pewien dowolny punkt P , który jest nieruchomy w układzie O_1 , będzie się poruszał w układzie O_2 z prędkością $-\vec{v}$. Zatem jego współrzędna x_2 będzie się zmieniała według zależności:

$$x_2 = x_1 - vt. \quad (3.1)$$

W ten sposób będzie się transformować współrzędna x dowolnego punktu przy przejściu od jednego do drugiego układu inercjalnego. Pozostałe współrzędne y i z pozostają bez zmian, ponieważ założyliśmy ruch układu O_2 wzdłuż osi x . Zatem

$$y_2 = y_1 \quad ; \quad z_2 = z_1. \quad (3.2)$$

Galileusz postulował, że czas biegnie jednakowo w obu układach, więc

$$t_2 = t_1. \quad (3.3)$$

Układ równań (3.1)+(3.3) wiążący współrzędne i czas dwóch układów inercjalnych nosi nazwę transformacji Galileusza. Są one transformacjami klasycznymi i jak się przekonamy, mogą być stosowane tylko w przypadku małych prędkości w porównaniu z prędkością światła. Czas we wszystkich układach inercjalnych jest taki sam, "płynie" tak samo.

Różniczkując stronami względem czasu, otrzymamy związki transformacyjne

$$\frac{dx_2}{dt} = \frac{dx_1}{dt} - v \frac{dt}{dt} \quad \text{czyli} \quad v_2 = v_1 - v.$$

W zapisie wektorowym

$$\vec{v}_1 = \vec{v}_2 + \vec{v}, \quad (3.4)$$

co opisuje klasyczne, galileuszowskie dodawanie prędkości. Różniczkując względem czasu związku dla wektorów prędkości, otrzymujemy ważną tezę mechaniki klasycznej: *przyśpieszenie jest niezmiennikiem względem transformacji Galileusza*

$$\frac{d\vec{v}_1}{dt} = \frac{d\vec{v}_2}{dt} + \frac{d\vec{v}}{dt},$$

czyli

$$\vec{a}_2 = \vec{a}_1 \quad \text{gdyż} \quad \frac{d\vec{v}}{dt} = 0. \quad (3.5)$$

Również prawo zachowania pędu pozostaje niezmiennicze we wszystkich układach inercjalnych. Na rys. 3.2 przedstawiono dwie cząstki o masach m oraz m' tworzące układ izolowany, na który nie działają siły zewnętrzne. Niech układ O_2 porusza się względem układu O_1 ze stałą prędkością \vec{v} wzdłuż osi x . Prawo zachowania pędu w układzie O_1 napiszemy w postaci

$$m\vec{v}_1 + m'\vec{v}'_1 = \text{const}, \quad (3.6)$$

gdzie \vec{v}_1 i \vec{v}'_1 są prędkościami odpowiednio masy m i m' . Niech teraz \vec{v}_2 i \vec{v}'_2 będą odpowiednio prędkościami tych samych dwóch cząstek względem układu O_2 .

Wiemy, że

$$\vec{v}_1 = \vec{v} + \vec{v}_2$$

$$\vec{v}'_1 = \vec{v} + \vec{v}'_2$$

Podstawienie tych wyrażeń do równania (3.6) daje, że

$$m(\vec{v} + \vec{v}_2) + m'(\vec{v} + \vec{v}'_2) = \text{const},$$

stąd

$$m\vec{v}_2 + m'\vec{v}'_2 = \text{const} - (m+m')\vec{v}.$$

Rys. 3.2. Całkowity pęd cząstek o masach m i m' jest wielkością niezmienniczą przy transformacji do układu inercjalnego O_2 .

Ponieważ $(m + m')\vec{v} = \text{const}$, więc

$$m\vec{v}_2 + m'\vec{v}'_2 = \text{const}. \quad (3.7)$$

Prawo zachowania pędu pozostaje niezmiennicze we wszystkich układach inercjalnych, poruszających się względem siebie ze stałymi prędkościami. Doszliśmy w ten sposób do bardzo ważnej zasady noszącej nazwę zasady względności Galileusza. Zasada ta brzmi: *istnieje nieskończenie wiele układów inercjalnych w których spełniona jest pierwsza i druga zasada dynamiki Newtona. Wszystkie te układy są równoważne i żaden z nich nie jest wyróżniony.*

3.2. Postulaty teorii względności

W końcu XIX w. Maxwell i Hertz zaproponowali koncepcję, że światło jest promieniowaniem elektromagnetycznym. Od kiedy stało się wiadome, że światło ma właściwości falowe, fizykom wydało się naturalną koniecznością zaproponowanie ośrodka, który ten ruch falowy miałby przenosić. Ośrodek ten był powszechnie zwany eterem świetlnym. W 1887 r. Michelson i Morley zaproponowali i przeprowadzili eksperyment w celu sprawdzenia natury eteru świetlnego i wyznaczenia prędkości światła względem niego. Gdyby eter istniał, to musiałby przenikać całą przestrzeń i musiałby być pierwotnym i bezwzględny układ odniesienia dla światła. Fizycy wyciągnęli zatem wniosek, że Ziemia musi albo pozostawać w spoczynku w tym eterze albo poruszać się względem niego, a w konsekwencji inercjalny układ odniesienia dla światła musi być w spoczynku lub poruszać się względem Ziemi.

Michelson i Morley użyli przyrządu zwanego obecnie interferometrem Michelsona, przedstawionego schematycznie na rys. 3.3. Ze źródła Z pada strumień światła na płytkę półprzezroczystą P , po czym rozdziela się na dwa promienie: 1 i 2. Promienie te odbijają się od zwierciadeł Z_1 i Z_2 . Promień 2 przechodzi częściowo przez płytkę P a promień 1, po częściowym odbiciu od niej, biegnie równoległe do promienia 2. Oba promienie wpadają do lunetki L , gdzie po przejściu przez soczewkę interferują ze sobą. Ponieważ odległości obu zwierciadeł od płytki P są jednakowe, więc oba promienie przechodzą tę samą drogę optyczną. Jeżeli jednak interferometr będzie się poruszał względem eteru w kierunku źródła światła Z - zwierciadło Z_1 , to czas przejścia promienia 1 nie będzie równy czasowi przejścia promienia 2. Sytuację tę przedstawia rys. 3.4. Prędkość światła względem eteru wynosi c . Zanim jednak światło dotrze do zwierciadła Z_1 , zwierciadło to zdąży przemieścić się o odcinek vt'_1 , gdzie t'_1 jest czasem biegu promienia 1 od punktu O do zwierciadła. Więc droga promienia w tym kierunku wyniesie:

$$ct'_1 = l + vt'_1,$$

stąd

$$t'_1 = \frac{l}{c - v}. \quad (3.8)$$

Podobnie czas powrotu t''_1 promienia 1 obliczymy z równania

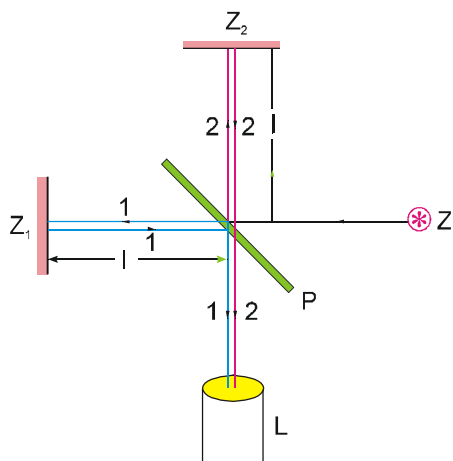
$$c t''_1 = l - vt''_1,$$

stąd

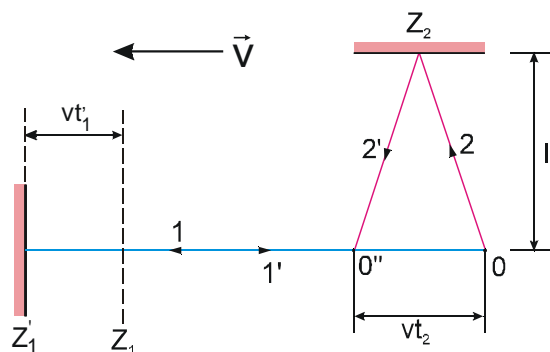
$$t''_1 = \frac{l}{c + v}. \quad (3.9)$$

Zatem, czas przebiegu promienia 1 wyniesie:

$$t_1 = t'_1 + t''_1 = \frac{2lc}{c^2 - v^2}. \quad (3.10)$$



Rys. 3.3. Schemat interferometru Michelsona użytego do wyznaczenia prędkości światła względem Ziemi.



Rys. 3.4. Względny ruch promienia światła odbijającego się od zwierciadeł Z_1 i Z_2 .

Podczas biegu promienia 2, punkt O przemieści się i zajmie położenie O'' . Promień ten przebiegnie więc drogę OZ_2O'' w czasie t_2 – droga ta wyniesie:

$$OZ_2O'' = 2 \cdot OZ_2 = 2\sqrt{l^2 + \left(\frac{OO''}{2}\right)^2} = 2\sqrt{l^2 + \left(\frac{vt_2}{2}\right)^2}.$$

Z drugiej strony droga ta jest równa:

$$OZ_2O'' = ct_2,$$

stąd otrzymamy

$$t_2 = \frac{2l}{\sqrt{c^2 - v^2}}. \quad (3.11)$$

Patrząc na obydwa wzory (3.10) i (3.11) widzimy, że $t_1 \neq t_2$. Promień 1 potrzebuje więcej czasu niż promień 2 na przebycie swojej drogi. Między tymi promieniami wystąpi różnica faz zależna od v . Przy odpowiedniej prędkości interferometru względem eteru i dobraniu odpowiedniej długości fali światła może wystąpić wygaszenie interferencyjne tam gdzie poprzednio było wzmocnienie.

Michelson i Morley zrealizowali doświadczenie w następujący sposób. Ustawili swój przyrząd ramieniem OZ_1 równoległe do orbitalnego biegu Ziemi wokół Słońca. Prędkość tego ruchu Ziemi względem gwiazd wynosi około 30 km/s. Przy długości ramion interferometru $l = 16$ m i tej prędkości, różnica dróg optycznych obu promieni powinna być równa około 1/4 długości fali światła zielonego. Jeżeli obrócić interferometr o 90° to ramiona zamieniają się miejscami i różnica dróg optycznych między promieniami w obu położeniach interferometru powinna wynieść połowę długości fali. Otóż, mimo wielokrotnego powtarzania tego doświadczenia w różnych warunkach, wynik był nieoczekiwany – nie udało się stwierdzić zmiany obrazu interferencyjnego, więc i ruchu Ziemi względem eteru. Dochodzimy zatem do

wniosku, że prędkość światła jest taka sama, niezależnie od tego czy jest ona mierzona przez obserwatora w układzie stacjonarnym, czy też przez obserwatora znajdującego się w układzie poruszającym się ze stałą prędkością względem światła. ”Negatywny” wynik doświadczenia Michelsona-Morleya spowodował przewrót w sposobie myślenia fizyków. Powstała konieczność głębszego spojrzenia w naturę przestrzeni i czasu.

Doświadczenie Michelsona-Morleya prowadzi również do wniosku, że nie ma wyróżnionego układu odniesienia, względem którego ośrodek w którym rozchodziłyby się fale elektromagnetyczne były w spoczynku. Z tego powodu Einstein odrzucił pojęcie eteru i w 1905 r. sformułował szczególną zasadę względności: *wszystkie prawa fizyki muszą być takie same we wszystkich układach inercjalnych poruszających się względem siebie ruchem jednostajnym prostoliniowym*. Nie można też stwierdzić bezwzględnego spoczynku jakiegokolwiek układu.

Teoria względności opiera się na dwóch postulatach:

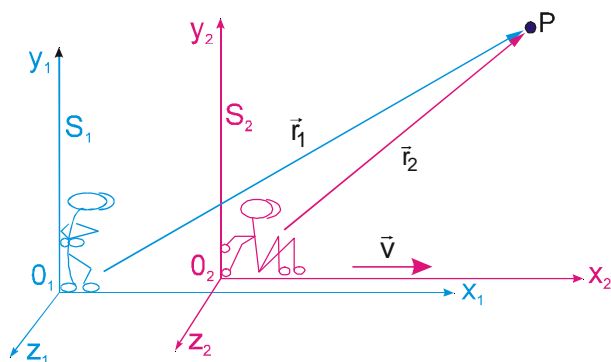
- szczególnej zasadzie względności;
- stałości prędkości światła w próżni.

3.3. Transformacje Lorentza

Transformacje Galileusza nie odpowiadają postulatowi teorii względności, ponieważ według nich prędkość światła jest równa w różnych układach inercjalnych. Później przedstawimy sposób wprowadzenia nowych transformacji, które będą odpowiadały postulatowi teorii względności; noszą one nazwę transformacji Lorentza.

Rozważmy rys. 3.5 na którym układ inercjalny O_1 jest w spoczynku, a układ inercjalny O_2 porusza się ruchem jednostajnym prostoliniowym z prędkością $v = \text{const}$ w kierunku osi x . W chwili $t_1 = t_2 = 0$ ze wspólnego początku układów O_1 i O_2 wysłany jest promień światła zdążający do punktu P o współrzędnych przestrzennych i czasowych równych odpowiednio:

$$x_1, y_2, z_1, t_1 \quad \text{i} \quad x_2, y_2, z_2, t_2$$



Rys. 3.5. Układ O_2 porusza się ze stałą prędkością względem stacjonarnego układu O_1 .

w układach O_1 i O_2 . Odległości r_1 i r_2 z odpowiednich początków układów do punktu P są określone wzorami:

$$r_1 = ct_1 \quad ; \quad r_2 = ct_2. \quad (3.12)$$

Jesteśmy więc zmuszeni przyjąć, że czasy t_1 i t_2 mierzone przez obserwatorów O_1 i O_2 są różne, pomimo że jest to sprzeczne z naszym życiowym doświadczeniem. Na podstawie wyrażeń (3.12) otrzymamy

$$x_1^2 + y_1^2 + z_1^2 = c^2 t_1^2 \quad (3.13)$$

$$x_2^2 + y_2^2 + z_2^2 = c^2 t_2^2$$

a z warunku symetrii $y_1 = y_2$ oraz $z_1 = z_2$. Uwzględniając to w wyrażeniach (3.13) mamy:

$$x_2^2 - c^2 t_2^2 = x_1^2 - c^2 t_1^2. \quad (3.14)$$

Związki transformacyjne pomiędzy układami O_1 i O_2 powinny spełniać pewne warunki formalne:

- transformacja musi być liniowa; tzn. zdarzeniu w jednym układzie inercyjnym musi odpowiadać pojedyncze zdarzenie w drugim układzie o jednoznacznie określonych współrzędnych;
- ruch jednostajny musi przekształcać się w ruch jednostajny;
- dla małych prędkości transformacja musi sprowadzić się do transformacji Galileusza.

W związku z powyższym

$$\begin{aligned} x_2 &= \gamma(x_1 - vt_1) \\ y_2 &= y_1 \\ z_2 &= z_1 \\ t_2 &= a(t_1 - b x_1) \end{aligned} \quad (3.15)$$

gdzie γ , a oraz b są stałymi. Wstawimy teraz wyrażenia (3.15) do równania (3.14), otrzymujemy

$$x_1^2 (\gamma^2 - a^2 b^2 c^2 - 1) + x_1 t_1 (-2\gamma^2 v + 2a^2 b c^2) + t_1^2 (\gamma^2 v^2 - a^2 c^2 + c^2) = 0.$$

Ponieważ wyrażenie to jest tożsamościowo równe zeru, więc

$$\begin{aligned} \gamma^2 - a^2 b^2 c^2 - 1 &= 0 \\ -2\gamma^2 v + 2a^2 b c^2 &= 0 \\ \gamma^2 v^2 - a^2 c^2 + c^2 &= 0 \end{aligned} \quad (3.16)$$

Powyższe równania rozwiązujemy względem stałych γ , a i b . Otrzymujemy wtedy

$$\gamma = a = \frac{1}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \quad (3.17)$$

oraz

$$b = \frac{v}{c^2}. \quad (3.18)$$

Wyrażenie $1/\sqrt{1-(v/c)^2}$ jest znane jako czynnik Lorentza i oznaczane zwykle symbolem γ . Natomiast stosunek v/c jest zazwyczaj oznaczany symbolem β .

Wzory transformacyjne (3.15) przybierają teraz postać:

$$\begin{aligned}x_2 &= \frac{x_1 - vt_1}{\sqrt{1 - \beta^2}} = \gamma(x_1 - vt_1) \\y_2 &= y_1 \\z_2 &= z_1 \\t_2 &= \frac{t_1 - \left(\frac{v}{c^2}\right)x_1}{\sqrt{1 - \beta^2}} = \gamma\left(t_1 - \frac{\beta}{c}x_1\right)\end{aligned}\tag{3.19}$$

Transformacje odwrotne są dane wzorami:

$$\begin{aligned}x_1 &= \frac{x_2 + vt_2}{\sqrt{1 - \beta^2}} = \gamma(x_2 + vt_2) \\y_1 &= y_2 \\z_1 &= z_2 \\t_1 &= \frac{t_2 + \left(\frac{v}{c^2}\right)x_2}{\sqrt{1 - \beta^2}} = \gamma\left(t_2 + \frac{\beta}{c}x_2\right)\end{aligned}\tag{3.20}$$

Transformacje te znane są jako transformacje Lorentza, od nazwiska holenderskiego fizyka H.A. Lorentza, który wprowadził je w 1890 r.

3.4. Konsekwencje transformacji Lorentza

3.4.1. Dodawanie prędkości

Łatwo możemy się teraz przekonać, że klasyczne prawo dodawania prędkości

$$\vec{v}_2 = \vec{v}_1 + \vec{v},\tag{3.21}$$

gdzie \vec{v}_2 oznacza prędkość w układzie O_2 , a \vec{v}_1 w układzie O_1 ; ma ograniczony zakres zastosowania tylko do małych prędkości. W celu wyprowadzenia prawa relatywistycznego musimy użyć transformacji Lorentza (3.19). Prędkość wzdłuż osi x w układzie O_1 wyraża się wzorem $v_1 = dx_1/dt_1$, zaś w układzie O_2 wzorem $v_2 = dx_2/dt_2$ (dla uproszczenia rozważamy przypadek jednowymiarowy). Różniczkując pierwsze i ostatnie równanie (3.19) otrzymujemy:

$$dx_2 = \frac{dx_1 - v dt_1}{\sqrt{1 - \beta^2}} \quad ; \quad dt_2 = \frac{dt_1 - \frac{v}{c^2} dx_1}{\sqrt{1 - \beta^2}},$$

stąd

$$v_2 = \frac{dx_2}{dt_2} = \frac{dx_1 - v dt_1}{dt_1 - \frac{v}{c^2} dx_1} = \frac{\frac{dx_1}{dt_1} - v}{1 - \frac{v}{c^2} \frac{dx_1}{dt_1}}.$$

Więc ostatecznie

$$v_2 = \frac{v_1 - v}{1 - \frac{v_1 v}{c^2}}. \quad (3.22)$$

Otrzymaliśmy wzór na dodawanie prędkości, który jest ogólniejszy od wzoru klasycznego. Wzór powyższy stosuje się w przypadku, gdy prędkości v_1 i v mają ten sam kierunek; w przeciwnym przypadku, musimy zmienić znaki "–" w tym wzorze na "+". Gdy $v_1/c \ll 1$, to $v_2 = v_1 - v$. Ze wzoru (3.22) wynika, że względna prędkość dwóch sygnałów świetlnych poruszających się w przeciwne strony wynosi: $(c + c)/(1 + c^2/c^2) = c$.

3.4.2. Skrócenie długości

Z teorii względności wynika, że ciała ruchome doznają skrócenia swoich wymiarów w kierunku ich ruchu. Załóżmy, że pewien pręt ma długość l_0 i porusza się wraz z układem O_2 . Pręt ten w układzie O_2 spoczywa i położony jest równoległe do osi x . Długość tego pręta jest wtedy zdefiniowana jako różnica między współrzędnymi jego końców A i B w układzie O_2

$$l_0 = x_{B2} - x_{A2}. \quad (3.23)$$

W układzie O_1 , pręt ten będzie się poruszał z prędkością v . W celu wyznaczenia długości l tego pręta w układzie O_1 należy wyznaczyć w tej samej chwili współrzędne jego końców x_{B1} i x_{A1} , gdy je pręt mija. Z transformacji Lorentza (3.19) mamy:

$$x_{B2} = \frac{x_{B1} - vt_1}{\sqrt{1 - (v/c)^2}} \quad ; \quad x_{A2} = \frac{x_{A1} - vt_1}{\sqrt{1 - (v/c)^2}}.$$

Wobec tego

$$x_{B1} - x_{A1} = (x_{B2} - x_{A2}) \sqrt{1 - (v/c)^2}.$$

Widzimy więc, że długość l_0 obserwator w układzie O_1 oceni jako

$$l = l_0 \sqrt{1 - \frac{v^2}{c^2}}, \quad (3.24)$$

a więc krótszą.

W sytuacji odwrotnej, gdy ten sam pręt jest nieruchomy w układzie O_1 , obserwator w układzie O_2 oceni jego długość jako krótszą. Ogólnie możemy powiedzieć, że *liniowe rozmiary ciała są największe w tym układzie, względem którego ciało spoczywa*. Skrócenie długości zachodzi tylko w przypadku długości mierzonych równoległe do kierunku ruchu względnego.

Należy tu wspomnieć, że Lorentz i Fitzgerald uważali, że zjawisko skrócenia (kontrakcji) przedmiotu znajdującego się w ruchu jest spowodowane pewną siłą działającą na ten przedmiot w czasie jego przechodzenia przez stacjonarny eter. Einstein przyjął zupełnie odmienny punkt widzenia, a mianowicie, że skrócenie to jest właściwością samej przestrzeni jako takiej i że bezwzględny lub wyróżniony spośród innych układ odniesienia nie istnieje.

3.4.3. Wydłużenie przedziałów czasowych

Następnym ważnym wnioskiem z teorii względności jest niejednakowy przebieg czasu w układach inercjalnych.

Wyobraźmy sobie zdarzenie zachodzące w punkcie A i chwili t_A (rys. 3.6). Rozważmy następnie inne zdarzenie zachodzące w tym samym punkcie, ale w innym czasie t_B . Oba zdarzenia są rejestrowane w układzie O_2 , w którym punkt A jest w spoczynku. Przedział czasu między dwoma zdarzeniami jest równy

$$T_2 = t_{B2} - t_{A2}. \quad (3.25)$$

Rozważmy teraz tę samą parę zdarzeń, ale obserwowaną z układu O_1 , poruszającego się równoległe do osi x układu O_2 , z prędkością względną $-v$. Oczywiście przedział czasowy obserwowany z układu O_1 jest określony wzorem

$$T_1 = t_{B1} - t_{A1}, \quad (3.26)$$

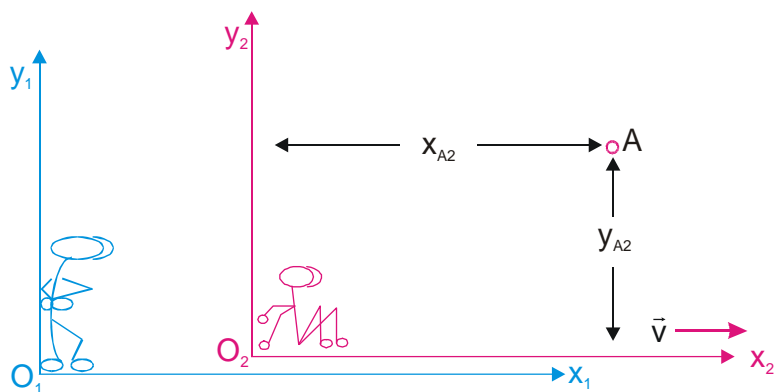
ale w układzie O_1 wartości współrzędnych przestrzennych pierwszego zdarzenia nie będą takie same jak drugiego zdarzenia, jak to było w układzie spoczynkowym O_2 . Chwili początkowej t_{A1} odpowiada współrzędna x_{A1} , chwili końcowej t_{B1} – współrzędna x_{B1} .

Z transformacji Lorentza (3.19) mamy

$$t_{A2} = \frac{t_{A1} - (v/c^2)x_{A1}}{\sqrt{1 - (v/c)^2}} \quad ; \quad t_{B2} = \frac{t_{B1} - (v/c^2)x_{B1}}{\sqrt{1 - (v/c)^2}}.$$

Po odjęciu stronami tych równości, otrzymujemy

$$T_2 = \frac{T_1 - (v/c^2)(x_{B1} - x_{A1})}{\sqrt{1 - (v/c)^2}}.$$



Rys. 3.6. Punkt A znajduje się w spoczynku względem układu O_2 . Dwa zdarzenia w punkcie A zachodzą według obserwatora O_2 w chwili t_{A2} i t_{B2} .

Ale $x_{B1} - x_{A1} = vT_1$ i wobec tego

$$T_2 = T_1 \sqrt{1 - (v/c)^2}, \quad (3.27)$$

czyli $T_2 < T_1$. Zatem przedział czasu, który rozdziela dwa następujące po sobie zdarzenia, w każdym układzie poruszającym się względem układu spoczywającego, jest dłuższy niż w układzie spoczywającym. Ponieważ mierzony przedział czasowy może zostać wydłużony jedynie wskutek zwolnienia chodu zegara używanego do pomiaru tego przedziału, więc powyższy wniosek jest równoważny stwierdzeniu, że poruszające się zegary chodzą wolniej niż zegary w spoczynku.

Powyższe wnioski stosują się do wszystkich zjawisk, również do procesów biologicznych. Dlatego efekt ten umożliwia w zasadzie "podróże w przyszłość". Kosmonauci wylatujący z prędkością bliską prędkości światła, po powrocie na Ziemię w czasie swojego życia stwierdziliby, że na Ziemi upłynął dużo dłuższy czas, tak że mogliby znaleźć się nawet w następnym stuleciu, w czasie gdy ich rówieśnicy dawno by już poumierali (paradoks bliźniąt).

Relatywistyczny efekt wydłużenia czasu został potwierdzony wielokrotnie doświadczalnie. Jednym z przykładów tego efektu jest rozpad nietrwałych cząstek elementarnych – mezonów. Średni czas życia spoczywającego mezonu μ wynosi $2,21 \times 10^{-6}$ s. Mezony te powstają w górnych warstwach atmosfery. Poruszają się one w kierunku Ziemi z prędkościami zbliżonymi do prędkości światła. Gdyby relatywistyczny efekt spowolnienia czasu nie wystąpił, to mogłyby one przebiegać stosunkowo niewielkie odległości rzędu 660 m, tymczasem przebywają one dużo większe odległości, docierając do powierzchni Ziemi.

3.5. Mechanika relatywistyczna

3.5.1. Masa i pęd

W mechanice klasycznej pęd ciała o masie bezwładnej m i prędkości \vec{v} jest definiowany jako

$$\vec{p} = m\vec{v}. \quad (3.28)$$

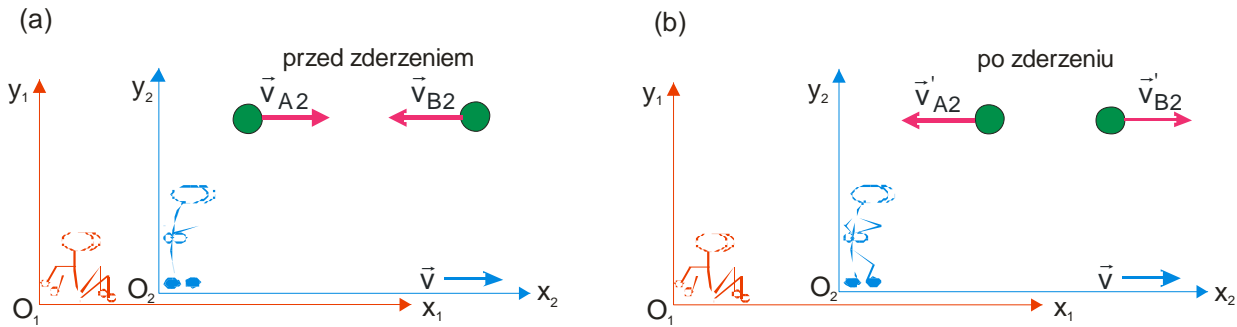
Prawo zachowania pędu izolowanego układu cząstek jest najbardziej fundamentalnym prawem fizyki. Izolowany układ cząstek m_1, m_2, \dots, m_n , nie podlegający działaniu sił zewnętrznych, będzie zachowywał się w czasie i przestrzeni w taki sposób, że

$$\sum_i m_i \vec{v}_i = \text{const}. \quad (3.29)$$

Kiedy zdarzenie obserwuje się z różnych, poruszających się układów odniesienia, nie ma powodu oczekiwać, że przestrzeń stanie się niejednorodna. Musimy zobaczyć, jak zachowuje się powyższe równanie gdy zastosujemy transformację Lorentza dla poruszających się układów współrzędnych. Przewidując komplikacje dotyczące masy jakie mogą powstać, przypisujemy masie specjalny symbol m_0 . Masa m_0 jest masą ciała w spoczynku, mierzona w naszym układzie odniesienia i nosi nazwę masy spoczynkowej.

Rozważmy dwie identyczne i idealnie sprężyste kule w poruszającym się układzie O_2 , każda o masie spoczynkowej m_0 (rys. 3.7). W poruszającym się układzie O_2 , kule A i B mają prędkości odpowiednio równe

$$\begin{aligned} \vec{v}_{A2} &= \vec{V} \\ \vec{v}_{B2} &= -\vec{V} \end{aligned} \quad (3.30)$$



Rys. 3.7 (a) Obserwator O_2 widzi dwie kule zbliżające się do siebie z jednakowymi prędkościami. (b) Obserwator O_2 będzie widział kule odskakujące z równymi ale przeciwnie skierowanymi prędkościami.

W wyrażeniach na prędkość zastosujemy transformację Lorentza, aby powiązać obserwacje tego samego zjawiska z obu układów. Z transformacji Lorentza wynika, że w układzie O_1

$$v_{A1} = \frac{v_{A2} + v}{1 + \beta(v_{A2}/c)} = \frac{V + v}{1 + \beta(V/c)}, \quad (3.31)$$

$$v_{B1} = \frac{v_{B2} + v}{1 + \beta(v_{B2}/c)} = \frac{-V + v}{1 - \beta(V/c)}, \quad (3.32)$$

gdzie $\beta = v/c$. Jeżeli suma mas obserwowana z układu O_1 jest równa M , to ta masa pozostanie stała w procesie zderzenia i w chwili zderzenia mamy

$$m_{A1}v_{A1} + m_{B1}v_{B1} = Mv, \quad (3.33)$$

gdzie $M = m_{A1} + m_{B1}$. Zatem, podczas gdy obserwator O_2 widzi masy chwilowo w spoczynku, obserwator O_1 widzi, że poruszają się one z prędkością v . Z ostatniego równania mamy

$$\begin{aligned} m_{A1}(v_{A1} - v_{B1}) &= M(v - v_{B1}) \\ m_{B1}(v_{B1} - v_{A1}) &= M(v - v_{A1}) \end{aligned} \quad (3.34)$$

Po zastosowaniu równań transformacyjnych (3.31) i (3.32) i uproszczeniu, stosunek wyrażeń (3.34) wynosi

$$\frac{m_{A1}}{m_{B1}} = \frac{1 + \beta(V/c)}{1 - \beta(V/c)}. \quad (3.35)$$

Korzystając z wyrażenia (3.31) otrzymujemy

$$1 - \frac{v_{A1}^2}{c^2} = 1 - \frac{(V + v)^2}{c^2[1 + \beta(V/c)]^2},$$

co można przekształcić algebraicznie, otrzymując

$$\sqrt{1 - \frac{v_{A1}^2}{c^2}} = \frac{\sqrt{1 - \beta^2} \sqrt{1 - (V/c)^2}}{1 + \beta(V/c)}.$$

Podobnie, korzystając z równania (3.32), otrzymujemy

$$\sqrt{1 - \frac{v_{B1}^2}{c^2}} = \frac{\sqrt{1 - \beta^2} \sqrt{1 - (V/c)^2}}{1 - \beta(V/c)}.$$

Czynniki $[1 + \beta(V/c)]$ oraz $[1 - \beta(V/c)]$ mogą być teraz wyznaczone z powyższych równań i podstawione do równania (3.35). Otrzymujemy w ten sposób następujące wyrażenie na stosunek dwóch mas widzianych z układu O_1

$$\frac{m_{A1}}{m_{B1}} = \frac{\sqrt{1 - (v_{B1}/c)^2}}{\sqrt{1 - (v_{A1}/c)^2}}. \quad (3.36)$$

Zatem masa widziana z poruszającego się układu odniesienia nie jest równa m_o , ale jest odwrotnie proporcjonalna do czynnika Lorentza $\gamma = 1/\sqrt{1 - (v/c)^2}$, który jest zawsze większy od 1, ale zbliża się do 1 gdy prędkość staje się bardzo mała w porównaniu z prędkością światła c . Pozwala to sformułować ogólne twierdzenie, że

$$m_{A1} \sqrt{1 - \frac{v_{A1}^2}{c^2}} = m_{B1} \sqrt{1 - \frac{v_{B1}^2}{c^2}} = m_o,$$

czyli po prostu

$$m = \frac{m_o}{\sqrt{1 - \beta^2}} = \gamma m_o. \quad (3.37)$$

Masa ciała nie jest zatem w ogólności stała ani taka sama dla wszystkich obserwatorów, ale jest wielkością która:

- zależy od układu odniesienia z jakiego jest obserwowana,
- jest równa m_o kiedy ciało jest w spoczynku w układzie odniesienia z którego jest obserwowane.

Właściwości czynnika Lorentza γ powodują, że masa staje się bardzo duża i w końcu zbliża się do nieskończoności, kiedy prędkość względna zbliża się do c .

Zgodnie ze wzorem na masę, wyrażenie relatywistyczne na pęd ma postać

$$\vec{p} = m\vec{v} = \gamma m_o \vec{v}, \quad (3.38)$$

a na zachowanie pędu układu izolowanego

$$\sum_{i=1}^n m_i \vec{v}_i = \sum_{i=1}^n \gamma m_{oi} \vec{v}_i = const. \quad (3.39)$$

3.5.2. Definicja siły

Chociaż prawa mechaniki klasycznej nie są na tyle uniwersalne, aby opisywały także zjawiska relatywistyczne, to drugie prawo Newtona

$$\vec{F} = \frac{d\vec{p}}{dt} = \frac{d}{dt}(m\vec{v}), \quad (3.40)$$

jest na tyle ogólne, że stosuje się również w mechanice relatywistycznej. Zróżniczkowanie równania (3.40) prowadzi do wyrażenia

$$\vec{F} = m \frac{d\vec{v}}{dt} + \vec{v} \frac{dm}{dt}, \quad (3.41)$$

gdzie m oznacza teraz γm_o .

3.5.3. Relatywistyczna energia kinetyczna

Aby znaleźć wyrażenie relatywistyczne na energię kinetyczną, obliczamy pracę wykonaną w celu zwiększenia prędkości cząstki od zera do końcowej wartości v . Żeby uprościć to zadanie, zakładamy, że siła i przesunięcie mają ten sam kierunek.

Energia kinetyczna, która jest rezultatem pracy wykonanej nad cząstką, jest więc równa

$$K = \int_0^r \vec{F} \cdot d\vec{r}. \quad (3.42)$$

Wyrażenie na działającą siłę ma postać

$$F = \frac{d}{dt}(mv) = \frac{d}{dt} \left[\frac{m_0 v}{\sqrt{1 - (v/c)^2}} \right] = \frac{m_0}{\sqrt{1 - (v/c)^2}} \frac{dv}{dt} + \frac{m_0 (v/c)^2}{[1 - (v/c)^2]^{3/2}} \frac{dv}{dt},$$

co upraszcza się do postaci

$$F = \frac{m_0}{[1 - (v^2/c^2)]^{3/2}} \frac{dv}{dt}. \quad (3.43)$$

Po uwzględnieniu tej zależności w wyrażeniu (3.42) mamy

$$K = \int_0^r \frac{m_0}{[1 - (v/c)^2]^{3/2}} \frac{d\vec{v}}{dt} d\vec{r}.$$

Ponieważ $d\vec{r} = \vec{v} dt$, $m = m_0 / \sqrt{1 - (v/c)^2}$, a $\vec{v} d\vec{v} = v dv$, więc

$$K = \int_0^r \frac{m_0}{[1 - (v/c)^2]^{3/2}} v dv,$$

co po scałkowaniu daje

$$K = \frac{m_0 c^2}{\sqrt{1 - (v/c)^2}} - m_0 c^2$$

i ostatecznie

$$K = (m - m_0) c^2. \quad (3.44)$$

Chociaż wyrażenie to wyprowadzono dla przypadku szczególnego, gdy siła ma ten sam kierunek co przesunięcie, to wyrażenie (3.44) jest ogólne i zawsze się stosuje.

Powyższe wyrażenie na energię kinetyczną możemy z łatwością zredukować do postaci $K = (1/2)mv^2$, kiedy $v \ll c$. Rozwijamy równanie (3.44) w szereg

$$K = m_0 c^2 \left[\left(1 - \frac{v^2}{c^2} \right)^{-1/2} - 1 \right] \approx m_0 c^2 \left(1 + \frac{1}{2} \frac{v^2}{c^2} + \frac{3}{8} \frac{v^4}{c^4} + \dots - 1 \right) \approx m_0 c^2 \left(\frac{v^2}{2c^2} + \frac{3}{8} \frac{v^4}{c^4} \right).$$

Kiedy $v/c \rightarrow 0$, to wyższe potęgi v/c można zaniedbać

$$K = m_o c^2 \left(\frac{1}{2} \frac{v^2}{c^2} \right) = \frac{1}{2} m_o v^2.$$

3.5.4. Energia całkowita

Z równania (3.44) wynika, że jeżeli ciało poruszające się z prędkością v_1 zwiększa swą prędkość do prędkości v_2 , to praca na to potrzebna, równa zmianie energii kinetycznej ciała, wynosi

$$K = (m_2 - m_o)c^2 - (m_1 - m_o)c^2 = (m_2 - m_1)c^2 = (\Delta m)c^2. \quad (3.45)$$

Zatem zmiana prędkości (energii kinetycznej) wywoła zmianę masy $\Delta m = m_2 - m_1$.

Dla ciała poruszającego się w polu sił zachowawczych, prawo zachowania energii ma postać

$$K_1 + U_1 = K_2 + U_2 = const, \quad (3.46)$$

gdzie K jest energią kinetyczną ciała w danym punkcie, a U jest energią potencjalną w tym samym punkcie. Na podstawie równań (3.45) i (3.46) otrzymujemy, że

$$K_2 - K_1 = U_1 - U_2 = (\Delta m)c^2,$$

czyli

$$\Delta m = \frac{K_2 - K_1}{c^2} = \frac{U_1 - U_2}{c^2}. \quad (3.47)$$

Ponieważ energia spoczynkowa jest zdefiniowana jako $E_o = m_o c^2$, energia całkowita będzie zdefiniowana jako

$$E = E_o + K, \quad (3.48)$$

a skoro $E = m_o c^2 + (m - m_o)c^2$, to

$$E = mc^2. \quad (3.49)$$

Należy zauważyć, że ta definicja energii całkowitej w mechanice relatywistycznej nie zawiera energii potencjalnej.

Równoważność masy i energii wyrażona wzorem (3.49) jest jedną z najważniejszych konsekwencji szczególnej teorii względności.

Inny użyteczny związek, który zawiera całkowitą energię E , może być otrzymany wprost z wyrażenia na masę $m\sqrt{1-(v/c)^2} = m_o$. Mnożąc obie strony tego wyrażenia przez c^2 , a następnie podnosząc do kwadratu i upraszczając, dostajemy

$$m^2 c^4 = m_o^2 c^4 + m^2 v^2 c^2.$$

Ponieważ $p = mv$, więc można to również zapisać w postaci

$$E^2 = E_o^2 + p^2 c^2. \quad (3.50)$$

Jeżeli ciało porusza się z bardzo dużą prędkością, to E_o^2 jest do pominięcia w porównaniu z $p^2 c^2$ i

$$E = pc. \quad (3.51)$$

Dla bardzo dużych prędkości, E_o jest też małe w porównaniu z K i z równania (3.49) wynika, że $E \approx K$, czyli na podstawie (3.48) mamy

$$K \approx pc. \quad (3.52)$$

Jeszcze inny związek zawierający całkowitą energię otrzymuje się różniczkując równanie (3.50)

$$\frac{dE}{dp} = \frac{pc^2}{E} = \frac{pc^2}{mc^2} = v. \quad (3.53)$$

Wobec tego, jeżeli ciało porusza się z prędkością światła, to $v = c$ oraz $dE = cdp$, czyli

$$E = pc + const.$$

Dla $p = 0$, $E = E_o$, a więc

$$E - E_o = pc. \quad (3.54)$$

Tymczasem z równania (3.50) wynika, że

$$E^2 - E_o^2 = p^2 c^2, \quad (3.55)$$

czyli biorąc pod uwagę obydwa równania, mamy

$$E + E_o = pc. \quad (3.56)$$

Porównanie równań (3.54) i (3.56) nasuwa wniosek, że $E_o = 0$, czyli $m_o = 0$. Inaczej mówiąc, jeżeli ciało porusza się z prędkością światła, to jego energia spoczynkowa oraz masa spoczynkowa muszą być równe zeru. Słuszne powinno być także stwierdzenie odwrotne: jeżeli jakaś materia nie ma masy spoczynkowej, czyli energii spoczynkowej, to musi poruszać się z prędkością światła. Chociaż z klasycznego punktu widzenia ciało nie może mieć masy równej zeru, to jest to poprawny relatywistyczny opis takich cząstek jak foton i neutrino.

3.6. Grawitacja a ogólna teoria względności

To co dotychczas rozważaliśmy, nazywamy szczególną teorią względności w odróżnieniu od ogólnej teorii względności. Pierwsza z nich była całkowicie rozwinięta przez Einsteina w 1905 r., podczas gdy druga – w zasadzie w 1911 r. Ogólna teoria względności stanowi współczesną teorię grawitacji.

Prawo powszechnego ciężenia, zwane również prawem grawitacji, zostało sformułowane przez Newtona w 1687 r. Prawo to mówi, że *dwa ciała o masach m_1 i m_2 przyciągają się wzajemnie siłą proporcjonalną do iloczynu ich mas i odwrotnie proporcjonalną do kwadratu ich odległości r , czyli*

$$F = G \frac{m_1 m_2}{r^2}. \quad (3.57)$$

Współczynnik proporcjonalności G nosi nazwę stałej grawitacji, która wynosi $6,67 \times 10^{-11} \text{ Nm}^2/\text{kg}^2$.

Ciężar ciała określimy z zależności

$$F = G \frac{Mm}{R^2}, \quad (3.58)$$

gdzie M jest masą Ziemi, m masą ciała, a R promieniem Ziemi. Korzystając z zależności $F = mg$ i wzoru (3.58), znajdujemy

$$g = G \frac{M}{R^2}. \quad (6.59)$$

Przyjmując $g \approx 9,81 \text{ m/s}^2$ i $R \approx 6400 \text{ km}$, z ostatniego wzoru otrzymujemy, że $M = 6 \times 10^{24} \text{ kg}$.

Siła grawitacyjna jest przykładem często spotykanej w fizyce siły centralnej, tzn. siły której kierunek działania przechodzi przez centrum i której wartość zależy od odległości od tego centrum. Innymi przykładami siły centralnej są siła elektrostatyczna i siła sprężystości.

Niezwykle ważną cechą siły centralnej jest to, że moment tej siły względem centrum wynosi zero, co oznacza, że w ruchu pod wpływem siły centralnej musi być zachowany moment pędu cząstki. Każda z sił centralnych jest siłą zachowawczą.

Można wykazać, że pod działaniem siły centralnej ciało może poruszać się po krzywej stożkowej: elipsie, paraboli lub hiperboli w zależności od warunków początkowych w jakich ciało to dostaje się w pole działania siły. Wynikają stąd sformułowane przez Keplera w latach 1609–1619, a więc jeszcze przed odkryciami Newtona, trzy prawa ruch planet. Kepler sformułował te prawa na podstawie obserwacji własnych oraz swoich poprzedników, M. Kopernika (1543 r.) i T. de Brahe (1576 r.). Nie będziemy ich jednak tutaj omawiać.

W teorii grawitacji Newtona przyjmuje się, że siła działa natychmiast, a to oznacza, że sygnał bądź energia przekazywane są natychmiast od masy m_1 do masy m_2 . W ten sposób naruszone jest

jedno z podstawowych założeń teorii względności, że żaden sygnał (żadna postać energii) nie może się rozchodzić z prędkością większą od prędkości światła. Równoważność masy i energii oznacza, że światło powinno ulegać przyciąganiu grawitacyjnemu. I tak np. światło przebiegające w pobliżu Słońca na drodze porównywalnej ze średnicą Słońca równej $1,44 \times 10^9$ m w czasie $t = 1,44 \times 10^9 / c = 5$ s powinno "spaść" o $h = (1/2)g_s t^2 = 3,7 \times 10^3$ m, gdzie g_s oznacza przyspieszenie w pobliżu Słońca. Wówczas kąt odchylenia promienia wynosi $\alpha = h/R_s = 5 \times 10^{-6}$ radiana czyli około $1''$. Zjawisko to zaobserwowano podczas zaćmienia Słońca jako zmianę położenia pewnych gwiazd.

Einstein stanął przed koniecznością sformułowania relatywistycznej teorii ciężkości. Sądził, że jego nowa teoria musi być zgodna z zasadą względności i automatycznie doprowadzić do równoważności mas grawitacyjnej i inercyjnej.

Ogólna teoria względności odgrywa dużą rolę w dziale astrofizyki zwanej kosmologią – nauce o powstaniu, rozmiarach i budowie Wszechświata (patrz roz. 20). Wyjaśnia ona takie zjawiska jak: zwiększenie długości fali przy emitowaniu światła przez ciała o dużej masie, zakrzywienie promienia świetlnego przechodzącego w pobliżu powierzchni Słońca w kierunku środka Słońca, czy też mechanizm powstawania tzw. "czarnych dziur" (specjalnego typu gwiazd).

ROZDZIAŁ 4

ELEKTROSTATYKA

W kolejnych rozdziałach zajmujemy się poznaniem, być może najważniejszego działu fizyki, dotyczącego oddziaływań elektromagnetycznych. Oddziaływania te nie tylko objaśniają wszystkie zjawiska elektryczne, ale również warunkują siły dzięki którym materia na poziomie atomowym i molekularnym istnieje jako całość. Również rozdziały poświęcone promieniowaniu i optyce wiążą się z oddziaływaniem elektromagnetycznym. Oddziaływania elektromagnetyczne badane są również w oparciu o mechanikę kwantową; pozwala to wyjaśnić istnienie i właściwości atomów, molekuł i ciał stałych. Wobec tego w pewnym sensie pozostała część skryptu poświęcona jest badaniu teorii elektromagnetyzmu i jego zastosowaniu.

4.1. Ładunek elektryczny

Wiemy, że materia zbudowana jest z elektronów, protonów i neutronów. Jeżeli obliczyć siłę oddziaływania grawitacyjnego między elektronem a protonem, znajdującymi się w odległości równej promieniowi atomu wodoru, to otrzymamy, że $F = Gm_p m_e / r^2 = 3,61 \times 10^{-47}$ N ($m_e = 9,11 \times 10^{-31}$ kg, $m_p = 1,67 \times 10^{-27}$ kg). Jednakże między elektronem i protonem działa jeszcze jedna siła – siła przyciągania równa $8,19 \times 10^{-8}$ N, to jest siła $2,27 \times 10^{39}$ razy większa. Ta dużo większa siła podlega również odwrotnie kwadratowej zależności od odległości. Siła ta nazywana jest siłą elektrostatyczną albo siłą elektryczną.

Odpychanie elektrostatyczne dwóch elektronów (lub dwóch protonów) dokładnie jest równe co do wartości sile przyciągania między elektronem i protonem, położonych w jednakowej odległości od siebie. W dużych obiektach ilość elektronów i protonów jest jednakowa i dlatego ogromne siły przyciągania i odpychania elektrostatycznego wzajemnie kompensują się i pozostaje jedynie słaba siła grawitacyjna. W tych warunkach oddziaływanie grawitacyjne dużych obiektów może okazać się silniejsze od oddziaływania elektrostatycznego (przykładem są "czarne dziury" we Wszechświecie).

Źródłem siły grawitacyjnej jest masa grawitacyjna, swego rodzaju ładunek grawitacyjny. Analogicznie, siła elektrostatyczna wywołana jest ładunkiem elektrycznym. W odróżnieniu od masy, ładunek elektryczny może być dodatni lub ujemny. Dwa ładunki o znakach przeciwnych przyciągają się, a dwa ładunki o jednakowych znakach – odpychają się.

Z eksperymentów wynika, że żadna z naładowanych cząstek nie charakteryzuje się ładunkiem, który byłby mniejszy od ładunku protonu lub elektronu. Ten elementarny ładunek

jest równy $1,60 \times 10^{-19}$ C i jest oznaczany symbolem e . Niektóre cząstki elementarne takie jak neutron, foton i neutrino, charakteryzują się zerowym ładunkiem elektrycznym. Naładowana cząstka ma ładunek skwantowany, tzn. równy całkowitej wielokrotności e .

W układzie zamkniętym całkowity ładunek pozostaje stały. Jest to prawo zachowania ładunku sformułowane przez Franklina w 1747 r. Prawo to jest spełnione nawet przy anihilacji naładowanych cząstek. Przy anihilacji elektronu z pozytonem znikają zarówno ładunek dodatni jak i ujemny, jednakże całkowity ładunek pozostaje zerowy zarówno przed jak i po anihilacji.

4.2. Prawo Coulomba

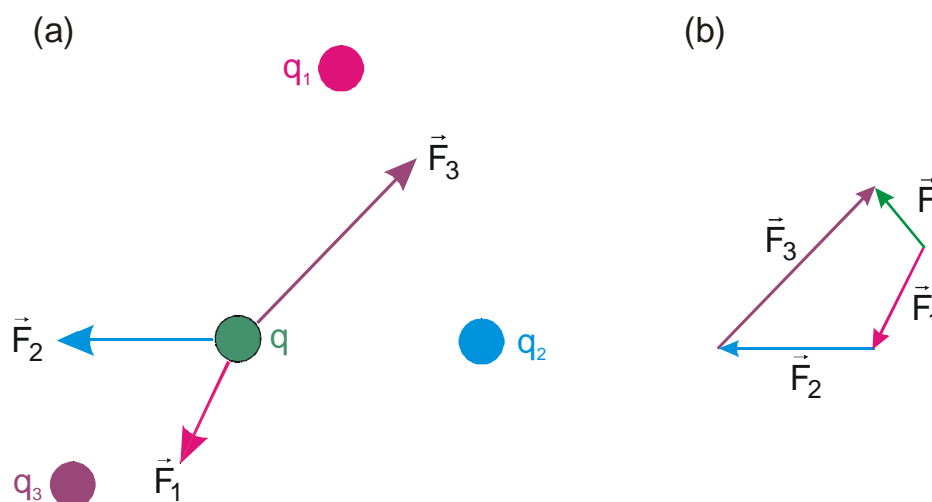
Podobnie jak dla siły grawitacyjnej określonej prawem powszechnego ciężenia, *siła działająca pomiędzy dwoma naładowanymi cząstkami jest proporcjonalna do iloczynu ładunków q_1 i q_2 i odwrotnie proporcjonalna do kwadratu odległości między nimi*

$$F = k_o \frac{q_1 q_2}{r^2}, \quad (4.1)$$

gdzie k_o jest współczynnikiem proporcjonalności zależnym od wyboru układu jednostek. W układzie SI ładunek określony jest przez siłę magnetyczną działającą pomiędzy dwoma jednakowymi elementami prądów (zob. pkt 6.5). Jednostką ładunku jest kulomb oznaczany C. Stała k_o w układzie SI wynosi $1/4\pi\epsilon_o$. Wówczas

$$F = \frac{1}{4\pi\epsilon_o} \frac{q_1 \cdot q_2}{r^2}, \quad (4.2)$$

gdzie $\epsilon_o = 1/4\pi k_o = 8,854 \times 10^{-12}$ C²/(Nm²). Wielkość tę nazywamy przenikalnością elektryczną próżni.



Rys. 4.1. (a) Siły działające na ładunek q za strony ładunków q_1, q_2, q_3 . (b) Wypadkowa siła otrzymana w wyniku dodania wektorowego sił działających na ładunek q .

Aby otrzymać siłę wypadkową działającą na wybrany ładunek w przypadku kilku naładowanych ciał, należy dodać wektorowo siły działające pomiędzy każdą parą ciał. Na rys. 4.1 na ładunek q działa siła $\vec{F} = \vec{F}_1 + \vec{F}_2 + \vec{F}_3$. Tę zasadę superpozycji dla przypadku sił elektrostatycznych potwierdzono eksperymentalnie (nie możemy jej wyprowadzić z bardziej fundamentalnych rozważań).

4.3. Pole elektryczne

Ażeby zmierzyć natężenie pola elektrycznego \vec{E} w pewnym punkcie P , należy umieścić w tym punkcie ładunek próbny q i zmierzyć siłę \vec{F} na niego działającą. Przy tym należy przekonać się, że obecność ładunku q nie zmieni położenia pozostałych ładunków

$$\vec{E} = \frac{\vec{F}}{q}. \quad (4.3)$$

Kierunek pola elektrycznego jest zgodny z kierunkiem siły działającej na ładunek dodatni. Wielkość E mierzona jest w niutonach na kulomb (N/C) lub, jak to pokażemy dalej, w woltach na metr (V/m).

Pole elektryczne ładunku punktowego Q w odległości r od niego określone jest wzorem:

$$\vec{E} = \frac{1}{q} \left(\frac{1}{4\pi \epsilon_0} \frac{Qq}{r^2} \frac{\vec{r}}{r} \right) = \vec{E}(x, y, z),$$

czyli

$$\vec{E} = \frac{1}{4\pi \epsilon_0} \frac{Q}{r^2} \vec{r}', \quad (4.4)$$

gdzie \vec{r}' jest wektorem jednostkowym skierowanym od ładunku Q do punktu $P(x, y, z)$.

Pole elektryczne pochodzące od n ładunków punktowych określone jest następującą sumą wektorową (zasada superpozycji)

$$\vec{E} = \frac{1}{4\pi \epsilon_0} \sum_{j=1}^n \frac{Q_j}{r_j^2} \vec{r}_j' = \sum_{j=1}^n \vec{E}_j(x, y, z). \quad (4.5)$$

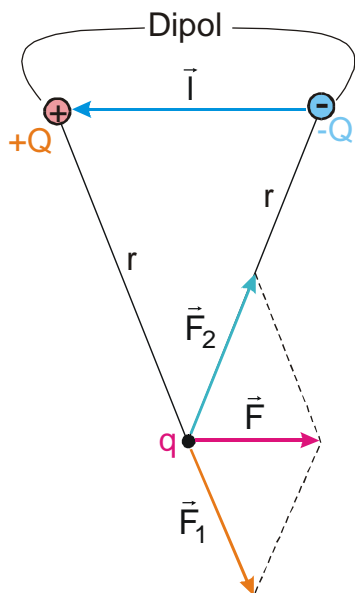
Jeżeli w przestrzeni mamy rozłożony w sposób ciągły ładunek o gęstości ładunku $\rho = dQ/dV = \rho(x, y, z)$ mierzony w jednostkach C/m³, to dla tego rozkładu ładunku

$$E = \frac{1}{4\pi \epsilon_0 \epsilon_r} \int_V \frac{\rho(x', y', z')}{r^2} dx' dy' dz'. \quad (4.6)$$

We wzorze tym ϵ_r określa względną przenikalność elektryczną ośrodka w stosunku do próżni (jest wielkością bezwymiarową).

Pojęciem gęstości ładunku możemy się posługiwać w odniesieniu do ładunków makroskopowych. W skali mikro (np. w atomie) gęstość ładunku zmienia się silnie od punktu do punktu i wtedy takie pojęcie traci sens.

Rozpatrzmy jeszcze pole elektryczne w punkcie P dla przypadku przedstawionego na rys. 4.2. Punkt P leży w jednakowej odległości od ładunków $+Q$ i $-Q$. Układ ładunków $+Q$ i $-Q$ położonych w odległości l od siebie nazywamy dipolem elektrycznym scharakteryzowanym



Rys. 4.2. Siły działające na ładunek q ze strony dipola o momencie $\vec{p} = Q\vec{l}$.

momentem dipolowym $\vec{p} = Q\vec{l}$. Zwrot wektora \vec{l} skierowany jest od ładunku ujemnego do dodatniego. Ze względu na podobieństwo trójkątów mamy $F / F_1 = l / r$, czyli

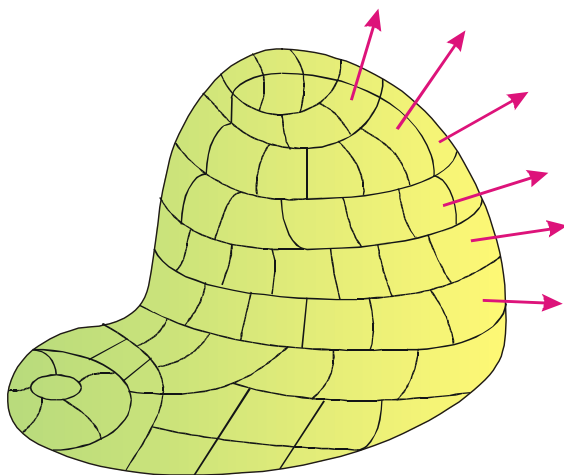
$$F = \frac{l}{r} F_1 = \frac{l}{r} \left(k_o \frac{Qq}{r^2} \right) = qk_o \frac{p}{r^3}. \quad (4.7)$$

Siła działająca ze strony dipola na ładunek q jest odwrotnie proporcjonalna do sześcianu odległości między nimi. Jeżeli uwzględnimy ten wzór w (4.3), to otrzymamy wyrażenie dla pola dipola

$$E = \frac{1}{4\pi \epsilon_o} \frac{p}{r^3}. \quad (4.8)$$

4.4. Strumień pola

Rozważmy jakieś pole elektryczne w przestrzeni. W przestrzeni tej wybieramy sobie pewną zamkniętą powierzchnię (rys. 4.3). Dzielimy następnie całą powierzchnię S na tak bardzo małe



Rys. 4.3. Strumień pola elektrycznego.

elementy dS , że powierzchnię dS można uważać za płaską. Każdemu elementowi dS przypisujemy wektor $d\vec{S}_j$ normalny do powierzchni i określający orientację elementu dS

$$|d\vec{S}_j| = dS; \quad \frac{d\vec{S}_j}{dS_j} = \vec{n}.$$

Niech \vec{E}_j będzie uśrednionym wektorem natężenia pola elektrycznego przez element powierzchni. Strumień natężenia pola elektrycznego przez powierzchnię $d\vec{S}_j$

$$d\Phi_E = \vec{E}_j \cdot d\vec{S}_j. \quad (4.9)$$

Całkowity strumień przez powierzchnię S będzie równy:

$$\Phi_E = \sum_j \vec{E}_j \cdot d\vec{S}_j = \int_S \vec{E} \cdot d\vec{S}. \quad (4.10)$$

W układzie jednostek SI jednostka strumienia natężenia pola elektrycznego ma wymiar Vm.

4.5. Prawo Gaussa

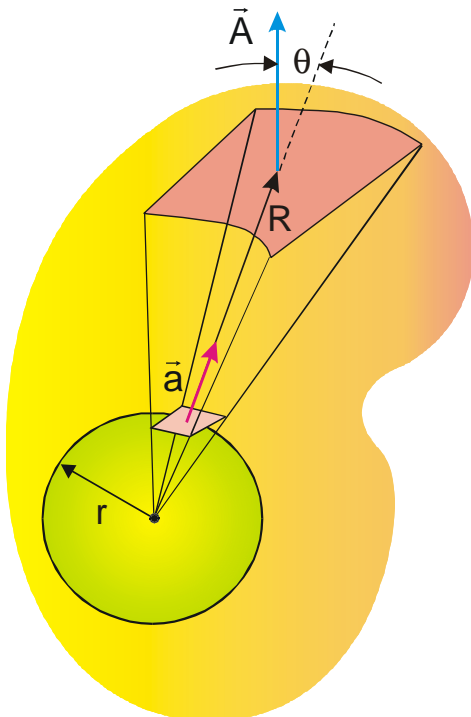
Wyprowadzimy prawo Gaussa w najprostszym przypadku, dla ładunku punktowego q otoczonego kulą o promieniu r i środkiem pokrywającym się z położeniem ładunku. Strumień Φ_E dla tego układu jest

$$\Phi_E = \int \vec{E} \cdot d\vec{S} = E4\pi r^2 = \frac{1}{4\pi \epsilon_0 \epsilon_r} \frac{q}{r^2} 4\pi r^2 = \frac{q}{\epsilon_0 \epsilon_r}. \quad (4.11)$$

Strumień pola nie zależy od wielkości powierzchni.

Pokażemy teraz, że całkowity strumień natężenia pola elektrycznego ładunku punktowego przez powierzchnię dowolnego kształtu będzie równy $q / \epsilon_0 \epsilon_r$.

Rozpatrzmy dowolną powierzchnię, która zawiera kulę wraz z ładunkiem (rys. 4.4) i udowodnimy, że całkowity strumień przez tę powierzchnię jest identyczny jak strumień przez powierzchnię kulistą.



Rys. 4.4 Strumień przez dowolną zamkniętą powierzchnię zawierającą ładunek q .

Stożek o wierzchołku w punkcie położenia ładunku q wycina mały element powierzchni \vec{a} z powierzchni kulistej i element powierzchni \vec{A} z powierzchni dowolnej. Powierzchnia elementu \vec{A} jest większa od powierzchni elementu \vec{a}

$$A = a \left(\frac{R}{r} \right)^2 \frac{1}{\cos \theta},$$

ze względu na ten sam kąt bryłowy

$$d\Omega = \frac{a}{r^2} = \frac{A}{R^2} \cos \theta,$$

oraz ze względu na nachylenie elementu \vec{A} do kierunku radialnego. Kąt θ jest kątem zawartym pomiędzy zewnętrzną normalną a kierunkiem radialnym. Strumień natężenia pola przez oba elementy jest równy

$$d\Phi_{E,a} = \vec{E}_r \cdot \vec{a} = E_r a,$$

oraz

$$d\Phi_{E,A} = \vec{E}_R \cdot \vec{A} = E_R A \cos \theta.$$

Wstawiając do równania na strumień $d\Phi_{E,A}$ wartości

$$E_R = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{q}{R^2} \quad \text{i} \quad A = a \left(\frac{R}{r} \right)^2 \frac{1}{\cos \theta},$$

dostajemy

$$d\Phi_{E,A} = \frac{q}{4\pi\epsilon_0\epsilon_r} a = E_r a. \quad (4.12)$$

Wynik ten oznacza, że strumienie przez oba elementy są równe. Również całkowity strumień przez obie powierzchnie będzie jednakowy, a więc strumień natężenia pola przez dowolną zamkniętą powierzchnię otaczającą ładunek q będzie równy $q/\epsilon_0\epsilon_r$.

Jeżeli ładunek leży na zewnątrz zamkniętej dowolnej powierzchni, to strumień przez tę powierzchnię znika. Jeżeli mamy n ładunków punktowych objętych powierzchnią, to strumień przez tę powierzchnię wynosi:

$$\Phi_E = \sum_{i=1}^n \frac{q_i}{\epsilon_0\epsilon_r}. \quad (4.13)$$

Skorzystaliliśmy z zasady superpozycji pól elektrycznych pochodzących od poszczególnych ładunków q_1, q_2, \dots, q_n .

W przypadku ładunku o gęstości objętościowej $\rho(x,y,z)$

$$\int_S \vec{E} \cdot d\vec{S} = \frac{1}{\epsilon_0\epsilon_r} \int_V \rho dV. \quad (4.14)$$

Prawo Gaussa brzmi: *strumień natężenia pola elektrycznego przez dowolną powierzchnię zamkniętą równa się iloczynowi całkowitego ładunku zamkniętego w tej powierzchni przez $q/\epsilon_0\epsilon_r$.*

4.6. Niektóre zastosowania twierdzenia Gaussa

4.6.1. Równomiernie naładowana kula

Ponieważ równomiernie naładowaną kulę można traktować jako składającą się z szeregu koncentrycznych warstw, więc przy obliczaniu pola wewnątrz kuli można stosować wzór (4.4). Zauważmy, że kładąc $r = R$ w równanie (4.4) mamy natężenie pola na powierzchni kuli

$$E = \frac{1}{4\pi\epsilon_o} \frac{Q}{R^2}, \quad (4.15)$$

gdzie Q jest całkowitym ładunkiem kuli.

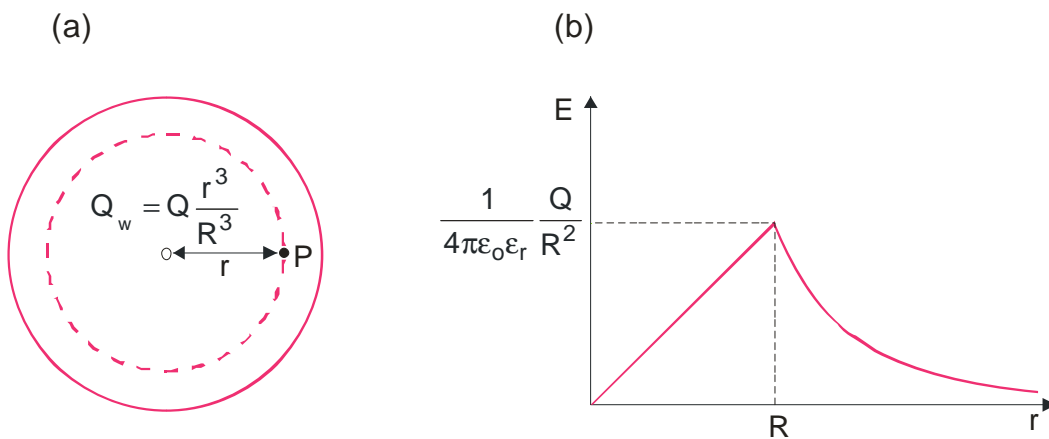
W celu obliczenia pola E w dowolnym punkcie P znajdującym się wewnątrz kuli wybieramy powierzchnię gaussowską przechodzącą przez ten punkt P , jak pokazano na rys. 4.5. Sfera ta obejmuje objętość $4\pi r^3/3$, która stanowi $(r/R)^3$ całej objętości kuli. Wobec tego ładunek wnętrza tej sfery wynosi $Q_w = Q(r/R)^3$. Stosując twierdzenie Gaussa

$$E4\pi r^2 = \frac{1}{\epsilon_o\epsilon_r} \frac{Q}{R^3} r^3,$$

otrzymujemy pole wnętrza równomiernie naładowanej kuli o promieniu R

$$E = \frac{1}{4\pi\epsilon_o\epsilon_r} \frac{Q}{R^3} r. \quad (4.16)$$

Na rys. 4.5b pokazano zależność tego pola od r .



Rys. 4.5. (a) Powierzchnia gaussowska przechodząca przez P obejmująca ładunek Q_w .
(b) Zależność pola elektrycznego od odległości od środka równomiernie naładowanej kuli.

4.6.2. Powierzchniowy rozkład ładunku

Rozważymy pole elektryczne wytworzone przez równomiernie naładowaną nieskończoną płaszczyznę o gęstości powierzchniowej ładunku σ (jednostką σ jest C/m^2). Powierzchnię gaussowską wybieramy w postaci prostopadłościanu lub cylindra o płaskich przekrojach poprzecznych S_o położonych w odległości a od powierzchni, jak pokazano na rys. 4.6. Ładunek znajdujący się wewnątrz powierzchni całkowania równy jest $Q_w = \sigma S_o$. Strumień pola wychodzący w obydwie strony naładowanej płaszczyzny jest jednakowy, więc całkowity strumień natężenia pola elektrycznego wynosi

$$\int \vec{E} \cdot d\vec{S} = 2ES_o$$

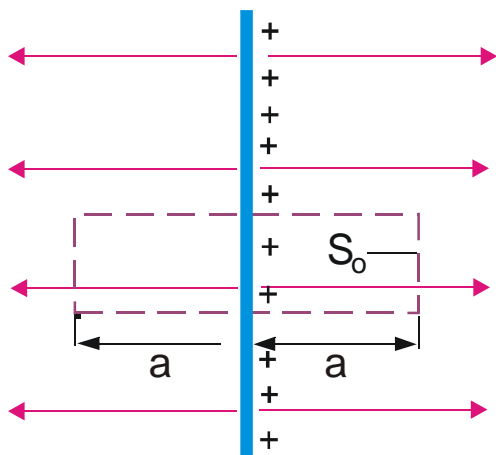


Fig. 4.6. Nieskończona powierzchnia metalowa o gęstości powierzchniowej ładunku σ .

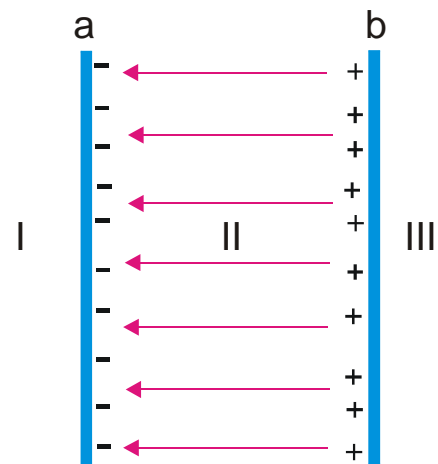


Fig. 4.7. Pole elektryczne między dwoma płaszczyznami o równych gęstościach ładunku powierzchniowego σ lecz przeciwnych znakach.

Zgodnie z twierdzeniem Gaussa

$$2ES_o = \frac{\sigma S_o}{\epsilon_o \epsilon_r},$$

czyli pole elektryczne naładowanej płaszczyzny jest równe

$$E = \frac{\sigma}{2\epsilon_o \epsilon_r}. \quad (4.17)$$

W praktyce często spotykamy się z przyrządami, w których znajdują się dwie równoległe płaszczyzny naładowane równymi lecz przeciwnymi ładunkami (rys. 4.7). Natężenie pola spowodowane ładunkiem płaszczyzny a wynosi $E_a = \sigma / 2\epsilon_o \epsilon_r$ i jest skierowane do tej płaszczyzny. Pole wytworzone przez płaszczyznę b wynosi $E_b = \sigma / 2\epsilon_o \epsilon_r$ i jest skierowane od tej płaszczyzny.

W obszarze I:

$$E_I = E_{aI} + E_{bI} = \frac{\sigma}{2\epsilon_o \epsilon_r} - \frac{\sigma}{2\epsilon_o \epsilon_r} = 0.$$

W obszarze II:

$$E_{II} = E_{aII} + E_{bII} = -\frac{\sigma}{2\epsilon_o \epsilon_r} - \frac{\sigma}{2\epsilon_o \epsilon_r} = -\frac{\sigma}{\epsilon_o \epsilon_r}. \quad (4.18)$$

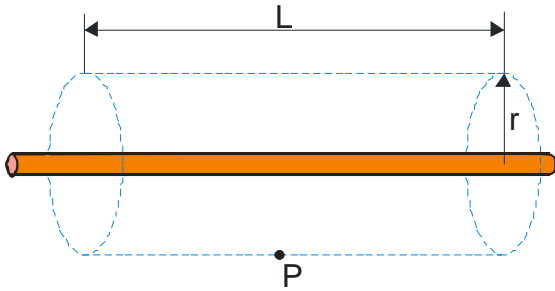
W obszarze III:

$$E_{III} = E_{aIII} + E_{bIII} = -\frac{\sigma}{2\epsilon_o \epsilon_r} + \frac{\sigma}{2\epsilon_o \epsilon_r} = 0.$$

Widzimy więc, że na zewnątrz płaszczyzn pole elektryczne znika, natomiast między płaszczyznami wynosi $\sigma / \epsilon_0 \epsilon_r$.

4.6.3. Liniowy rozkład ładunku

Rozpatrzmy teraz pole elektryczne wytworzone w odległości r przez równomiernie naładowany prostoliniowy przewodnik lub pręt, którego długość wyraźnie przewyższa odległość r . Niech λ oznacza ładunek przypadający na jednostkę długości pręta. Jako powierzchnię gaussowską wybieramy walec o długości L (rys. 4.8). Wewnątrz powierzchni walcowej znajduje się ładunek $Q_w = \lambda L$.



Rys. 4.8. Odcinek długiego naładowanego pręta. Powierznię gaussowską stanowi walec o długości L i promieniu r .

Zgodnie z prawem Gaussa

$$\oint \vec{E} \cdot d\vec{S} = \frac{\lambda L}{\epsilon_0 \epsilon_r}.$$

Ze względu na symetrię, linie sił pola mają kształt prostych radialnych. Dlatego wektory \vec{E} i $d\vec{S}$ są wzajemnie prostopadłe na bocznej powierzchni zamykającej walec i równoległe na powierzchni walcowej. Z tego powodu możemy napisać

$$\oint \vec{E} \cdot d\vec{S} = E 2\pi r L.$$

Porównując to wyrażenie do $\lambda L / \epsilon_0 \epsilon_r$, mamy

$$2\pi r L E = \frac{\lambda L}{\epsilon_0 \epsilon_r},$$

stąd natężenie pola elektrycznego liniowo rozłożonego ładunku ma postać

$$E = \frac{\lambda}{2\pi \epsilon_0 \epsilon_r r}. \quad (4.19)$$

4.6.4. Powierzchnia przewodnika

Większość ciał stałych możemy podzielić na dwie grupy: przewodniki i izolatory (dielektryki). Dodatkowy ładunek umieszczony na powierzchni lub wewnątrz dielektryka pozostaje nieruchomy. Inaczej jest w przewodnikach, które zawierają dużą liczbę swobodnych elektronów nie związanych z konkretnymi atomami. Dlatego w przewodniku pole elektryczne może istnieć jedynie w ciągu krótkiego okresu czasu dopóki swobodne elektrony nie zgromadzą się na powierzchni przewodnika pod wpływem działania zewnętrznego pola i nie utworzą przeciwnie skierowanego pola.

Na rys. 4.9 przedstawiono przekrój przewodnika o dowolnym kształcie. Bezpośrednio pod powierzchnią przewodnika wybieramy powierzchnię gaussowską S zaznaczoną linią przerywaną.

Zgodnie z twierdzeniem Gaussa

$$\int_S \vec{E} \cdot d\vec{S} = \frac{Q_w}{\epsilon_0 \epsilon_r}.$$

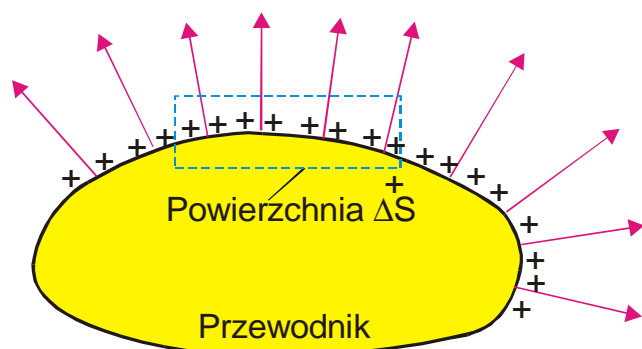
Ustalenie się stanu równowagi ładunkowej w przewodniku oznacza, że w jego wnętrzu pole elektryczne jest równe zero, a w związku z tym ładunek wewnętrzny przewodnika $Q_w = 0$. Linie sił pola elektrycznego na powierzchni przewodnika są skierowane prostopadle do powierzchni. Gdyby składowa \vec{E} wzdłuż powierzchni nie zniknęła to pojawiłby się prąd elektryczny.

Wydzielmy na powierzchni (rys. 4.9) nieduży prostopadłościan o podstawie ΔS . Zgodnie z twierdzeniem Gaussa

$$E \Delta S = \frac{\sigma \Delta S}{\epsilon_0 \epsilon_r},$$

czyli natężenie pola na powierzchni przewodnika wynosi

$$E = \frac{\sigma}{\epsilon_0 \epsilon_r}. \quad (4.20)$$



Rys. 4.9. Wewnątrz prostopadłościanu o podstawie ΔS znajduje się ładunek $\sigma \Delta S$.

4.7. Potencjał elektryczny

Rozpatrzmy całkę natężenia pola elektrycznego \vec{E} po krzywej łączącej punkty A i B . Pokażemy, że całka

$$\int_A^B \vec{E} \cdot d\vec{s} = const ,$$

przybiera tę samą wartość dla wszystkich dróg łączących punkty A i B .

Rozpatrzmy najpierw pole elektrostatyczne pochodzące od ładunku punktowego. Ponieważ zależność pomiędzy siłą \vec{F} a wektorem \vec{E} jest $\vec{F} = q\vec{E}$, więc praca sił pola elektrostatycznego na drodze od punktu A do punktu B

$$W_{AB} = \int_A^B \vec{F} \cdot d\vec{s} = q \int_A^B \vec{E} \cdot d\vec{s} = -q \int_B^A \vec{E} \cdot d\vec{s} = U_A - U_B \quad (4.21)$$

i jest równa zmianie energii potencjalnej pola elektrostatycznego. Analogicznie jak dla energii potencjalnej pola grawitacyjnego (zob. pkt. 2.3) przyjmujemy $U = 0$, gdy ładunek znajduje się w nieskończoności. Wówczas

$$U_A = -q \int_{\infty}^A \vec{E} \cdot d\vec{s}. \quad (4.22)$$

Jeżeli przesuniemy ładunek q z nieskończoności do punktu położonego w odległości r od ładunku punktowego Q , to energia potencjalna będzie równa pracy przeciwko sile elektrostatycznej

$$U = -q \int_{\infty}^r \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{r^2} dr = -\frac{1}{4\pi\epsilon_0\epsilon_r} qQ \left[-\frac{1}{r} \right]_{\infty}^r.$$

Wobec tego energia potencjalna ładunku punktowego q umieszczonego w polu ładunku Q wynosi

$$U = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{qQ}{r}. \quad (4.23)$$

Potencjał elektryczny V określamy jako energię potencjalną jednostkowego ładunku

$$V = \frac{U}{q}. \quad (4.24)$$

Jednostką potencjału elektrycznego jest J/C; w układzie jednostek SI nazywamy ją woltem i oznaczamy przez V. Na podstawie wyrażenia (4.24) otrzymujemy wzór określający potencjał ładunku punktowego Q jako

$$V = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{Q}{r}. \quad (4.25)$$

Potencjał elektryczny jest więc pracą jaką należy wykonać aby przesunąć ładunek jednostkowy z nieskończoności na odległość r od ładunku punktowego Q . Potencjał elektryczny równy jest energii potencjalnej przypadającej na jednostkowy ładunek, podobnie jak natężenie pola elektrycznego równe jest sile działającej na jednostkowy ładunek.

Różnica potencjałów pomiędzy dwoma punktami stanowi pracę jaką należy wykonać w celu przesunięcia jednostkowego ładunku z jednego punktu pola do drugiego. Różnicę potencjałów przyjęto także nazywać napięciem elektrycznym. Ogólne wyrażenie na różnicę potencjałów można otrzymać dzieląc obie części wyrażenia (4.21) przez q :

$$V_A - V_B = - \int_B^A \vec{E} \cdot d\vec{s}. \quad (4.26)$$

Rozważmy jako przykład różnicę potencjałów pomiędzy powierzchnią jednorodnie naładowanej sfery a jej środkiem. Ponieważ wzdłuż drogi całkowania $\vec{E} = 0$, mamy więc $V_A - V_B = 0$; inaczej mówiąc, potencjał w środku jest identyczny jak na powierzchni.

Na rys. 4.10 przytoczono wykresy zależności V i E od r . Zauważmy, że wartość pola elektrycznego (rys. 4.10b) równa jest wziętej ze znakiem przeciwnym pochodnej potencjału elektrycznego. Wynika to z wyrażenia (4.26) zgodnie z którym możemy napisać

$$dV = -\vec{E} \cdot d\vec{s}, \quad (4.27)$$

gdzie dV jest przyrostem funkcji V przy przejściu od punktu (x, y, z) do punktu $(x+dx, y+dy, z+dz)$

$$dV = \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz. \quad (4.28)$$

Z kolei wektor przesunięcia $d\vec{s}$ można zapisać w postaci:

$$d\vec{s} = \vec{i} dx + \vec{j} dy + \vec{k} dz.$$

Przyjmując teraz, że $\vec{E} = -\text{grad}V$ [patrz równania (1.40) i (1.41)]

$$\vec{E} = -\left(\vec{i} \frac{\partial V}{\partial x} + \vec{j} \frac{\partial V}{\partial y} + \vec{k} \frac{\partial V}{\partial z}\right), \quad (4.29)$$

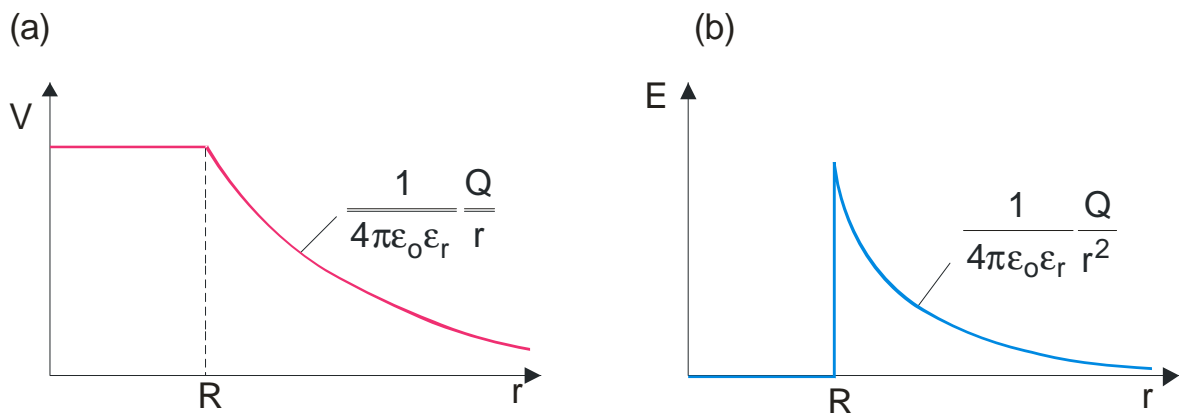
widzimy, że iloczyn skalarny

$$\vec{E} \cdot d\vec{s} = -\left(\frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy + \frac{\partial V}{\partial z} dz\right) = -dV,$$

co potwierdza relację (4.27). Pokazaliśmy zatem, że

$$\vec{E} = -\text{grad}V. \quad (4.30)$$

Znak minus oznacza, że wektor natężenia pola elektrycznego skierowany jest od większego do mniejszego potencjału. Wektor $\text{grad}V$ pokrywa się z kierunkiem wzrostu funkcji V .

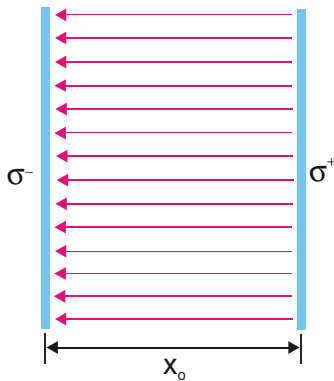


Rys. 4.10. Potencjał (a) i natężenie pola elektrycznego (b) wytworzonego przez naładowaną sferę o promieniu R .

Rozważmy teraz pole elektryczne i różnicę potencjałów pomiędzy dwiema przeciwnie naładowanymi równoległymi płytkami o powierzchni S , położonymi w odległości x_0 jedna od drugiej (rys. 4.11). Jeżeli całkowity ładunek jednej płytki wynosi $+Q$, a drugiej $-Q$, to gęstość ładunku σ równa jest odpowiednio Q/S i $-Q/S$. Zgodnie z (4.27)

$$V = -Ex_0.$$

Ponieważ linie sił pola elektrycznego skierowane są od ładunków dodatnich do ujemnych, to znak minus wskazuje, że dodatnia płytka charakteryzuje się wyższym potencjałem. Ze wzoru (4.18) mamy $E = -\sigma / \epsilon_0 \epsilon_r$, tak więc różnica potencjałów między płytkami wynosi



$$\Delta V = \frac{\sigma x_0}{\epsilon_0 \epsilon_r} = \frac{x_0 Q}{\epsilon_0 \epsilon_r S}. \quad (4.31)$$

Rys. 4.11. Dwie równoległe płytki naładowane równymi co do wartości lecz przeciwnymi ładunkami.

Jeżeli kilka naładowanych ciał położonych jest w odległościach odpowiednio r_1, r_2, \dots, r_n od punktu P , to potencjał elektryczny w tym punkcie jest równy sumie potencjałów od poszczególnych ciał. Jest to wynikiem zasady superpozycji:

$$V = -\int \vec{E} \cdot d\vec{s} = -\int (\vec{E}_1 + \vec{E}_2 + \dots + \vec{E}_n) d\vec{s} = V_1 + V_2 + \dots + V_n.$$

Różnica potencjałów między dwoma punktami określona jest jednoznacznie, ponieważ siły elektrostatyczne są zachowawcze. A w związku z tym (patrz pkt. 2.3) całka $\int \vec{F} \cdot d\vec{s}$ po konturze zamkniętej jest równa zero. Tak więc w elektrostatyce

$$\oint \vec{E} \cdot d\vec{s} = 0. \quad (4.32)$$

Powyższa całka po konturze zamkniętej nazywana jest cyrkulacją wektora natężenia pola elektrycznego.

Wzór (4.32) nie jest słuszny w przypadku zmiennych w czasie pól elektrycznych. Pola takie nie są potencjalne.

4.8. Pojemność elektryczna

Prawie każdy układ elektryczny posiada kondensator. Kondensator zbudowany jest z dwóch płytek i posiada właściwości utrzymywania i gromadzenia ładunku elektrycznego, jeżeli do jego

okładek przyłożyć różnicę potencjałów. Stosunek nagromadzonego ładunku do różnicy potencjałów V nazywamy pojemnością:

$$C = \frac{Q}{\Delta V}. \quad (4.33)$$

Jednostką pojemności jest C/V zwana faradem (F). Farad jest bardzo dużą jednostką, dlatego przyjęto stosować mniejsze jednostki jak mikrofarad (μF), nanofarad (nF), pikofarad (pF).

Kondensator zwykle zbudowany jest z płytek przewodzących oddzielonych cienką warstwą izolatora. Ładunki zgromadzone na płytkach są jednakowe co do wartości lecz przeciwnego znaku. Kondensator taki pokazano schematycznie na rys. 4.11. Ze wzoru (4.31) znajdujemy, że różnica potencjałów pomiędzy tymi dwoma płytkami wynosi $\Delta V = x_o Q / \epsilon_o \epsilon_r S$. Stąd wynika, że pojemność kondensatora płaskiego wynosi

$$C = \frac{Q}{\Delta V} = \frac{\epsilon_o \epsilon_r S}{x_o}. \quad (4.34)$$

4.9. Gęstość energii pola elektrycznego

Założmy, że początkowo nie naładowany kondensator stopniowo ładowano, przy czym różnica potencjałów wzrastała od 0 do V_o . Ładunek na okładkach kondensatora będzie wzrastał od 0 do Q_o , gdzie $Q_o = CV_o$. Praca wykonana przy przemieszczaniu ładunku dq od ujemnie naładowanej płytki do naładowanej dodatnio wynosi

$$dW = Vdq.$$

Całkowita praca, czyli energia zgromadzona w kondensatorze

$$W = \int_0^{V_o} Vdq = \int_0^{Q_o} \frac{q}{C} dq = \frac{1}{2} \frac{Q_o^2}{C}. \quad (4.35)$$

Interesujące jest aby przekształcić wzór (4.35) i zapisać energię zgromadzoną w kondensatorze nie w zależności od ładunku, ale w zależności od natężenia pola elektrycznego. Dla kondensatora płaskiego, uwzględniając (4.31), mamy

$$E = \frac{\Delta V}{x_o} = \frac{Q_o}{\epsilon_o \epsilon_r S}, \quad \text{czyli} \quad Q_o = \epsilon_o \epsilon_r SE.$$

Podstawiając to wyrażenie do (4.35) otrzymamy

$$W = \frac{1}{2} \frac{(\epsilon_o \epsilon_r SE)^2}{C}.$$

Uwzględniając z kolei (4.34) mamy

$$W = \frac{\varepsilon_o \varepsilon_r E^2}{2} Sx_o.$$

Teraz dzieląc obie części przez objętość kondensatora Sx_o , otrzymujemy gęstość energii pola elektrycznego

$$w = \frac{1}{2} \varepsilon_o \varepsilon_r E^2. \quad (4.36)$$

Energia zużyta na przemieszczenie ładunku gromadzona jest w polu elektrycznym kondensatora, a gęstość energii pola elektrycznego wynosi $\varepsilon_o \varepsilon_r E^2/2$ (J/m³). Z bardziej ogólnych ale zarazem bardziej złożonych rozważań wynika, że całkowita energia konieczna do uformowania dowolnego rozkładu ładunków, jest równa dokładnie całce po $\varepsilon_o \varepsilon_r E^2/2$ liczonej po całej przestrzeni V , gdzie E jest polem utworzonym przez taki rozkład ładunku

$$W = \frac{\varepsilon_o \varepsilon_r}{2} \int E^2 dV. \quad (4.37)$$

Wobec tego wyrażenie (4.36) ma bardziej ogólne znaczenie i pozwala przyjąć fizyczną interpretację energii zgromadzonej w jednostce objętości pola elektrycznego.

4.10. Dielektryki

W poprzednich punktach rozważaliśmy pola utworzone przez ładunki w przewodnikach znajdujących się w próżni. Wiadomo, że jeżeli między okładkami kondensatora umieścimy substancję, to pojemność kondensatora wzrasta do C' . Wówczas biorąc stosunek C' do C możemy określić względną przenikalność dielektryczną substancji

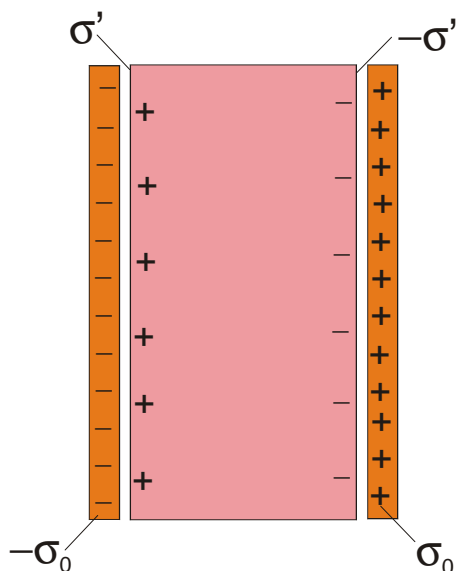
$$\varepsilon_r = \frac{C'}{C}. \quad (4.38)$$

We wzorze tym C jest pojemnością kondensatora próżniowego.

Poprzednio wspomniano o grupie ciał zwanych dielektrykami, w których ładunki nie mają możliwości swobodnego przemieszczania. Jeżeli dielektryk umieścimy w zewnętrznym polu elektrycznym, to na jego granicach indukują się ładunki (rys. 4.12) na skutek ograniczonego przesunięcia ładunków w skali mikroskopowej. Zjawisko to nazywa się polaryzacją dielektryka. Efekt polaryzacji jest jakościowo podobny do powstania łańcucha dipoli. Na jednym końcu łańcucha dipole mają ładunki dodatnie, a na drugim ujemne, a więc dielektryk jako całość wykazuje istnienie ładunków na swoich powierzchniach prostopadłych do kierunku linii sił pola. Ładunki te nazywa się ładunkami nie związanymi. Po usunięciu zewnętrznego pola elektrycznego ładunek na powierzchniach dielektryka znika.

Wskutek zjawiska polaryzacji zmienia się wartość natężenia pola w ośrodku dielektrycznym w stosunku do tego natężenia pola, jakie istnieje w danym obszarze "wypełnionym" próżnią. Jest

to wynik nałożenia się na pole zewnętrzne dodatkowego pola wytworzonego przez ładunki związane. Przed opisem ilościowym tego zjawiska, omówimy rodzaje polaryzowalności dielektryka.



Rys. 4.12. Powstanie ładunku indukowanego σ' na powierzchni dielektryka umieszczonego między okładkami kondensatora.

W cząsteczkach niektórych dielektryków (np. H_2 , Cl_2 , CCl_4 , węglowodory) elektrony są rozmieszczone niejednorodnie dookoła jąder. W cząsteczkach tych środki ciężkości ładunków dodatnich i ujemnych, przy braku zewnętrznego pola elektrycznego, pokrywają się i moment dipolowy równa się zeru. Z tego powodu cząsteczki takich dielektryków nazywamy niespolaryzowanymi. Jeśli niespolaryzowaną cząstkę dielektryka umieścimy w zewnętrznym polu elektrycznym, to następuje rozsuniecie środków ciężkości ładunku dodatniego i ujemnego cząsteczki i wzbudzi się w niej moment dipolowy \vec{p}_e . Moment elektryczny dipola takiego dielektryka równa się

$$\vec{p}_e = \epsilon_0 \alpha \vec{E}, \quad (4.39)$$

gdzie α jest współczynnikiem polaryzowalności atomu. Kierunek wektora \vec{p}_e pokrywa się z kierunkiem wektora \vec{E} natężenia zewnętrznego pola elektrycznego.

Liczną grupę stanowią cząsteczki o samoistnym momencie dipolowym \vec{p}_e , w których środek ciężkości średniego ładunku dodatniego nie pokrywa się z środkiem ciężkości średniego ładunku ujemnego. Przykładem mogą być cząsteczka H_2O (również NH_3 , HCl , CH_3Cl) w której atomy wodoru i tlenu rozłożone są niesymetrycznie. Takie cząsteczki nazywamy spolaryzowanymi. W wyniku nieuporządkowanego ruchu cieplnego cząsteczek, wektory ich momentów dipolowych wykazują chaotyczną orientację i wypadkowy moment dipolowy w dowolnej objętości dielektryka równa się zeru. Jednak pod wpływem zewnętrznego pola elektrycznego, cząsteczki dielektryka dążą do zajęcia takiego położenia, aby kierunek wektorów ich momentów

dipolowych \vec{p}_e był zgodny z kierunkiem wektora \vec{E} . Pojawia się więc orientacja momentów dipolowych cząsteczek przeważnie wzdłuż linii sił pola. Orientacja ta jest tym większa, im silniejsze jest pole elektryczne w dielektryku oraz im słabszy jest ruch cieplny cząsteczek, tj. im niższa jest temperatura. Powyższe zjawisko nosi nazwę polaryzacji skierowanej dielektryka o cząsteczkach spolaryzowanych.

Cząsteczki niespolaryzowane uzyskują w polu elektrycznym momenty dipolowe indukowane w wyniku odkształcenia orbit elektronowych. Zachodzi wówczas tzw. polaryzacja elektronowa dielektryka.

W dielektrykach jonowych (krystalicznych) typu NaCl, CsCl; wszystkie jony dodatnie przesuwają się pod wpływem pola elektrycznego w kierunku odpowiadającym kierunkowi natężenia \vec{E} , natomiast wszystkie jony ujemne w kierunku przeciwnym. Ten rodzaj polaryzacji nosi nazwę polaryzacji jonowej.

Jako wskaźnik ilościowy polaryzacji dielektryka służy wektor polaryzacji \vec{P}_e . *Wektorem polaryzacji nazywamy granicę stosunku momentu elektrycznego określonej objętości dielektryka do tej objętości, gdy ta ostatnia dąży do zera*

$$\vec{P}_e = \lim_{V \rightarrow 0} \frac{1}{V} \sum_{i=1}^N \vec{p}_{ei}, \quad (4.40)$$

gdzie N oznacza liczbę dipoli zawartych w objętości V dielektryka, a \vec{p}_{ei} moment elektryczny i -tego dipola.

W przypadku dielektryka jednorodnego o cząsteczkach niespolaryzowanych, umieszczonego w jednorodnym polu elektrycznym,

$$\vec{P}_e = N_o \vec{p}_e, \quad (4.41)$$

gdzie N_o oznacza liczbę cząsteczek w jednostce objętości. Wynika stąd, że wektory \vec{p}_{ei} wszystkich cząsteczek wykazują jednakowy kierunek, odpowiadający kierunkowi wektora natężenia \vec{E} pola w dielektryku. Stosując wzór (4.39) otrzymujemy

$$\vec{P}_e = N_o \epsilon_o \alpha \vec{E} = \epsilon_o \chi_e \vec{E}. \quad (4.42)$$

Współczynnik $\chi_e = N_o \alpha$ nazywamy podatnością dielektryczną substancji.

4.11. Twierdzenie Gaussa w przypadku obecności dielektryków. Wektor indukcji elektrycznej

W poprzednim punkcie stwierdziliśmy, że w dielektryku na pole elektryczne ładunków swobodnych nakłada się dodatkowe pole elektryczne. Z tego względu wektor natężenia pola elektrycznego \vec{E} powinien zależeć od właściwości elektrycznych dielektryka. Okazuje się, że wartość liczbową \vec{E} jest zawsze odwrotnie proporcjonalna do stałej dielektrycznej ϵ ośrodka. Z

tego względu w celu jednoznacznego scharakteryzowania pola elektrycznego celowe jest wprowadzenie takiej wielkości \vec{D} , która by nie zależała od stałej dielektrycznej danej substancji. Można z łatwością wykazać, że warunek ten spełnia wielkość wektorowa zdefiniowana następująco:

$$\vec{D} = \epsilon \epsilon_0 \vec{E}. \quad (4.43)$$

Wielkość \vec{D} nazywamy wektorem indukcji elektrycznej. Wektor \vec{D} charakteryzuje zatem pole elektryczne, które wytwarzają w danej substancji same tylko ładunki swobodne. Ładunki związane, powstające w dielektryku, mogą jednak wywołać zmianę rozkładu w przestrzeni ładunków swobodnych wytwarzających pole.

W układzie jednostek SI indukcję elektryczną mierzy się w C/m².

Strumień indukcji elektrycznej w dowolnym środowisku przez element powierzchni jest określony przez iloczyn skalarny:

$$d\Phi_D = \vec{D}_j \cdot d\vec{S}_j,$$

gdzie wektor $d\vec{S}_j$ określa pole i orientację j -tego elementu powierzchni, a \vec{D}_j jest uśrednionym wektorem indukcji elektrycznej dla j -tego elementu. Całkowity strumień przez powierzchnię będzie równy:

$$\Phi_D = \int_S \vec{D} \cdot d\vec{S} = \sum q_{swob}, \quad (4.45)$$

gdzie zgodnie z definicją wektora indukcji elektrycznej uwzględniono tylko ładunki swobodne.

W próżni $D = \epsilon_0 E$, a zatem równanie (4.45) przybiera postać

$$\int_S \epsilon_0 \vec{E} \cdot d\vec{S} = \sum q_{swob}. \quad (4.46)$$

Pole w dowolnym środowisku różni się od pola w próżni tym, że wytwarzają je ładunki zarówno swobodne, jak i związane. W najogólniejszym przypadku do prawej strony równania (4.46) należy dodać sumę algebraiczną ładunków związanych objętych przez powierzchnię zamkniętą S

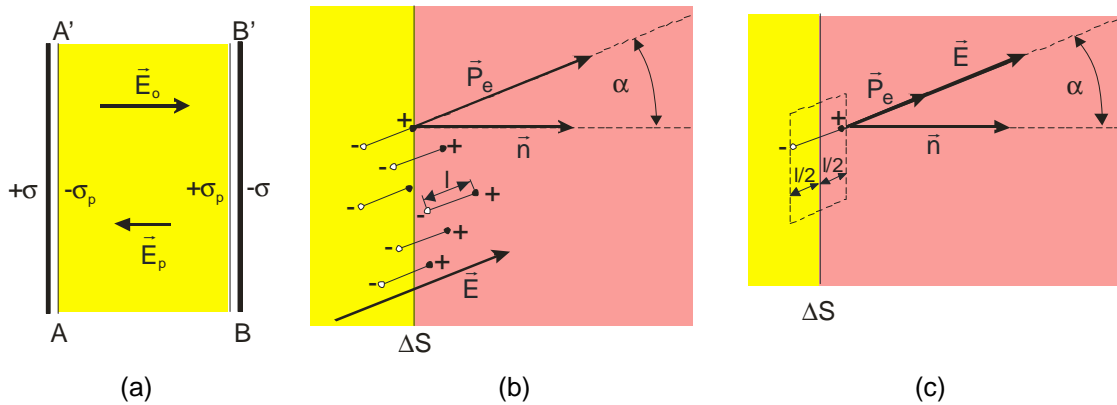
$$\int_S \epsilon_0 \vec{E} \cdot d\vec{S} = \sum q_{swob} + \sum q_{związ}. \quad (4.47)$$

Ładunki swobodne wytwarzają zewnętrzne pole elektryczne, natomiast ładunki związane wytwarzają pole wewnętrzne spolaryzowanego dielektryka.

Rozpatrzmy warstwę jednorodnego dielektryka zawartą między dwoma nieskończonymi równoległymi płaszczyznami, naładowanymi do stałych gęstości powierzchniowych ładunków swobodnych $+\sigma$, $-\sigma$ [rys. 4.13(a)]. W wyniku polaryzacji dielektryka na jego powierzchniach AA' i BB' powstają ładunki związane, których gęstości powierzchniowe są równe odpowiednio

$-\sigma_p$ i $+\sigma_p$. Na skutek tego pole elektryczne E_p ładunków związanych jest skierowane przeciwnie względem pola zewnętrznego \vec{E}_o , wytworzonego przez ładunki swobodne. Natężenie pola wypadkowego $\vec{E} = \vec{E}_o + \vec{E}_p$.

Znajdziemy teraz sumę ładunków związanych, które powstały w wyniku polaryzacji dielektryka, objętego zamkniętą powierzchnią S . Na rys. 4.13(b) przedstawiony jest tak mały element ΔS tej powierzchni, że można go uważać za płaski. Ładunki dipoli znajdujących się na zewnątrz powierzchni zamkniętej nie wywierają w ogóle żadnego wpływu na strumień natężenia pola przez tę powierzchnię. Suma algebraiczna wszystkich ładunków dipoli całkowicie objętych powierzchnią, równa się zero. Przy obliczaniu $\sum q_{zwiąq}$ uwzględnia się zatem tylko te dipole, które przecinają powierzchnię S . Jest rzeczą oczywistą, że przecinane są jedynie takie dipole, których środki ciężkości leżą po lewej i po prawej stronie powierzchni S w odległościach mniejszych niż $(l/2)\cos\alpha$, gdzie l oznacza długość dipola, a α jest kątem zawartym między zewnętrzną normalną \vec{n} do elementu ΔS powierzchni a momentem \vec{p}_e dipola. Warunek ten spełniają wszystkie dipole, których środki leżą wewnątrz objętości $l\Delta S\cos\alpha$ [rys. 4.13(c)]. Jeżeli liczba cząsteczek dielektryka w jednostce objętości równa się N_o , to liczba dipoli przeciętych przez element ΔS powierzchni wynosi $N_o l\Delta S\cos\alpha$. Każdy przecięty dipol ma wewnątrz zamkniętej powierzchni nie zubożniony ładunek $-q$.



Rys. 4.13. Powstawanie ładunku związanego.

Całkowity ładunek związany $\Delta q_{zwiąq}$, odpowiadający powierzchni ΔS , równa się zatem

$$\Delta q_{zwiąq} = -N_o q l \cos \alpha \Delta S = -N_o p_e \cos \alpha \Delta S .$$

Dzięki założeniu równoległości wszystkich dipoli, iloczyn $N_o p_e$ równy jest modułowi wektora polaryzacji. W związku z tym

$$\Delta q_{zwiąq} = -\vec{P}_e \cos \alpha \Delta S = -\vec{P}_e \cdot \vec{n} \Delta S = -\vec{P}_e \cdot d\vec{S} , \quad (4.48)$$

gdzie \vec{n} jest wektorem jednostkowym normalnym do powierzchni ΔS .

W celu uzyskania ogólnej sumy ładunków związanych, znajdujących się wewnątrz zamkniętej powierzchni S , należy wyrażenie (4.48) scałkować po powierzchni S

$$\sum q_{zwiąq} = - \int_S \vec{P}_e \cdot d\vec{S}. \quad (4.49)$$

W związku z tym, twierdzenie Gaussa dla dowolnej substancji spolaryzowanej, zgodnie ze wzorem (4.47), zapisujemy w postaci

$$\int_S \epsilon_o \vec{E} \cdot d\vec{S} = \sum q_{swob} - \int_S \vec{P}_e \cdot d\vec{S},$$

stąd

$$\int_S (\epsilon_o \vec{E} + \vec{P}_e) d\vec{S} = \sum q_{swob}. \quad (4.50)$$

Wstawiając tu $\sum q_{swob}$ z równania (4.45) otrzymujemy

$$\int_S (\epsilon_o \vec{E} + \vec{P}_e) d\vec{S} = \int_S \vec{D} \cdot d\vec{S}.$$

Ponieważ równanie to powinno być spełnione dla dowolnej zamkniętej powierzchni S , przeto

$$\vec{D} = \epsilon_o \vec{E} + \vec{P}_e. \quad (4.51)$$

Uwzględniając (4.42) mamy

$$\vec{D} = \epsilon_o \vec{E} + \epsilon_o \chi_e \vec{E} = \epsilon_o (1 + \chi_e) \vec{E}. \quad (4.52)$$

Z drugiej strony, w myśl definicji (4.43), wektor \vec{D} równy jest

$$\vec{D} = \epsilon_o \epsilon_r \vec{E}.$$

Zatem

$$\epsilon_r = 1 + \chi_e. \quad (4.53)$$

Stała dielektryczna równa się podatności dielektrycznej zwiększonej o 1. Obydwie te wielkości są bezwymiarowe. Dla próżni $\epsilon_r = 1$, a $\chi_e = 0$.

ROZDZIAŁ 5

PRĄD ELEKTRYCZNY I SIŁA MAGNETYCZNA

Oprócz siły elektrostatycznej na ładunek działa "nowa siła" proporcjonalna do prędkości ładunku v . Siłę tę nazywamy magnetyczną. W teorii elektromagnetyzmu, magnetyzm i zrozumienie istoty pola magnetycznego odgrywa bardzo ważną rolę. W tym rozdziale zobaczymy, że siła działająca na magnes to siła działająca na poruszające się elektrony atomów magnesu. Pokażemy również, że istnienie siły magnetycznej jest prostym następstwem uwzględnienia szczególnej teorii względności w prawie Coulomba.

5.1. Prąd elektryczny

Zjawiska magnetyczne pojawiają się podczas ruchu ładunków, czyli przy przepływie prądu. Płynący w przewodniku prąd określa się jako ilość ładunku przepływającego przez dany przekrój w jednostce czasu

$$I = \frac{Q}{t}. \quad (5.1)$$

Jednostką natężenia prądu jest amper $A = C/s$.

Z prądem bezpośrednio związana jest gęstość prądu, którą określa się jako iloczyn gęstości ładunku przez jego prędkość \vec{v} .

$$\vec{j} = \rho \vec{v}. \quad (5.2)$$

Gęstość prądu to prąd przepływający przez jednostkową powierzchnię; mierzona jest w C/m^2s czyli A/m^2 .

Jeżeli gęstość prądu pomnożymy przez przekrój przewodnika, to otrzymamy prąd

$$I = \vec{j} \cdot \vec{S}, \quad (5.3)$$

gdzie \vec{S} jest wektorem normalnym do przekroju o długości równym wartości S . Jeżeli na przekroju S gęstość prądu \vec{j} zmienia się, to

$$I = \int_S \vec{j} \cdot d\vec{S}. \quad (5.4)$$

W przewodniku metalicznym dodatnie ładunki (jądra atomów) nie mogą się przemieszczać; jądra tworzą sieć krystaliczną. Jednakże zewnętrzne elektrony, czyli elektrony przewodnictwa,

nie są związane z określonymi atomami i mogą swobodnie poruszać się w przewodniku. W nieobecności zewnętrznego pola elektrycznego elektrony przewodnictwa poruszają się chaotycznie i ich średnia prędkość jest zerowa. Niech n oznacza liczbę elektronów przewodnictwa w jednostce objętości (jest to tzw. koncentracja elektronów). Wówczas gęstość ładunku $\sigma = ne$, a gęstość prądu $\vec{j} = ne\vec{v}_d$, gdzie v_d jest prędkością dryfu elektronów. W związku z tym zgodnie z (5.3), natężenie prądu

$$I = ne\vec{v}_d \cdot \vec{S}. \quad (5.5)$$

Według propozycji Franklina przyjęto, że prąd elektryczny płynący do okładki kondensatora dostarcza jej ładunek dodatni. Wiemy, że okładka kondensatora naładowana jest dodatnim ładunkiem jeżeli opuszczą ją elektrony przewodnictwa. Wynika z tego, że elektrony przewodnictwa poruszają się w kierunku przeciwnym do kierunku prądu. Przyjmujemy, że strzałka oznaczająca kierunek prądu wskazuje kierunek ruchu ładunków dodatnich. Prąd faktycznie uwarunkowany jest ruchem elektronów, a elektrony poruszają się w kierunku przeciwnym do wskazanego przez strzałkę.

Typowa prędkość dryfu elektronów przewodnictwa w metalach wynosi 10^{-4} m/s. Prądy mogą również płynąć w gazach i cieczech. Przykładem prądu płynącego w gazie jest prąd w lampach neonowych. Prąd ten uwarunkowany jest zarówno ruchem dodatnich jonów jak i elektronów. Ponieważ elektrony są lżejszymi cząstkami, ich ruchliwość jest dużo wyższa i dlatego wnoszą większy wkład w wielkość prądu. Elektron, przy zderzeniu z jonom lub atomem w gazie, przekazuje część energii kinetycznej, która następnie wydziela się w postaci promieniowania elektromagnetycznego.

5.2. Prawo Ohma

Jeżeli do przewodnika przyłożyć różnicę potencjałów V , to będzie w nim płynął prąd I . Na początku XIX w. G. Ohm odkrył prawo, zgodnie z którym prąd w metalach jest proporcjonalny do przyłożonego napięcia pod warunkiem, że temperatura przewodnika jest stała. Określił on rezystancję jako

$$R = \frac{V}{I}. \quad (5.6)$$

Prawo Ohma możemy więc sformułować następująco: *stosunek $R = V/I$ nie zależy od natężenia prądu I dla metali przy stałej temperaturze.*

W układzie SI stosunek V/A ma specjalną nazwę om (Ω).

Prawo Ohma nie jest nowym, fundamentalnym prawem przyrody lecz wynika z innych praw oddziaływania i budowy materii. Wiemy, że elektrony przewodnictwa mogą przemieszczać się w materii na odległości wielokrotnie przewyższające rozmiary atomu przed ich zderzeniem z atomem. Niech l oznacza średnią odległość między zderzeniami zwaną średnią drogą swobodną.

Wówczas średni czas pomiędzy zderzeniami wynosi $\Delta t = l/u$, gdzie u – średnia prędkość elektronów przewodnictwa. Kierunek prędkości u zmienia się chaotycznie i ruch ten nie powoduje pojawienia się wypadkowego prądu. Jeżeli do odcinka przewodnika przyłożyc napięcie (różnicę potencjałów) to na każdy elektron przewodnictwa będzie działać siła eE . Pod wpływem tej siły w czasie Δt każdy elektron przewodnictwa przyjmuje prędkość $v_d = \Delta u$, którą określamy w następujący sposób:

$$m \frac{\Delta u}{\Delta t} = eE \quad \text{a stąd} \quad \Delta u = v_d = \frac{eE\Delta t}{m}.$$

Zmieniając Δt na średni czas l/u i uśredniając po czasie, otrzymujemy:

$$\langle v_d \rangle = \frac{e l E}{2 m u}. \quad (5.7)$$

Kierunek prędkości dryfu wszystkich elektronów jest jednakowy (zgodny z $-\vec{E}$) i dlatego powstaje prąd wypadkowy, przy czym $v_d \ll u$.

Podstawiając do wzoru (5.5) wyrażenie na $\langle v_d \rangle$, otrzymamy natężenie prądu w przewodniku o przekroju S

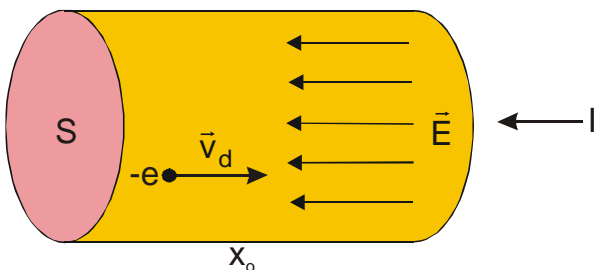
$$I = \frac{ne^2 l S}{2 m u} E.$$

Teraz możemy obliczyć rezystancję przewodnika o długości x_0 (rys. 5.1) W tym przypadku różnica potencjałów na długości przewodnika wynosi Ex_0 . Podstawiając w miejsce E wielkość V/x_0 , mamy

$$I = \frac{ne^2 l S}{2 m u x_0} V,$$

stąd

$$R = \frac{V}{I} = \frac{2 m u x_0}{ne^2 l S}. \quad (5.8)$$



Z wyrażenia tego wynika, że rezystancja jest wprost proporcjonalna do długości przewodnika i odwrotnie proporcjonalna do powierzchni przekroju poprzecznego. Wobec tego wyrażenie określające rezystancję możemy napisać w postaci

$$R = \rho \frac{x_0}{S}, \quad (5.9)$$

Rys. 5.1. Odcinek przewodnika o długości x_0 .

gdzie współczynnik proporcjonalności ρ nazywamy rezystywnością. Jednostką rezystywności jest Ωm .

5.3. Ciepło Joule'a

Elektron przewodnictwa przy zderzeniu z atomem traci energię pobraną od pola elektrycznego. Energia ta przekształca się w chaotyczny ruch atomów, tzn. w ciepło. Ponieważ energia kinetyczna elektronów nie wzrasta, stratę energii ładunku dq na skutek zderzeń przy pokonaniu różnicy potencjałów V , zapiszemy w postaci

$$dE_{cie} = Vdq .$$

Dzieląc teraz obie strony tego wyrażenia przez dt , mamy

$$\frac{dE_{cie}}{dt} = V \frac{dq}{dt} = VI ,$$

czyli straty mocy elektrycznej wynoszą

$$P = VI . \tag{5.10}$$

Wyrażenie (5.10) możemy zapisać w postaci $P = I^2 R$ zamieniając V na IR , lub jako $P = V^2/R$ zamieniając I na V/R . Wielkość P przedstawia moc elektryczną, która przekształca się w ciepło. Energia prądu elektrycznego w lampach przekształca się w ciepło i światło.

5.4. Obwody prądu stałego

W celu podtrzymania w przewodniku prądu stałego, okazuje się konieczne posiadanie źródła energii elektrycznej. Rozpowszechnionymi typami tych źródeł są baterie elektryczne i generatory elektryczne. Źródła te nazywamy źródłami siły elektromotorycznej, którą w skrócie będziemy oznaczać jako *SEM*. W tych źródłach energia elektryczna otrzymywana jest z innych form energii: z energii chemicznej w bateriach lub energii mechanicznej w generatorach elektrycznych. W bateriach słonecznych energia świetlna przekształcana jest w energię elektryczną. Ładunek Δq przechodząc od ujemnego bieguna baterii do dodatniego przyjmuje energię $\Delta W = E\Delta q$, gdzie E oznacza *SEM* źródła.

Chociaż teoria obwodów elektrycznych jest przedmiotem zainteresowań fizyki stosowanej i techniki, przedstawimy pewne jej wyniki.

Większość obwodów zawiera kombinację szeregowego i równoległego łączenia rezystancji. Przyłączeniu szeregowym (rys. 5.2a) przez wszystkie rezystory płynie ten sam prąd, a całkowita różnica potencjałów

$$V = V_1 + V_2 + V_3 .$$

Dzieląc obie strony tego wyrażenia przez wartość prądu I w obwodzie, otrzymujemy

$$\frac{V}{I} = \frac{V_1}{I} + \frac{V_2}{I} + \frac{V_3}{I} ,$$

czyli rezystancja całkowita

$$R_c = R_1 + R_2 + R_3. \quad (5.11)$$

Przy równoległym łączeniu (rys. 5.2b) mamy

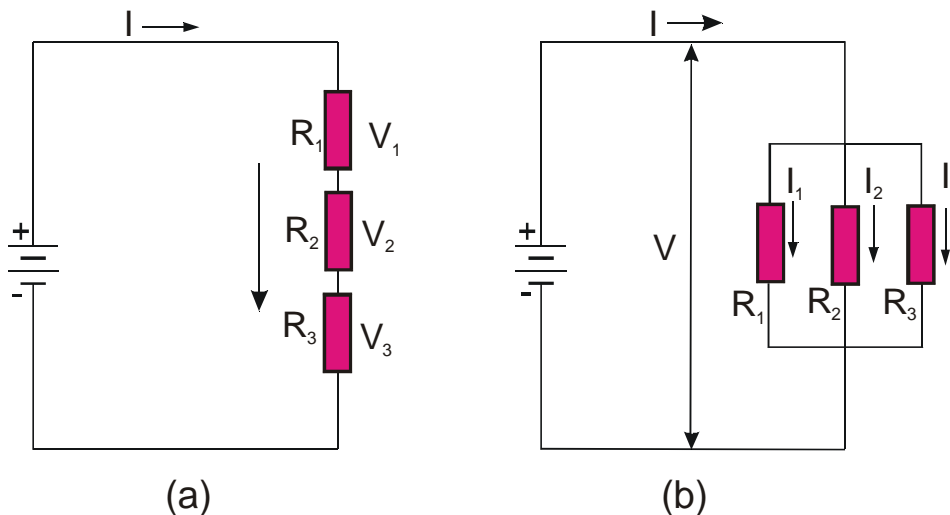
$$I = I_1 + I_2 + I_3.$$

Dzieląc przez V obie strony tego wyrażenia mamy

$$\frac{I}{V} = \frac{I_1}{V} + \frac{I_2}{V} + \frac{I_3}{V},$$

czyli

$$\frac{1}{R_c} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}. \quad (5.12)$$



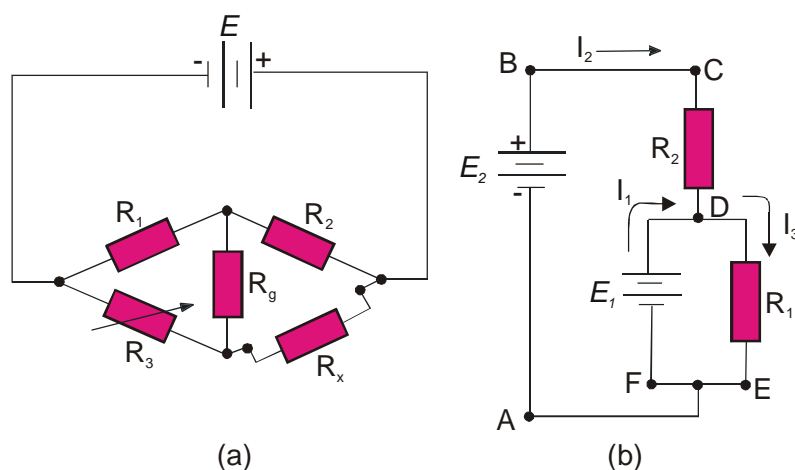
Rys. 5.2. Szeregowe (a) i równoległe (b) łączenie rezystorów R_1 , R_2 i R_3 .

Przy równoległym łączeniu rezystancji dodajemy więc odwrotności poszczególnych rezystancji obwodu aby uzyskać odwrotność rezystancji wypadkowej.

Wiele złożonych obwodów elektrycznych można obliczyć rozbijając je na równoległe i szeregowe połączenia rezystancji.

5.4.1. Prawa Kirchhoffa

W bardziej złożonych obwodach rezystory mogą być tak połączone, że nie uda się ich sprowadzić do połączenia równoległego i szeregowego. Przykład takiego obwodu pokazany jest na rys. 5.3a. Obwody mogą również zawierać kilka źródeł prądu (rys. 5.3b). Istnieje ogólny sposób obliczania obwodów stosując prawa Kirchhoffa.



Rys. 5.3. (a) Schemat mostka Wheatstone'a stosowanego do pomiaru R_x , (b) Stabilizator napięcia.

Prawo dla konturu: *Suma algebraiczna spadków napięć liczonych wzdłuż zamkniętego konturu jest równa zero.*

Prawo dla węzła: *Suma algebraiczna wszystkich prądów węzła (wchodzących i wychodzących) jest równa zero.*

Prawo dla konturu jest wynikiem prawa zachowania energii, natomiast prawo dla węzła - wynikiem prawa zachowania ładunku elektrycznego.

Każdej części obwodu przypisujemy wartość i kierunek prądu. Stosując pierwsze prawo przyjęto, że spadek napięcia jest dodatni, jeżeli kierunek obejścia zamkniętego konturu jest zgodny z kierunkiem prądu; natomiast jest ujemny, jeżeli kierunek obejścia jest przeciwny do kierunku prądu. SEM będziemy traktować jako dodatnią przechodząc od "+" do "-". Jeżeli po rozwiązaniu zadania wartość prądu okaże się ujemna, to oznacza że faktyczny kierunek prądu jest przeciwny do obranego.

Jako przykład znajdziemy prądy płynące w obwodzie z rys. 5.3b. Zastosujemy prawo dla konturu $ABCDEA$:

$$E_2 - I_2 R_2 - I_3 R_1 = 0$$

i dla konturu $EFDE$:

$$E_1 - I_3 R_1 = 0.$$

Uwzględniając drugie równanie w pierwszym mamy

$$E_2 - E_1 - I_2 R_2 = 0, \quad \text{czyli} \quad I_2 = \frac{E_2 - E_1}{R_2}.$$

Prąd I_1 możemy określić stosując prawo dla węzła w punkcie D

$$I_1 + I_2 - I_3 = 0,$$

czyli

$$I_1 = I_3 - I_2 = \frac{E_1}{R_1} - \frac{E_2 - E_1}{R_2} = E_1 \left(\frac{1}{R_1} + \frac{1}{R_2} \right) - \frac{E_2}{R_2}.$$

Zauważmy, że jeżeli $E_1(1/R_1 + 1/R_2) = E_2/R$, to $I_1 = 0$, co oznacza że z baterii E_1 nie jest pobierany. Układ w którym $I_1 \approx 0$ ma ważne znaczenie w praktyce. Załóżmy, że na rezystorze R_1 należy uzyskać spadek napięcia równy dokładnie E_1 , chociaż przez R_1 płynie duży prąd I_3 . Układ przedstawiony na rys. 5.3b działa w tym przypadku jak stabilizator napięcia. Nieduża zmiana napięcia źródła mocy E_2 nie wpływa na napięcie E_1 przyłożone do R_1 . E_1 może być źródłem o małej mocy, chociaż przez R_1 płynie duży prąd (prąd ten w zasadzie pobierany jest od źródła E_2 , a napięcie określone jest źródłem E_1). Jeżeli zastosować tylko jedną baterię E_1 to ulegnie szybko wyczerpaniu.

Schemat obwodu pokazany na rys. 5.3a stosowany jest w praktyce do pomiaru nieznannej rezystancji R_x z dużą dokładnością. Jeżeli w tym schemacie R_g oznacza rezystancję galwanometru (bardzo czułego przyrządu do pomiaru prądu) to otrzymujemy schemat zwany mostkiem Wheatstone'a. Przy nieznannej rezystancji R_x dobiera się taką rezystancję R_3 aby prąd płynący przez R_g był zerowy. Wówczas $R_2/R_1 = R_x/R_3$ oraz $R_x = R_3(R_2/R_1)$.

5.5. Siła magnetyczna

W życiu codziennym często spotykamy się z siłą magnetyczną. Oto przykłady: magnesy trwałe, cewki indukcyjne, elektromagnesy, przekaźniki elektryczne, układy odchyłania kineskopów telewizyjnych, itp. Wszystkie te przejawy siły magnetycznej mogą być wytłumaczone w oparciu o fundamentalne oddziaływania pomiędzy poruszającymi się ładunkami, czyli prądami. W rozdziale 6 wykażemy, że z punktu widzenia budowy atomowej w magnesach istnieją stałe prądy wirowe i że siła działająca na magnes uwarunkowana jest fundamentalnym oddziaływaniem poruszających się ładunków.

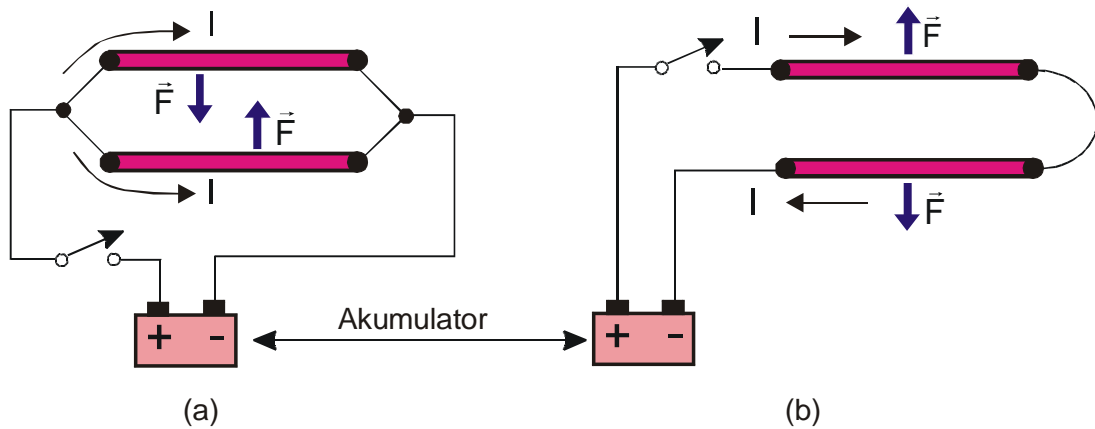
Niezależnie od tego czy ładunek porusza się, czy jest w spoczynku, działa na niego siła elektryczna $\vec{F}_E = q\vec{E}$. Jednakże jeżeli ładunek porusza się z prędkością v , to działa na niego dodatkowa siła \vec{F}_{mag} , która jak wynika z pomiarów, jest proporcjonalna do iloczynu qv . W niektórych popularnonaukowych książkach ta dodatkowa siła (nazywana magnetyczną) wprowadzona jest jako wynik obserwacji eksperymentalnych i traktowana jako nowa fundamentalna siła przyrody. W następnym punkcie wykażemy, że istnienie tej "dodatkowej" siły proporcjonalnej do qv wynika z zasady względności. Siła magnetyczna jest prostym, relatywistycznym wynikiem prawa Coulomba. Wobec tego, jeżeli \vec{F} oznacza relatywistyczną siłę elektromagnetyczną (siłę Lorentza) działającą na ładunek q , to możemy napisać

$$\vec{F} = q\vec{E} + \vec{F}_{mag}. \quad (5.13)$$

Ponieważ \vec{F}_{mag} jest proporcjonalna do qv , można określić taką wielkość wektorową \vec{B} , aby

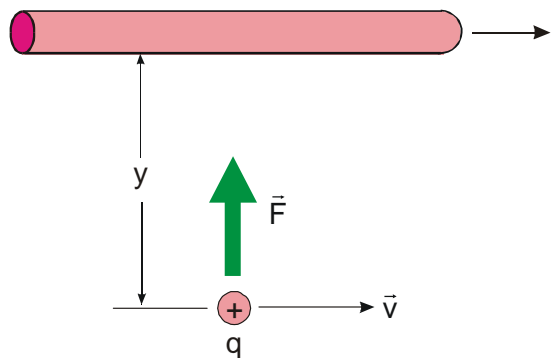
$$\vec{F}_{mag} = q\vec{v} \times \vec{B}. \quad (5.14)$$

W następnym punkcie udowodnimy, że wyrażenie to stosowane jest do określenia pola magnetycznego \vec{B} i nauczymy się obliczać wielkość \vec{B} poprzez prądy, które ją powodują.



Rys. 5.4. Położone w odległości kilku centymetrów od siebie dwa nie naładowane przewody przyciągają się (a) lub odpychają się (b) w zależności od kierunków płynących w nich prądów.

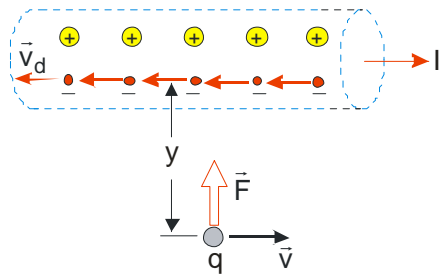
Znany jest fakt eksperymentalny, że dwa przewody równoległe położone względem siebie, przyciągają się gdy kierunki prądów w nich płynące są jednakowe; lub odpychają się, gdy kierunki tych prądów są przeciwne (rys. 5.4). Jeżeli prąd w niżej położonym przewodniku zamienić poruszającym się ładunkiem q , to ładunek będzie przyciągany do przewodnika co pokazano na rys. 5.5.



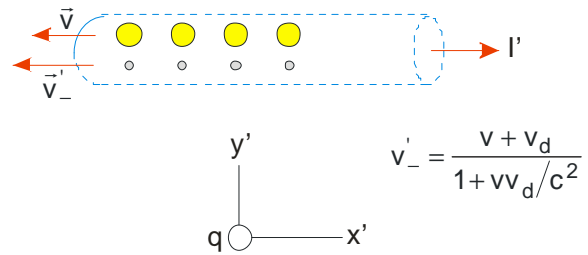
Rys. 5.5. Ładunek punktowy q poruszając się z prędkością \vec{v} równoległe do przewodnika z prądem I , przyciągany jest przez przewodnik.

Rozpatrzmy bardziej wnikliwie sytuację zilustrowaną na rys. 5.5. Prąd I uwarunkowany jest elektronami przewodnictwa poruszającymi się w lewo z prędkością dryfu v_d . Aby przewodnik był elektrycznie obojętny, powinien istnieć "łańcuszek" dodatnio naładowanych jonów przenoszących ładunek równy co do wartości lecz przeciwny co do znaku (rys. 5.6). Niech λ_- oznacza liniową gęstość ładunku elektronów przewodnictwa, a λ_+ – liniową gęstość dodatnich

jonów, przy czym $\lambda_+ = \lambda_-$. Chcemy obliczyć siłę działającą na ładunek q , jeżeli taka siła istnieje.



Rys. 5.6. Powiększony przewódnik z rys. 5.5.



Rys. 5.7. Przewódnik z prądem pokazany na rys. 5.5, widziany przez obserwatora poruszającego się z ładunkiem q .

Najpierw udowodnimy, że w układzie odniesienia związanym z poruszającym się ładunkiem q , obojętny przewódnik, jak pokazano na rys. 5.5, będzie miał ładunek o gęstości liniowej

$$\lambda' = -\frac{1}{\sqrt{1 - (v/c)^2}} \frac{v}{c^2} I,$$

widziany przez obserwatora poruszającego się razem z ładunkiem q .

Niech z poruszającym się ładunkiem związany jest "kreskowany" układ odniesienia $Ox'y'$. Z rys. 5.7 widzimy, że w tym układzie odniesienia dodatnie jony poruszają się w lewo z prędkością v i w tym samym kierunku lecz z większą prędkością v'_- poruszają się elektrony przewodnictwa.

Gęstość ładunku w przewodniku

$$\lambda' = \lambda'_+ + \lambda'_-. \quad (5.15)$$

Ze względu na lorentzowskie skrócenie

$$\lambda'_+ = \frac{1}{\sqrt{1 - (v/c)^2}} (\lambda_+)_o, \quad (5.16)$$

$$\lambda'_- = \frac{1}{\sqrt{1 - (v'_-/c)^2}} (\lambda_-)_o, \quad (5.17)$$

gdzie $(\lambda_-)_o$ jest gęstością ładunku elektronów przewodnictwa pozostających w spoczynku, przy czym na wskutek skrócenia lorentzowskiego mamy

$$\lambda_- = \frac{1}{\sqrt{1 - (v_d/c)^2}} (\lambda_-)_o, \quad \text{czyli} \quad (\lambda_-)_o = \sqrt{1 - (v_d/c)^2} \lambda_-.$$

Podstawiając ostatnie wyrażenie do (5.17) otrzymujemy

$$\lambda'_- = \frac{\sqrt{1 - (v_d/c)^2}}{\sqrt{1 - (v_-/c)^2}} \lambda_-.$$

Niech $\beta = v/c$, $\beta_d = v_d/c$ i $\beta' = v_-/c$. Wówczas

$$\lambda'_- = \frac{\sqrt{1 - \beta_d^2}}{\sqrt{1 - \beta'^2}} \lambda_-.$$

Ażeby wyłączyć β' wykorzystamy relatywistyczne prawo dodawania prędkości (3.22)

$$\beta' = \frac{\beta_d + \beta}{1 + \beta_d \beta}.$$

Wówczas

$$\lambda'_- = \frac{\sqrt{1 - \beta_d^2}}{\sqrt{1 - \left(\frac{\beta_d + \beta}{1 + \beta_d \beta}\right)^2}} \lambda_- = \frac{(1 + \beta_d \beta) \sqrt{1 - \beta_d^2}}{\sqrt{(1 - \beta^2)(1 - \beta_d^2)}} \lambda_- = \frac{1 + \beta_d \beta}{\sqrt{1 - \beta^2}} \lambda_-. \quad (5.18)$$

Podstawiając (5.16) i (5.18) do (5.15) otrzymujemy

$$\lambda' = \frac{1}{\sqrt{1 - \beta^2}} \lambda_+ + \frac{1 + \beta_d \beta}{\sqrt{1 - \beta^2}} \lambda_-.$$

Zamieniamy λ_+ na $(-\lambda_-)$

$$\lambda' = \frac{\beta_d \beta}{\sqrt{1 - \beta^2}} \lambda_- = \frac{1}{\sqrt{1 - \beta^2}} \frac{v}{c^2} \lambda_- v_d.$$

Podstawiając wyrażenie $I = -\lambda_- v_d$ otrzymujemy

$$\lambda' = -\frac{1}{\sqrt{1 - (v/c)^2}} \frac{vI}{c^2}. \quad (5.19)$$

Faktycznie wykazaliśmy, że w układzie odniesienia związanym z poruszającym się ładunkiem, obserwator widzi ładunek obojętnego przewodnika o gęstości liniowej określonej powyższym wzorem. Uzyskany rezultat jest konsekwencją skrócenia lorentzowskiego obserwowanego przez obserwatora związanego z poruszającym się ładunkiem. Ponieważ elektrony poruszają się szybciej od dodatnich jonów, odległość między nimi ulega większemu skróceniu niż między jonami. Wobec tego z punktu widzenia poruszającego się obserwatora moduł λ_- okazuje się większy od modułu λ_+ i wypadkowa gęstość ładunku będzie ujemna, a nie równa zero jak w układzie związanym z przewodnikiem. W związku z tym ujemnie naładowany przewodnik przyciąga ładunek q z siłą [zobacz wzór (4.19)]

$$F' = qE' = q \left(\frac{\lambda'}{2\pi\epsilon_0 y} \right).$$

Podstawiając teraz wyrażenie na λ'_- z (5.19) znajdujemy siłę jaką rejestruje poruszający się obserwator

$$F' = \frac{qv}{\sqrt{1-(v/c)^2}} \frac{I}{2\pi\epsilon_0 c^2 y}. \quad (5.20)$$

Udowodniliśmy więc, że z punktu widzenia teorii względności na poruszający się ładunek q działa siła. Jeżeli siłę tę rejestruje poruszający się obserwator, to powinien ją również rejestrować każdy obserwator poruszający się bez przyśpieszenia. Zgodnie z teorią względności, spoczywający obserwator zmierzy siłę $F = (1 - v^2/c^2)^{1/2} F'$. Wobec tego, spoczywający obserwator zmierzy siłę

$$F = \frac{1}{2\pi\epsilon_0 c^2} \frac{I}{y} qv. \quad (5.21)$$

Na zakończenie możemy stwierdzić, że stosując prawo Coulomba i wyniki teorii względności, otrzymaliśmy wyrażenie (5.21) określające siłę działającą na ładunek poruszający się w odległości y od przewodnika z prądem (rys. 5.5). Siła ta jest wprost proporcjonalna do wielkości ładunku, jego prędkości i natężenia prądu w przewodniku; zaś odwrotnie proporcjonalna do odległości ładunku od przewodnika. Wartość tej siły zgodna jest z wynikami eksperymentalnymi

$$F = 2 \times 10^{-7} \frac{I}{y} qv. \quad (5.22)$$

Współczynnik proporcjonalności wynosi 2×10^{-7} jeżeli wszystkie wielkości wyrażone są w jednostkach układu SI.

5.6. Indukcja magnetyczna

Natężenie pola elektrycznego określamy jako stosunek siły elektrostatycznej do wartości ładunku. Analogicznie określamy pole magnetyczne B , jako stosunek siły magnetycznej do qv

$$B = \frac{F_{mag}}{qv}. \quad (5.23)$$

Wprowadzona wielkość B opisująca pole magnetyczne nazywana jest (ze względów historycznych) indukcją magnetyczną. Obecnie coraz częściej stosowana jest nazwa – pole magnetyczne. Wielkość \vec{B} mierzona jest w N/(Am); jednostce tej nadano nazwę tesla, która w układzie SI oznaczana jest przez T.

Dzieląc obydwie strony wyrażenia (5.21) przez qv , otrzymamy wartość pola magnetycznego w odległości y od izolowanego przewodnika z prądem I

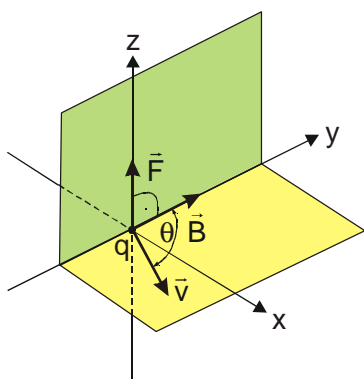
$$B = \frac{1}{4\pi\epsilon_0 c^2} \frac{2I}{y}.$$

Przy takim określeniu B często można spotkać inny zapis

$$B = \frac{\mu_0}{4\pi} \frac{2I}{y}, \quad (5.24)$$

gdzie μ_0 oznacza przenikalność magnetyczną próżni wynoszącą $4\pi \times 10^{-7} \text{ N/m}^2$. Zauważmy, że

$$\mu_0 \epsilon_0 = \frac{1}{c^2}. \quad (5.25)$$



Rys. 5.8. Siła magnetyczna \vec{F} działająca na ładunek q poruszający się z prędkością \vec{v} w polu magnetycznym \vec{B} jest prostopadła do obu wektorów.

Dotychczas rozważaliśmy tylko siłę działającą na ładunek q poruszający się równoległe do kierunku prądu I . Jeżeli ładunek q porusza się w innym kierunku, to teoria względności pozwala również określić kierunek siły magnetycznej działającej na ładunek. W tym przypadku zależność \vec{F}_{mag} od kierunku prędkości \vec{v} można opisać w postaci iloczynu wektorowego

$$\vec{F}_{mag} = q\vec{v} \times \vec{B}. \quad (5.26)$$

Zgodnie z określeniem iloczynu wektorowego, wartość bezwzględna siły wyraża się wzorem

$$F_{mag} = qvB \sin \theta, \quad (5.27)$$

gdzie θ jest kątem między \vec{v} i \vec{B} .

Związki między wektorami \vec{v} , \vec{B} i \vec{F} przedstawione są na rys. 5.8. Wektor \vec{F} jest prostopadły do wektorów \vec{v} i \vec{B} . Wartość siły jest maksymalna gdy $\vec{v} \perp \vec{B}$. Gdy \vec{v} i \vec{B} są do siebie równoległe, to siła $\vec{F} = 0$.

W jednorodnym polu magnetycznym ładunek porusza się po okręgu. W tym przypadku przyspieszenie wynosi v^2/R a siła nań działająca

$$F = \frac{mv^2}{R} = qvB,$$

stąd

$$R = \frac{mv}{qB}.$$

Prędkość kątowna ładunku $\omega = v/R = qB/m$, natomiast częstość

$$f = \frac{\omega}{2\pi} = \frac{qB}{2\pi m},$$

nie zależy od prędkości cząstki i jest nazywana częstością cyklotronową, ponieważ z taką częstością krąży naładowana cząstka w cyklotronie. Szybkie cząstki poruszają się po dużych okręgach, wolne po małych, ale każda z nich potrzebuje tego samego okresu T aby wykonać jeden obieg. Pole magnetyczne nie zmienia energii kinetycznej poruszającego się ładunku.

Częstość cyklotronowa jest parametrem charakterystycznym dla naładowanej cząstki w polu magnetycznym i można ją porównać z częstością charakterystyczną wahadła w ziemskim polu grawitacyjnym, czy też z częstością charakterystyczną układu drgającego masa-sprężynka.

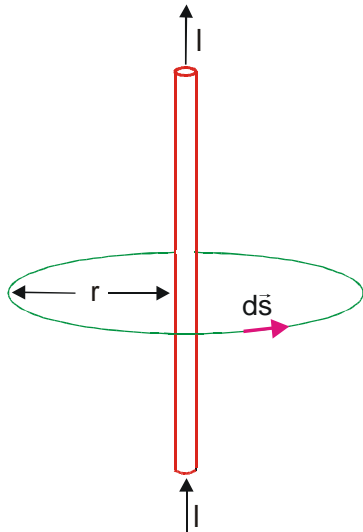
ROZDZIAŁ 6

POLE MAGNETYCZNE

Znajdziemy teraz pola magnetyczne utworzone przez typowe konfiguracje prądów. Do tego celu potrzebne jest nam równanie, które w przypadku pola magnetycznego odgrywałoby taką rolę jak twierdzenie Gaussa dla pola elektrycznego – takimi prawami są prawo Ampere'a i prawo Biota-Savarta-Laplace'a.

6.1. Prawo Ampere'a

W poprzednim rozdziale wykazaliśmy, że prąd I płynący w nieskończonym prostoliniowym przewodniku wytwarza pole magnetyczne. Kierunek wektora \vec{B} można określić za pomocą reguły prawej ręki. Jeżeli duży palec prawej ręki orientujemy w kierunku prądu, wówczas pozostałe palce w zgiętym położeniu wskazują kierunek linii sił pola magnetycznego \vec{B} . Linie sił pola magnetycznego mają postać koncentrycznych okręgów, w których środku znajduje się przewodnik z prądem (rys. 6.1).



Rys. 6.1. Całkę krzywoliniową oblicza się po zaznaczonym okręgu o promieniu r . Zaznaczono element konturu całkowania $d\vec{s}$.

Obliczymy teraz całkę z \vec{B} po konturze zamkniętej. Taką całkę nazywamy cyrkulacją i zapiszemy w postaci

$$\oint \vec{B} \cdot d\vec{s}.$$

W przypadku długiego prostoliniowego przewodnika (rys. 6.1) wybieramy zamknięty kontur w postaci okręgu o promieniu r . Ponieważ \vec{B} i $d\vec{s}$ są zawsze równoległe więc

$$\oint \vec{B} \cdot d\vec{s} = B \oint d\vec{s} = \frac{2I}{4\pi\epsilon_0 c^2 r} 2\pi r = \mu_0 I.$$

Wynik ten nie zależy od r i jest słuszny dla dowolnego okręgu, jeżeli przewodnik przechodzi przez jego środek.

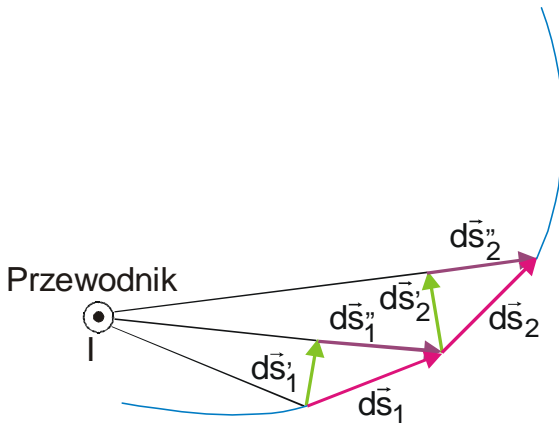
Można wykazać, że związek ten jest słuszny dla konturu dowolnego kształtu obejmującego przewodnik. Na rys. 6.2 zaznaczono elementy dowolnego konturu $d\vec{s}_1$ i $d\vec{s}_2$. Składowe $d\vec{s}'_1$ i $d\vec{s}''_1$ są prostopadłe do promienia-wektora wychodzącego z przewodnika. Widzimy, że

$$\vec{B}_1 \cdot d\vec{s}_1 = \vec{B}_1 \cdot (d\vec{s}'_1 + d\vec{s}''_1) = \vec{B}_1 \cdot d\vec{s}'_1,$$

ponieważ \vec{B} i $d\vec{s}''_1$ są wzajemnie prostopadłe, to $\vec{B} \cdot d\vec{s}''_1 = 0$. Podobnie

$$\vec{B}_1 \cdot d\vec{s}_1 + \vec{B}_2 \cdot d\vec{s}_2 = \vec{B}_1 \cdot d\vec{s}'_1 + \vec{B}_2 \cdot d\vec{s}'_2.$$

Zauważmy, że $d\vec{s}'_1$ i $d\vec{s}'_2$ są elementami odpowiednich okręgów obejmującymi jednakowe kąty o wierzchołkach na przewodniku. Ponieważ dla wszystkich łuków okręgów tworzących jednakowe kąty, wartości $\vec{B} \cdot d\vec{s}$ są jednakowe, to



Rys. 6.2. Dowolny kontur wokół przewodnika z prądem skierowanym zza płaszczyzny rysunku.

$$\oint_{\text{po konturze}} \vec{B} \cdot d\vec{s} = \oint_{\text{po okręgu}} \vec{B} \cdot d\vec{s} = \mu_o I.$$

Jeżeli dowolny kontur obejmuje n przewodników o różnych prądach I_1, I_2, \dots, I_n , to

$$\oint (\vec{B}_1 + \vec{B}_2 + \dots + \vec{B}_n) \cdot d\vec{s} = \mu_o (I_1 + I_2 + \dots + I_n),$$

gdzie \vec{B}_i oznacza pole wytworzone przez i -ty prąd. Stąd

$$\oint_C \vec{B} \cdot d\vec{s} = \mu_o I_c, \tag{6.1}$$

gdzie I_c jest całkowitym prądem objętym przez zamknięty kontur C .

Prawo Ampera brzmi: *cyrkulacja wektora indukcji magnetycznej jest równa sumie algebraicznej natężeń prądów płynących wewnątrz konturu całkowania pomnożonych przez przenikalność magnetyczną ośrodka.*

W ogólnym przypadku przenikalność magnetyczna ośrodka

$$\mu = \mu_o \mu_r, \tag{6.2}$$

gdzie μ_r jest liczbą bezwymiarową, zwaną względną przenikalnością magnetyczną.

W przypadku prądu niejednorodnego $I = \int \vec{j} \cdot d\vec{S}$, prawo Ampera zapisujemy w postaci

$$\oint_C \vec{B} \cdot d\vec{s} = \mu_o \int_S \vec{j} \cdot d\vec{S}, \tag{6.3}$$

gdzie powierzchnia S jest rozpięta na konturze C .

Chociaż przy wyprowadzaniu prawa Ampere'a rozważaliśmy tylko prądy prostoliniowe, prawo to można wyprowadzić dla dowolnych prądów stałych, tak prostoliniowych jak i krzywoliniowych.

Zwróćmy uwagę na podstawową różnicę między polem magnetycznym i polem elektrostatycznym. Wiemy, że pole elektrostatyczne jest zachowawcze co oznacza, że

$$\oint_C \vec{E} \cdot d\vec{s} = 0.$$

Natomiast *pole magnetyczne nie jest polem zachowawczym, ponieważ cyrkulacja wektora \vec{B} po konturze zamkniętej jest różna od zera. Pole takie nazywamy wirowym.*

6.2. Strumień magnetyczny

Strumień magnetyczny określamy w sposób analogiczny jak w przypadku strumienia elektrycznego, a mianowicie

$$\Phi_B = \int_S \vec{B} \cdot d\vec{S}. \quad (6.4)$$

Z rys. 6.1 widzimy, że linie sił pola magnetycznego \vec{B} zamykają się wokół prądu. Dlatego całka powyższa po dowolnej zamkniętej powierzchni jest równa zero (do wnętrza powierzchni wchodzi dokładnie taki sam strumień, jaki z niej wychodzi). Jeżeli mamy n prądów to wytworzony przez nie strumień jest równy

$$\oint \vec{B} \cdot d\vec{S} = \oint \vec{B}_1 \cdot d\vec{S} + \dots + \oint \vec{B}_n \cdot d\vec{S},$$

gdzie \vec{B}_i ($i = 1, 2, \dots, n$) pole magnetyczne wytworzone przez i -ty prąd. Ponieważ każdy człon prawej części równania jest równy zero, to mamy

$$\oint \vec{B} \cdot d\vec{S} = 0. \quad (6.5)$$

Tak jak w przypadku prawa Ampere'a, można udowodnić, że wzór ten jest słuszny dla prądów krzywoliniowych dowolnego kształtu.

Strumień magnetyczny przez dowolną zamkniętą powierzchnię równa się zero. Twierdzenie to, zwane twierdzeniem Gaussa dla pola magnetycznego \vec{B} , jest wynikiem faktu, że w przyrodzie nie występują "ładunki" magnetyczne. Oznacza to, że pole magnetyczne jest bezźródłowe. Jednostką strumienia jest weber; $\text{Wb} = \text{Tm}^2$.

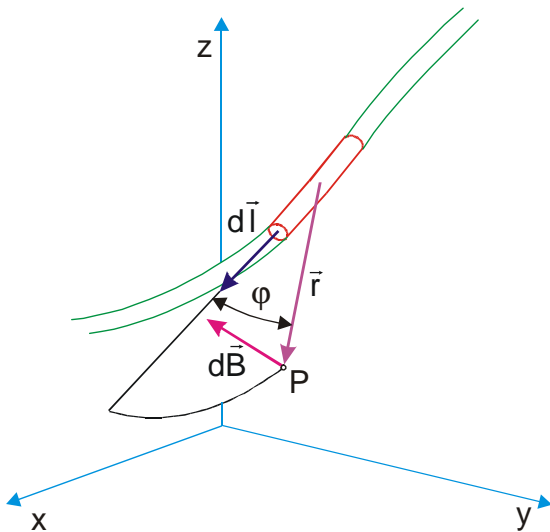
6.3. Prawo Biota-Savarta-Laplace'a

Pole magnetyczne wytworzone przez dowolny rozkład prądów można obliczyć stosując równanie zwane prawem Biota-Savarta-Laplace'a. Prawo to sformułował francuski matematyk i fizyk P. Laplace'a na podstawie wyników eksperymentalnych J. Biota i F. Savarta. Matematycznie równanie to jest ekwiwalentne prawu Ampere'a, chociaż w niektórych przypadkach wygodniej posługiwać się prawem Biota-Savarta-Laplace'a.

Element przewodnika $d\vec{l}$ z prądem I daje przyczynek $d\vec{B}$ do indukcji magnetycznej w punkcie P (rys. 6.3) określony wzorem

$$d\vec{B} = \frac{\mu_0 \mu_r}{4\pi} \frac{I}{r^3} d\vec{l} \times \vec{r}, \quad (6.6)$$

gdzie $d\vec{l}$ oznacza wektor (skierowany w kierunku prądu elektrycznego) o długości dl elementu przewodnika, \vec{r} promień wodzący przeprowadzony od rozpatrywanego elementu $d\vec{l}$ do punktu pola P, a r moduł promienia wodzącego. Wektor indukcji magnetycznej $d\vec{B}$ w dowolnym punkcie pola skierowany jest prostopadle do płaszczyzny, w której leżą wektory $d\vec{l}$ i \vec{r} , tak że obrót wektora $d\vec{l}$ do pokrycia się z wektorem \vec{r} wzdłuż najkrótszej drogi zgodnie z obrotem prawej śruby powoduje przesuw śruby wyznaczający zwrot $d\vec{B}$.



Rys. 6.3. $d\vec{B}$ jest indukcją pola magnetycznego jakie wytwarza element $d\vec{l}$ przewodnika z prądem w odległości \vec{r} od tego elementu.

Ze wzorów (6.6) i (6.7) wynika, że

$$d\vec{H} = \frac{I}{4\pi r^3} d\vec{l} \times \vec{r}. \quad (6.8)$$

Porównanie wektorowych wielkości charakteryzujących pole elektrostatyczne (\vec{E} oraz \vec{D}) i magnetyczne (\vec{B} i \vec{H}) wskazuje, że odpowiednikiem wektora \vec{E} pola elektrycznego jest wektor

Obok indukcji magnetycznej wprowadza się inną wielkość wektorową charakteryzującą pole magnetyczne; mianowicie natężenie pola magnetycznego \vec{H} związane z indukcją \vec{B} następującą zależnością

$$\vec{H} = \frac{\vec{B}}{\mu_0 \mu_r}. \quad (6.7)$$

Wzór ten jest słuszny dla przypadku jednorodnego, izotropowego ośrodka dla którego wektor \vec{H} jest równoległy do wektora \vec{B} . Sens fizyczny natężenia pola magnetycznego bliżej omówimy w pkt. 6.7. Natężenie pola \vec{H} mierzymy w jednostkach A/m.

indukcji magnetycznej \vec{B} ponieważ zarówno wielkość \vec{E} jak i wielkość \vec{B} określają dynamiczne oddziaływania tych pól i zależą od właściwości ośrodka, w którym wytwarzane są pola. Z kolei odpowiednikiem wektora indukcji elektrycznej \vec{D} jest natężenie pola magnetycznego \vec{H} .

Prawo Biota-Savarta-Laplace'a umożliwia określenie natężenia \vec{H} i indukcji \vec{B} pola magnetycznego prądu płynącego w przewodniku o dowolnym kształcie. Zgodnie z zasadą superpozycji, indukcja magnetyczna \vec{B} w dowolnym punkcie pola magnetycznego przewodnika przez który płynie prąd I równa się sumie wektorowej indukcji $\Delta\vec{B}_i$ elementarnych pól magnetycznych wytwarzanych przez poszczególne odcinki Δl_i tego przewodnika

$$\vec{B} = \sum_{i=1}^n \Delta\vec{B}_i, \quad (6.9)$$

gdzie n oznacza ogólną liczbę odcinków, na jakie podzielony został przewodnik. W granicy, gdy n dąży do nieskończoności

$$\vec{B} = \int_l d\vec{B}. \quad (6.10)$$

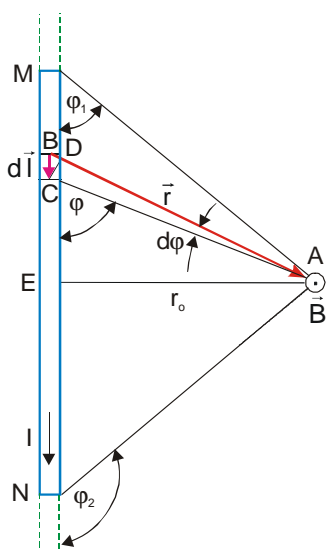
6.4. Pole magnetyczne niektórych konfiguracji prądów

6.4.1. Prostoliniowy przewodnik z prądem

Prawo Biota-Savarta-Laplace'a w postaci skalarnej ma postać

$$dB = \frac{\mu_o\mu_r}{4\pi} \frac{Idl \sin\varphi}{r^2}, \quad (6.11)$$

gdzie φ oznacza kąt zawarty między wektorami $d\vec{l}$ i \vec{r} .



Rys. 6.4. Pole magnetyczne prostoliniowego przewodnika z prądem.

Na rys. 6.4 pokazano część prostoliniowego przewodnika z prądem. Wektory $d\vec{l}$ i \vec{r} leżą w płaszczyźnie rysunku, co oznacza, że wszystkie wektory $d\vec{B}$ uwzględniające wpływ poszczególnych odcinków przewodnika na indukcję \vec{B} mają kierunek prostopadły do płaszczyzny rysunku i są zwrócone ku nam. W tej sytuacji wektor \vec{B} jest równy bezwzględnej algebraicznej sumie modułów wektorów $d\vec{B}$.

$$B = \int \frac{\mu_o\mu_r Idl \sin\varphi}{4\pi r^2} = \frac{\mu_o\mu_r}{4\pi} I \int \frac{dl \sin\varphi}{r^2} \quad (6.12)$$

Jak wynika z rys. 6.4

$$r = \frac{r_o}{\sin \varphi}, \quad \text{oraz} \quad dl = \frac{CD}{\sin \varphi}.$$

Tymczasem $CD = r d\varphi$, a stąd

$$dl = \frac{r d\varphi}{\sin \varphi} = \frac{r_o d\varphi}{\sin^2 \varphi}.$$

Podstawiając wyrażenie na r i dl do (6.12) uzyskamy

$$B = \frac{\mu_o \mu_r}{4\pi r_o} I \int_{\varphi_1}^{\varphi_2} \sin \varphi d\varphi = \frac{\mu_o \mu_r I}{4\pi r_o} (\cos \varphi_1 - \cos \varphi_2), \quad (6.13)$$

gdzie φ_1 i φ_2 oznaczają wartości kąta φ dla skrajnych punktów M i N przewodnika.

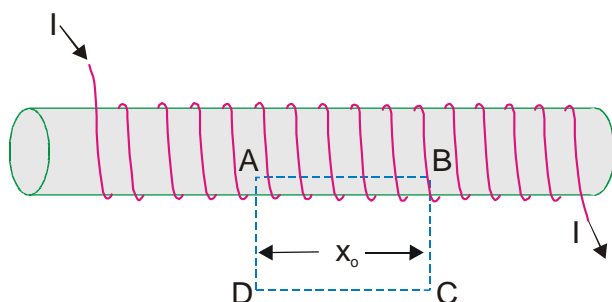
Jeżeli przewodnik jest nieskończenie długi, wówczas $\varphi_1 = 0$, a $\varphi_2 = \pi$, co prowadzi do wyniku

$$B = \frac{\mu_o \mu_r}{4\pi} \frac{2I}{r_o}.$$

Potwierdziłmy więc uprzednio uzyskane wyrażenie (5.24).

6.4.2. Solenoid

Solenoidem nazywamy cewkę cylindryczną składającą się z dużej liczby zwojów drutu, tworzących linię śrubową (rys. 6.5). Jeżeli zwoje solenoidu są rozmieszczone dostatecznie blisko siebie, wówczas solenoid można rozpatrywać jako układ połączonych szeregowo prądów kołowych o jednakowym promieniu, mających wspólną oś. Niech n oznacza ilość zwojów na jednostkę długości.



Aby obliczyć pole magnetyczne wewnątrz nieskończenie długiego solenoidu, skorzystamy z prawa Ampere'a dla prostokątnego konturu $ABCD$

$$\int_{ABCD} \vec{B} \cdot d\vec{s} = \mu_o \mu_r I_{cat}.$$

W naszym przypadku $I_{cat} = nx_o I$, gdzie I jest natężeniem prądu płynącego przez solenoid.

W związku tym możemy dalej napisać

Rys. 6.5. Solenoid. Prostokąt zaznaczony linią przerywaną jest konturem całkowania w prawie Ampere'a.

$$\vec{B}_{wew} \int_{AB} d\vec{s} + \int_{BC} \vec{B} \cdot d\vec{s} + \vec{B}_{zew} \int_{CD} d\vec{s} + \int_{DA} \vec{B} \cdot d\vec{s} = \mu_0 \mu_r n x_0 I .$$

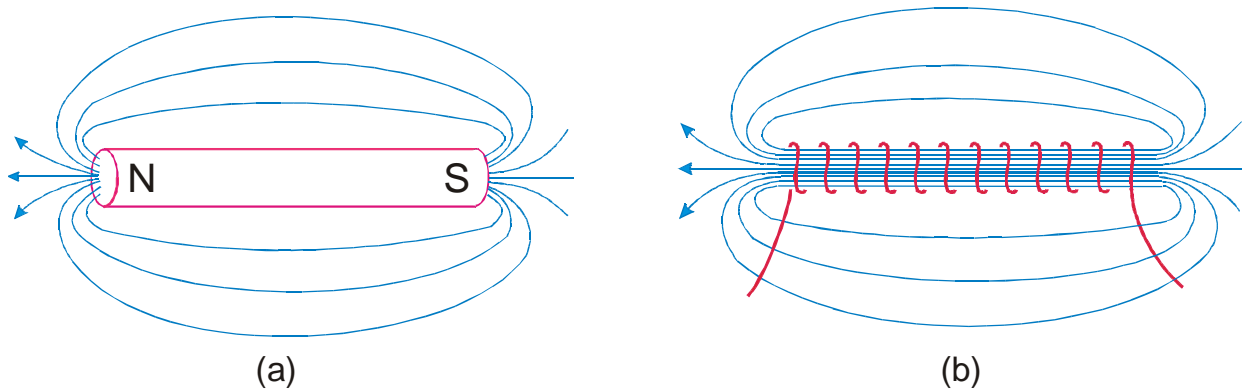
W wyrażeniu tym pierwsza całka jest równa Bx_0 ponieważ \vec{B} i $d\vec{s}$ są równoległe. Druga i czwarta całka są zerowe ze względu na wzajemną prostopadłość wektorów \vec{B} i $d\vec{s}$. Wartość trzeciej całki można przyjąć jako równą zero, ponieważ \vec{B}_{zew} jest bliskie zera. Tak więc

$$Bx_0 = \mu_0 \mu_r n x_0 I ,$$

czyli

$$B = \mu_0 \mu_r n I . \quad (6.14)$$

Zauważmy, że wartość pola magnetycznego nie zależy od położenia punktu wewnątrz solenoidu, tak że odcinek AB niekoniecznie musi leżeć na osi solenoidu. Pole wewnątrz solenoidu jest jednorodne i nie zależy od kształtu zwojów, jeżeli długość solenoidu jest dostatecznie długa. Kształt linii sił pola magnetycznego solenoidu jest podobny do linii sił pola wokół magnesu stałego o podobnym kształcie (rys. 6.6).



Rys. 6.6. Linie sił pola magnetycznego magnesu stałego (a) i solenoidu (b) o podobnym kształcie.

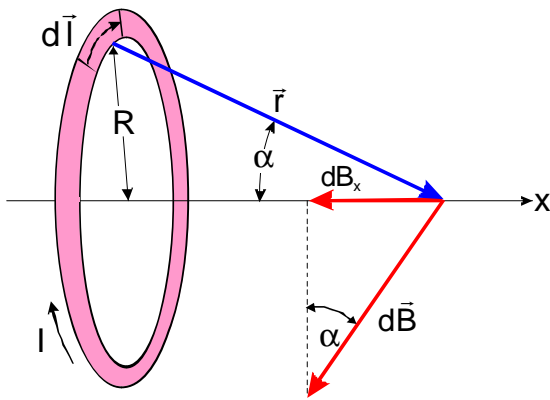
6.4.3. Prąd kołowy

Rysunek 6.7 przedstawia drut w kształcie kołowym o promieniu R , przez który płynie prąd I . Każdy element o długości dl wnosi przyczynek $d\vec{B}$ prostopadły do \vec{r} . Zgodnie z równaniem (6.6)

$$dB = \frac{\mu_0 \mu_r}{4\pi} I \frac{dl}{r^2} .$$

Składowa tego pola magnetycznego wzdłuż osi x wynosi

$$dB_x = \frac{\mu_0 \mu_r}{4\pi} I \frac{dl}{r^2} \sin \alpha = \frac{\mu_0 \mu_r}{4\pi} I \frac{R}{r^3} dl .$$



Całkując po obwodzie koła mamy

$$B = \frac{\mu_0 \mu_r}{2} \frac{IR^2}{r^3}.$$

Z powyższego wzoru wynika, że indukcja wypadkowego pola w środku kołowego przewodnika wynosi

$$B = \mu_0 \mu_r \frac{I}{2R}. \quad (6.15)$$

Rys. 6.7. Pole magnetyczne $d\vec{B}$ utworzone przez element prądu kołowego $I dl$ o promieniu R .

6.5. Oddziaływanie przewodników z prądem

Jeżeli w polu magnetycznym o indukcji \vec{B} znajduje się przewodnik, w którym płynie prąd I , to w oparciu o wzór (5.14) można obliczyć siłę która działa na element przewodnika $d\vec{l}$. Mianowicie, na ładunek dq poruszający się z prędkością \vec{v} działa siła

$$d\vec{F} = dq\vec{v} \times \vec{B} = dq \frac{d\vec{l}}{dt} \times \vec{B} = \frac{dq}{dt} d\vec{l} \times \vec{B},$$

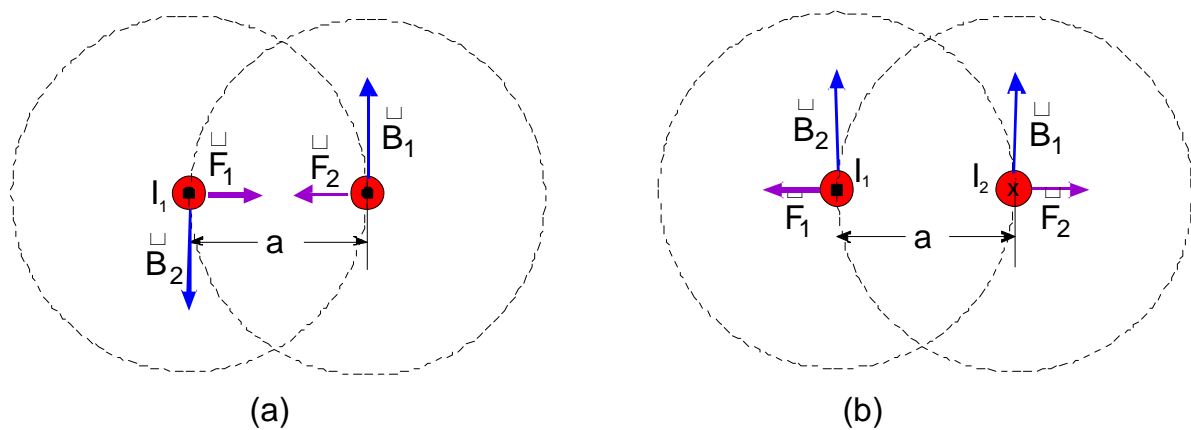
czyli

$$d\vec{F} = I d\vec{l} \times \vec{B}. \quad (6.16)$$

Jest to wzór Ampere'a na siłę elektrodynamiczną. Kierunek działającej siły jest zgodny z kierunkiem iloczynu wektorowego $d\vec{l} \times \vec{B}$ (reguła prawej śruby). Aby policzyć siłę działającą na cały przewodnik z prądem, należy podzielić go na elementy o długości $d\vec{l}$ i zsumować wektorowo przyczynki dla każdego elementu.

Zwróćmy uwagę na osobliwość oddziaływania elektromagnetycznego wyrażoną wzorem (6.16). W elektrostatyce mieliśmy do czynienia z siłami centralnymi, ponieważ siła oddziaływania dwóch ładunków punktowych skierowana jest wzdłuż prostej łączącej ładunki. Tymczasem siła oddziaływania elektromagnetycznego skierowana jest prostopadle do linii sił pola magnetycznego.

W punkcie 5.5 (rys. 5.4) wspomnieliśmy, że dwa równoległe przewodniki umieszczone w niedużej odległości odpychają się, gdy kierunki płynących w nich prądów są przeciwne; lub przyciągają się, gdy kierunki tych prądów są jednakowo skierowane. Oddziaływanie wzajemne równoległych prądów można wytłumaczyć uwzględniając fakt, że każdy z przewodników wytwarza pole magnetyczne, oddziaływujące na drugi przewodnik z prądem. Rysunek 6.8 ilustruje to oddziaływanie.



Rys. 6.8. Oddziaływanie między dwoma równoległymi przewodnikami z prądem.

Jeżeli druty są bardzo długie, to przewodnik 1 wytwarza pole B_1 w odległości a od siebie (wzór (5.24))

$$B_1 = \frac{\mu_o \mu_r}{4\pi} \frac{2I_1}{a}.$$

Pod wpływem tego pola na przewodnik 2 działa siła dF_2

$$dF_2 = \frac{\mu_o \mu_r}{4\pi} \frac{2I_1 I_2}{a} dl.$$

gdzie dl jest elementem drugiego przewodnika. W sposób analogiczny można pokazać, że

$$dF_1 = \frac{\mu_o \mu_r}{4\pi} \frac{2I_1 I_2}{a} dl,$$

Można zatem napisać, że $dF_1 = dF_2 = dF$. W celu określenia modułu wektora siły \vec{F} działającej na odcinku przewodnika o skończonej długości l , scałkujemy to równanie względem l od 0 do l . Otrzymamy

$$F = \frac{\mu_o \mu_r}{4\pi} \frac{2I_1 I_2}{a} l,$$

a z ostatniego wyrażenia mamy siłę działającą na jednostkę długości każdego z przewodników

$$\frac{F}{l} = \frac{\mu_o \mu_r}{4\pi} \frac{2I_1 I_2}{a}. \quad (6.17)$$

Wzór (6.17) pozwala zdefiniować jednostkę natężenia prądu 1A w układzie SI.

Amper jest natężeniem prądu, który płynąc w dwóch równoległych, prostoliniowych, nieskończenie długich przewodnikach o przekroju okrągłym, znikomo małym; umieszczonych w odległości 1m jeden od drugiego, wywołałby między tymi przewodnikami siłę oddziaływania 2×10^{-7} N na każdy metr długości.

6.6. Efekt Halla

Zjawisko Halla polega na powstaniu w metalu lub półprzewodniku, pola elektrycznego skierowanego prostopadle do wektora magnetycznego \vec{B} i wektora gęstości prądu \vec{j} płynącego w próbce. Zjawisko to zostało odkryte przez amerykańskiego fizyka Halla w 1879 r.

Rozważmy płytkę metalu (lub półprzewodnika) umieszczoną w polu magnetycznym, w którym płynie prąd o gęstości \vec{j} (rys. 6.9). Przy danym kierunku \vec{j} prędkość elektronów skierowana jest w lewo, a pod wpływem siły Lorentza elektrony odchylane są do góry. Wobec tego na górnej powierzchni płytki metalu pojawia się zwiększona koncentracja elektronów (powierzchnia ładuje się ujemnie), natomiast na dolnej – niedobór elektronów (dolna powierzchnia ładuje się dodatnio). Na skutek tego, pomiędzy tymi dwoma powierzchniami powstaje dodatkowe, poprzeczne pole elektryczne skierowane z dołu do góry. Kiedy natężenie \vec{E}_B tego poprzecznego pola elektrycznego osiągnie wielkość równoważącą działanie siły Lorentza, to ustali się stacjonarny rozkład ładunków w kierunku poprzecznym. Wówczas:

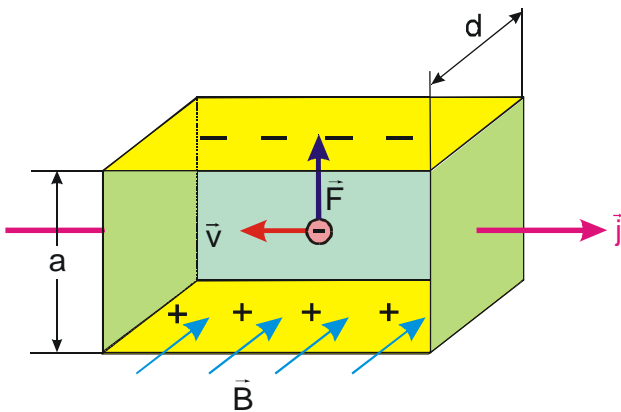
$$eE_B = e \frac{\Delta V}{a} = evB,$$

czyli

$$\Delta V = vBa,$$

gdzie a jest szerokością płytki, a ΔV – poprzeczną hallowską różnicą potencjałów. Uwzględniając, że natężenie prądu $I = jS = nevS$, otrzymujemy

$$\Delta V = \frac{I}{nead} Ba = \frac{1}{en} \frac{IB}{d} = R \frac{IB}{d}. \quad (6.18)$$



Rys. 6.9. Efekt Halla

Widzimy, że napięcie Halla (halowska poprzeczna różnica potencjałów) jest wprost proporcjonalne do indukcji magnetycznej B , natężenia prądu I i odwrotnie proporcjonalne do grubości płytki d . We wzorze (6.18) $R = 1/en$ nazwane jest stałą Halla. Z pomiarów eksperymentalnych można określić koncentrację nośników prądu w przewodniku, określić typ przewodnictwa półprzewodnika (znak stałej Halla jest zgodny ze znakiem ładunku e nośników prądu). Z tego powodu pomiar efektu Halla jest efektywną metodą badania typu i koncentracji nośników w metalach i półprzewodnikach.

6.7. Magnetyzm

Namagnesowany pręt (magnes stały) w jednorodnym polu magnetycznym dąży do położenia wzdłuż linii sił pola magnetycznego (rys. 6.10). Kompas magnetyczny również stanowi nieduży namagnesowany pręcik, orientujący się wzdłuż linii pola magnetycznego. Zachowanie magnesu można opisać zakładając, że na jednym końcu znajduje się ładunek magnetyczny q_m , a na drugim końcu – ładunek magnetyczny $-q_m$. Gdyby istniał ładunek magnetyczny, to w polu magnetycznym działałaby na niego siła $\vec{F} = q_m \cdot \vec{B}$ analogiczna do siły działającej na ładunek elektryczny w polu elektrycznym.

Chociaż w przyrodzie ładunki magnetyczne nie istnieją, pojęcie to wprowadza się ze względu na wygodny matematyczny sposób opisanie właściwości magnesów.

Moment sił działających na magnes (rys. 6.10) wynosi

$$T = Fl \sin \alpha ,$$

czyli

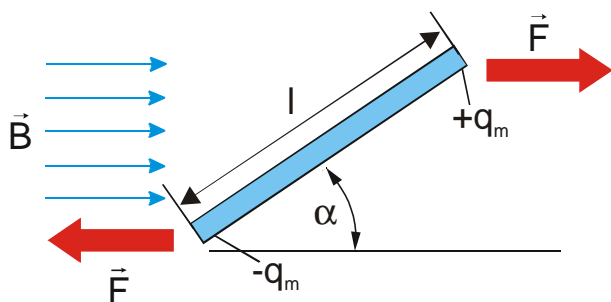
$$T = q_m Bl \sin \alpha . \quad (6.19)$$

Iloczyn $q_m l = \mu$ określa się jako moment magnetyczny. Wobec tego

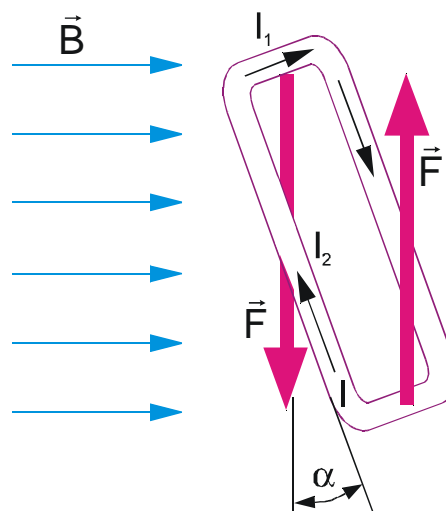
$$T = \mu B \sin \alpha ,$$

a w zapisie wektorowym

$$\vec{T} = \vec{\mu} \times \vec{B} . \quad (6.20)$$



Rys. 6.10. Namagnesowana próbka o długości l położona pod kątem α do linii sił zewnętrznego pola magnetycznego \vec{B} .



Rys. 6.11. Prostokątna ramka o powierzchni $l_1 l_2$ w jednorodnym polu magnetycznym.

Należy zauważyć, że w analogiczny sposób zachowuje się pętla z prądem (rys. 6.11). Obliczmy iloczyn wektorowy $\vec{l} \times \vec{B}$ dla każdego z czterech boków ramki. Siły magnetyczne przyłożone do dwóch przeciwległych boków o długości l_1 tworzą moment obrotowy

$$T = Fl_2 \sin \alpha .$$

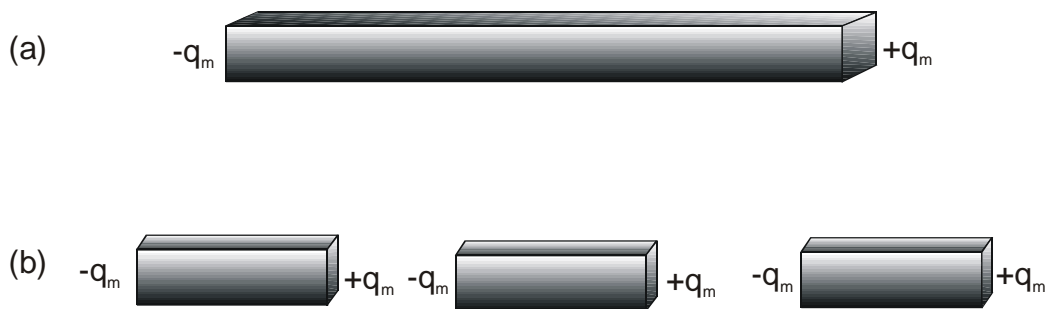
Ponieważ siła działająca ze strony pola magnetycznego na przewodnik z prądem wynosi $F = Il_1B$, stąd

$$T = (Il_1B)(l_2 \sin \alpha) = Il_1l_2B \sin \alpha = ISB \sin \alpha . \quad (6.21)$$

Siły przyłożone do dwóch stron ramki o długości l_2 działają w przeciwnych kierunkach wzdłuż jednej osi i dlatego wzajemnie kompensują się.

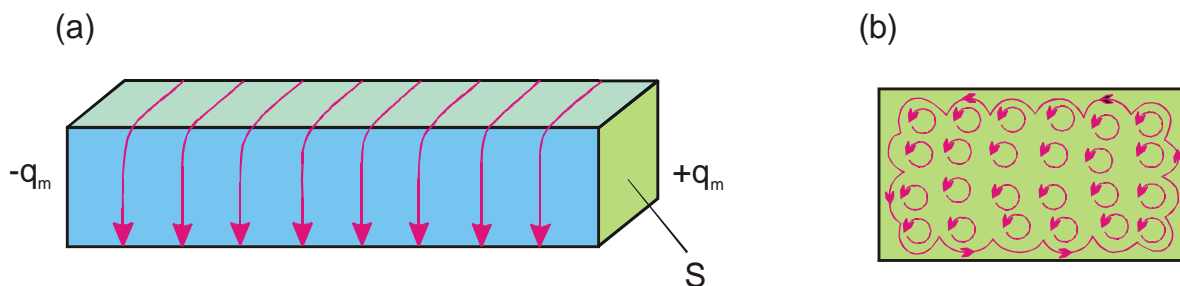
Wynika z tego, że pętla z prądem wytwarza pole magnetyczne identycznie jak magnes. Przyrównując prawe strony wyrażeń (6.19) i (6.21) znajdziemy moment magnetyczny pętli z prądem

$$\mu = IS . \quad (6.22)$$



Rys. 6.12. (a) Magnes. (b) Magnes podzielony na trzy części.

Do chwili odkrycia elektronu i prądu elektronowego w metalach zachowanie magnesów w polach magnetycznych wyjaśniano na podstawie hipotezy o ładunkach magnetycznych. Gdyby tak było, można byłoby wydzielić wyizolowany biegun magnetyczny dodatni, bądź ujemny (tak jak w przypadku ładunków elektrycznych). Wydawałoby się, że można to zrobić łamiąc jeden z końców magnesu. Jednak przy tym na drugim końcu pojawia się nowy, przeciwny biegun (rys. 6.12). Do chwili obecnej nie udało się odkryć w przyrodzie izolowanego bieguna magnetycznego. Fizycy są zgodni, że właściwości magnesów można wyjaśnić zamkniętymi, wewnątrzatomowymi prądami. Hipotezę tę wysunął Amper'e. Obecnie wiemy, że te mikroprądy są spowodowane orbitalnym i spinowym ruchem elektronów. Wewnątrz magnesu, prądy w sąsiadujących ze sobą częściach obwodów elementarnych, mają kierunki przeciwne więc znoszą się (rys. 6.13b). Nie znoszą się natomiast prądy płynące w częściach zewnętrznych elementarnych obwodów położonych na obwodzie przekroju, dając jakby wypadkowy prąd opływający zewnętrzny obwód przekroju.



Rys. 6.13. (a) Magnes po którego powierzchni cyrkulują prądy Ampere'a. (b) Widok magnesu w przekroju – pokazano prądy atomowe, wypadkowy prąd zaznaczono grubą linią.

Tak więc dowolny magnes stanowi jak gdyby solenoid, po powierzchni którego cyrkuluje prąd Ampere'a. Obliczmy wielkość tego prądu dla magnesu o ładunku magnetycznym q_m . Na każdy zwój solenoidu działa moment sił $T_1 = ISB \sin \alpha$. Na solenoid zawierający N zwojów będzie działał moment sił $T = NISB \sin \alpha$. Porównując to wyrażenie z (6.19) mamy

$$q_m Bl \sin \alpha = NISB \sin \alpha ,$$

czyli

$$q_m = \frac{N}{l} IS = n_l IS , \quad (6.23)$$

gdzie $n_l I$ ma sens fizyczny prądu powierzchniowego. Wzór (6.23) potwierdza fakt, że solenoid o prądzie powierzchniowym $n_l I$ zachowuje się podobnie jak magnes o ładunku $q_m = n_l IS$.

Wprowadziliśmy uprzednio dwie wielkości charakteryzujące pole magnetyczne \vec{B}, \vec{H} i relację (6.7) wiążącą te wielkości. Zastanówmy się bliżej nad fizycznym sensem wprowadzania tych dwóch wektorów. Otóż pola magnetyczne wytworzone przez prądy płynące w przewodnikach określają natężenie pola magnetycznego \vec{H} . Tylko ta wielkość jest niezmiennicza przy otaczaniu przewodów z prądem kolejno różnymi ośrodkami. O natężeniu pola decydują tylko prądy – ich natężenie, kształt przewodnika oraz położenie punktu w którym się je wyznacza.

Jeżeli obszar w którym istnieje pole magnetyczne wypełnia jakaś substancja, to okazuje się, że wartość sił działających na przewodnik z prądem zależy od rodzaju substancji. Wynika stąd, że indukcja magnetyczna, decydująca o działaniu dynamicznym pola, zależy od rodzaju tej substancji. W celu wyjaśnienia tego faktu należy przyjąć, że ośrodek (substancja) "magnesuje się" w pewien sposób.

Podobnie jak w dielektrykach, tak w magnetykach istnieją, lub też powstają pod wpływem pola magnetycznego, momenty magnetyczne (dipole magnetyczne). Po przyłożeniu zewnętrznego pola magnetycznego dipole te porządkują się w pewien sposób i ciało jako całość stanowi pewien dipol magnetyczny. Ale dipol wytwarza pole własne, które może dodawać się lub odejmować od pola zewnętrznego. Dlatego wypadkowe pola w ośrodkach będą na ogół różne od pola w próżni.

Wielkością charakteryzującą pole magnetyczne związane jedynie z magnetycznymi właściwościami ośrodka jest magnetyzacja. Jest to wielkość wektorowa zdefiniowana jako moment magnetyczny jednostki objętości.

$$\vec{J} = \frac{1}{V} \sum \vec{p}_m, \quad (6.24)$$

a więc iloraz wypadkowego momentu magnetycznego dowolnej objętości ciała V przez tę objętość.

Wypadkową indukcję w ośrodku tworzą więc dwie składowe: indukcja \vec{B}_o związana jedynie z prądem makroskopowym wytwarzającym pole magnetyczne o natężeniu \vec{H} (składowa niezależna od ośrodka) i indukcja \vec{B}_J związana jedynie z prądami magnesującymi, które na ogół (z wyjątkiem magnesów stałych) powstają pod wpływem zewnętrznego pola \vec{H} . Mamy więc

$$\vec{B} = \vec{B}_o + \vec{B}_J, \quad (6.25)$$

a każda ze składowych wiąże się z wielkością bezpośrednio odpowiedzialną za jej powstanie

$$\vec{B}_o = \mu_o \vec{H} \quad ; \quad \vec{B}_J = \mu_o \vec{J}. \quad (6.26)$$

Ponieważ magnetyzacja \vec{J} zależy od natężenia pola magnetycznego

$$\vec{J} = \chi \vec{H}, \quad (6.27)$$

gdzie χ jest wielkością bezwymiarową nazwaną podatnością magnetyczną. Ze wzorów powyższych otrzymujemy

$$\vec{B} = \mu_o (1 + \chi) \vec{H} = \mu_o \mu_r \vec{H}. \quad (6.28)$$

Względna przenikalność magnetyczna ośrodka μ_r określa ile razy pole magnetyczne makroprądów \vec{H} zwiększa się na skutek pola mikroprądów ośrodka. Indukcja magnetyczna \vec{B} w ośrodku może być co do modułu większa lub mniejsza niż indukcja w próżni, zależnie od tego czy natężenie pola magnetycznego \vec{H} i magnetyzacja \vec{J} mają te same zwroty czy też przeciwne. Z tego powodu wszystkie ciała można zaliczyć do jednej z trzech podstawowych grup: mogą być one diamagnetykami, paramagnetykami lub ferromagnetykami. Ponieważ właściwości magnetyczne ciał uwarunkowane są głównie orbitalnym i spinowym momentem magnetycznym elektronów, więc właściwości te omówione zostaną po wprowadzeniu elementów mechaniki kwantowej.

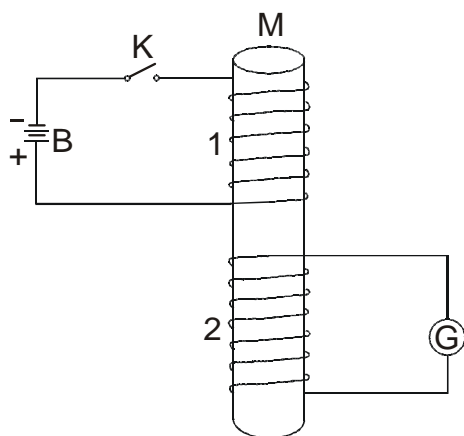
ROZDZIAŁ 7

INDUKCJA ELEKTROMAGNETYCZNA

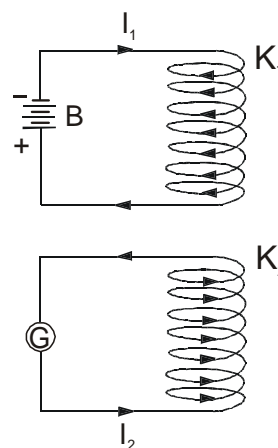
Dotychczas poznaliśmy prawa rządzące elektrycznością, magnetyzmem oraz związek między ruchem ładunków elektrycznych i polem magnetycznym. Udało nam się sformułować prawa z tym związane: prawo Gaussa, prawo Ampere'a i prawo Biota-Savarta-Laplace'a. Prawa te nie opisują jednak bezpośredniego związku pola magnetycznego z polem elektrycznym. Czy pole magnetyczne powoduje powstanie pola elektrycznego i jakie prawa rządzą tym procesem? W 1831 roku, po dziesięciu latach wytrwałych prób, M. Faradayowi udało się rozwiązać to zagadnienie. Wykonał eksperyment, który miał w następstwie olbrzymie znaczenie dla rozwoju fizyki i techniki.

7.1. Prawo Faradaya

Schemat doświadczenia przeprowadzonego przez Faradaya przedstawiono na rys. 7.1. Na pręt drewniany M nawinięte były dwa długie druty miedziane. Przy nie zmieniającym się natężeniu prądu w pierwszym obwodzie, w drugim obwodzie galwanometr nie wykazywał prądu, natomiast w czasie zwiernania i rozwierania wyłącznika K wskazówka galwanometru odchyłała się nieco, a następnie szybko wracała do położenia równowagi. Świadczy to o powstawaniu w danym obwodzie krótkotrwałego prądu nazywanego prądem indukowanym.



Rys. 7.1. Schemat doświadczenia Faradaya prowadzącego do odkrycia zjawiska indukcji.



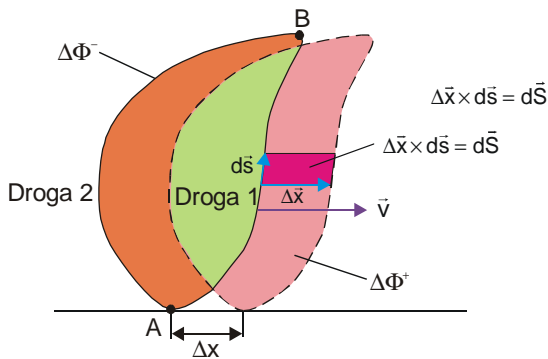
Rys. 7.2. Powstanie prądu indukcyjnego I_2 w czasie ruchu cewki z prądem I_1 .

Kierunki prądów indukowanych były przeciwne dla przypadku zwierania i rozwierania wyłącznika K . Zamiast stosować gwałtowne zmiany prądu przy użyciu klucza K , Faraday wykazał, że prąd indukowany wytwarza się również przy łagodnych zmianach prądu w obwodzie, uzyskanych przy pomocy rezystorów o zmiennej rezystancji. Faraday uzyskał również prądy indukowane nieco innymi metodami. Na rys. 7.2 przedstawione są dwie cewki: jedna z prądem stałym, druga połączona z galwanometrem. Prąd w drugiej cewce płynie wówczas, gdy cewki są we wzajemnym ruchu. Przy zbliżaniu i oddalaniu, prądy indukowane w cewce 2 mają kierunki przeciwne. Podobne zjawiska zachodzą, gdy obwód I z prądem zostanie zastąpiony magnesem stałym. W obu przypadkach prądy indukowane płyną jedynie w czasie ruchu względnego obwodów lub magnesu i obwodu. W czasie spoczynku prąd indukowany przestaje istnieć.

W celu wyjaśnienia odkrytych przez Faradaya zjawisk indukcji elektromagnetycznej w możliwie jasny sposób, odbiegniemy nieco od historycznego toku wydarzeń. Wykażemy, że w zamkniętym przewodzącym konturze dowolnego kształtu poruszającym się w polu magnetycznym (rys. 7.3) powstaje SEM. Pole magnetyczne \vec{B} może być dowolną funkcją współrzędnych. Praca wykonana przeciwko siłom magnetycznym przy przemieszczeniu ładunku q na odległość $d\vec{s}$ wynosi

$$dW = \vec{F}_{mag} \cdot d\vec{s} = (q\vec{v} \times \vec{B})d\vec{s} = q \left(\frac{d\vec{x}}{dt} \times \vec{B} \right) d\vec{s},$$

gdzie $d\vec{s}$ jest wektorem o długości ds . Stosując tożsamość wektorową



Rys. 7.3. Zamknięty kontur porusza się z prędkością \vec{v} wzdłuż kierunku osi x . Linia przerywaną zaznaczono położenie konturu po czasie Δt .

Całkowaną pracę wykonaną przy przemieszczaniu ładunku q z punktu A do punktu B po drodze I konturu zapiszemy w postaci

$$W_{AB1} = \int_A^B dW = -q \frac{\int_A^B \vec{B} \cdot d\vec{S}}{dt} = -q \frac{d\Phi^+}{dt}$$

$$\vec{a} \times \vec{b} \cdot \vec{c} = \vec{a} \cdot \vec{b} \times \vec{c},$$

piszemy wyrażenie na pracę w następujący sposób

$$dW = -q \frac{(\vec{B} \times d\vec{x})d\vec{s}}{dt} = -\frac{\vec{B}(d\vec{x} \times d\vec{s})}{dt}.$$

Z rys 7.3 widać, że $d\vec{x} \times d\vec{s}$ można zastąpić elementem powierzchni $d\vec{S}$, wówczas

$$dW = -q \frac{\vec{B} \cdot d\vec{S}}{dt}.$$

a $d\Phi^+$ oznacza wzrost strumienia w prawym obszarze zakreskowanym. Analogiczna praca wykonana przy przemieszczeniu ładunku po drodze 2 z punktu B do punktu A

$$W_{BA2} = \int_B^A dW = q \frac{d\Phi^-}{dt}.$$

Z kolei $d\Phi^-$ oznacza zmniejszanie strumienia w lewym obszarze zakreskowanym. Praca zużyta na przemieszczenie jednostkowego ładunku po całym konturze

$$W = W_{AB1} + W_{BA2} = -q \left(\frac{d\Phi^+}{dt} - \frac{d\Phi^-}{dt} \right),$$

a ponieważ $d\Phi^+ - d\Phi^-$ oznacza zmianę strumienia magnetycznego przez powierzchnię ograniczoną konturem ABA , więc praca zużyta na przemieszczenie ładunku wynosi

$$W = -q \frac{d\Phi_B}{dt}.$$

Ponieważ siłę elektromotoryczną określamy jako pracę zużyta na przemieszczenie jednostkowego ładunku (patrz pkt. 5.4), więc

$$SEM = \frac{W}{q} = -\frac{d\Phi_B}{dt}. \quad (7.1)$$

Siła elektromotoryczna nie jest siłą w dosłownym tego słowa znaczeniu. Mierzona jest w voltach (J/C), a więc *przedstawia energię przypadającą na jednostkowy ładunek, dostarczoną elektronowi przewodnictwa przy obejściu obwodu*. W obwodzie zamkniętym umieszczonym w zmiennym polu magnetycznym pojawia się więc siła elektromotoryczna indukcji elektromagnetycznej. Jeżeli kontur jest nieruchomy, to siła magnetyczna znika ($\vec{v} = 0$) i $SEM = 0$. Jednakże, jeżeli źródło pola magnetycznego porusza się i powoduje zmianę strumienia obejmowanego przez kontur, to w konturze powstaje pole elektryczne

$$\oint \vec{E} \cdot d\vec{s} = -\frac{d\Phi_B}{dt}. \quad (7.2)$$

Formułę tę można bezpośrednio otrzymać z (7.1) stosując zasadę względności do konturu i źródła pola magnetycznego. Obserwator nieruchomy względem konturu zauważy taką samą siłę działającą na jednostkowy ładunek q , co obserwator poruszający się razem ze źródłem pola magnetycznego. Dla obserwatora nieruchomego, siła działająca na jednostkowy ładunek, według definicji, przedstawia pole elektryczne. Formuła (7.2) jest słuszna również dla przypadku nieruchomych obwodów. Zmiana prądu w jednym obwodzie wywołuje zmianę pola magnetycznego obejmowanego przez kontur drugiego obwodu i indukuje w nim siłę elektromotoryczną.

W rzeczywistości jednak istnienie drugiego obwodu nie jest konieczne. Pole elektryczne pojawia się niezależnie od tego czy obwód istnieje, czy też go nie ma. Równanie (7.2) ma bardziej ogólny charakter i jest słuszne dla dowolnego domniemanego obwodu zamkniętego w przestrzeni. Równanie to można również zapisać w postaci

$$\oint_C \vec{E} \cdot d\vec{s} = -\frac{d}{dt} \left(\int_S \vec{B} \cdot d\vec{S} \right),$$

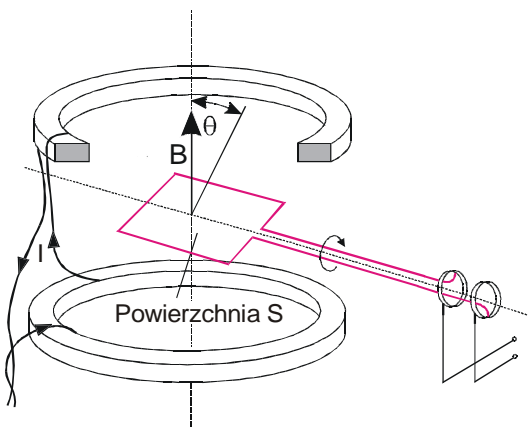
gdzie S oznacza dowolną powierzchnię rozpiętą na konturze C . Ponieważ granice całkowania po $d\vec{S}$ nie zmieniają się w czasie, można przejść z różniczkowaniem pod znak całki

$$\oint_C \vec{E} \cdot d\vec{s} = -\int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}. \quad (7.3)$$

Zwróćmy uwagę na to, że całka okrężna wektora natężenia pola elektrycznego wzdłuż obwodu zamkniętego nie jest równa zero. Wynika stąd, że pole elektryczne wzbudzone przez zmienne pole magnetyczne jest polem wirowym.

Zjawisko odkryte przez Faradaya stanowiło podstawę, która umożliwiła zbudowanie w następnych latach silników, prądnic i transformatorów elektrycznych. Z tego powodu Faraday uważany jest za jednego z twórców elektrotechniki.

Najpospolitszą częścią urządzeń elektrycznych jest pętla lub cewka obracająca się ze stałą prędkością w jednorodnym polu magnetycznym o indukcji B (rys. 7.4).



Rys. 7.4. Dwie cewki wytwarzają w przybliżeniu jednorodne pole magnetyczne o indukcji \vec{B} . Pętla obraca się z prędkością kątową ω . Indukuje się w niej sinusoidalna SEM.

Niech prędkość kątowa pętli wynosi ω . Położenie pętli określa kąt $\theta = \omega t + \alpha$, gdzie α określa położenie pętli w chwili $t = 0$. Składowa indukcji \vec{B} prostopadła do powierzchni pętli wynosi $B \sin \theta$. W związku z tym strumień indukcji płynący przez pętlę w chwili t jest równy

$$\Phi_B(t) = SB \sin(\omega t + \alpha),$$

gdzie S jest powierzchnią pętli. Indukowana siła elektromotoryczna wynosi

$$SEM = -SB\omega \cos(\omega t + \alpha). \quad (7.4)$$

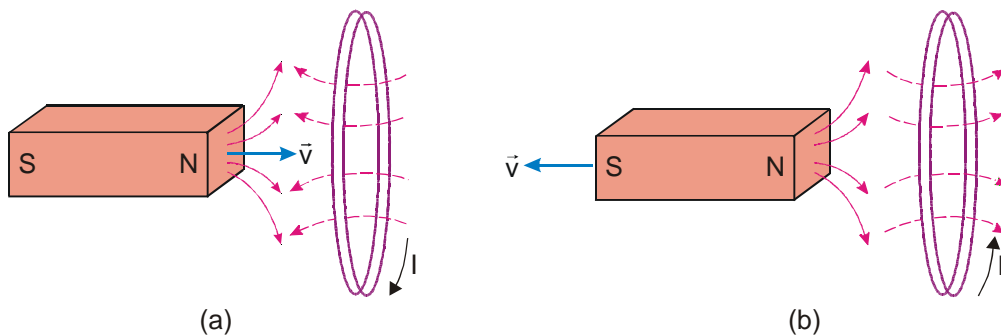
7.2. Reguła Lenza

W 1834 roku Lenz ustalił następującą regułę: *prąd indukowany w obwodzie ma zawsze taki kierunek, że wytworzony przezeń strumień magnetyczny przez powierzchnię ograniczoną przez*

ten obwód przeciwdziała zmianom strumienia, które wywołały pojawienie się prądu indukowanego.

Matematycznym wyrazem reguły Lenza jest znak ”-” w równaniach (7.1)–(7.3). Zauważmy, że zwiększenie strumienia ($d\Phi_B / dt > 0$) wywołuje $SEM < 0$, to jest pole indukowanego prądu skierowane jest przeciwko strumieniowi. Z kolei zmniejszenie strumienia ($d\Phi_B / dt < 0$) wywołuje $SEM > 0$, t.j. kierunki strumienia i pola indukowanego prądu są zgodne.

Reguła Lenza jest zilustrowana poglądowo na rys. 7.5. Gdy magnes stały porusza się w prawo (rys. 7.5a), zwiększa się strumień magnetyczny przez zamkniętą pętlę i prąd indukowany I wytwarza pole skierowane przeciwnie do pierwotnego strumienia. Na rys. 7.5b z kolei pokazano początkowo nieruchomy magnes, który zaczyna poruszać się w lewo, co prowadzi do zmniejszenia strumienia magnetycznego przechodzącego przez pętlę. Prąd indukowany wytwarza pole (linie przerywane) przeciwdziałające przyczynie, która go spowodowała.



Rys. 7.5. Ilustracja reguły Lenza.

Reguła Lenza jest konsekwencją spełnienia prawa zachowania energii. Wróćmy na chwilę do obwodu poruszającego się w polu magnetycznym (rys. 7.3). Jeżeli rezystancja obwodu wynosi R , to zgodnie z prawem zachowania energii, na pracę źródła prądu w czasie dt ($EIdt$) składa się praca na ciepło Joule'a (I^2Rdt) i praca związana z przemieszczeniem obwodu w polu magnetycznym ($Id\Phi_B$). Mamy więc

$$E I dt = I^2 R dt + I d\Phi_B,$$

stąd

$$I = \frac{1}{R} \left(E - \frac{d\Phi_B}{dt} \right) = \frac{1}{R} (E + E_i),$$

gdzie $E_i = -d\Phi_B / dt$ jest indukowaną siłą elektromotoryczną.

Dotychczas rozważaliśmy prądy indukowane w obwodach liniowych. Prądy te mogą jednak powstawać również w przewodnikach masywnych. Obwód zamknięty prądu indukowanego tworzy się samorzutnie w przewodniku. Nazywamy je prądami wirowymi (prądy Foucaulta). Wywołują one silne nagrzewanie przewodników.

7.3. Indukcyjność. Samoindukcja

Zgodnie z prawem Biota-Savarta-Laplace'a (pkt. 6.3), prąd płynący w obwodzie wytwarza pole magnetyczne proporcjonalne do natężenia prądu I . Z tego powodu

$$\Phi_B = LI, \quad (7.5)$$

gdzie współczynnik proporcjonalności L nazywamy indukcyjnością obwodu. Przy zmianie natężenia prądu w obwodzie, będzie również zmieniać się wytworzony przez niego strumień magnetyczny, co z kolei prowadzi do zaindukowania się SEM . Powstanie SEM w przewodzącym obwodzie, na skutek zmiany natężenia prądu w tym obwodzie, nazywamy samoindukcją.

Jednostką indukcyjności jest henr (H). Z równania (7.5) wynika, że 1H jest to indukcyjność takiego obwodu, kiedy przy prądzie 1A strumień magnetyczny samoindukcji wynosi 1Wb, bowiem $1 \text{ H} = 1 \text{ Wb}/1\text{A} = 1 \text{ Vs}/\text{A}$.

Obliczymy indukcyjność nieskończenie długiego solenoidu. Zgodnie z (6.14) mamy

$$B = \mu_o \mu_r n I,$$

gdzie $n = N/l$ (N jest całkowitą liczbą uzwojeń solenoidu). Wobec tego, całkowity strumień płynący przez solenoid jest równy BSN , czyli

$$\Phi_B = \mu_o \mu_r \frac{N^2 I}{l} S.$$

Uwzględniając (7.5)

$$L = \mu_o \mu_r \frac{N^2 S}{l}, \quad (7.6)$$

czyli indukcyjność solenoidu zależy od liczby zwojów solenoidu N , jego długości l , pola przekroju S i przenikalności magnetycznej rdzenia solenoidu μ_r .

W ogólnym przypadku można pokazać, że indukcyjność obwodu zależy tylko od jego kształtu, rozmiarów i przenikalności magnetycznej ośrodka, w którym się znajduje. W tym sensie indukcyjność obwodu jest odpowiednikiem pojemności elektrycznej przewodnika, która także zależy od kształtu przewodnika, jego rozmiarów i przenikalności dielektrycznej ośrodka.

Z prawa Faradaya otrzymujemy, że SEM samoindukcji

$$E_s = -\frac{d\Phi_B}{dt} = -\frac{d}{dt}(LI) = -\left(L\frac{dI}{dt} + I\frac{dL}{dt}\right).$$

Jeżeli obwód nie ulega deformacji i przenikalność magnetyczna nie zmienia się, to $L = const$ i

$$E_s = -L\frac{dI}{dt}. \quad (7.7)$$

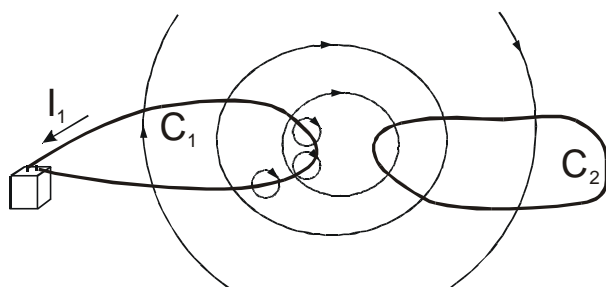
Znak „-” uwarunkowany regułą Lenza wskazuje, że obecność indukcyjności w obwodzie prowadzi do zwalniania zmian prądu, co przejawia się w bezwładności elektrycznej obwodu. W ten sposób indukcyjność obwodu stanowi miarę jego bezwładności wobec zmian prądu.

7.3.1. Indukcyjność wzajemna

Wyobraźmy sobie dwa nieruchome obwody (pętle) C_1 i C_2 umieszczone względem siebie, na przykład jak na rys. 7.6. Niech w obwodzie C_1 płynie prąd o natężeniu I_1 . Strumień indukcji B_1 przez obwód C_2 wynosi

$$\Phi_{21} = \int_{S_2} \vec{B} \cdot d\vec{S} = M_{21} I_1,$$

gdzie S_2 jest powierzchnią obwodu C_2 . Stałą M_{21} nazywamy indukcyjnością wzajemną wyrażoną w henrach. W obwodzie C_2 powstaje indukowana siła elektromotoryczna



$$E_{21} = -M_{21} \frac{dI_1}{dt}. \quad (7.8)$$

Podobnie w celu obliczenia siły elektromotorycznej indukowanej w obwodzie C_1 na skutek zmian natężenia prądu w obwodzie C_2 musimy wprowadzić nowy współczynnik indukcji wzajemnej M_{12}

Rys. 7.6. Prąd o natężeniu I_1 płynący w pętli C_1 wywołuje strumień Φ_{21} w pętli C_2 .

$$E_{12} = -M_{12} \frac{dI_2}{dt}. \quad (7.9)$$

Okazuje się, że dla dowolnych dwóch obwodów

$$M_{12} = M_{21}.$$

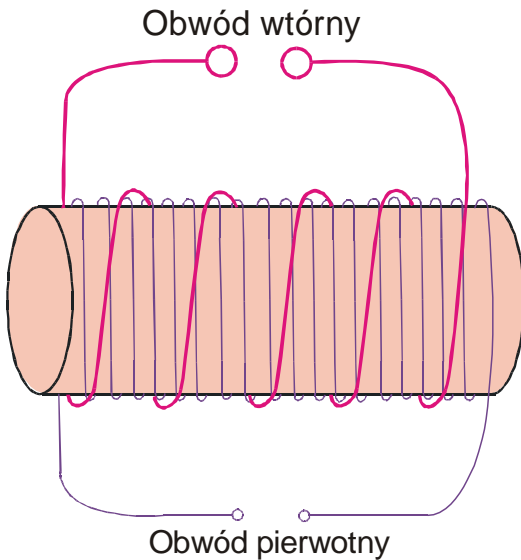
Dzięki temu nie musimy pamiętać o rozróżnianiu M_{12} od M_{21} . Możemy więc mówić o indukcyjności wzajemnej M dowolnych dwóch obwodów i o przenikalności magnetycznej ośrodka otaczającego obwody.

7.4. Transformator

Zjawisko indukcyjności wzajemnej zostało wykorzystane w konstrukcji transformatorów (rys. 7.7). Jeżeli na rdzeń nawinięte są dwie cewki, to zmiana prądu w jednej z nich powoduje indukowanie prądu w drugiej cewce. Wartość indukowanej SEM możemy obliczyć z prawa Faradaya. W większości przypadków uzwojenie wtórne nawijane jest na uzwojenie pierwotne, tak aby obydwa uzwojenia obejmowały jednakowe strumienie pola magnetycznego. Niech n_1

oznacza ilość zwojów uzwojenia pierwotnego, a n_2 – ilość zwojów uzwojenia wtórnego. Wówczas zgodnie z (7.1) indukowane napięcie (SEM) w uzwojeniu wtórnym zapiszemy w postaci

$$V_2 = -n_2 \frac{d\Phi_B}{dt}.$$



Analogicznie SEM w obwodzie pierwotnym

$$V_1 = -n_1 \frac{d\Phi_B}{dt}.$$

Stosunek napięć jest równy

$$\frac{V_2}{V_1} = \frac{n_2}{n_1}.$$

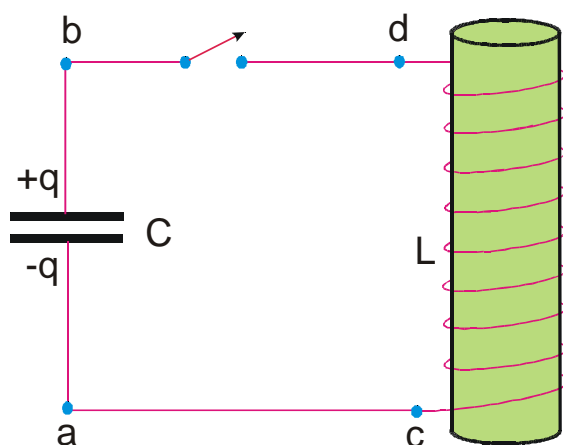
Kiedy do obwodu pierwotnego przykładamy napięcie zmienne V_{zm} , prąd wzrasta do chwili dopóki $n_1 d\Phi_B / dt$ nie osiągnie wartości V_{zm} . Tak więc $V_{zm} = V_1$.

Rys. 7.7. Transformator

Napięcie w obwodzie wtórnym można zmieniać dobierając odpowiedni stosunek liczby zwojów. Jest to wygodnym sposobem transformacji niskich napięć na wysokie i odwrotnie. Widzimy w tym zaletę stosowania prądu zmiennego w porównaniu ze stałym. Ma to ogromne znaczenie praktyczne przy przesyłaniu energii elektrycznej na duże odległości. Najbardziej ekonomiczne generatory wytwarzają stosunkowo niskie napięcie zmienne. Transformator pozwala podwyższyć napięcie przy nieznacznej stracie mocy. Na końcu linii przesyłowej, w celu obniżenia napięcia do bezpiecznego i bardziej dogodnego poziomu, stosuje się drugi transformator.

7.5. Energia pola magnetycznego

Kondensatory stosowane są nie tylko do gromadzenia ładunku elektrycznego, ale w połączeniu z indukcyjnością stosowane są do generacji zmiennego prądu i napięcia. Rozważymy prosty obwód elektryczny, w którym pojemność i indukcyjność są połączone równolegle (rys. 7.8). Jest to tzw. obwód drgający LC (bliższe informacje patrz pkt. 8.7). Załóżmy, że rezystancja obwodu jest zerowa.



Rys. 7.8. Drgający obwód LC

Niech w chwili $t = 0$ ładunek kondensatora wynosi q_0 . Energia początkowa układu zmagazynowana jest w kondensatorze. Zgodnie z równaniem (4.35)

$$W = \frac{1}{2} \frac{q_0^2}{C} = \frac{1}{2} C V_0^2,$$

gdzie $V_0 = q_0 / C$. Zgodnie z prawem zachowania energii, ta początkowa energia nie może zniknąć. Wykażemy, że jest ona gromadzona w polu magnetycznym cewki indukcyjnej.

Ładunek dq płynący przez cewkę przyjmuje energię Vdq , gdzie $V = -L di/dt$. Wobec tego energia tracona przez ładunek i przyjmowana przez cewkę wynosi

$$dW = \left(L \frac{di}{dt} \right) dq = L di \frac{dq}{dt} = L i di.$$

Jeżeli prąd rośnie od zera do I_0 , to energia gromadzona w cewce indukcyjnej wynosi

$$W = \int_0^{I_0} L i di = \frac{1}{2} L I_0^2. \quad (7.12)$$

Interesującym jest przekształcić wzór (7.12) wyrażając prawą stronę przez wielkość pola magnetycznego w cewce indukcyjnej. Jest to łatwo wykonać w przypadku długiego solenoidu dla którego $B = \mu_0 \mu_r N I / l$ i $L = \mu_0 \mu_r N^2 S / l$ (patrz odpowiednio wzory (6.14) i (7.6)). Uzależniając I od B i wstawiając wzór na L , z wyrażenia (7.12) otrzymujemy

$$W = \frac{1}{2 \mu_0 \mu_r} B^2 S l.$$

Dzieląc teraz obie strony tego wyrażenia przez objętość solenoidu $S l$ otrzymujemy wzór na gęstość energii pola magnetycznego

$$w = \frac{1}{2 \mu_0 \mu_r} B^2. \quad (7.13)$$

Pomimo tego, że powyższe obliczenia gęstości energii pola magnetycznego dotyczą solenoidu, można w ogólnym przypadku udowodnić, że dla cewki indukcyjnej dowolnego kształtu całka po $B^2 / 2 \mu_0 \mu_r$ w całej przestrzeni jest równa $L I^2 / 2$, gdzie L jest indukcyjnością cewki.

Analogicznie do wielkości $\epsilon_o \epsilon_r E^2 / 2$ (wzór (4.36)) interpretowanej jako energia zmagazynowana w jednostce objętości pola elektrycznego, możemy powiedzieć, że $B^2 / 2\mu_o \mu_r$ jest energią zmagazynowaną w jednostce objętości pola magnetycznego. W przypadku ogólnym, pola elektryczne i magnetyczne mogą jednocześnie występować w przestrzeni, a wówczas całkowita gęstość energii pola elektromagnetycznego wynosi

$$w = \frac{1}{2} \left(\epsilon_o \epsilon_r E^2 + \frac{B^2}{\mu_o \mu_r} \right). \quad (7.14)$$

7.6. Równania Maxwella

Z poprzednich rozważań wiemy, że zmiana pola magnetycznego powoduje powstawanie pola elektrycznego (prawo Faradaya)

$$\oint_C \vec{E} \cdot d\vec{s} = - \frac{d\Phi_B}{dt}.$$

Z kolei z prawa Ampere'a (wzór (6.3)) wynika, że

$$\oint_C \vec{B} \cdot d\vec{s} = \mu_o \int \vec{j} \cdot d\vec{S},$$

przy czym \vec{j} oznacza gęstość prądu przewodzenia. W następnym punkcie wykazujemy, że w przypadku zmieniającego się pola elektrycznego, do prawej strony ostatniego równania należy dodać człon $(1/c^2) d\Phi_E / dt$, a więc człon analogiczny do $d\Phi_B / dt$ występujący w prawie Faradaya.

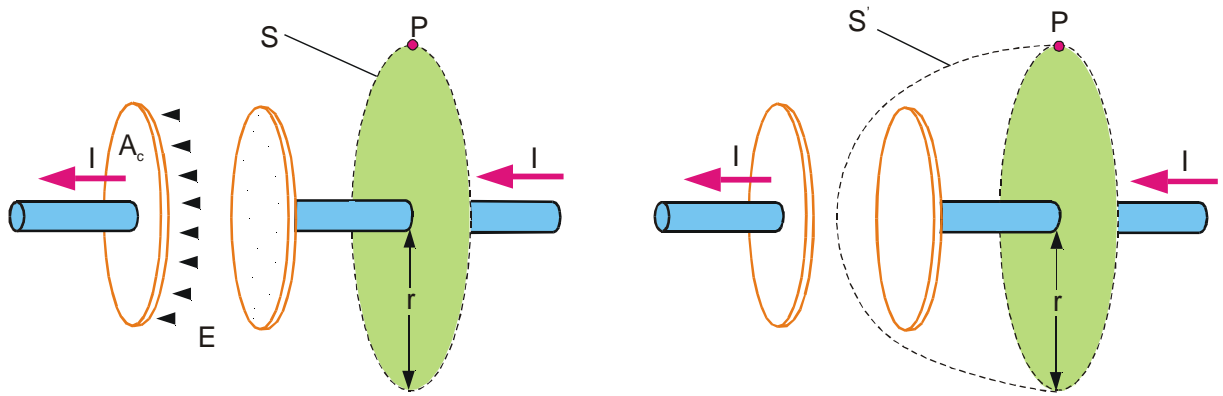
7.6.1. Prąd przesunięcia

Rozważmy przykład zilustrowany na rys. 7.9. Kondensator z płytkami o kształcie kołowym ładowany jest prądem I , który przenosi ładunki z lewej płytki na prawą. Pole magnetyczne w punkcie P możemy obliczyć prowadząc przez ten punkt okrąg o promieniu r i stosując prawo Ampere'a. Na rys. 7.9a przez płaszczyznę ograniczoną tym okręgiem płynie prąd I . Zgodnie z prawem Ampere'a

$$\oint_{\text{po okręgu}} \vec{B} \cdot d\vec{s} = \mu_o \int \vec{j} \cdot d\vec{S} = \mu_o I,$$

czyli $B2\pi r = \mu_o I$, a stąd

$$B = \frac{\mu_o I}{2\pi r}. \quad (7.15)$$



Rys. 7.9. Przez kondensator o płytkach kołowych płynie prąd. (a) Pole elektryczne między okładkami kondensatora, prąd I przecina powierzchnię S ograniczoną przerywaną linią. (b) Wygięta powierzchnia S' napięta na tej linii nie przecinana prądem I .

Jednakże prawo Ampere'a powinno być spełnione dla dowolnej powierzchni rozpiętej na tym okręgu, w szczególności na powierzchni S' na rys. 7.9b. Jednakże w tym przypadku mamy

$$\int_{S'} \vec{j} \cdot d\vec{S} = 0,$$

ponieważ przez powierzchnię S' prąd nie płynie. Wówczas zgodnie z prawem Ampera $\oint \vec{B} \cdot d\vec{s} = 0$, a to przeczy poprzedniemu wynikowi (7.15).

W 1860 roku Maxwell opierając się na analogicznych przykładach doszedł do wniosku, że wyrażenie na prawo Ampere'a przytoczone w rozdziale 6 (pkt. 6.1) jest niesłuszne w przypadku zmiennego pola elektrycznego. Jednocześnie Maxwell odkrył, że niepoprawność zapisu można usunąć dodając do prawej strony równania (6.3) wyrażenie $(1/c^2) (\partial \vec{E} / \partial t) \cdot d\vec{S}$. W poprawionej formie prawo Ampere'a zapisujemy następująco

$$\oint_C \vec{B} \cdot d\vec{s} = \mu_0 \oint_S \vec{j} \cdot d\vec{S} + \frac{1}{c^2} \int_S \frac{\partial \vec{E}}{\partial t} \cdot d\vec{S}. \quad (7.16)$$

Teraz udowodnimy, że równanie to prowadzi do jednoznacznej wartości B w punkcie P niezależnie od postaci powierzchni całkowania S lub S' . Dla części powierzchni S' położonej pomiędzy płytkami kondensatora pole elektryczne $E = Q / \epsilon_0 A_c$ (patrz wzór 4.18). Wobec tego różniczkując to wyrażenie względem t mamy

$$\frac{\partial E}{\partial t} = \frac{1}{\epsilon_0 A_c} \frac{\partial Q}{\partial t} = \frac{1}{\epsilon_0 A_c} I.$$

Całkowanie po powierzchni S' daje

$$\int_{S'} \frac{\partial \vec{E}}{\partial t} \cdot d\vec{S} = \frac{I}{\epsilon_0},$$

co dalej prowadzi do związku

$$B2\pi r = \frac{1}{c^2} \frac{I}{\epsilon_0}.$$

Ponieważ $1/c^2 = \mu_0 \epsilon_0$, więc

$$B = \frac{\mu_0 I}{2\pi r}.$$

Otrzymaliśmy więc wynik identyczny jak przy całkowaniu po powierzchni S.

Pierwszy człon po prawej stronie wzoru (7.16) przedstawia realny prąd płynący przez powierzchnię rozpiętą na zamkniętym konturze. Drugi człon można interpretować jako prąd związany ze zmianą natężenia pola elektrycznego. Maxwell nazwał go prądem przesunięcia. Prąd ten jest przedłużeniem prądu przewodzenia wpływającego do kondensatora i jest mu równy. Prąd przesunięcia zapewnia więc ciągłość obwodów zawierających kondensatory.

Odcinki bezprzewodowe obwodów elektrycznych mogą być wypełnione dielektrykiem, wtedy w miejsce pola elektrycznego \vec{E} wprowadzamy wektor indukcji elektrycznej \vec{D} i równanie (7.16) przyjmuje postać.

$$\oint_C \vec{H} \cdot d\vec{s} = \int_S \left(\vec{j} + \frac{\partial \vec{D}}{\partial t} \right) \cdot d\vec{S}, \quad (7.17)$$

a więc gęstość prądu przesunięcia ma ogólną postać

$$\vec{j}_p = \frac{\partial \vec{D}}{\partial t}. \quad (7.18)$$

Ponieważ zgodnie ze wzorem (4.51) $\vec{D} = \epsilon_0 \vec{E} + \vec{P}_e$, więc

$$\vec{j}_p = \epsilon_0 \frac{\partial \vec{E}}{\partial t} + \frac{\partial \vec{P}_e}{\partial t}. \quad (7.19)$$

Składnik $\partial \vec{P}_e / \partial t$ wyraża część gęstości prądu w dielektryku (przesunięcie ładunków lub obrót dipoli) i nosi nazwę gęstości prądu polaryzacyjnego. Zatem \vec{j}_p stanowi sumę gęstości prądu przesunięcia w próżni $\epsilon_0 \partial \vec{E} / \partial t$ i prądu polaryzacyjnego.

7.6.2. Równania Maxwella w postaci całkowej

Dotychczas zapoznaliśmy się z poszczególnymi fragmentami równań Maxwella. Po wprowadzeniu prądu przesunięcia możemy je przedstawić w najbardziej ogólnej formie zwanej równaniami Maxwella.

Tabela 7.1. Równania Maxwella

1. Uogólnione prawo Faradaya (7.3)	$\oint_C \vec{E} \cdot d\vec{s} = - \int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}$	zmienne pole magnetyczne wytwarza wirowe pole elektryczne, które może wywoływać prąd elektryczny
2. Uogólnione prawo Ampere'a (7.17)	$\oint_C \vec{H} \cdot d\vec{s} = \int_S \left(\vec{j} + \frac{\partial \vec{D}}{\partial t} \right) \cdot d\vec{S}$	prąd elektryczny lub zmienne pole elektryczne wytwarzają wirowe pole magnetyczne
3. Prawo Gaussa dla pola elektrycznego (4.45)	$\oint_S \vec{D} \cdot d\vec{S} = \int \rho dV$	ładunek elektryczny wytwarza pole elektryczne
4. Prawo Gaussa dla pola magnetycznego (6.5)	$\oint_S \vec{B} \cdot d\vec{S} = 0$	nie istnieje w przyrodzie ładunek magnetyczny, pole magnetyczne jest bezźródłowe

Dla uzyskania pełnego układu równań Maxwella należy dołączyć jeszcze podstawowe związki (4.43) i (6.28) między wektorami elektrycznymi i magnetycznymi

$$\vec{D} = \epsilon_o \epsilon_r \vec{E},$$

$$\vec{B} = \mu_o \mu_r \vec{H}.$$

Równania Maxwella stanowią fundamentalną podstawę teorii zjawisk elektromagnetycznych, podobnie jak zasady dynamiki Newtona są podstawą mechaniki. Przy pomocy tych równań można znaleźć pola \vec{E} i \vec{B} w dowolnym punkcie przestrzeni i w dowolnej chwili czasu, jeżeli znane są współrzędne i prędkości ładunków wytwarzających pola. Równania Maxwella są niesymetryczne względem pól elektrycznego i magnetycznego. Związane jest to z istnieniem ładunków elektrycznych i brakiem ładunków magnetycznych.

Dla pól stacjonarnych (niezależnych od czasu) równania Maxwella przyjmują postać

$$\begin{array}{ll} \oint_C \vec{E} \cdot d\vec{s} = 0 & \oint_S \vec{D} \cdot d\vec{S} = \int \rho dV \\ \int_C \vec{H} \cdot d\vec{s} = \int_S \vec{j} \cdot d\vec{S} & \oint_S \vec{B} \cdot d\vec{S} = 0 \end{array}$$

W danym przypadku pola elektryczne i magnetyczne są niezależne od siebie, co pozwala badać niezależnie stałe pole elektryczne i magnetyczne.

Teoria Maxwella jest teorią makroskopową. Nie jest w stanie wyjaśnić tych zjawisk, w których przejawia się wewnętrzna budowa ciała.

7.6.3. Równania Maxwella w postaci różniczkowej

W pkt. 1.2.5 przedstawiono dwa twierdzenia analizy wektorowej: twierdzenie Stokesa (wzór (1.49)) i twierdzenie Gaussa-Ostrogradzkiego [wzór (1.50)]

$$\oint_C \vec{a} \cdot d\vec{s} = \int_S \text{rot} \vec{a} \cdot d\vec{S},$$

$$\oint_S \vec{a} \cdot d\vec{S} = \int_V \text{div} \vec{a} \cdot d\vec{V}.$$

Stosując te twierdzenia i uwzględniając związki podane w tabeli 7.1 otrzymujemy pełny układ równań Maxwella w postaci różniczkowej:

$$\text{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}, \quad (7.20)$$

$$\text{rot} \vec{H} = \vec{j} + \frac{\partial \vec{D}}{\partial t}, \quad (7.21)$$

$$\text{div} \vec{D} = \rho, \quad (7.22)$$

$$\text{div} \vec{B} = 0. \quad (7.23)$$

Jeżeli ładunek i prądy w danym ośrodku rozmieszczone są w sposób ciągły, to obydwie formy równań Maxwella (całkowa i różniczkowa) są ekwiwalentne. Jeżeli jednak istnieją powierzchnie, na których zachodzi skokowa zmiana tych wielkości, to całkowa forma równań jest bardziej ogólna.

ROZDZIAŁ 8

DRGANIA I FALE

Ruch w przyrodzie jest zjawiskiem powszechnym. Formami ruchu są drgania i fale. Ruchem drgającym, lub wprost *drGANIAMI*, nazywamy każdy ruch lub zmianę stanu, które charakteryzuje powtarzalność w czasie wartości wielkości fizycznych, określających ten ruch lub stan. Z drGANIAMI spotykamy się przy badaniu różnych zjawisk fizycznych: dźwięku, światła, prądów zmiennych, fal radiowych, wahań wahadeł, itp. Okazuje się, że zarówno prawa rządzące tymi zjawiskami jak i metody ich badania są ogólne. Łatwo zauważymy analogię między drGANIAMI mechanicznymi i elektrycznymi. Stąd wynika konieczność jednakowego podejścia do drGAN o odmiennej naturze fizycznej.

Proces rozchodzenia się zaburzeń (drGAN) w ośrodku jest przykładem ruchu falowego. Falami nazywamy również różnego rodzaju zaburzenia stanu materii lub pola rozchodzące się w przestrzeni. Większość informacji jakie mamy o świecie dociera do naszej świadomości poprzez organa zmysłowe słuchu i wzroku za pośrednictwem fal. Informacje te przychodzą do nas z pewnym opóźnieniem, gdyż do przebycia drogi między źródłem fal a obserwatorem jest potrzebny pewien okres czasu. Wspólną cechą wszystkich zjawisk falowych jest zdolność przenoszenia energii, przy czym w procesie tym występuje w sposób ciągły okresowa zmiana jednego rodzaju energii w drugi, np. kinetycznej w potencjalną, energii pola elektrycznego w energię pola magnetycznego i na odwrót.

8.1. Drganie harmoniczne

Jeżeli układ, na który nie działają zmienne siły zewnętrzne, zostaje wprowadzony w drGANIA na skutek jakiegokolwiek początkowego odchylenia od położenia równowagi, to takie drGANIA nazywamy swobodnymi. W drgającym układzie zachowawczym nie występuje rozproszenie energii; drGANIA swobodne w tym przypadku nazywamy drGANIAMI niegasnącymi.

Ruch drgający nazywamy okresowym (periodycznym), jeżeli wartości wielkości fizycznych zmieniające się podczas drGAN, powtarzają się w pewnych odstępach czasu. Najprostszym rodzajem drGAN okresowych są tzw. drGANIA harmoniczne, dla których zależność drGAN wielkości fizycznej od czasu t opisujemy równaniem typu

$$s = A \cos(\omega_0 t + \varphi), \quad (8.1)$$

gdzie A - maksymalna wartość drgającej wielkości nazywana amplitudą, ω_0 - częstość kątowna, φ - faza początkowa drGAN w chwili czasu $t = 0$, $(\omega_0 t + \varphi)$ - faza drGAN w chwili czasu t .

Określone stany układu wykonującego drgania harmoniczne powtarzają się w odstępie czasu T nazywanego okresem drgań, w którym faza drgań wzrasta o 2π , t.j.

$$\omega_o(t+T) + \varphi = (\omega_o t + \varphi) + 2\pi,$$

stąd

$$T = \frac{2\pi}{\omega_o}. \quad (8.2)$$

Wielkość odwrotna do okresu drgań

$$\nu = \frac{1}{T}, \quad (8.3)$$

równa liczbie okresów drgań (liczbie drgań) wykonywanych w jednostce czasu, nazywamy częstotliwością drgań. Porównując (8.2) i (8.3) otrzymujemy

$$\omega_o = 2\pi\nu. \quad (8.4)$$

Jednostką częstotliwości jest hertz (1Hz) – jest to częstotliwość periodycznych drgań w których jeden cykl wykonywany jest w 1s.

Obliczmy pierwszą i drugą pochodną po czasie wielkości s wykonującej drgania harmoniczne (obliczymy więc prędkość i przyspieszenie w przypadku drgań harmonicznyc)

$$\frac{ds}{dt} = -A\omega_o \sin(\omega_o t + \varphi) = A\omega_o \cos\left(\omega_o t + \varphi + \frac{\pi}{2}\right), \quad (8.5)$$

$$\frac{d^2s}{dt^2} = -A\omega_o^2 \cos(\omega_o t + \varphi) = A\omega_o^2 \cos(\omega_o t + \varphi + \pi). \quad (8.6)$$

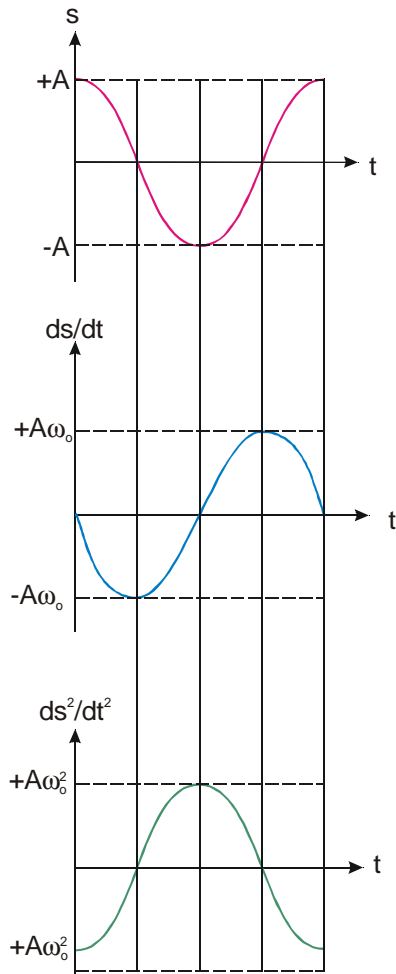
Amplitudy wielkości (8.5) i (8.6) równe są odpowiednio $A\omega_o$ i $A\omega_o^2$. Faza prędkości (8.5) różni się od fazy wielkości (8.1) o $\pi/2$, a faza przyspieszenia o π . Tak więc w chwili czasu gdy $s = 0$, ds/dt przyjmuje największą wartość; podczas gdy s przyjmuje maksymalną ujemną wartość, to d^2s/dt^2 przyjmuje największą dodatnią wartość (rys. 8.1). Ze związku (8.6) wynika równanie różniczkowe drgań harmonicznyc

$$\frac{d^2s}{dt^2} + \omega_o^2 s = 0. \quad (8.7)$$

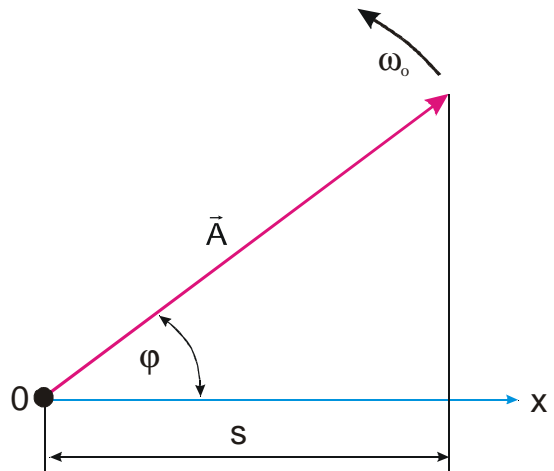
Rozwiązaniem tego równania jest wyrażenie (8.1).

Drgania harmoniczne można przedstawić graficznie za pomocą obracającego się wektora amplitudy. Jest to metoda wektorowa. W tym celu z dowolnego punktu 0 osi x pod kątem φ równym fazie początkowej drgań wykreślamy wektor \vec{A} , którego moduł jest równy amplitudzie A rozważanego drgania (rys. 8.2). Jeżeli wektor wprawimy w obrót z prędkością kątową ω_o , to

rzut końca wektora będzie się przemieszczać wzdłuż osi x i przyjmować wartości od $+A$ do $-A$, a drgająca wielkość będzie zmieniać się w czasie według wzoru $s = A \cos(\omega_0 t + \varphi)$. A zatem obracający się wektor amplitudy w zupełności charakteryzuje drganie harmoniczne.



Rys. 8.1. Zależność s , ds/dt i ds^2/dt^2 od czasu dla drgań harmonicznycch o okresie $T = 2\pi/\omega_0$.



Rys. 8.2. Rzut obracającego się wektora \vec{A} na oś x charakteryzuje drganie harmoniczne.

Przedstawienie drgań harmonicznycch za pomocą obracających się wektorów ma szerokie zastosowanie przy analizie złożonych drgań.

W fizyce często stosowana jest druga metoda opisu drgań harmonicznycch. W metodzie tej drgającą wielkość opisuje się liczbą zespoloną (patrz pkt. 1.2.3)

$$\bar{s} = A e^{i(\omega_0 t + \varphi)}. \quad (8.8)$$

Część rzeczywista tego wyrażenia

$$Re \bar{s} = A \cos(\omega_0 t + \varphi) = s,$$

przedstawia drganie harmoniczne. W teorii drgań przyjmuje się, że drgająca wielkość s równa jest części rzeczywistej wyrażenia zespolonego (8.8), czyli

$$s = Re A e^{i(\omega_0 t + \varphi)}. \quad (8.9)$$

8.1.1. Mechaniczne drgania harmoniczne

Oscylatorem harmonicznym nazywamy układ opisywany równaniem typu (8.7). Drgania oscylatora harmonicznego stanowią ważny przykład ruchu periodycznego i służą jako ścisły lub przybliżony model w wielu zagadnieniach fizyki klasycznej i kwantowej. Przykładem oscylatora

harmonicznego są wahadła sprężynowe, fizyczne, matematyczne, czy też drgające obwody elektryczne.

Wahadłem sprężynowym nazywamy ciężar o masie m zawieszony na nieważkiej sprężynie, wykonujący drgania harmoniczne pod wpływem siły sprężystości $F = -kx$, gdzie k nazywamy współczynnikiem sprężystości. Równanie ruchu dla wahadła

$$m \frac{d^2 x}{dt^2} = -kx,$$

czyli

$$\frac{d^2 x}{dt^2} + \frac{k}{m} x = 0.$$

Porównując to wyrażenie do równania (8.7) wynika, że wahadło sprężynowe wykonuje drgania harmoniczne $x = A \cos(\omega_o t + \varphi)$ z częstością kołową

$$\omega_o = \sqrt{\frac{k}{m}} \quad (8.10)$$

i okresem

$$T = 2\pi \sqrt{\frac{m}{k}}. \quad (8.11)$$

Wzór (8.11) jest słuszny dla drgań sprężystych w których spełnione jest prawo Hooke'a, kiedy masa sprężyny jest mała w porównaniu z masą ciała.

Siła $F = m(d^2 x / dt^2)$ działająca na masę m wynosi

$$F = -mA\omega_o^2 \cos(\omega_o t + \varphi) = -m\omega_o^2 x$$

i faktycznie jest proporcjonalna do przesunięcia x ciężarka z położenia równowagi i skierowana w przeciwną stronę.

Energia kinetyczna ciężarka wykonującego drgania harmoniczne

$$K = \frac{mv^2}{2} = \frac{mA^2\omega_o^2}{2} \sin^2(\omega_o t + \varphi) = \frac{mA^2\omega_o^2}{4} [1 - \cos 2(\omega_o t + \varphi)]. \quad (8.12)$$

Energia potencjalna ciężarka wykonującego drgania harmoniczne pod wpływem siły F jest równa

$$U = -\int_0^x F dx = \frac{m\omega_o^2 x^2}{2} = \frac{mA^2\omega_o^2}{2} \cos^2(\omega_o t + \varphi) = \frac{kx^2}{2},$$

czyli

$$U = \frac{mA^2\omega_o^2}{4} [1 + \cos 2(\omega_o t + \phi)]. \quad (8.13)$$

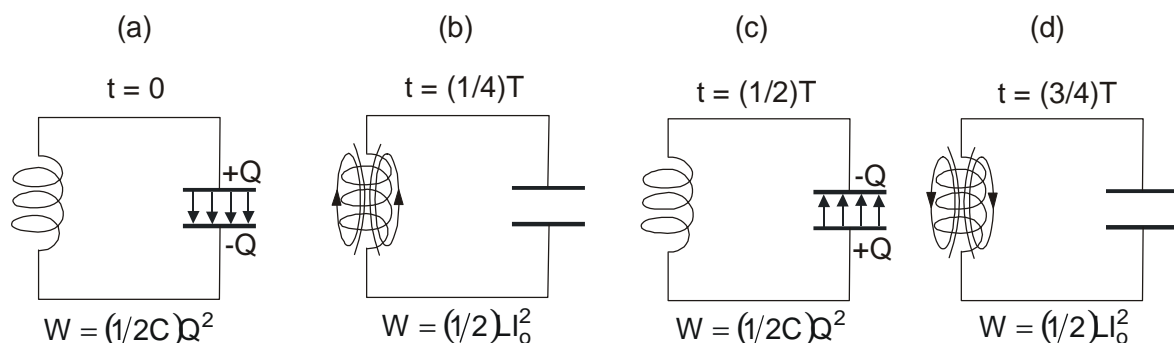
Dodając stronami równania (8.12) i (8.13), otrzymujemy energię całkowitą

$$W = K + U = \frac{mA^2\omega_o^2}{2}. \quad (8.14)$$

Z wyrażen (8.12) i (8.13) wynika, że U i K zmieniają się z częstością $2\omega_o$, tzn. z częstością dwa razy wyższą niż częstość drgań harmoniczných. Ponieważ wartości średnie $\langle \sin^2 \alpha \rangle$ i $\langle \cos^2 \alpha \rangle$ są równe $1/2$, więc z wyrażen (8.12), (8.13) i (8.14) wynika, że $\langle K \rangle = \langle U \rangle = W/2$.

8.1.2. Elektryczne drgania harmoniczne

Wśród różnych zjawisk elektrycznych, szczególne miejsce zajmują drgania elektromagnetyczne, w których takie wielkości elektryczne jak ładunek, prąd, pole elektryczne i magnetyczne zmieniają się periodycznie.



Rys. 8.3. Przebieg drgań elektrycznych w obwodzie LC.

Rozważymy bardziej szczegółowo obwód drgający LC przedstawiony na rys. 7.8. W chwili początkowej ładunek kondensatora ma wartość maksymalną (rys. 8.3a). Różnica potencjałów między jego okładkami jest wówczas największa, natomiast prąd równa się zero. Następnie zaczyna się rozładowanie kondensatora. Dzięki zjawisku samoindukcji prąd w obwodzie stopniowo wzrasta, osiągając wartość maksymalną w chwili $t = T/4$, gdy $q = 0$ (rys. 8.3b). Z kolei dalej obserwuje się stopniowe zmniejszanie wartości prądu (przy zachowaniu kierunku), aż do zera osiąganego w chwili $t = T/2$ (rys. 8.3c). Ładunek kondensatora i różnica potencjałów między jego okładkami osiąga wtedy ponownie wartość maksymalną. Znaki ładunków okładek i kierunek pola elektrycznego między nimi są jednak teraz przeciwne względem tych, jakie występowały w chwili początkowej. W ten sposób w wyniku zjawiska samoindukcji następuje zmiana kierunku ładowania kondensatora. Dalej procesy przebiegają w odwrotnym kierunku (rys. 8.3d). W obwodzie występuje więc okresowe przenoszenie energii pola elektrycznego

kondensatora w energię pola magnetycznego prądu. W chwilach $t = 0, T/2, T,$ itd. energia pola elektrycznego osiąga wartości maksymalne CV_o^2 , jednocześnie energia pola magnetycznego równa się zero. Odwrotnie, w chwilach $t = T/4, 3T/4,$ itd. energia pola magnetycznego jest maksymalna i równa $LI_o^2 / 2$, natomiast energia pola elektrycznego równa się zero.

Rozważmy teraz obwód elektryczny zawierający cewkę o indukcyjności L , kondensator o pojemności C i rezystor o rezystancji R . Zgodnie z prawem Kirchoffa

$$IR + V_c = E_s,$$

gdzie IR oznacza napięcie na rezystorze, $V_c = Q/C$ napięcie na kondensatorze, a $E_s = -LdI/dt$ SEM indukcji powstającą w cewce gdy płynie prąd zmienny. Tak więc

$$L \frac{dI}{dt} + IR + \frac{Q}{C} = 0.$$

Uwzględniając, że $I = dQ/dt$ oraz $dI/dt = d^2Q/dt^2$, otrzymamy równanie różniczkowe okresowych zmian ładunku Q w obwodzie

$$\frac{d^2Q}{dt^2} + \frac{R}{L} \frac{dQ}{dt} + \frac{1}{LC} Q = 0. \quad (8.15)$$

W przypadku tego obwodu nie istnieją zewnętrzne SEM, w związku z tym rozważane drgania stanowią drgania swobodne. Jeżeli $R = 0$, to drgania są harmonicznymi.

Z równania (8.15) otrzymamy równanie różniczkowe swobodnych drgań harmonicznymi

$$\frac{d^2Q}{dt^2} + \frac{1}{LC} Q = 0.$$

Podobnie jak w przypadku drgań mechanicznych, ładunek wykonuje drgania harmoniczne według prawa

$$Q = Q_o \cos(\omega_o t + \varphi), \quad (8.16)$$

gdzie Q_o jest amplitudą drgań ładunku kondensatora z częstością ω_o nazywaną częstością drgań własnych obwodu

$$\omega_o = \frac{1}{\sqrt{LC}}, \quad (8.17)$$

a

$$T = 2\pi\sqrt{LC} \quad (8.18)$$

jest okresem drgań.

Natężenie prądu w obwodzie drgającym

$$I = dQ / dt = -\omega_o Q_o \sin(\omega_o t + \varphi) = I_o \cos\left(\omega_o t + \varphi + \frac{\pi}{2}\right), \quad (8.19)$$

gdzie $I_o = \omega_o Q_o$ jest amplitudą natężenia prądu.

Napięcie na kondensatorze

$$V_c = \frac{Q}{C} = \frac{Q_o}{C} \cos(\omega_o t + \varphi) = V_o \cos(\omega_o t + \varphi), \quad (8.20)$$

gdzie $V_o = Q_o / C$ jest amplitudą napięcia.

Z wyrażeń (8.16) i (8.19) wynika, że drgania prądu I wyprzedzają w fazie drgania ładunku Q o $\pi/2$, tj. kiedy prąd osiąga maksymalną wartość, ładunek (a także napięcie – zob. (8.20)) przyjmuje zerową wartość i na odwrót.

8.2. Składanie drgań harmoniczných równoległych o jednakowej częstotliwości.

Dudnienie

Ciało drgające może brać udział w kilku procesach drgających, a wówczas należy określić wypadkowy ruch drgający ciała. Dokonamy złożenia drgań harmoniczných o jednakowych częstotliwościach

$$x_1 = A_1 \cos(\omega_o t + \varphi_1),$$

$$x_2 = A_2 \cos(\omega_o t + \varphi_2),$$

posługując się wykresem wektorowym (rys. 8.4). Wektory \vec{A}_1 i \vec{A}_2 obracają się z

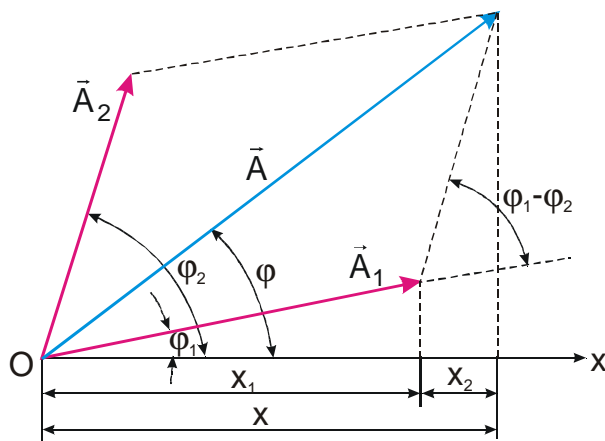
jednakową częstotliwością kątową ω_o , a różnica faz $\varphi_2 - \varphi_1$ pomiędzy nimi pozostaje stała. Wówczas równanie drgania wypadkowego ma postać

$$x = x_1 + x_2 = A \cos(\omega_o t + \varphi), \quad (8.21)$$

gdzie amplituda A i faza φ są określone wyrażeniami

$$A^2 = A_1^2 + A_2^2 + 2A_1 A_2 \cos(\varphi_2 - \varphi_1), \quad (8.22)$$

$$\operatorname{tg} \varphi = \frac{A_1 \sin \varphi_1 + A_2 \sin \varphi_2}{A_1 \cos \varphi_1 + A_2 \cos \varphi_2}. \quad (8.23)$$



Rys. 8.4. Wektorowa metoda składania drgań.

Wobec tego ciało biorące udział w dwóch drganiach harmoniczných o jednakowych kierunkach wykonuje także drgania harmoniczne w tym kierunku i o tej samej częstotliwości co drgania składowe. Amplituda drgania wypadkowego zależy od różnicy faz ($\varphi_2 - \varphi_1$) drgań składowych:

- gdy $\varphi_2 - \varphi_1 = \pm 2\pi m$ ($m = 0, 1, 2, \dots$), wówczas $A = A_1 + A_2$,
- gdy $\varphi_2 - \varphi_1 = \pm \pi(2m + 1)$ ($m = 0, 1, 2, \dots$), wówczas $A = |A_1 + A_2|$.

Interesujący jest przypadek, gdy dwa dodawane drgania równoległe nieznacznie różnią się częstotliwościami drgań. W wyniku dodania tych drgań otrzymujemy drgania o okresowej zmianie amplitudy zwane dudnieniem.

Niech amplitudy składanych drgań będą równe A , a ich częstotliwości kołowe ω i $\omega + \Delta\omega$ przy czym $\Delta\omega \ll \omega$. Przyjmijmy, że fazy początkowe drgań są zerowe, wówczas

$$x_1 = A \cos \omega t,$$

$$x_2 = A \cos(\omega + \Delta\omega)t.$$

Dodając te wyrażenia i uwzględniając że $\Delta\omega / 2 \ll \omega$, znajdujemy

$$x = \left(2A \cos \frac{\Delta\omega}{2} t \right) \cos \omega t. \quad (8.24)$$

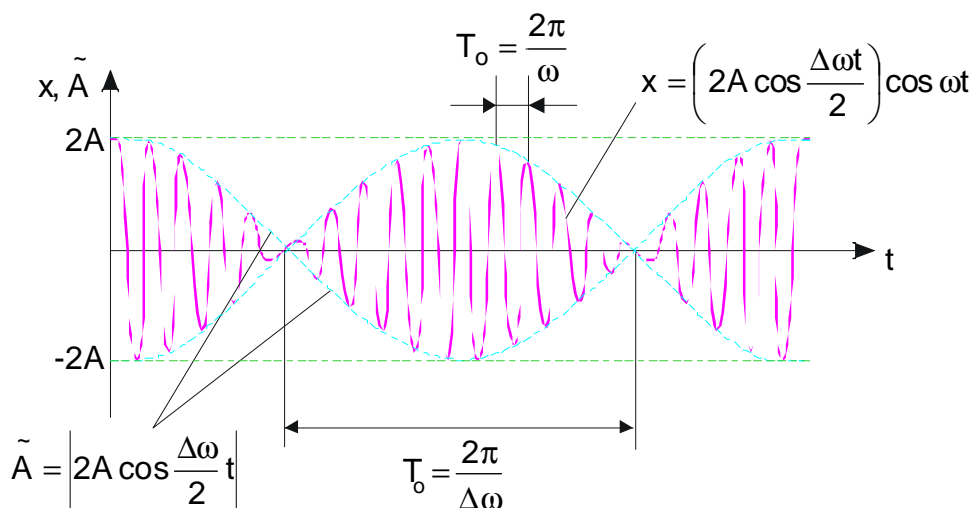
Otrzymane wyrażenie jest iloczynem czynnika stanowiącego modulowaną amplitudę

$$\tilde{A} = \left| 2A \cos \frac{\Delta\omega}{2} t \right| \quad (8.25)$$

o okresie dudnień

$$T_o = \frac{2\pi}{\Delta\omega}$$

i szybko zmieniającego się członu $\cos \omega t$. Charakter zależności x od t pokazany jest na rys. 8.5, gdzie ciągłe linie przedstawiają wykres drgania wypadkowego (8.24), a obwiednie – wolno zmieniające się w czasie wyrażenie (8.25) określające amplitudę.



Rys. 8.5. Nałożenie się drgań harmoniczných o zbliżonych częstotliwościach kołowych daje w wyniku dudnienie.

8.3. Składanie drgań wzajemnie prostopadłych

Rozważmy przypadek złożenia dwóch drgań harmonicznym o jednakowej częstotliwości ω , zachodzących w kierunkach wzajemnie prostopadłych wzdłuż osi x i y . Dla prostoty przyjmijmy, że faza początkowa pierwszego drgania jest zerowa:

$$\begin{aligned}x &= A \cos \omega t \\y &= B \cos(\omega t + \varphi)\end{aligned}\tag{8.26}$$

Równanie trajektorii drgania wypadkowego znajdujemy poprzez wyłączenie z wyrażen (8.26) parametru t . Zapisując drganie składowe w postaci

$$\frac{x}{A} = \cos \omega t ; \quad \frac{y}{B} = \cos \omega t \cos \varphi - \sin \omega t \sin \varphi$$

i zmieniając w drugim wyrażeniu $\cos \omega t$ na x/A i $\sin \omega t$ na $\sqrt{1 - (x/A)^2}$, otrzymujemy po prostych przekształceniach równanie elipsy

$$\frac{x^2}{A^2} - \frac{2xy}{AB} + \frac{y^2}{B^2} = \sin^2 \varphi.\tag{8.27}$$

Drgający punkt porusza się po elipsie, więc otrzymaliśmy przypadek tak zwanych drgań eliptycznie spolaryzowanych.

Orientacja osi elipsy i jej rozmiary zależą od amplitud drgań składowych i różnicy faz φ . Rozpatrzmy niektóre szczególnie przypadki:

- $\varphi = m\pi$ ($m = 0, \pm 1, \pm 2, \dots$), w tym przypadku elipsa degeneruje się do odcinka prostej

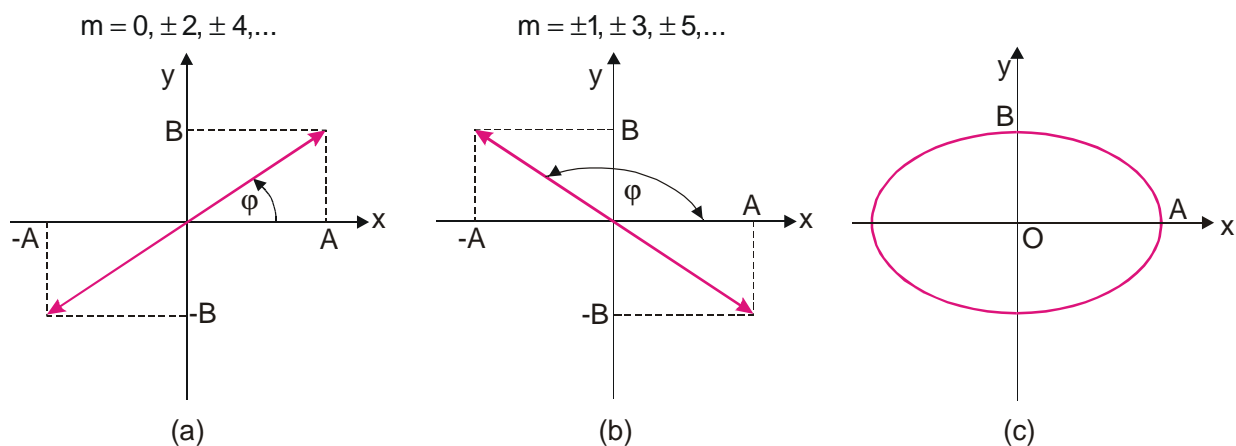
$$y = \pm \frac{B}{A} x,\tag{8.28}$$

gdzie znak $+$ odpowiada zeru i parzystym wartościom m (rys. 8.6a), a znak $-$ nieparzystym wartościom m (rys. 8.6b). Wypadkowe drganie stanowi drganie harmoniczne o częstotliwości ω i amplitudzie $\sqrt{1 - (x/A)^2}$ zachodzące wzdłuż prostej nachylonej pod kątem $\varphi = \arctg[(B/A)\cos m\pi]$. W tym przypadku mamy drgania liniowo spolaryzowane.

- $\varphi = (2m + 1)\pi/2$ ($m = 0, \pm 1, \pm 2, \dots$), w tym przypadku otrzymujemy

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1.\tag{8.29}$$

Jest to równanie elipsy, której osie pokrywają się z osiami współrzędnych, a jej półosie są równe odpowiednim amplitudom (rys. 8.6c).



Rys. 8.6. Superpozycja drgań harmoniczných wzajemnie prostopadłych o różných amplitudach i jednakowych częstościach.

Złożenie drgań harmoniczných o różných częstościach daje w wyniku skomplikowane krzywe, zwane krzywymi Lissajous. Kształt tych krzywych zależy od stosunku amplitud, częstości i początkowych faz drgań.

8.4. Drgania swobodne tłumione

Wszystkie rzeczywiste układy drgające są układami rozpraszającymi energię. Rozpraszanie energii zachodzi poprzez wykonanie pracy przeciw siłom tarcia, czy też poprzez zamianę w energię cieplną. W wielu przypadkach zanikania drgań mechaniczných (w pierwszym przybliżeniu) siły tarcia są proporcjonalne do prędkości

$$F_t = -rv,$$

gdzie r jest współczynnikiem oporu. Wówczas równanie ruchu dla tłumioných drgań ciała (np. masy zawieszonęj na sprężynie) ma postać

$$m \frac{d^2 s}{dt^2} = -ks - r \frac{ds}{dt},$$

gdzie m jest masą ciała drgającego, ks siłą zwrotną sprężyny a $r(ds/dt)$ siłą tarcia. Otrzymujemy więc

$$\frac{d^2 s}{dt^2} + \frac{r}{m} \frac{ds}{dt} + \frac{k}{m} s = 0. \quad (8.30)$$

Zwykle równanie różniczkowe tłumioných drgań swobodných zapisujemy w postaci

$$\frac{d^2 s}{dt^2} + 2\delta \frac{ds}{dt} + \omega_o^2 s = 0, \quad (8.31)$$

gdzie s jest drgającą wielkością opisującą dany proces fizyczny, $\delta = const$ jest współczynnikiem pochłaniania, a ω_0 częstością nietłumionych drgań swobodnych tego układu, gdy $\delta = 0$.

Rozwiązaniem równania (8.31) jest funkcja postaci

$$s = e^{-\delta t} u. \quad (8.32)$$

Po znalezieniu pierwszej i drugiej pochodnej wyrażenia (8.32) i podstawieniu do (8.31) otrzymujemy

$$\frac{d^2 u}{dt^2} + (\omega_0^2 - \delta^2) u = 0. \quad (8.33)$$

Rozwiązanie równania (8.33) zależy od znaku współczynnika $(\omega_0^2 - \delta^2)$. Kiedy ten współczynnik jest dodatni

$$\omega^2 = (\omega_0^2 - \delta^2), \quad (8.34)$$

to rozwiązaniem równania (8.33) jest

$$u = A_0 \cos(\omega t + \varphi).$$

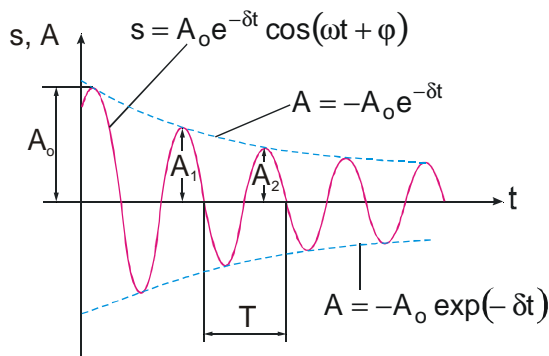
Wobec tego rozwiązaniem równania (8.33) w przypadku małego tłumienia ($\delta^2 < \omega_0^2$) jest

$$s = A_0 e^{-\delta t} \cos(\omega t + \varphi), \quad (8.35)$$

gdzie

$$A = A_0 e^{-\delta t} \quad (8.36)$$

jest amplitudą drgań tłumionych, a A_0 – amplitudą początkową. Zależność (8.35) pokazana jest linią ciągłą na rys. 8.7, a zależność (8.36) – liniami przerywanymi. Okres czasu $\tau = 1/\delta$ w ciągu którego amplituda drgań tłumionych zmniejsza się e razy, nazywamy czasem relaksacji.



Rys. 8.7. Drgania swobodne tłumione.

Jeżeli $A(t)$ i $A(t+T)$ są amplitudami dwóch kolejnych drgań odpowiadających chwilom czasu różniących się o umowny okres drgań, to stosunek

Drgania tłumione nie są harmonicznymi, ponieważ drgania nie powtarzają się; np. największe wartości wychylenia, prędkości i przyspieszenia. Dlatego wielkość ω możemy tylko umownie nazwać częstością kołową drgań tłumionych

$$T = \frac{2\pi}{\omega} = \frac{2\pi}{\sqrt{\omega_0^2 - \delta^2}}.$$

$$\frac{A(t)}{A(t+T)} = e^{\delta T}$$

nazywamy dekrementem tłumienia, a jego logarytm

$$\Theta = \ln \frac{A(t)}{A(t+T)} = \delta T = \frac{T}{\tau} \quad (8.37)$$

logarymicznym dekrementem tłumienia.

W celu scharakteryzowania drgającego układu wprowadzono pojęcie dobroci Q , która dla małych wartości logarymicznego dekrementu tłumienia jest równa

$$Q = \frac{\pi}{\Theta} = \frac{\pi}{\delta T_o} = \frac{\omega_o}{2\delta}. \quad (8.38)$$

Dla wahadła sprężynowego, porównując równanie (8.30) z równaniem (8.31) mamy

$$\delta = \frac{r}{2m}, \quad (8.39)$$

oraz

$$s = A_o e^{-\delta t} \cos(\omega t + \varphi),$$

gdzie

$$\omega = \sqrt{\omega_o^2 - \frac{r^2}{4m^2}}.$$

Dobroć wahadła sprężynowego zgodnie z (8.38) wnosi $Q = \sqrt{km} / r$.

Analogiczne wyrażenie możemy otrzymać dla swobodnych tłumionych drgań ładunku w obwodzie zawierającym R , L i C . Równanie to ma postać (8.15), w związku z tym współczynnik tłumienia

$$\delta = \frac{R}{2L}. \quad (8.40)$$

Równanie różniczkowe (8.15) możemy zapisać w sposób analogiczny do (8.31)

$$\frac{d^2 Q}{dt^2} + 2\delta \frac{dQ}{dt} + \omega_o^2 Q = 0.$$

Drgania ładunku zachodzą według prawa

$$Q = Q_o e^{-\delta t} \cos(\omega t + \varphi)$$

z częstością kołową

$$\omega = \sqrt{\frac{1}{LC} - \frac{R^2}{4L^2}} \quad (8.41)$$

mniejszą od częstości drgań swobodnych (porównaj z (8.17)). Dla $R = 0$ wyrażenie (8.41) przechodzi w (8.17). Dobroć obwodu drgającego

$$Q = \frac{1}{R} \sqrt{\frac{L}{C}}. \quad (8.42)$$

Na zakończenie zauważmy, że gdy współczynnik tłumienia δ wzrasta, to okres drgań tłumionych również rośnie i przy $\delta = \omega_o$ przyjmuje wartość nieskończoną, tj. ruch drgający przestaje być periodycznym. W tym przypadku wielkość drgająca asymptotycznie zbliża się do zera, kiedy $t \rightarrow \infty$. Proces ten nazywamy aperiodycznym.

8.5. Drgania wymuszone

Ażeby w realnym układzie drgającym otrzymać drgania nietłumione, należy kompensować straty energii. Taka kompensacja jest możliwa gdy stosuje się odpowiednio działający czynnik wymuszający

$$X = X_o \cos \omega t.$$

W przypadku drgań mechanicznych rolę $X(t)$ odgrywa siła wymuszająca

$$F = F_o \cos \omega t. \quad (8.43)$$

Z uwzględnieniem siły (8.43) prawo ruchu dla wahadła sprężynowego zapiszemy w postaci

$$m \frac{d^2 s}{dt^2} = -ks - r \frac{ds}{dt} + F_o \cos \omega t.$$

Stąd otrzymujemy równanie w bardziej ogólnej formie

$$\frac{d^2 s}{dt^2} + 2\delta \frac{ds}{dt} + \omega_o^2 s = \frac{F_o}{m} \cos \omega t. \quad (8.44)$$

Jeżeli rozważamy elektryczny obwód drgający, to rolę $X(t)$ odgrywa włączona do obwodu zewnętrzna, periodycznie zmieniająca się SEM lub zmienne napięcie $V = V_o \cos \omega t$. Wówczas równanie (8.15) możemy napisać w postaci

$$\frac{d^2 Q}{dt^2} + \frac{R}{L} \frac{dQ}{dt} + \frac{1}{LC} Q = \frac{V_o}{L} \cos \omega t. \quad (8.45)$$

Stosując (8.17) i (8.40) otrzymujemy

$$\frac{d^2 Q}{dt^2} + 2\delta \frac{dQ}{dt} + \omega_o^2 Q = \frac{V_o}{L} \cos \omega t. \quad (8.46)$$

Drgania powstające pod wpływem zewnętrznej, periodycznie zmieniającej się siły, lub pod wpływem periodycznie zmieniającej się SEM, nazywamy odpowiednio wymuszonymi drganiami mechanicznymi lub elektrycznymi.

Równania (8.44) i (8.46) można sprowadzić do liniowego niejednorodnego równania różniczkowego

$$\frac{d^2s}{dt^2} + 2\delta \frac{ds}{dt} + \omega_o^2 s = x_o \cos \omega t. \quad (8.47)$$

Rozwiązanie tego równania stanowi suma ogólnego rozwiązania (8.35) równania jednorodnego (8.31) i szczególnego rozwiązania równania niejednorodnego. Zamienimy prawą stronę równania na wielkość zespoloną

$$\frac{d^2s}{dt^2} + 2\delta \frac{ds}{dt} + \omega_o^2 s = x_o e^{i\omega t}. \quad (8.48)$$

Szczególnego rozwiązania tego równania będziemy poszukiwać w postaci

$$s = s_o e^{i\eta t}.$$

Podstawiając wyrażenia dla s i jego pochodnych do równania (8.48), otrzymujemy

$$s_o e^{i\eta t} (-\eta^2 + 2i\delta\eta + \omega_o^2) = x_o e^{i\omega t}.$$

Ponieważ równość ta musi być spełniona dla każdej chwili czasu, więc $\eta = \omega$. Z równania tego można również określić s_o . Dodatkowo mnożąc licznik i mianownik wyrażenia s_o przez $(\omega_o^2 - \omega^2 - 2i\delta\omega)$, otrzymujemy

$$s_o = \frac{x_o}{(\omega_o^2 - \omega^2) + 2i\delta\omega} = x_o \frac{(\omega_o^2 - \omega^2) - 2i\delta\omega}{(\omega_o^2 - \omega^2)^2 + 4\delta^2\omega^2}.$$

To złożone wyrażenie wygodniej jest przedstawić w formie

$$s_o = A e^{-i\varphi},$$

gdzie

$$A = \frac{x_o}{\sqrt{(\omega_o^2 - \omega^2)^2 + 4\delta^2\omega^2}} \quad (8.49)$$

oraz

$$\varphi = \arctg \frac{2\delta\omega}{\omega_o^2 - \omega^2}. \quad (8.50)$$

Tak więc rozwiązanie równania (8.48) w postaci zespolonej ma postać

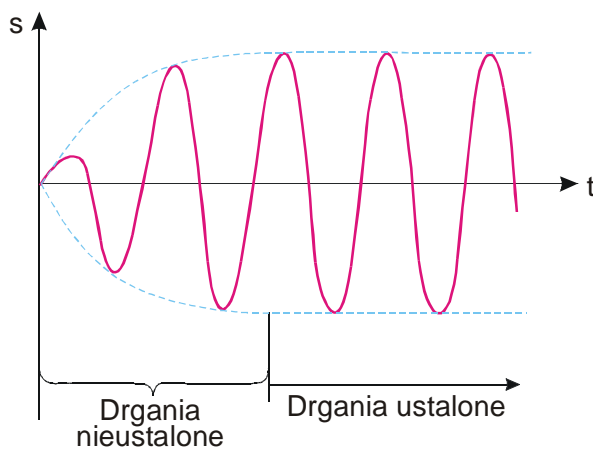
$$s = Ae^{i(\omega t - \varphi)},$$

natomiast jego rzeczywista część jest równa

$$s = A \cos(\omega t - \varphi). \quad (8.51)$$

Składowa ogólnego rozwiązania równania (8.48) stanowiąca ogólne rozwiązanie równania jednorodnego (patrz (8.31)) odgrywa istotną rolę tylko w początkowym stadium procesu (przy ustalaniu się drgań) do chwili kiedy amplituda drgań wymuszonych nie osiąga wartości określonej równością (8.49). Graficznie drgania wymuszone przedstawione są na rys. 8.8. Tak więc w stanie ustalonym, drgania wymuszone przebiegają z częstością ω i są drganiami harmonicznymi, a ich amplituda i faza określone są wzorami (8.49) i (8.50) zależnymi od ω .

Zapišemy teraz wyrażenia (8.49), (8.50) i (8.51) dla drgań elektromagnetycznych uwzględniając, że $\omega_o^2 = 1/(LC)$ (zob. (8.17)) i $\delta = R/(2L)$ (zob. (8.40)):



Rys. 8.8. Drgania wymuszone

$$Q_o = \frac{V_o}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}, \quad (8.52)$$

$$\operatorname{tg} \varphi = \frac{R}{1/(\omega C) - \omega L}. \quad (8.53)$$

Różniczkując

$$Q = Q_o \cos(\omega t - \alpha)$$

względem t , otrzymamy natężenie prądu w obwodzie dla drgań ustalonych

$$I_o = -\omega Q_o \sin(\omega t - \alpha) = I_o \cos(\omega t - \alpha + \pi/2), \quad (8.54)$$

gdzie

$$I_o = \omega Q_o = \frac{V_o}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}. \quad (8.55)$$

Wyrażenie (8.54) można zapisać w postaci

$$I = I_o \cos(\omega t - \varphi),$$

gdzie $\varphi = \alpha - \pi/2$ oznacza przesunięcie fazowe pomiędzy prądem i napięciem. Zgodnie z wyrażeniem (8.53)

$$\operatorname{tg} \varphi = \operatorname{tg} \left(\alpha - \frac{\pi}{2} \right) = -\frac{1}{\operatorname{tg} \alpha} = \frac{\omega L - 1/(\omega C)}{R}. \quad (8.56)$$

Z wyrażenia (8.56) wynika, że prąd opóźniony jest w fazie względem napięcia ($\varphi > 0$) gdy $\omega L > 1/(\omega C)$ i wyprzedza napięcie ($\varphi < 0$) gdy $\omega L < 1/(\omega C)$.

8.6. Amplituda i faza drgań wymuszonych. Rezonans

Rozważmy zależność amplitudy A drgań wymuszonych od częstości traktując w sposób ogólny drgającą wielkość – bądź przesunięcie (x) drgającego ciała z położenia równowagi lub ładunku (Q) kondensatora.

Z wyrażenia (8.49) wynika, że amplituda może przyjąć pewną maksymalną wartość. Ażeby określić częstość rezonansową ω_{rez} to jest częstość, przy której amplituda A osiąga maksimum, należy znaleźć maksimum funkcji (8.49), czyli minimum wyrażenia pod pierwiastkiem mianownika. Różniczkując to wyrażenie względem ω i przyrównując do zera, określimy ω_{rez}

$$-4(\omega_o^2 - \omega^2)\omega + 8\delta^2\omega = 0.$$

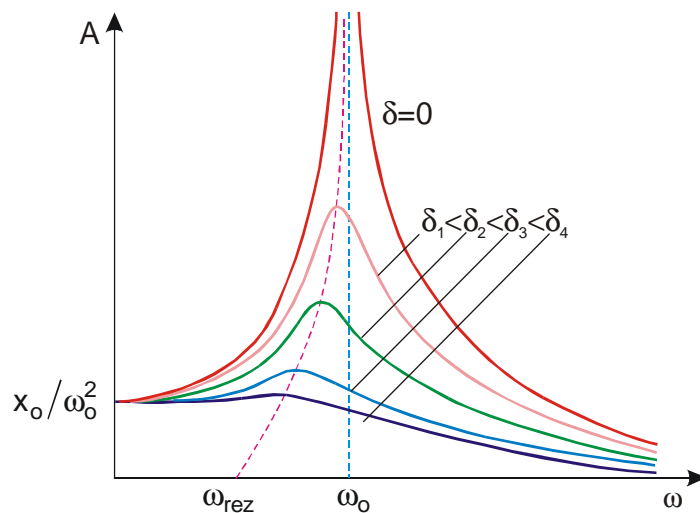
Równość ta jest spełniona dla $\omega = 0$ i $\omega = \pm\sqrt{\omega_o^2 - 2\delta^2}$. Sens fizyczny mają jedynie wartości dodatnie. Tak więc, częstotliwość rezonansowa

$$\omega_{rez} = \sqrt{\omega_o^2 - 2\delta^2}. \quad (8.57)$$

Zjawisko silnego wzrastania amplitudy drgań wymuszonych przy zbliżaniu się częstości siły wymuszającej do częstości ω_{rez} nazywamy rezonansem. W przypadku kiedy $\delta^2 \ll \omega_o^2$ wartość ω_{rez} praktycznie pokrywa się z częstością drgań własnych ω_o układu drgającego. Podstawiając (8.57) do (8.49), otrzymujemy

$$A_{rez} = \frac{x_o}{2\delta\sqrt{\omega_o^2 - \delta^2}}. \quad (8.58)$$

Na rys. 8.9 pokazano zależność amplitudy drgań wymuszonych od częstości dla różnych wartości δ . Z wyrażen (8.57) i (8.58) wynika, że im mniejsze δ tym wyżej i bardziej na prawo leży maksimum danej krzywej. Jeżeli $\omega \rightarrow 0$, to wszystkie krzywe przyjmują jedną wartość, różną od zera, x_o / ω_o^2 , nazywaną odchyleniem statycznym. Jeżeli $\omega \rightarrow \infty$ to wszystkie krzywe asymptotycznie dążą do zera. Przytoczony zbiór krzywych nazywamy krzywymi rezonansowymi.



Rys. 8.9. Krzywe rezonansu dla różnych wartości współczynnika tłumienia.

Zjawisko rezonansu może być zjawiskiem pożytecznym jak i szkodliwym. I tak, rezonans jest wykorzystywany w akustyce przy stosowaniu pudeł rezonansowych w rozmaitych instrumentach. Działanie odbiorników radiowych i telewizyjnych jest możliwe dzięki zastosowaniu rezonansu elektrycznego. Z drugiej strony np., jeżeli wirujące części urządzenia nie są dokładnie wyważone, wymuszają drgania innych części urządzenia i jeżeli jest spełniony przy tym warunek rezonansu, to amplituda drgań wymuszonych może być tak duża, że doprowadzi do zniszczenia drgających części a więc i urządzenia.

8.7. Prąd zmienny

Rozważymy wymuszone drganie elektromagnetyczne zachodzące w obwodzie prądu elektrycznego zawierającego rezystor, cewkę indukcyjną i kondensator. Prąd zmienny można traktować jako kwasistacyjny, co oznacza, że chwilowe wartości natężenia prądu we wszystkich przekrojach obwodu są praktycznie jednakowe (zmiana prądu zachodzi dostatecznie wolno, a zaburzenia elektromagnetyczne w obwodzie rozprzestrzeniają się z prędkością światła). Dla chwilowych kwasistacyjnych prądów spełnione jest prawo Ohma i prawa Kirchhoffa.

Rozpatrzmy w kolejności procesy zachodzące w obwodzie zawierającym rezystor, cewkę indukcyjną i kondensator po przyłożeniu do nich napięcia zmiennego

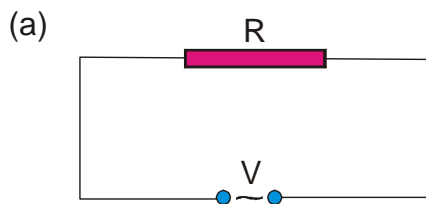
$$V = V_o \cos \omega t . \quad (8.59)$$

8.7.1. Obwód zawierający rezystancję

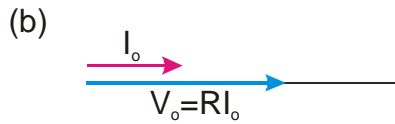
Prąd płynący przez rezystor określony jest prawem Ohma

$$I = \frac{V}{R} = \frac{V_o}{R} \cos \omega t = I_o \cos \omega t ,$$

gdzie amplituda natężenia prądu



$$I_o = \frac{V_o}{R}.$$



Dla poglądowego przedstawienia związków pomiędzy prądami zmiennymi i napięciami wygodniej jest posługiwać się metodą wektorową. Na rys. 8.10b pokazano wektory amplitud prądu i napięcia na rezystorze. Przesunięcie fazowe pomiędzy I_o i V_o jest zerowe.

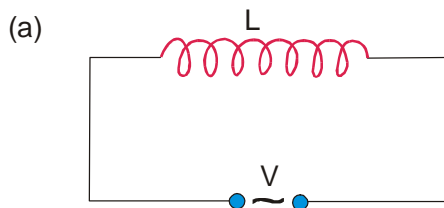
Rys. 8.10. Obwód zawierający rezystancję: (a) schemat obwodu, (b) wykres fazowy.

8.7.2. Obwód zawierający indukcyjność

Jeżeli do obwodu zawierającego cewkę przyłożymy napięcie zmienne [wzór (8.59)], to płynie prąd zmienny w wyniku czego powstanie *SEM* samoindukcji

$$E_s = -L \frac{dI}{dt}.$$

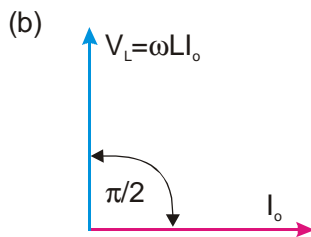
Wówczas prawo Ohma dla rozważanego obwodu ma postać



$$V_o \cos \omega t - L \frac{dI}{dt} = 0,$$

stąd

$$L \frac{dI}{dt} = V_o \cos \omega t. \quad (8.60)$$



Tak więc spadek napięcia na cewce indukcyjnej

$$V_L = L \frac{dI}{dt}. \quad (8.61)$$

Rys. 8.11. Obwód zawierający indukcyjność: (a) schemat obwodu, (b) wykres fazowy.

Z równania (8.60) wynika, że

$$dI = \frac{V_o}{L} \cos \omega t dt.$$

lub po scałkowaniu, uwzględniając, że stała całkowania jest równa zero (nie istnieje składowa stała prądu), otrzymujemy

$$I = \frac{V_o}{\omega L} \sin \omega t = \frac{V_o}{\omega L} \cos \left(\omega t - \frac{\pi}{2} \right) = I_o \cos \left(\omega t - \frac{\pi}{2} \right), \quad (8.62)$$

gdzie

$$I_o = \frac{V_o}{\omega L}.$$

Wielkość

$$R_L = \omega L \quad (8.63)$$

nazywamy reaktancją indukcyjną. Podstawiając $V_o = \omega L I_o$ w wyrażenie (8.60) i uwzględniając (8.61), otrzymujemy spadek napięcia na cewce indukcyjnej

$$V_L = \omega L I_o \cos \omega t. \quad (8.64)$$

Porównanie wyrażen (8.62) i (8.64) sprowadza się do wniosku, że spadek napięcia V_L wyprzedza w fazie prąd I płynący przez cewkę o kąt $\pi/2$, co pokazano na wykresie fazowym (rys. 8.11b).

8.7.3. Obwód zawierający pojemność

Jeżeli napięcie zmienne (8.59) przyłożymy do kondensatora to z upływem czasu kondensator będzie przeładowywał się a w obwodzie popłynie prąd zmienny. Jeżeli rezystancję przewodów można zaniedbać, to

$$\frac{Q}{C} = V_c = V_o \cos \omega t.$$

Natężenie prądu

$$I = \frac{dQ}{dt} = \omega C V_o \sin \omega t = I_o \cos\left(\omega t + \frac{\pi}{2}\right), \quad (8.65)$$

gdzie

$$I_o = \omega C V_o = \frac{V_o}{1/(\omega C)}.$$

Wielkość

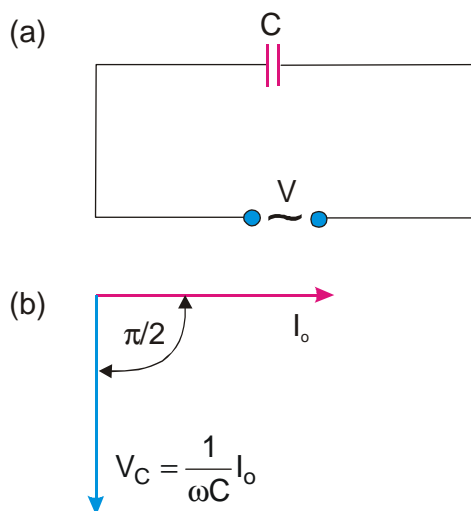
$$R_C = \frac{1}{\omega C}$$

nazywamy reaktancją pojemnościową. Dla prądu stałego ($\omega = 0$) $R_C = \infty$, co oznacza, że prąd stały nie płynie przez kondensator.

Spadek napięcia na kondensatorze

$$V_c = \frac{1}{\omega C} I_o \cos \omega t. \quad (8.66)$$

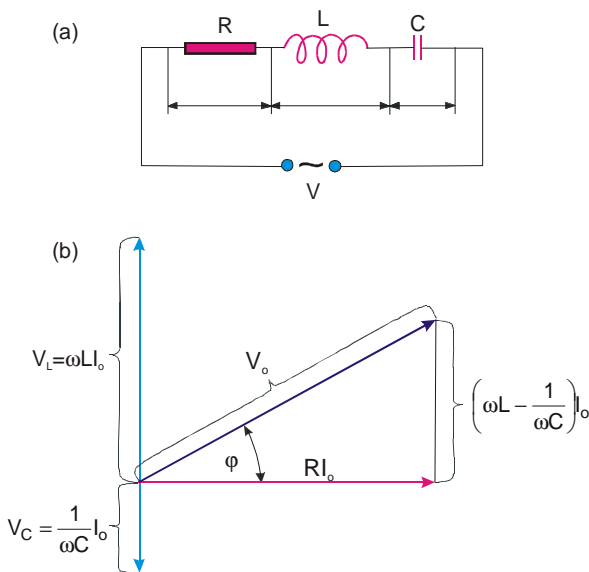
Porównanie wyrażen (8.65) i (8.66) wskazuje, że spadek napięcia V_c opóźniony jest w fazie o $\pi/2$ w porównaniu z prądem I . Pokazano to na wykresie fazowym (rys. 8.12b).



Rys. 8.12. Obwód zawierający kondensator: (a) schemat obwodu, (b) wykres fazowy.

8.7.4. Obwód RLC

Na rys. 8.13a pokazano obwód zawierający rezystor R , cewkę indukcyjną L , kondensator o pojemności C , do którego przyłożono napięcie zmienne. W obwodzie płynie prąd zmienny powodujący spadek napięcia na poszczególnych elementach obwodu V_R , V_L i V_C . Na rys. 8.13b pokazano z kolei wykres fazowy amplitud spadku napięć na rezystorze (V_R), cewce (V_L) i kondensatorze (V_C). Amplituda V_o przyłożonego napięcia powinna być równa sumie geometrycznej amplitud tych spadków napięć. Jak widać z rys. 8.13b, kąt φ określa różnicę faz pomiędzy napięciem i natężeniem prądu.



Rys. 8.13. Obwód RLC: (a) schemat obwodu, (b) wykres fazowy.

Z rysunku wynika, że

$$\operatorname{tg} \varphi = \frac{\omega L - 1/(\omega C)}{R}. \quad (8.67)$$

Z trójkąta prostokątnego otrzymujemy

$$(R I_o)^2 + \left[\left(\omega L - \frac{1}{\omega C} \right) I_o \right]^2 = V_o^2,$$

stąd amplituda prądu ma wartość

$$I_o = \frac{V_o}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C} \right)^2}}, \quad (8.68)$$

co jest zgodne z (8.55).

Tak więc, jeżeli napięcie w obwodzie zmienia się według prawa

$$V = V_o \cos \omega t,$$

to w obwodzie płynie prąd

$$I = I_o \cos(\omega t - \varphi), \quad (8.69)$$

gdzie φ i I_o określone są wzorami (8.67) i (8.68). Wielkość

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C} \right)^2} = \sqrt{R^2 + (R_L - R_C)^2} \quad (8.70)$$

nazywamy impedancją obwodu, a wielkość

$$X = R_L - R_C = \omega L - \frac{1}{\omega C}$$

nazywamy reaktancją.

Zauważmy, że impedancja obwodu RLC osiąga minimum, gdy

$$\omega L - \frac{1}{\omega C} = 0,$$

czyli gdy

$$\omega_o = \frac{1}{\sqrt{LC}}. \quad (8.71)$$

Częstość tę nazywamy rezonansową i oznaczamy przez ω_o .

8.7.5. Moc wydzielana w obwodzie prądu zmiennego

Chwilowa wartość mocy rozpraszanej w obwodzie równa jest

$$P(t) = V(t)I(t),$$

gdzie $V(t) = V_o \cos \omega t$, $I(t) = I_o \cos(\omega t - \varphi)$ (zobacz wyrażenie (8.59) i (8.69)). Korzystając z wzoru trygonometrycznego dla $\cos(\omega t - \varphi)$, otrzymamy

$$P(t) = I_o V_o \cos(\omega t - \varphi) \cos \omega t = I_o V_o (\cos^2 \omega t \cos \varphi + \sin \omega t \cos \omega t \sin \varphi).$$

Praktyczne znaczenie ma nie chwilowa wartość mocy, ale jej średnia wartość za okres drgań. Uwzględniając, że $\langle \cos^2 \omega t \rangle = 1/2$, $\langle \sin \omega t \cos \omega t \rangle = 0$, otrzymamy

$$\langle P \rangle = \frac{1}{2} I_o V_o \cos \varphi. \quad (8.72)$$

Z wykresu fazowego (rys. 8.13) wynika, że $V_o \cos \varphi = R I_o$. Dlatego

$$\langle P \rangle = \frac{1}{2} R I_o^2.$$

Taką moc wydziela prąd stały $I = I_o / \sqrt{2}$

Wielkości

$$I = \frac{I_o}{\sqrt{2}}; \quad V = \frac{V_o}{\sqrt{2}}$$

nazywamy odpowiednio wartościami skutecznymi prądu i napięcia.

Uwzględniając skuteczne wartości prądu i napięcia, wyrażenie dla średniej mocy (8.72) można zapisać w postaci

$$\langle P \rangle = IV \cos \varphi, \quad (8.73)$$

gdzie czynnik $\cos \varphi$ nazywamy "współczynnikiem mocy". Wyrażenie (8.73) pokazuje, że w ogólnym przypadku moc wydzielająca się w obwodzie prądu zmiennego zależy nie tylko od natężenia prądu i napięcia, ale również od przesunięcia fazowego między nimi.

8.8. Procesy falowe

Rozpatrzmy dokładniej mechanizm powstawania i rozchodzenia się drgań w ośrodku sprężystym, między którego cząsteczkami istnieją siły przeciwdziałające jego odkształcaniu. Między cząsteczkami gazów działają siły wzajemnego przyciągania i odpychania. Dzięki temu gazy zdolne są przeciwdziałać zmianie ich objętości, czyli charakteryzują się sprężystością objętościową. Jednocześnie zmieniają swą postać, inaczej mówiąc nie mają sprężystości postaciowej. Analogicznymi własnościami charakteryzują się ciecze. Ciała stałe wykazują zarówno sprężystość postaci jak i objętości.

Ciało drgające w ośrodku sprężystym oddziałuje okresowo na przylegające do niego cząsteczki ośrodka, wytrącając je ze stanu równowagi i zmuszając do wykonywania drgań wymuszonych. Ośrodek w pobliżu ciała odkształca się i pojawiają się siły sprężystości. Siły te z kolei działają na bardziej odległe cząsteczki, wyprowadzając je ze stanu równowagi. A zatem, wszystkie bardziej oddalone od ciała obszary ośrodka są stopniowo wprowadzane w ruch drgający.

Rozpatrzony powyżej proces rozchodzenia się drgań w ośrodku sprężystym jest przykładem ruchu falowego. W fizyce *falami nazywamy różnego rodzaju rozchodzące się w przestrzeni zaburzenia stanu materii lub pola*. I tak np. fale akustyczne w gazach lub cieczech są drganiami ciśnienia rozchodzącego się w tych ośrodkach, natomiast fale elektromagnetyczne są to rozchodzące się w przestrzeni drgania natężeń pola elektrycznego i magnetycznego.

Mechaniczne zaburzenia (odkształcenia) rozchodzące się w ośrodku sprężystym nazywamy falami sprężystymi. Jeżeli zaburzenia te (to jest amplitudy drgań odpowiednich cząstek ośrodka) są małe, to fale takie nazywamy dźwiękowymi lub akustycznymi; przy częstościach od 16 do 20000 Hz fale te oddziałują na organy słuchu człowieka wywołując wrażenie dźwiękowe (zob. pkt. 8.14).

Fala sprężysta nazywa się podłużną, jeżeli drgania cząstek ośrodka są równoległe do kierunku rozchodzenia się fal. Jeżeli cząsteczki ośrodka drgają w płaszczyznach prostopadłych do kierunku rozchodzenia się fali, to fala taka nazywa się poprzeczną. Fale poprzeczne mogą propagować się tylko w ośrodku który charakteryzuje się sprężystością postaci. A więc fale poprzeczne mogą istnieć tylko w ciałach stałych. Fale podłużne związane są z odkształceniem objętościowym ośrodka i dlatego mogą się rozchodzić zarówno w ciałach stałych, jak i w cieczech i gazach.

Istotna różnica między falami sprężystymi w ośrodku, a innym dowolnym uporządkowanym ruchem jego cząsteczek polega na tym, że rozchodzenie się fal nie jest związane z przenoszeniem materii.

Jeżeli rozpatrzeć proces falowy bardziej szczegółowo, to jest oczywistym, że drgają nie tylko cząsteczki rozmieszczone wzdłuż jakiejś osi, ale drga zbiór cząsteczek rozmieszczonych w pewnej objętości. Oznacza to, że fala rozprzestrzeniająca się od źródła drgań obejmuje coraz to

nowe obszary przestrzeni. *Miejsce geometryczne punktów do których dochodzą drgania w danej chwili t , nazywamy czołem fali. Z kolei miejsce geometryczne punktów drgających w jednakowej fazie nazywamy powierzchnią falową.* Powierzchni falowych można przeprowadzić bardzo dużo, natomiast w danej chwili czasu jest to tylko jedno czoło fali.

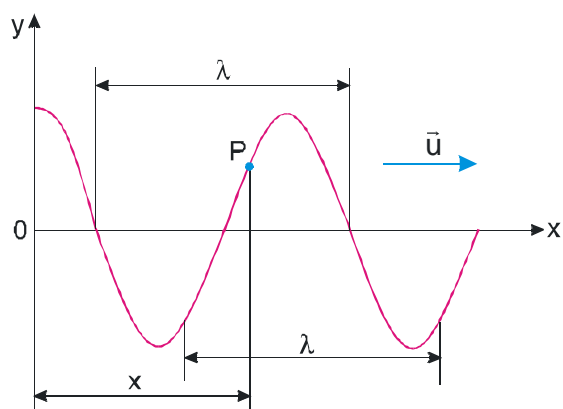
Powierzchnie falowe mogą mieć dowolne kształty. W prostszych przypadkach przedstawiają one zbiór płaszczyzn równoległych do siebie lub zbiór koncentrycznych sfer. W tych przypadkach fale nazywamy odpowiednio płaską i kulistą.

8.9. Fale biegnące

Falami biegnącymi nazywamy fale, które (w odróżnieniu od fal stojących) przenoszą energię. Prostym przykładem fali biegnącej jest fala na strunie. Dla fal biegnących dowolnego rodzaju, formalizm matematyczny jest identyczny.

Jeżeli jeden z końców rozciągniętej struny wykonuje drgania harmoniczne, to wzdłuż struny będzie rozprzestrzeniać się fala sinusoidalna. Udowodnimy, że prędkość rozchodzenia się fali po strunie określona jest siłą naciągu struny T i masą struny μ przypadającą na jednostkę jej długości. Składową prędkości ruchu punktu na strunie w kierunku pionowym oznaczymy przez v , a prędkość fali, to jest prędkość przemieszczania się grzbietu fali po strunie, oznaczymy przez u .

Najpierw jednak wyprowadzimy wyrażenie określające ruch falowy. Założymy, że oś x jest zgodna z kierunkiem propagacji fali. W tym przypadku powierzchnie falowe są prostopadłe do osi x , a w związku z tym przemieszczenie y będzie zależało tylko od x i t , tj. $y = y(x, t)$.



Rys. 8.14. Fala biegnąca po strunie.

Rozważmy pewien punkt struny P znajdujący się w odległości x od źródła drgań 0 . Jeżeli drganie cząstek znajdujących się w płaszczyźnie $x = 0$ opisujemy funkcją $y = A \cos \omega t$, to punkt P drga według tego prawa, jednakże jego drgania są opóźnione o czas $\tau = x/u$ potrzebny dla przemieszczenia fali. Wówczas równanie drgań cząstek leżących w odległości x ma postać

$$y(x, t) = A \cos \omega \left(t - \frac{x}{u} \right). \quad (8.74)$$

Jest to równanie fali biegnącej. Jeżeli fala płaska propaguje się w kierunku przeciwnym, to

$$y(x, t) = A \cos \omega \left(t + \frac{x}{u} \right).$$

W ogólnym przypadku równanie płaskiej sinusoidalnej fali, propagującej się wzdłuż dodatniego kierunku osi x , ma postać

$$y(x,t) = A \cos \left[\omega \left(t - \frac{x}{u} \right) + \varphi_0 \right], \quad (8.75)$$

gdzie A – amplituda fali, ω – częstość kołowa, φ_0 – faza początkowa określona wyborem początkowych wielkości x i t , natomiast $[\omega(t - x/u) + \varphi_0]$ oznacza całkowitą fazę fali.

Dla scharakteryzowania fali sinusoidalnej wprowadza się pojęcie liczby falowej

$$k = \frac{2\pi}{\lambda} = \frac{2\pi}{uT} = \frac{\omega}{u}. \quad (8.76)$$

Uwzględniając to, równanie (8.75) można zapisać w postaci

$$y(x,t) = A \cos(\omega t - kx + \varphi_0). \quad (8.77)$$

Opierając się na wzorze Eulera, równanie płaskiej sinusoidalnej fali przyjmuje postać

$$y(x,t) = A e^{i(\omega t - kx + \varphi_0)},$$

gdzie sens fizyczny ma jedynie część rzeczywista.

Założmy, że w procesie falowym faza jest stała, to jest

$$\omega \left(t - \frac{x}{u} \right) + \varphi_0 = const. \quad (8.78)$$

Różniczkując to wyrażenie otrzymujemy $dt - (1/u)dx = 0$, skąd

$$\frac{dx}{dt} = u. \quad (8.79)$$

Tak więc, prędkość u propagacji fali w równaniu (8.78) jest niczym innym jak prędkością przemieszczania się fazy fali. Z tego powodu prędkość tę nazywamy prędkością fazową.

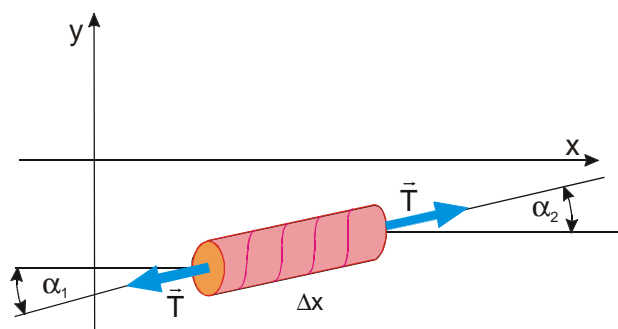
Z wyrażenia (8.76) wynika, że

$$u = \frac{\omega}{k}, \quad (8.80)$$

co oznacza, że prędkość fal sinusoidalnych zależy od ich częstości. To zjawisko nazywamy dyspersją fal, a ośrodek w którym obserwowana jest dyspersja fal nazywamy ośrodkiem dyspersyjnym.

Teraz znajdziemy relację określającą zależność prędkości fali na strunie od T i μ . Rozpatrzmy element struny Δx , którego końce tworzą małe kąty α_1 i α_2 z osią x (rys. 8.15). Przez małe kąty rozumiemy takie, dla których słuszne jest przybliżenie $\sin \alpha = \alpha = \partial y / \partial x$. Siła wypadkowa działająca na element struny w kierunku pionowym jest równa $F_{wyp} = T\alpha_1 - T\alpha_2$.

Siła ta jest równa iloczynowi masy elementu struny $\mu\Delta x$ przez przyspieszenie pionowe $\partial^2 y / \partial t^2$. Wobec tego



$$F_{\text{wyp}} = T\alpha_1 - T\alpha_2 = (\mu\Delta x)\frac{\partial^2 y}{\partial t^2},$$

$$T \frac{\Delta\alpha}{\Delta x} = \mu \frac{\partial^2 y}{\partial t^2},$$

$$\frac{\Delta\alpha}{\Delta x} = \frac{\mu}{T} \frac{\partial^2 y}{\partial t^2}.$$

Rys. 8.15. Siły działające na element Δx struny.

Podstawiając $\alpha = \partial y / \partial x$ w lewą część ostatniego wyrażenia, otrzymamy

$$\frac{\partial^2 y}{\partial x^2} = \frac{\mu}{T} \frac{\partial^2 y}{\partial t^2}. \quad (8.81)$$

Jest to różniczkowe równanie falowe struny. Prędkość fali można określić podstawiając w (8.81) odpowiednie pochodne funkcji $y(x,t)$ określonej równaniem (8.77).

$$\frac{\partial^2 y}{\partial x^2} = -Ak^2 \cos(\omega t - kx + \varphi_0),$$

$$\frac{\partial^2 y}{\partial t^2} = -A\omega^2 \cos(\omega t - kx + \varphi_0).$$

Uwzględniając te wyrażenia w (8.81), mamy

$$k^2 = \frac{\mu}{T} \omega^2,$$

skąd znajdujemy

$$u = \sqrt{\frac{T}{\mu}}. \quad (8.82)$$

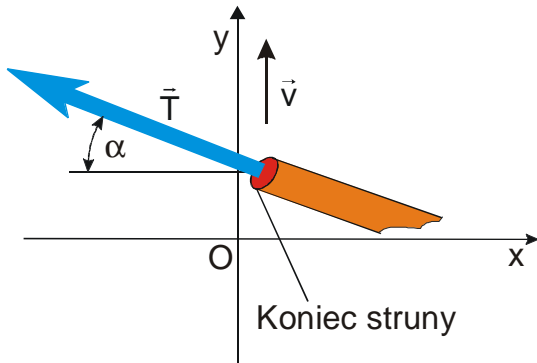
Podstawiając do równania (8.81) w miejsce wyrażenia μ/T parametr $1/u^2$, otrzymamy

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{u^2} \frac{\partial^2 y}{\partial t^2}. \quad (8.83)$$

Równanie to znane jest jako równanie falowe i jest słuszne dla wszystkich rodzajów fal biegnących: akustycznych, sprężystych, elektrycznych, elektromagnetycznych. Rozwiązaniem tego równania jest wyrażenie (8.77).

8.10. Przenoszenie energii przez fale

Działając ręką na koniec długiej, rozciągniętej struny i wymuszając vibracje, wykonujemy pracę, która ujawnia się w postaci energii kinetycznej i potencjalnej oddzielnych elementów struny. Z upływem czasu zaczynają drgać coraz bardziej oddalone elementy struny. Energia dostarczona do jednego końca struny, przenoszona z prędkością fali może być przyjęta i pochłonięta na drugim końcu.



Rys. 8.16. Koniec struny odciągnięty do góry celem wzbudzenia fali biegnącej.

Prędkość przenoszenia energii można znaleźć obliczając siłę z jaką ciągniemy strunę do góry i w dół. W tym celu skorzystamy z wzoru określającego moc $P = \vec{F} \cdot \vec{v}$ i zastosujemy go dla struny z rys. 8.16

$$P = \vec{F} \cdot \vec{v} = T \left(\frac{\partial y}{\partial t} \right) \sin \alpha$$

kładąc $v = \partial y / \partial t$. Kąt jest mały więc $\sin \alpha = \partial y / \partial x$. Tak więc

$$P = -T \left(\frac{\partial y}{\partial t} \right) \left(\frac{\partial y}{\partial x} \right),$$

a ponieważ $y = A \cos \omega(t - x/u)$, więc

$$\frac{\partial y}{\partial t} = -A\omega \sin \omega \left(t - \frac{x}{u} \right)$$

$$\frac{\partial y}{\partial x} = A \frac{\omega}{u} \sin \omega \left(t - \frac{x}{u} \right)$$

i możemy napisać wyrażenie

$$P = TA^2 \frac{\omega^2}{u} \sin^2 \omega t$$

określające chwilową moc P przekazywaną w chwili czasu t w punkcie $x = 0$.

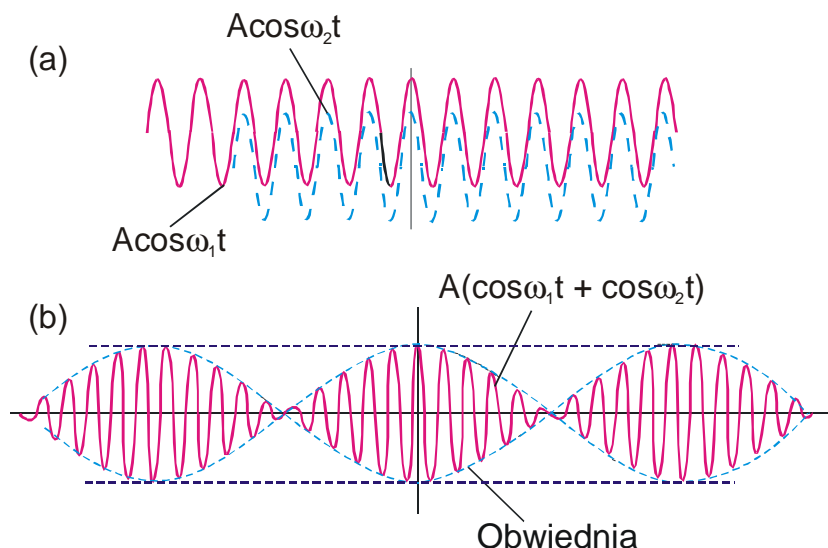
Średnia wartość mocy jest dwukrotnie mniejsza, ponieważ $\langle \sin^2 \omega t \rangle = 1/2$

$$\langle P \rangle = \frac{T\omega^2}{2u} A^2. \quad (8.84)$$

Średnia wartość przenoszonej mocy nazywana jest natężeniem fali. Natężenie fali jest proporcjonalne do kwadratu jej amplitudy, a w przestrzeni trójwymiarowej (fale akustyczne i fale elektromagnetyczne) równe jest średniej mocy przenoszonej przez jednostkową powierzchnię czoła fali.

8.11. Paczka falowa. Prędkość grupowa

Jeżeli ośrodek, w którym propagują się fale jest liniowy, tzn. właściwości ośrodka nie zmieniają się pod wpływem zaburzeń spowodowanych przez fale, to spełniona jest zasada superpozycji (nakładania) fal: *wypadkowe wzbudzenie w dowolnym punkcie ośrodka liniowego przy jednoczesnej propagacji kilku fal jest równe sumie wzbudzeń wywołanych przez każdą z tych fal z osobna.*



Rys. 8.17. (a) Dwie fale monochromatyczne o nieznacznie różniących się częstościach, będących w fazie w początku układu współrzędnych. (b) Suma dwóch fal monochromatycznych.

W niniejszym punkcie zajmiemy się dodawaniem fal monochromatycznych o różnych lecz zbliżonych częstościach. W tym przypadku z upływem czasu fale będą "rozbiegać się" fazowo jedna względem drugiej. Rozpatrzmy uproszczony przypadek złożenia dwóch jednakowych fal monochromatycznych, z których jedna ma częstość ω_1 , a druga – ω_2 (rys. 8.17). Przyjmijmy oznaczenia

$$\bar{\omega} = \frac{\omega_1 + \omega_2}{2} \quad , \quad \Delta\omega = \frac{\omega_2 - \omega_1}{2} .$$

Wówczas suma

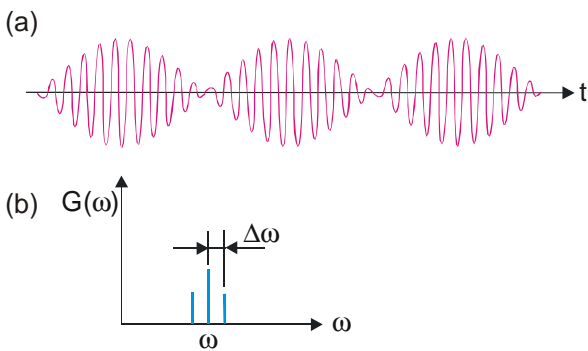
$$S(t) = A \cos(\bar{\omega} + \Delta\omega)t + A \cos(\bar{\omega} - \Delta\omega)t .$$

Wyrażenie to możemy przekształcić stosując znane z trygonometrii reguły dodawania cosinusów $\cos \alpha + \cos \beta = 2 \cos[(\alpha + \beta)/2] \cos[(\alpha - \beta)/2]$. Wówczas

$$S(t) = 2A \cos[(\Delta\omega)t] \cos \bar{\omega}t = A(t) \cos \bar{\omega}t , \quad (8.85)$$

gdzie $A(t) = 2A \cos(\Delta\omega)t$ jest obwiednią, czyli funkcją modulującą (rys. 8.17). W naszym przykładzie obwiednię stanowi fala monochromatyczna o mniejszej częstości.

Dodając większą liczbę fal monochromatycznych o nieznacznie różniących się częstościach, można otrzymać funkcję modulacji o dowolnej formie. Jako przykład rozpatrzmy pojedynczy impuls, który odgrywa ważną rolę w procesach falowych. Taki impuls ma kształt podobny do środkowego przebiegu na rys. 8.17 (dudnienia) i nazywa się paczką falową. Zobaczymy jak poprzez dobór fal monochromatycznych o zbliżonych częstościach można zbudować paczkę falową. Wykorzystując obraz przedstawiony na rys. 8.17b można stłumić boczne dudnienia dodając trzecią falę monochromatyczną o częstości $\bar{\omega}$ i amplitudzie równej wysokości dudnienia (zauważmy, że $A(t)$ zmienia znak przy kolejnych dudnieniach). Taka fala monochromatyczna dodaje się do centralnego dudnienia i okazuje się przesunięta w fazie o 180° w obszarze sąsiednich dudnień (rys. 8.18). Funkcja $G(\omega)$ charakteryzuje względne intensywności trzech sumowanych fal monochromatycznych.



Rys. 8.18. (a) Suma trzech fal monochromatycznych. (b) Względny rozkład amplitud.

Po to aby stłumić kolejne boczne dudnienia należy dodać dwie fale monochromatyczne

$$\cos(\bar{\omega} - \Delta\omega/2) \quad \text{i} \quad \cos(\bar{\omega} + \Delta\omega/2) \quad (8.86)$$

z odpowiednimi amplitudami, jak pokazano na rys. 8.19. Dodanie jeszcze dwóch fal monochromatycznych nie ma wpływu na środkowe dudnienie, lecz ich suma będzie przesunięta w fazie o 180° względem kolejnych dudnień.

W rzeczywistości w celu uformowania pojedynczej paczki falowej dokonuje się sumowania nieskończenie wielu fal monochromatycznych o zbliżonych częstościach. Odpowiadająca temu sytuacja zilustrowana jest na rys. 8.20. Na tym rysunku funkcja $G(\omega)$ charakteryzuje względne amplitudy pojedynczych składowych monochromatycznych. Taka funkcja $G(\omega)$ nazwana jest funkcją gaussowską i jest opisana wzorem

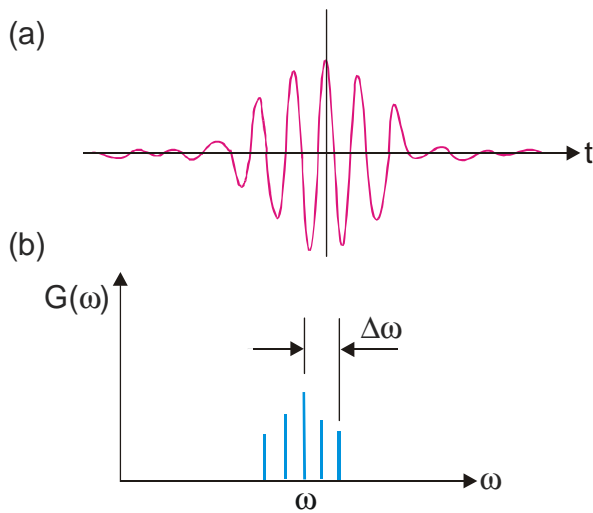
$$G(\omega) = \exp\left[-\frac{(\bar{\omega} - \omega)^2}{2(\Delta\omega)^2}\right], \quad (8.87)$$

gdzie $\Delta\omega$ jest odchyleniem średniokwadratowym (lub standardowym) wielkości względem $\bar{\omega}$. Wielkość $\Delta\omega$ będziemy nazywali rozrzutem częstości.

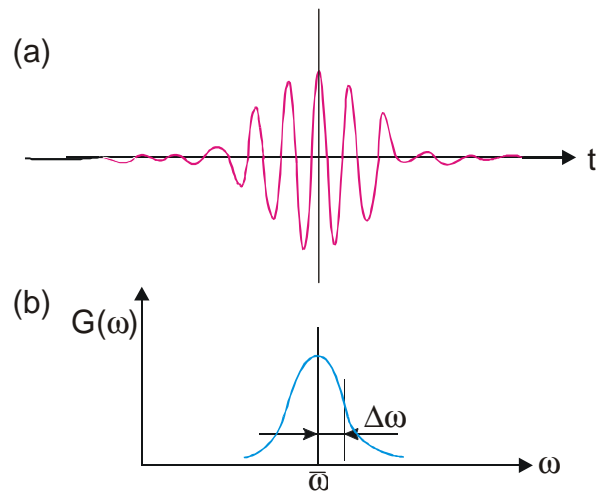
Aby znaleźć sumę nieskończonej liczby fal monochromatycznych, należy obliczyć całkę

$$\int G(\omega) \cos \omega t d\omega. \quad (8.88)$$

Całkę tę można obliczyć korzystając z tablic całek



Rys. 8.19. (a) Suma pięciu fal monochromatycznych. (b) Względny rozkład amplitud.



Rys. 8.20. (a) Suma nieskończonej liczby fal monochromatycznych. (b) Względny rozkład amplitud. $G(\omega)$ – funkcja gaussowska ze średnią wartością $\bar{\omega}$ i odchyleniem średniokwadratowym $\Delta\omega$.

$$\int \exp\left[-\frac{(\bar{\omega} - \omega)^2}{2(\Delta\omega)^2}\right] \cos \omega t d\omega = \sqrt{2\pi} \exp\left[-\left(\frac{t^2}{2}\right)(\Delta\omega)^2\right] \cos \bar{\omega} t, \quad (8.89)$$

Widzimy, że po prawej stronie tej równości napisana jest fala monochromatyczna $\cos \bar{\omega} t$ zmodulowana gaussowską obwiednią $\exp\left[-\left(t^2/2\right)(\Delta\omega)^2\right]$. Odchylenie średniokwadratowe dla tej funkcji jest równe

$$\Delta t = \frac{1}{\Delta\omega} \quad (8.90)$$

i nazywane jest szerokością paczki falowej. Wobec tego, *rozrzut częstości składowych monochromatycznych równy jest odwrotności szerokości paczki falowej*. Funkcja $G(\omega)$ nazywana jest transformatą (obrazem) Fouriera paczki falowej.

Po zsumowaniu fal o różnych częstościach można otrzymać interesujący wynik – prędkość rozchodzenia się obwiedni może znacznie różnić się od prędkości fazowej, z którą rozchodzą się składowe monochromatyczne. Na skutek tego prędkość rozchodzenia się paczki falowej jako całości, może istotnie różnić się od prędkości rozchodzenia się fal monochromatycznych wchodzących w jej skład. Prędkość paczki falowej, czyli prędkość obwiedni, nazwana jest prędkością grupową.

Na rys. 8.21 pokazano na przykładzie dwóch fal monochromatycznych o zbliżonych częstościach i długościach fal, dlaczego obwiednia może propagować się z inną prędkością. Superpozycję takich dwóch fal biegnących możemy zapisać w postaci

$$y(x, t) = A \cos[(\bar{\omega} + \Delta\omega)t - (\bar{k} + \Delta k)x] + A \cos[(\bar{\omega} - \Delta\omega)t - (\bar{k} - \Delta k)x],$$

gdzie $\bar{k} = 2\pi / \lambda$ jest średnią liczbą falową. W przypadku z rys. 8.21, obwiednia przemieszcza się w prawo dwukrotnie szybciej od każdej składowej monochromatycznej.

Stosując dla $\cos \alpha + \cos \beta$ trygonometryczną formę dodawania, otrzymujemy

$$y(x,t) = 2A \cos[(\Delta\omega)t - (\Delta k)x] \cos(\bar{\omega}t - \bar{k}x).$$

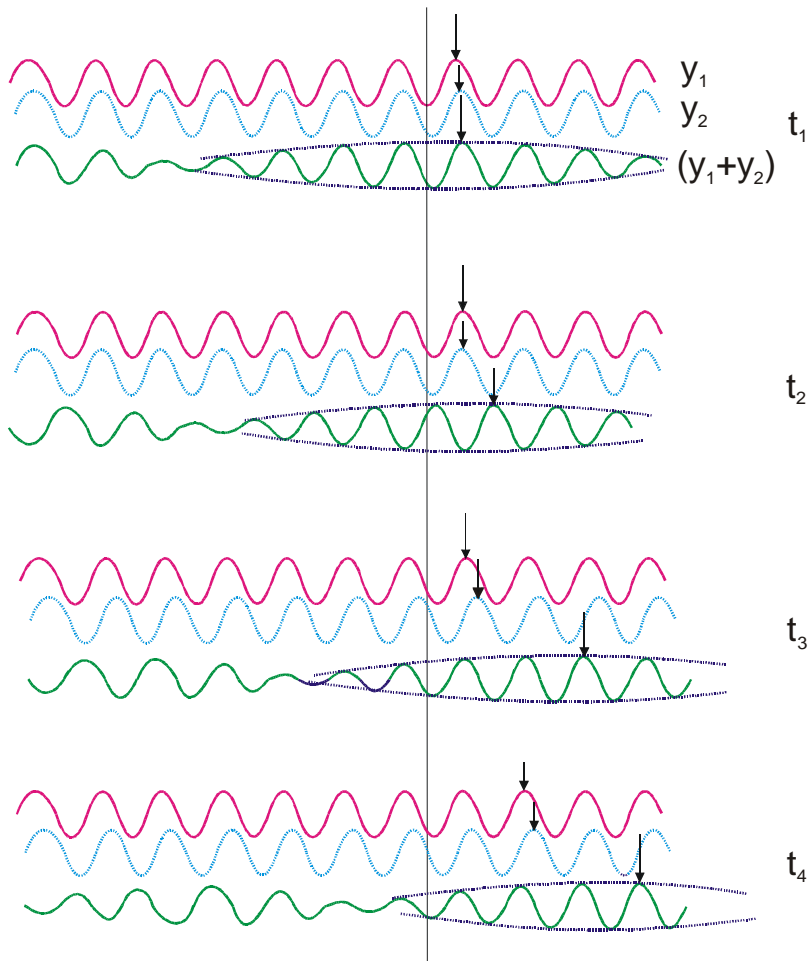
Widzimy, że obwiednie możemy opisać w postaci $A(x,t) = 2A \cos[(\Delta\omega)t - (\Delta k)x]$. Maksimum funkcji jest przyjmowane gdy

$$(\Delta\omega)t - (\Delta k)x = 0,$$

to jest przy

$$\frac{x}{t} = \frac{\Delta\omega}{\Delta k}.$$

Wielkość ta określa prędkość przemieszczania się grzbietu obwiedni. W ten sposób określona jest prędkość grupowa.



Rys. 8.21. Dwie fale monochromatyczne y_1 i y_2 poruszają się w prawo z nieznacznie różniącymi się prędkościami. W tym przypadku obwiednia sumy y_1 i y_2 przemieszcza się w prawo z podwójną prędkością. Przytoczono cztery kolejne położenia odpowiadające chwilom czasu t_1, \dots, t_4 . Strzałkami zaznaczono położenie grzbietów w różnych chwilach czasu.

Jeżeli mamy zbiór fal monochromatycznych o zbliżonych częstościach, przy czym ω stanowi funkcję zależną od k , tak że $\omega = \omega(k)$, to prędkość grupowa określona jest w następujący sposób

$$v_g = \frac{d\omega}{dk}. \quad (8.91)$$

Znajdziemy związek między prędkością fazową a prędkością grupową. Uwzględniając, że $\lambda = 2\pi/k$, mamy

$$v_g = \frac{d(uk)}{dk} = u + k \frac{du}{dk} = u + k \left(\frac{du}{d\lambda} \frac{d\lambda}{dk} \right) = u + k \left(-\frac{\lambda}{k} \right) \frac{du}{d\lambda},$$

czyli

$$v_g = u - \lambda \frac{du}{d\lambda}. \quad (8.92)$$

Ze związku (8.92) wynika, że v_g może być mniejsze jak i większe od u w zależności od znaku $du/d\lambda$. W ośrodkach niedispersyjnych $du/d\lambda = 0$ i prędkość grupowa jest równa prędkości fazowej. W teorii względności udowadnia się, że prędkość grupowa $v_g \leq c$, podczas gdy dla prędkości fazowej nie ma ograniczenia.

Najbardziej charakterystyczny przykład rozchodzenia się fali z prędkością grupową to przechodzenie światła przez dielektryk (patrz pkt. 9.3.5). Drugie bardzo ważne zastosowanie pojęcia prędkości grupowej związane jest z mechaniką kwantową, w której cząstkom przypisuje się paczki falowe. Prędkość cząstki jest zgodna z prędkością grupową paczki falowej, a nie z prędkościami oddzielnych składowych, które zwykle różnią się od siebie.

8.12. Interferencja fal

Nakładanie się w czasie i w przestrzeni kilku procesów falowych jest związane z pojęciem koherencji. Fale sinusoidalne których częstości są jednakowe są zawsze koherentne. *Fale nazywamy koherentnymi jeżeli różnica ich faz pozostaje stała w czasie.* Fale których częstości różnią się, są koherentne tylko w czasie koherentności (zob. pkt. 9.6). *Przy nakładaniu w przestrzeni kilku fal, w różnych punktach następuje wzmocnienie lub osłabienie wypadkowej fali w zależności od relacji fazowych pomiędzy nimi. Zjawisko to nazywamy interferencją fal.*

Rozpatrzmy superpozycję dwóch koherentnych fal sferycznych wzbudzonych źródłami punktowymi S_1 i S_2 (rys. 8. 22) drgającymi z jednakowymi amplitudami A_0 i częstościami ω_0 ze stałą różnicą faz.

Równanie fali sferycznej ma postać

$$S(r,t) = \frac{A_0}{r} \cos(\omega t - kr + \varphi_0), \quad (8.93)$$

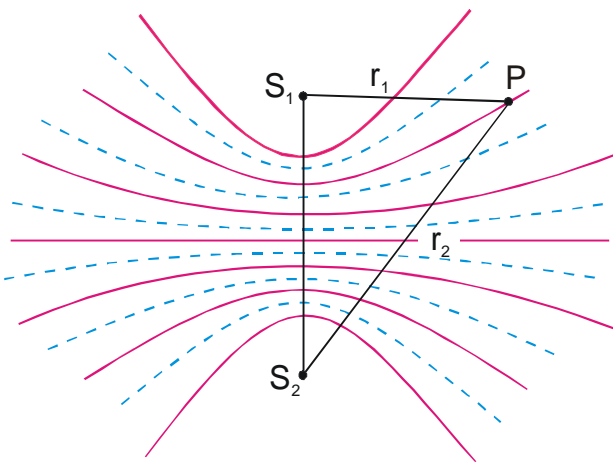
gdzie r – odległość od źródła fali do określonego punktu ośrodka. Zauważmy, że amplituda fali sferycznej maleje według prawa $1/r$ (gdy nie ma miejsca dysypacja energii). W wyniku superpozycji dwóch fal

$$S_1(r, t) = \frac{A_o}{r_1} \cos(\omega t - kr_1 + \varphi_1),$$

$$S_2(r, t) = \frac{A_o}{r_2} \cos(\omega t - kr_2 + \varphi_2),$$

amplituda fali wypadkowej w punkcie P wynosi

$$A^2 = A_o^2 \left\{ \frac{1}{r_1^2} + \frac{1}{r_2^2} + \frac{2}{r_1 r_2} \cos[k(r_1 - r_2) + (\varphi_1 - \varphi_2)] \right\}.$$



Ponieważ dla koherentnych źródeł różnica faz początkowych $\varphi_1 - \varphi_2 = \text{const}$, więc wynik interferencji fal w różnych punktach ośrodka zależy od wielkości $\Delta = r_1 - r_2$ nazywanej różnicą dróg fal.

W punktach gdzie

$$k(r_1 - r_2) - (\varphi_1 - \varphi_2) = \pm 2m\pi, \quad (8.94)$$

($m = 0, 1, 2, \dots$) obserwuje się maksima interferencyjne, a amplituda wypadkowa wynosi

$$A = (A_o / r_1) + (A_o / r_2).$$

Rys. 8.22. Interferencja koherentnych fal sferycznych.

Natomiast w punktach gdzie

$$k(r_1 - r_2) - (\varphi_1 - \varphi_2) = \pm(2m + 1)\pi \quad m = 1, 2, \dots \quad (8.95)$$

obserwuje się minima interferencyjne – amplituda wypadkowa $A = |(A_o / r_1) - (A_o / r_2)|$. $m = 0, 1, 2, \dots$ nazywamy odpowiednio rzędem interferencyjnego maksimum lub minimum.

Warunki (8.94) i (8.95) sprowadzają się do tego, że

$$r_1 - r_2 = \text{const}, \quad (8.96)$$

Wyrażenie (8.96) przedstawia równanie hiperboli o ogniskach w punktach S_1 i S_2 . Między dwoma maksimami interferencyjnymi (rys. 8.22) znajdują się minima interferencyjne. Obraz interferencyjny z rys. 8.22 uzyskuje się w przypadku gdy $\varphi_1 - \varphi_2 = 0$.

8.13. Fale stojące

Szczególnym przypadkiem interferencji są fale stojące. Fale stojące powstają przy nakładaniu się dwu sinusoidalnych fal biegnących, propagujących się w kierunkach przeciwnych z jednakowymi prędkościami i amplitudami.

Przyjmijmy, że dwie fale sinusoidalne propagujące się w przeciwnych kierunkach wzdłuż osi x mają jednakowe amplitudy i częstości. Oprócz tego, początek układu odniesienia wybieramy w punkcie, w którym obie fale mają jednakowe fazy, a czas mierzymy z chwilą kiedy fazy obu fal są zerowe. Wówczas równania fal mają postać

$$y_1 = A \cos(\omega t - kx),$$

$$y_2 = A \cos(\omega t + kx).$$

Dodając te równania i uwzględniając, że $k = 2\pi / \lambda$, otrzymujemy równanie fali stojącej

$$y_1 + y_2 = 2A \cos kx \cos \omega t = 2A \cos(2\pi x / \lambda) \cos \omega t. \quad (8.97)$$

Z równania fali stojącej (8.97) wynika, że w każdym punkcie fali stojącej zachodzą drgania o tej samej częstości ω z amplitudą $A_{st} = |2A \cos(2\pi x / \lambda)|$, zależną od współrzędnej x .

W punktach gdzie

$$\frac{2\pi x}{\lambda} = \pm m\pi, \quad (m = 0, 1, 2, \dots)$$

amplituda fali stojącej osiąga maksymalną wartość równą $2A$.

W punktach gdzie

$$\frac{2\pi x}{\lambda} = \pm \left(m + \frac{1}{2}\right)\pi, \quad (m = 0, 1, 2, \dots)$$

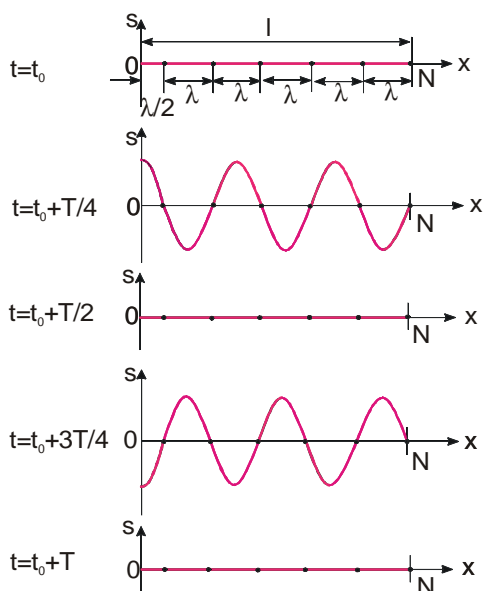
amplituda fali stojącej jest zerowa.

Punkty w których amplituda fali stojącej jest maksymalna ($A_{st} = 2A$) nazywamy strzałkami fali stojącej, a punkty w których amplituda jest zerowa ($A_{st} = 0$), nazywamy węzłami fali stojącej.

Z ostatnich wyrażen otrzymujemy współrzędne strzałek i węzłów

$$x_s = \pm m \frac{\lambda}{2}, \quad (m = 0, 1, 2, \dots)$$

$$x_w = \pm \left(m + \frac{1}{2}\right) \frac{\lambda}{2}. \quad (m = 0, 1, 2, \dots)$$



Rys. 8.23. Fala stojąca.

Z powyższych relacji wynika, że odległość między dwiema sąsiednimi strzałkami i dwoma sąsiednimi węzłami jest jednakowa i wynosi $\lambda/2$. Na rys. 8.23 pokazano ruch różnych punktów rozciągniętej sprężystej nici o długości l , w której wzbudzona została stojąca fala poprzeczna. Lewy koniec nici O wprowadzono w drgania harmoniczne, zaś prawy N zamocowano nieruchomo. W tym przypadku przy odbiciu fali od miejsca zamocowania, faza ulega zmianie o π i w miejscu tym tworzy się węzeł fali stojącej.

8.14. Fale dźwiękowe

Falami dźwiękowymi lub akustycznymi nazywamy podłużne fale sprężyste, mogące rozchodzić się w ciałach stałych, cieczech i gazach. Zakres częstotliwości tych fal obejmuje przedział od około 20 Hz do 20 kHz. Fale o mniejszych częstotliwościach (poniżej 20 Hz) nazywamy infradźwiękami, a fale o częstotliwościach większych niż słyszalne – ultradźwiękami. Fale infradźwiękowe są zwykle generowane przez źródła o wielkich masach, np. podłużne fale sejsmiczne. Ultradźwięki są wytwarzane przez specjalne generatory dużej częstotliwości i mają różne znaczenie praktyczne, np. w defektoskopach, hydrolokacji i medycynie.

Fale dźwiękowe okresowe dzielimy na tony i dźwięki złożone. Tony wywołują zmiany ciśnienia w ośrodku o przebiegu drgań harmonicznym prostym. Dźwięki złożone powstają w wyniku wzajemnego nakładania się różnych drgań harmonicznym. Dźwięki charakteryzujemy wysokością, barwą i natężeniem. Wysokość dźwięku rośnie ze wzrostem częstotliwości. Barwa dźwięku jest związana z zawartością wielu drgań o różnych przebiegach i częstotliwościach. Natężenie dźwięku zależy od wielkości amplitudy fali.

Jeżeli amplituda fali dźwiękowej nie jest zbyt duża, to prędkość dźwięku nie zależy od częstotliwości. W tabeli 8.1 zebrano prędkości dźwięku w kilku materiałach.

Tabela 8.1. Prędkość dźwięku

Ośrodek	Prędkość (m/s)
Powietrze (20° C)	340
Wodór	1286
Woda	1450
Żelazo	5130
Aluminium	5100
Guma	54

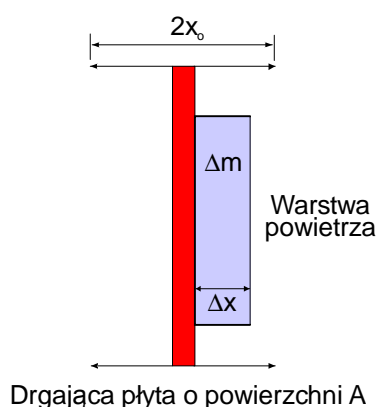
Rozważmy falę dźwiękową harmoniczną rozchodzącą się wzdłuż osi x (patrz rys. 8.24). Załóżmy, że cienka płaska płyta o powierzchni A wykonuje drgania z amplitudą x_0 i częstotliwością $\omega/2\pi$. Przekazuje ona energię warstwie powietrza o masie Δm . Maksymalna energia kinetyczna tej warstwy powietrza wynosi

$$\frac{1}{2} \Delta m v_0^2 = \frac{1}{2} \Delta m \omega^2 x_0^2,$$

czyli

$$\Delta E = \frac{1}{2} (\rho A \Delta x) \omega^2 x_0^2, \quad (8.98)$$

gdzie ρ jest gęstością powietrza.



Rys. 8.24. Drgająca płyta powoduje, że warstwa powietrza drga z amplitudą x_0 .

Natężenie fali jest definiowane jako moc na jednostkę powierzchni; czyli z ostatniego równania mamy

$$I = \frac{1}{2} \rho \omega^2 x_0^2 u. \quad (8.100)$$

Natężenie fali jest więc proporcjonalne do kwadratu amplitudy i kwadratu częstotliwości.

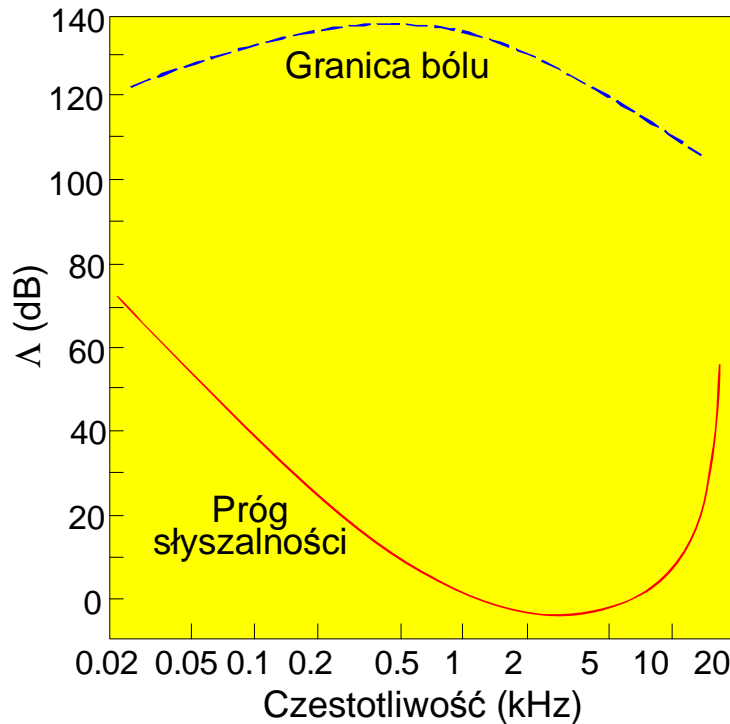
Natężenie najśłabszego słyszalnego dźwięku wynosi 10^{-12} W/m², a najsilniejszego – 1 W/m² (patrz rys. 8.25). Wobec tak szerokich granic natężeń dźwięku, które ucho ludzkie może odbierać, przyjęto natężenie najśłabszego dźwięku $I_0 = 10^{-12}$ W/m² za natężenie odniesienia i wszystkie inne wartości natężenia porównuje się z I_0 w skali logarytmicznej.

Poziomem natężenia dźwięku nazywamy wielkość

$$L = 10 \lg \frac{I}{I_0}, \quad (8.101)$$

wyrażaną w decybelach (dB). Ze wzoru (8.101) wynika, że najśłabszy słyszalny dźwięk ma zerowy poziom natężenia, dźwięk o natężeniu 100 razy większym ma poziom 20 dB, a najsilniejszy dźwięk o natężeniu 1 W/m² (typowy koncert rockowy) ma poziom natężenia 120

dB. Gdy głośność przekracza 120–135 dB, ucho nie odczuwa już wrażenia dźwięku lecz bólu. Jak widać z rys. 8.25, ucho jest najbardziej wrażliwe dla tonów o częstotliwości około 3500 Hz.



Rys. 8.25. Czulość ucha i granica bólu.

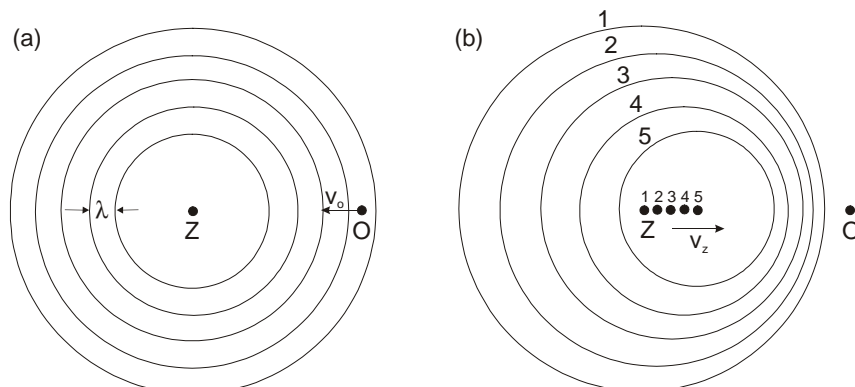
8.15. Zjawisko Dopplera

Kiedy obserwator porusza się w kierunku do źródła dźwięku to słyszy, że częstota dźwięku rośnie, gdy zaś oddala się – częstota dźwięku maleje. Ta zmiana częstoty wynikająca z wzajemnego ruchu obserwatora i źródła nazywa się zjawiskiem Dopplera. Typowym przykładem tego zjawiska jest dźwięk gwizdka zbliżającego się pociągu. Gdy pociąg mija nas, częstota (wysokość tonu) spada. W przypadku fal elektromagnetycznych, gdy źródło porusza się w kierunku obserwatora (lub obserwator porusza się w kierunku źródła), rejestrowana częstotliwość fali rośnie (przesunięcie ”w kierunku ultrafioletu”); natomiast gdy źródło i obserwator oddalają się od siebie, obserwowana częstotliwość światła maleje (przesunięcie ”w kierunku podczerwieni”).

Zjawisko Dopplera jest poglądowo przedstawione na rys. 8.26. Gdyby obserwator nie poruszał się, to w ciągu czasu t rejestrowałby ut/λ fal, gdzie u jest prędkością fazową fali, a λ – długością fali. Ponieważ jednak porusza się w kierunku źródła z prędkością v_o , to w tym samym czasie t odbiera on $v_o t$ dodatkowych fal. Z tego powodu częstotliwość f' odbierana przez obserwatora, równa liczbie fal odbieranych w jednostce czasu, wynosi

$$f' = \left(\frac{ut}{\lambda} + \frac{v_o t}{\lambda} \right) \frac{1}{t} = \frac{u + v_o}{\lambda} = \frac{(u + v) f}{u} = f \left(1 + \frac{v_o}{u} \right). \quad (8.102)$$

Zatem częstotliwość odbierana przez obserwatora jest większa od częstotliwości źródła. Gdy obserwator oddala się, to we wzorze (8.102) należy zmienić znak prędkości v_o , co powoduje, że częstotliwość ulega zmniejszeniu.



Rys. 8.26. Zjawisko Dopplera: a) źródło jest nieruchome (zbliżający się obserwator odbiera fale o większej częstotliwości), b) źródło zbliża się do obserwatora (fala ma mniejszą długość z przodu, a większą z tyłu).

Gdy źródło porusza się względem nieruchomego obserwatora, występuje efekt skrócenia długości fali w kierunku obserwatora oraz wydłużenia długości fali w kierunku przeciwnym, co pokazano na rys. 8.26b. Jeżeli częstotliwość źródła jest f , a jego prędkość v_z , to w czasie jednego okresu drgań źródło to przesuwa się o odcinek v_z/f i każda długość fali ulega skróceniu o ten odcinek. Długość fali dochodzącej do obserwatora wynosi więc

$$\lambda' = \frac{u}{f} - \frac{v_z}{f}.$$

A stąd częstotliwość odbierana przez obserwatora

$$f' = \frac{u}{\lambda'} = \frac{uf}{u - v_z} = f \frac{1}{1 - (v_z/u)}. \quad (8.103)$$

Mamy więc efekt podobny jak w przypadku ruchu obserwatora; zbliżanie się źródła powoduje wzrost częstotliwości, a oddalanie ($v_z < 0$) zmniejszenie się częstotliwości.

W przypadku gdy porusza się zarówno źródło dźwięku jak i obserwator, to należy traktować, że ruch źródła odbywa się niezależnie od ruchu obserwatora.

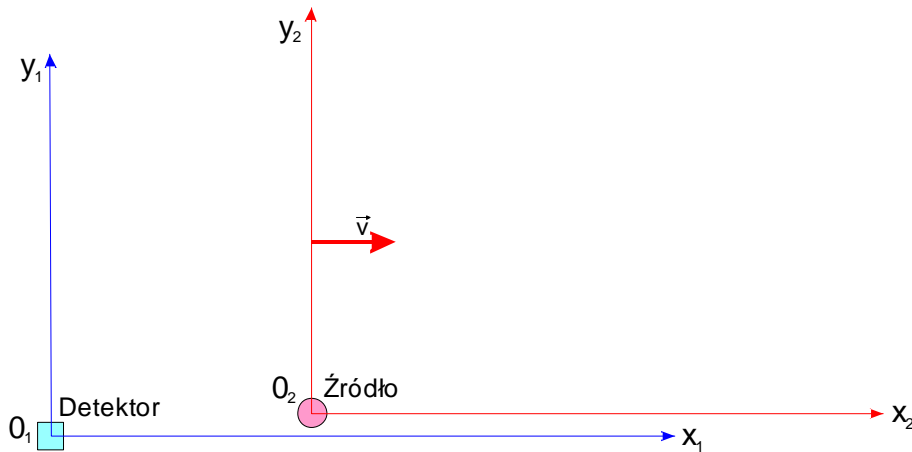
W obydwu ostatnich wzorach [(8.102) i (8.103)], wartość zmiany częstości zależy od stosunku prędkości obserwatora (lub źródła) do prędkości fazowej fali. W przypadku fal akustycznych, ten stosunek nazywa się liczbą Macha

$$m = \frac{v}{u}.$$

Gdy źródło porusza się z prędkością przewyższającą prędkość dźwięku, to wytwarza ono falę uderzeniową o właściwościach odmiennych od poprzednio omawianych fal.

W powyższych rozważaniach stosowaliśmy wzory mechaniki klasycznej. Ażeby jednak obliczyć wielkość efektu Dopplera dla fal elektromagnetycznych, musimy stosować teorię względności.

Na rys. 8.27 pokazano źródło światła i detektor oddalające się ze względną prędkością v . Załóżmy, że w obu układach mamy identyczne zegary wskazujące zero w chwili gdy układy O_1 i O_2 znajdowały się w tym samym miejscu. Przypuśćmy, że źródło emituje impuls światła, gdy jego zegar wskazuje czas T_2 . Chcemy obliczyć czas T_1 , gdy impuls dochodzi do detektora w układzie O_1 . Z rozważań przeprowadzonych w punkcie 3.4.3 wiemy, że obserwator w układzie spoczywającym O_1 widzi, że zegar biegnie szybciej niż poruszający się zegar w układzie O_2 .



Rys. 8.27. Według obserwatora w układzie O_1 źródło światła porusza się w prawo z prędkością v .

Według obserwatora O_1

$$t_1 = \frac{T_2}{\sqrt{1 - (v/c)^2}}. \quad (8.104)$$

W układzie odniesienia O_1 czas przelotu impulsu jest równy x_1/c . Jeżeli t_1 jest momentem przybycia impulsu światła do O_1 , to w układzie odniesienia O_1 mamy

$$T_1 = t_1 + (\text{czas przelotu}) = t_1 + \frac{x_1}{c}.$$

Wyliminujemy x_1 biorąc pod uwagę relację (8.104) i $x_1 = vt_1$. Wówczas

$$T_1 = \frac{T_2}{\sqrt{1 - (v/c)^2}} + \frac{vT_2}{c\sqrt{1 - (v/c)^2}},$$

czyli

$$T_1 = \frac{1 + v/c}{\sqrt{1 - (v/c)^2}} T_2.$$

Okres czasu między dwoma kolejnymi impulsami, czyli okres drgań, wynosi

$$\tau_1 = \frac{1 + v/c}{\sqrt{1 - (v/c)^2}} \tau_2, \quad (8.105)$$

gdzie τ_2 jest okresem zmierzonym przy źródle. Związek między częstotliwością a okresem jest dany wzorem $f = 1/\tau$. W związku z tym na podstawie (8.105) otrzymujemy

$$f_1 = \sqrt{\frac{1 - v/c}{1 + v/c}} f_2. \quad (8.106)$$

W przypadku gdy źródło się przybliża, znak przy v/c będzie odwrócony

$$f_1 = \sqrt{\frac{1 + v/c}{1 - v/c}} f_2. \quad (8.107)$$

Przy małych prędkościach względnych ($v \ll c$), po rozwinięciu dwóch ostatnich zależności w szereg, możemy pominąć $(v/c)^2$ oraz człony o wyższych potęgach. Otrzymamy wtedy

$$f_1 = \left(1 \pm \frac{v}{c}\right) f_2, \quad (8.108)$$

przy czym znak + odnosi się do $v < 0$, zaś znak – do $v > 0$.

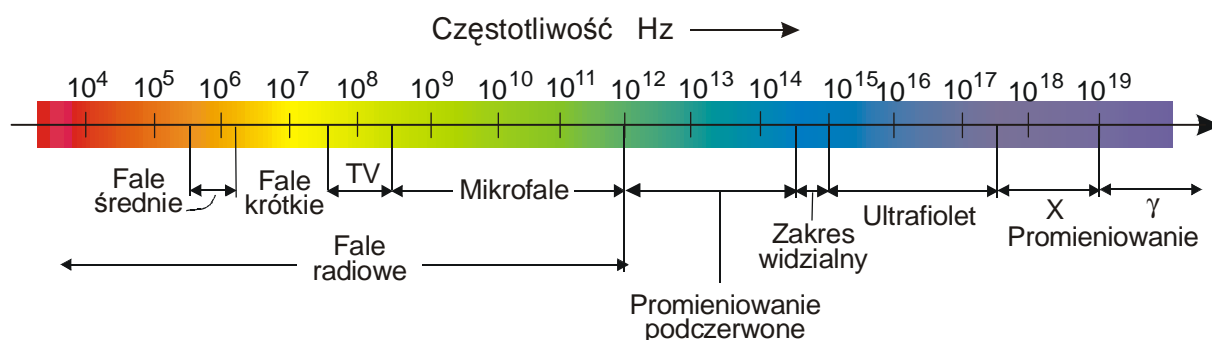
Zjawisko Dopplera dla fal elektromagnetycznych ma liczne praktyczne zastosowania, np. w fizyce atomowej, w astronomii do określenia prędkości odległych świecących ciał niebieskich, w radiolokacji do pomiaru prędkości i odległości ruchomych obiektów.

ROZDZIAŁ 9

FALE ELEKTROMAGNETYCZNE

W punkcie 7.6 przytoczone są równania Maxwella dla próżni i dla dowolnego ośrodka pod warunkiem, że wyrażenia na ρ i \vec{j} zawierają ładunki wewnętrzne i prądy. Równania te mają jednoznaczne rozwiązania dla \vec{E} i \vec{B} przy danym rozkładzie ładunku i prądu. Z prostej analizy równań Maxwella wynika, że pola elektryczne i magnetyczne mogą istnieć także, gdy źródła będą wyłączone. Ładunki w stanie spoczynku i stałe prądy tworzą stałe pola (pole \vec{E} opisywane jest prawem Coulomba, a pole \vec{B} – prawem Ampere’a). Prąd zmienny lub ładunek poruszający się z pewnym przyspieszeniem powodują pojawienie się zmiennego pola magnetycznego; inaczej mówiąc $\partial\vec{B}/dt \neq 0$. W tym przypadku zgodnie z równaniem (7.3) pole elektryczne powstaje nawet wtedy, gdy wszędzie $\rho = 0$. Przy tym pochodna $\partial\vec{E}/dt \neq 0$ i na skutek tego zgodnie z równaniem (7.16), pojawia się pole \vec{B} nawet po wyłączeniu źródła prądu. Naturalnie, że wkład \vec{B} jest taki, że $\partial\vec{B}/dt \neq 0$ co powoduje dodatkowy wkład w pole \vec{E} , itd.

Maxwell nie tylko opisał wszystkie zjawiska elektryczne za pomocą czterech prostych równań, ale przewidział konsekwencje tych równań, których poprzednio nie wiązano z elektrycznością. W 1864 roku udowodnił, że ładunek poruszający się z przyspieszeniem emituje pola elektryczne i magnetyczne propagujące się z prędkością $c = \sqrt{\epsilon_0 \mu_0}$. Te wypromieniowane pola elektryczne i magnetyczne są wzajemnie prostopadłe, a także prostopadłe do kierunku propagacji fali. Jeżeli ładunek wykonuje drgania, to częstotliwość fali jest zgodna z częstotliwością drgań. Maxwell przewidział, że światło stanowi fale elektromagnetyczne o zakresie częstotliwości $(4-7) \times 10^{14}$ Hz i że istnieją fale elektromagnetyczne o dużo niższych i dużo wyższych częstotliwościach. Na rys. 9.1 przytoczono widmo fal elektromagnetycznych.



Rys. 9.1. Widmo fal elektromagnetycznych.

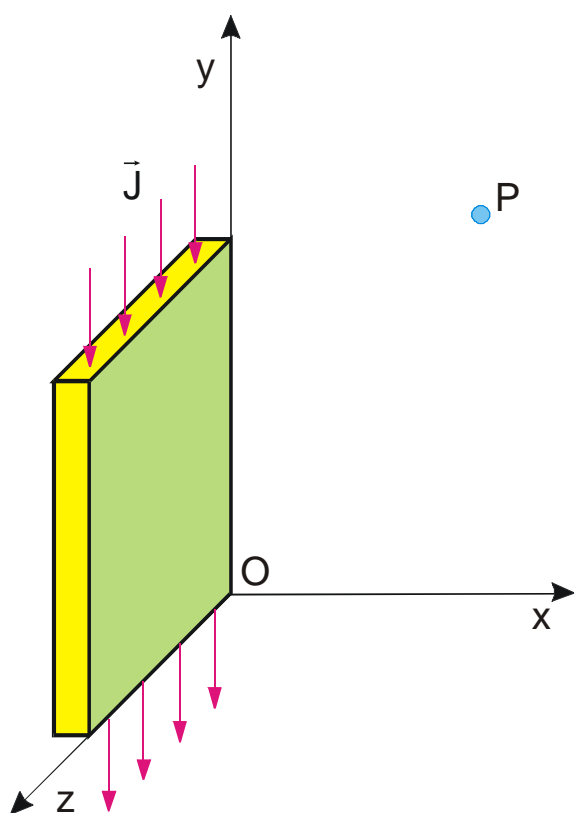
Maxwell nie tylko odkrył wielką tajemnicę natury światła, lecz również przewidział, że drgania ładunku w obwodzie rezonansowym prowadzą do emisji fal elektromagnetycznych.

Wobec tego przewidział możliwość stosowania łączności radiowej. Bez wątpienia, Maxwellowi udało się dokonać tego, czego dokonał Newton w teorii powszechnego ciężenia. Jednakże znaczenie pracy Maxwella jest większe, ponieważ w większości zjawisk fizycznych występują oddziaływania elektromagnetyczne, a nie grawitacyjne.

W rozdziale tym zapoznamy się z podstawowymi właściwościami fal elektromagnetycznych.

9.1. Równanie różniczkowe fali elektromagnetycznej

Rozważmy prąd powierzchniowy J płynący w nieskończonej płaszczyźnie Oyz w ujemnym kierunku osi y (rys. 9.2). Wielkość J to prąd powierzchniowy przypadający na jednostkę długości wzdłuż osi z .



Rys. 9.2. Prostokątny element nieskończonej powierzchni z prądem powierzchniowym J .

lub

$$2Bb = \mu_0 Jb.$$

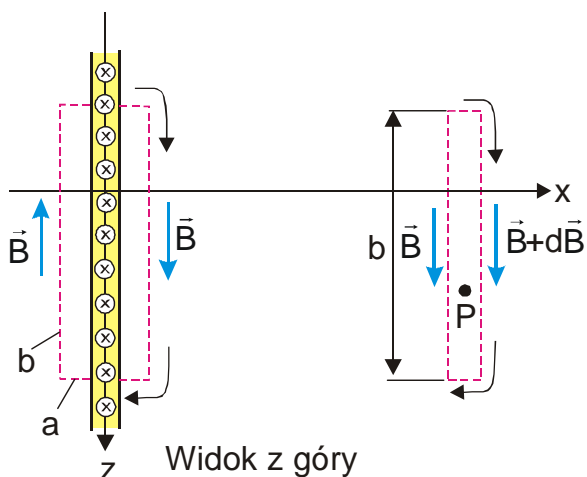
Stąd znajdujemy

$$B = \frac{\mu_0 J}{2} \quad (9.1)$$

w pobliżu płaskiego prądu.

Podobnie jak w przypadku prądu stałego, pole magnetyczne w pobliżu płaszczyzny z prądem zmiennym można obliczyć całkując po konturze prostokątnym obejmującym prąd, jak to pokazano na rys. 9.3. Niech a oznacza szerokość, a b wysokość prostokąta. Interesuje nas pole \vec{B} w odległości $a/2$ od powierzchni. Jeżeli a dąży do zera, to do zera dąży powierzchnia prostokąta; wówczas w równaniu (7.16) można zaniedbać człon $\int \left(\frac{\partial \vec{E}}{\partial t}\right) d\vec{S}$. Ponieważ prąd J skierowany jest za płaszczyznę rysunku, kontur obchodzimy zgodnie z kierunkiem wskazówek zegara. Wówczas równanie (7.16) napiszemy w postaci

$$\oint \vec{B} \cdot d\vec{s} = \mu_0 Jb,$$



Rys. 9.3. Widok z góry elementu prądu przedstawionego na rys. 9.2. Całki krzywoliniowe liczone są w kierunku ruchu wskazówek zegara wokół prądu i wokół punktu P.

W tym przypadku równanie (7.16) przyjmie postać

$$\oint_C \vec{B} \cdot d\vec{s} = 0 + \frac{1}{c^2} \int \frac{\partial \vec{E}}{\partial t} d\vec{s},$$

lub

$$(B_z + dB_z)b - B_z b = -\frac{1}{c^2} \frac{\partial E_y}{\partial t} (b dx),$$

gdzie $B = B_z$ na lewej stronie oraz $B = B_z + dB_z$ na prawej stronie prostokątnego konturu (górna i dolna krawędź nie dają wkładu do $\int \vec{B} \cdot d\vec{s}$). Wobec tego

$$dB_z = -\frac{1}{c^2} \frac{\partial E_y}{\partial t} dx,$$

$$\left(\frac{dB_z}{dx} \right)_{t=const} = -\frac{1}{c^2} \frac{\partial E_y}{\partial t} dx,$$

$$\frac{\partial B_z}{\partial x} = -\frac{1}{c^2} \frac{\partial E_y}{\partial t}. \quad (9.2)$$

Lewa strona tego równania zawiera pochodną cząstkową, ponieważ czas t jest traktowany jako stały. Na rys. 9.3 pokazano sytuację odpowiadającą tej chwili czasu.

Ostatnie wyrażenie jest również słuszne dla prądu stałego J . Jednakże w naszym przypadku prąd J zmienia się w czasie, a otrzymany wynik słuszny jest jedynie w pobliżu źródła.

Ażeby znaleźć pole magnetyczne w punkcie P na rys. 9.3, posłużymy się prostokątnym konturem całkowania wokół punktu P pokazanym na tym rysunku. Jeżeli całkujemy po konturze w kierunku zgodnym z ruchem wskazówek zegara, to wektor $d\vec{s}$ będzie skierowany za płaszczyznę rysunku w ujemnym kierunku osi y . Wówczas

$$\vec{E} \cdot d\vec{s} = -E_y dS = -E_y b dx.$$

Z uogólnionego prawa Faradaya [wzór (7.3)] można otrzymać jeszcze jeden związek między polami B i E . Zgodnie z rys. 9.4, całkujemy w kierunku przeciwnym do ruchu wskazówek zegara po prostokątnej konturze wokół punktu P w płaszczyźnie Oxy

$$\oint_C \vec{E} \cdot d\vec{s} = - \int \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S},$$

$$(E_y + dE_y)h - E_y h = - \frac{\partial B_z}{\partial t} (h dx),$$

czyli

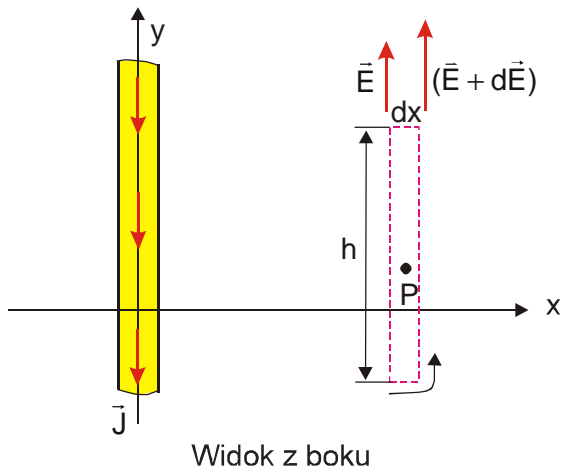
$$dE_y = - \frac{\partial B_z}{\partial t} dx,$$

a dalej

$$\left(\frac{dE_y}{dx} \right)_{t=const} = - \frac{\partial B_z}{\partial t},$$

i ostatecznie

$$\frac{\partial E_y}{\partial x} = - \frac{\partial B_z}{\partial t}. \quad (9.3)$$



Widok z boku

Chcemy teraz obliczyć pole B w punkcie P . Mamy dwa równania, (9.2) i (9.3), z dwiema niewiadomymi B_z i E_y . Różniczkując pierwsze z nich po x , a drugie po t , można wyłączyć E_y

$$\frac{\partial}{\partial x} \left(\frac{B_z}{\partial x} \right) = \frac{\partial}{\partial x} \left(- \frac{1}{c^2} \frac{\partial E_y}{\partial t} \right),$$

$$\frac{\partial^2 B_z}{\partial x^2} = - \frac{1}{c^2} \frac{\partial^2 E_y}{\partial x \partial t}. \quad (9.4)$$

Rys. 9.4. Widok z boku na element płaskiego prądu przedstawionego na rys. 9.2.

Podobnie

$$\frac{\partial}{\partial t} \left(\frac{\partial E_y}{\partial x} \right) = \frac{\partial}{\partial t} \left(\frac{\partial B_z}{\partial t} \right)$$

$$\frac{\partial^2 E_y}{\partial x \partial t} = - \frac{\partial^2 B_z}{\partial t^2}$$

Podstawiając to wyrażenie w prawą stronę równania (9.4), mamy

$$\frac{\partial^2 B_z}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 B_z}{\partial t^2}. \quad (9.5)$$

Równanie (9.5) to słynne równanie różniczkowe równanie falowe Maxwella. Rozwiązanie tego równania przedstawia falę biegnącą propagującą się z prędkością c . Równanie (9.3) zawiera uzupełniającą informację wskazującą, że wielkość pola elektrycznego jest równa $E = cB$ i że pola \vec{E} i \vec{B} są wzajemnie prostopadłe.

9.2. Promieniowanie płaskiego prądu

Założmy, że prąd powierzchniowy (rys. 9.2) ma postać

$$J = J_o \cos \omega t,$$

przy czym prąd J_o płynie w kierunku przeciwnym do osi y . Taki prąd sinusoidalny jest łatwo wzbudzić stosując układ elektryczny opisany w punkcie 8.1.2. Określimy $B(x,t)$ dla dowolnej wartości x i t . Przy małych wartościach x rozwiązanie określone jest wyrażeniem (9.1)

$$B_z(x,t) = \frac{\mu_o}{2} J_o \cos \omega t.$$

W przypadku dużych wartości x spróbujemy odgadnąć rozwiązanie, które byłoby zgodne z powyższym rozwiązaniem dla małych wartości x i udowodnimy, że spełnia równanie falowe (9.5). Jednoznaczne rozwiązanie dla rozważanego źródła prądu ma postać

$$B_z(x,t) = \frac{\mu_o}{2} J_o \cos \omega \left(t - \frac{x}{c} \right). \quad (9.6)$$

Podstawiając to rozwiązanie do lewej strony równania (9.5) mamy

$$\frac{\partial^2 B_z}{\partial x^2} = -\left(\frac{\omega}{c}\right)^2 \frac{\mu_o}{2} J_o \cos \omega \left(t - \frac{x}{c} \right) = -\frac{\omega^2}{c^2} B_z,$$

a do prawej

$$\frac{1}{c^2} \frac{\partial^2 B_z}{\partial t^2} = \frac{1}{c^2} (-\omega^2) \frac{\mu_o}{2} J_o \cos \omega \left(t - \frac{x}{c} \right) = -\frac{\omega^2}{c^2} B_z.$$

Widzimy więc, że przy wyborze rozwiązania (9.6) równanie (9.5) jest spełnione. Oprócz tego przy $x \rightarrow 0$ otrzymujemy $B_z = (\mu_o / 2) J_o \cos \omega t$. Rozwiązanie to spełnia warunek brzegowy i jest jednoznaczny rozwiązaniem problemu.

W poprzednim rozdziale wykazaliśmy, że funkcja $\cos \omega(t - x/c)$ opisuje falę monochromatyczną propagującą się wzdłuż osi x z prędkością c , przy czym długość fali

$\lambda = 2\pi c/\omega$. Teraz znając B możemy obliczyć pole E podstawiając rozwiązanie na B do równania (9.3)

$$\frac{\partial E_y}{\partial x} = -\frac{\partial}{\partial t} \left[B_o \cos \omega \left(t - \frac{x}{c} \right) \right] = \omega B_o \sin \omega \left(t - \frac{x}{c} \right),$$

stąd

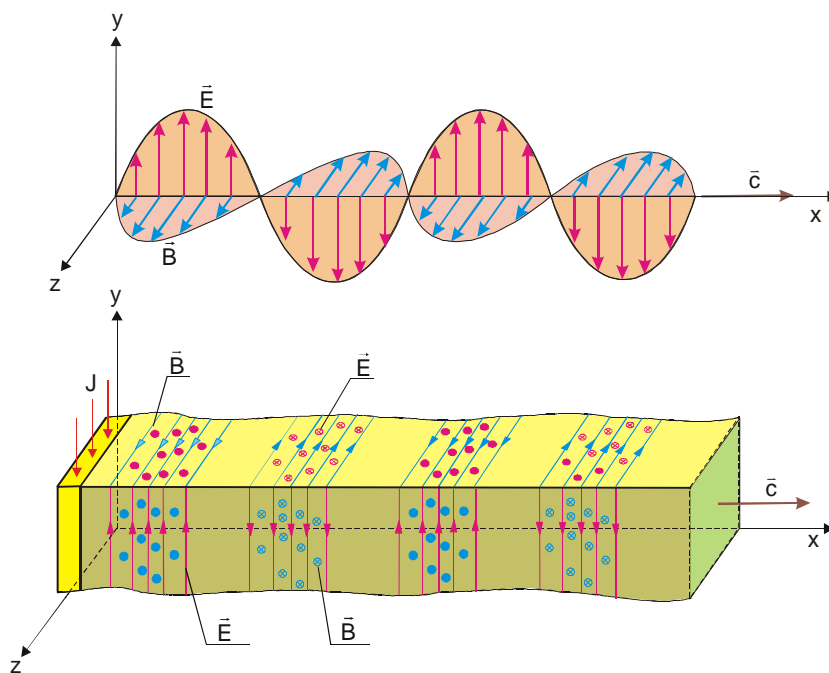
$$E_y = \omega B_o \int \sin \omega \left(t - \frac{x}{c} \right) dx = c B_o \cos \omega \left(t - \frac{x}{c} \right) + const.$$

Stała całkowania jest równa zero, ponieważ ładunki tworzące stałe pole elektryczne nie występują. Tak więc

$$E_y = c B_z = \frac{c\mu_o}{2} J_o \cos \omega \left(t - \frac{x}{c} \right) \quad \text{pole promieniowania} \quad (9.7)$$

Za dodatni kierunek prądu J_o przyjęto kierunek przeciwny do kierunku osi y . Dlatego dodatnie znaki wielkości E_y i B_z oznaczają, że w pobliżu źródła pole E_y skierowane jest przeciwnie do prądu J . Wygodnie jest to zapamiętać w następujący sposób. Popatrzmy co zachodzi na dolnej krawędzi elementu płaskiego prądu. Dodatnie ładunki będą gromadzić się na dolnej krawędzi, a ujemne na górnej. Linie sił pola skierowane są z dołu do góry, tj. przeciwnie do kierunku prądu.

Wykazaliśmy, że $E = cB$ a także, że pola elektryczne i magnetyczne są wzajemnie prostopadłe. Otrzymane rozwiązania dla \vec{E} i \vec{B} są prawdziwe dla dowolnego konturu całkowania i są jednoznaczne. Na rys. 9.5 pokazano linie sił pól elektrycznego i magnetycznego dla przypadku płaskiego sinusoidalnego prądu.

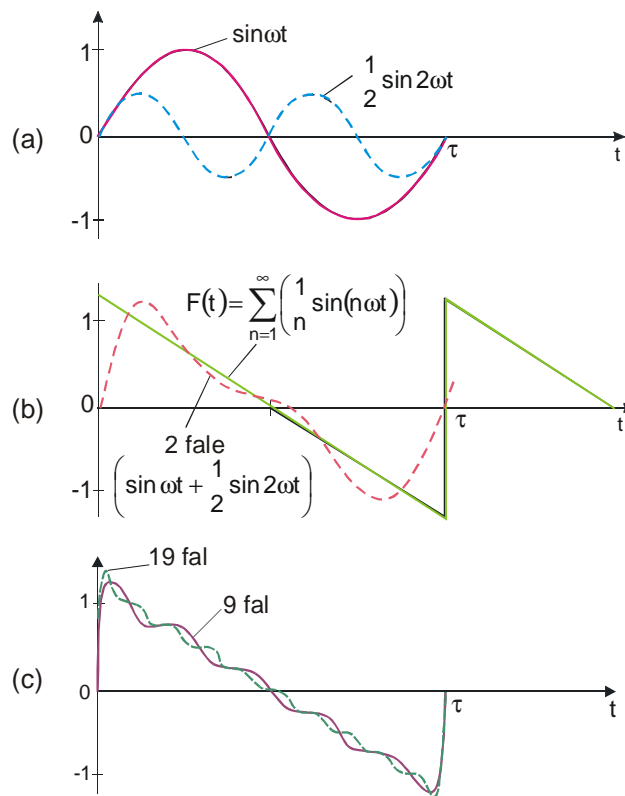


Rys. 9.5. Płaska fala monochromatyczna propagująca się w prawo z prędkością c . Fala emitowana jest przez sinusoidalny prąd J płynący w płaszczyźnie Oyz .

Rozważmy teraz zamiast prądu sinusoidalnego, prąd dowolnie zmieniający się w czasie. Taką funkcją będą opisywane zmiany czasowe pól elektrycznych i magnetycznych wytwarzanych przez ten prąd. Rozważmy przypadek, kiedy prąd powierzchniowy określany jest funkcją piłokształtną o okresie τ . Wówczas $\omega = 2\pi/\tau$. Na rys. 9.6 pokazano w jaki sposób funkcję piłokształtną można przedstawić w postaci sumy nieskończonej liczby fal sinusoidalnych

$$F(t) = \sum_{n=1}^{\infty} \left(\frac{1}{n} \sin(n\omega t) \right).$$

Wyrażenie to nazywane jest rozkładem Fouriera periodycznej funkcji $F(t)$. W ogólnym przypadku dowolną funkcję periodyczną o częstotliwości $1/\tau$ można zapisać w postaci sumy fal monochromatycznych o częstotliwościach $n(1/\tau)$, gdzie n przyjmuje liczby całkowite od 1 do ∞ . Dodając sinusoidy pozbawione odcinków prostoliniowych, można otrzymać absolutnie proste linie z ostrymi załamaniem. Do tego konieczna jest jednak nieskończona liczba sinusoid.



Rys. 9.6. Przedstawienie piłokształtnej funkcji w postaci sumy nieskończonej liczby fal sinusoidalnych: (a) dwie fale sinusoidalne; (b) wynik złożenia dwóch fal sinusoidalnych; (c) suma pierwszych dziewięciu i dziewiętnastu fal sinusoidalnych.

Dla generacji piłokształtnej fali elektromagnetycznej prąd powierzchniowy określany jest wzorem

$$J = J_o \sum_{n=1}^{\infty} \left(\frac{1}{n} \sin(n\omega t) \right),$$

gdzie $\omega = 2\pi/\tau$. Każdemu członowi tej sumy odpowiadają rozwiązania określone wyrażeniami (9.6) i (9.7). Ponieważ równania Maxwella są liniowe odnośnie E , B i J , więc pełne rozwiązanie jest równe sumie oddzielnych rozwiązań. Wynika to z zasady superpozycji. Pełne rozwiązanie zapiszemy w postaci

$$E = cB = \frac{c\mu_o}{2} J_o \sum \frac{1}{n} \left\{ \sin \left[n\omega \left(t - \frac{x}{c} \right) \right] \right\}.$$

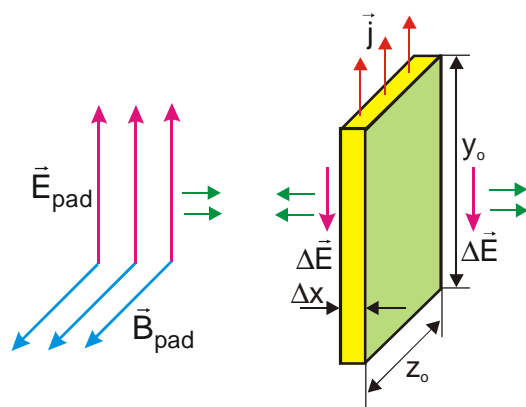
Ponieważ wszystkie fale odpowiadające pojedynczym członom sumy propagują się wzdłuż osi x z prędkością c , więc będą rozpatrywane tak, jak pokazano na rys. 9.6. Wobec tego pole w dowolnym punkcie przestrzeni powtarza piłokształtną zależność prądu od czasu i jedynie opóźnia się względem źródła o czas x/c .

9.3. Oddziaływanie promieniowania z materią

W tym punkcie zajmujemy się oddziaływaniem fali elektromagnetycznej z materią. Dokładniej rozważymy przypadek przewodnika i dielektryka. Okazuje się, że słaby przewodnik częściowo pochłania energię i pęd fali, co powoduje zmianę energii i pędu fali elektromagnetycznej. Dobry przewodnik odbija falę ze 100% efektywnością. Przez dielektryk taki jak gaz, fala propaguje się nie ulegając pochłanianiu; jednakże propagacja fali w tym przypadku zachodzi wolniej niż w próżni. Te pozorne paradoksy rozwiązuje się stosując podejście mikroskopowe z uwzględnieniem budowy atomowej materii.

9.3.1. Energia promieniowania

Na rys. 9.7 pokazano falę elektromagnetyczną padającą na prostokątny element płytki o nieskończonych rozmiarach. Jeżeli gęstość prądu indukowanego jest równa j , to w prostokątnym elemencie będzie płynął prąd $I = jz_o\Delta x$.



Rys. 9.7. Fala padająca \vec{E}_{pad} przemieszcza się w prawo i pada na płytkę indukując prąd, który na prawo i na lewo od płytki promieniuje własne pole $\Delta \vec{E}$.

Różnica potencjałów pomiędzy górną a dolną krawędzią wynosi $V = Ey_o$. Tak więc energia tracona w jednostce czasu na ciepło Joule'a wynosi

$$\frac{dW}{dt} = IV = (jz_o\Delta x)Ey_o = jE(y_o z_o \Delta x),$$

gdzie $y_o z_o \Delta x$ oznacza objętość elementu płytki. Wykazaliśmy, że moc tracona w jednostce objętości przewodnika wynosi jE .

Jeżeli na przewodzącą płytkę pada płaska fala monochromatyczna, to przy tym nie tylko wydziela się ciepło z szybkością jE w jednostce objętości, lecz zgodnie z formułą (9.7) indukowany prąd j także promieniuje falę elektromagnetyczną. Na rys. 9.7 pokazano pole promieniowania cienkiej płytki o grubości Δx . Pole padającej fali płaskiej będziemy oznaczać \vec{E}_{pad} . Odpowiadający mu prąd powierzchniowy $J = j\Delta x$. Tak więc zgodnie z (9.7) możemy napisać

$$\Delta E = -\frac{c\mu_o}{2} j\Delta x. \quad (9.8)$$

Znak minus wskazuje, że wewnątrz płytki pole $\Delta\vec{E}$ skierowane jest przeciwnie do prądu \vec{j} . Niech ΔP_S oznacza straty mocy na jednostkową powierzchnię. W przypadku cienkiej płytki o grubości Δx

$$\Delta P_S = \frac{1}{y_o z_o} \frac{dW}{dt} = jE\Delta x. \quad (9.9)$$

Podstawiając tutaj w miejsce wielkości Δx wyrażenie otrzymane z (9.8), mamy

$$\Delta P_S = -\frac{2}{c\mu_o} E\Delta E. \quad (9.10)$$

Rozpatrzmy teraz stos takich cienkich płytek. Jeżeli stos płytek jest nieskończenie gruby, to pole \vec{E}_{pad} zostanie całkowicie pochłonięte (co ma miejsce w przypadku słabego przewodnika) i całkowitą moc promieniowania z jednostki powierzchni można znaleźć całkując wyrażenie (9.10)

$$P_S = -\frac{2}{c\mu_o} \int_{E_{pad}}^0 E dE = \frac{1}{c\mu_o} E_{pad}^2 = \frac{1}{\mu_o} E_{pad} B_{pad}.$$

Moc promieniowana przez jednostkową powierzchnię charakteryzowana jest wektorem Poytinga. Jego wartość oznaczana jest przez P_S . Ponieważ kierunek strumienia energii określany jest iloczynem wektorowym $\vec{E} \times \vec{B}$, to dla \vec{P}_S mamy

$$\vec{P}_S = \frac{1}{\mu_o} \vec{E} \times \vec{B}. \quad (9.11)$$

Sprawdzimy teraz czy otrzymany wzór nie przeczy otrzymanemu wcześniej wyrażeniu dla energii przypadającej na jednostkę objętości pola (wzór (7.14)). Rozpatrzmy falę płaską padającą na powierzchnię A . Zgodnie z określeniem P_S , strumień energii w czasie dt wynosi

$$dW = P_S A dt,$$

gdzie dW oznacza energię w objętości $dV = A dx$. Ponieważ $dt = dx/c$, stąd

$$dW = P_S A \frac{dx}{c} = \frac{P_S}{c} dV ,$$

czyli

$$\frac{dW}{dV} = \frac{P_S}{c} .$$

Stosując teraz wyrażenie (9.11) otrzymamy

$$\frac{dW}{dV} = \frac{1}{c\mu_o} EB .$$

Zamieniając E na cB

$$\frac{dW}{dV} = \frac{B^2}{2\mu_o} + \frac{B^2}{2\mu_o} ,$$

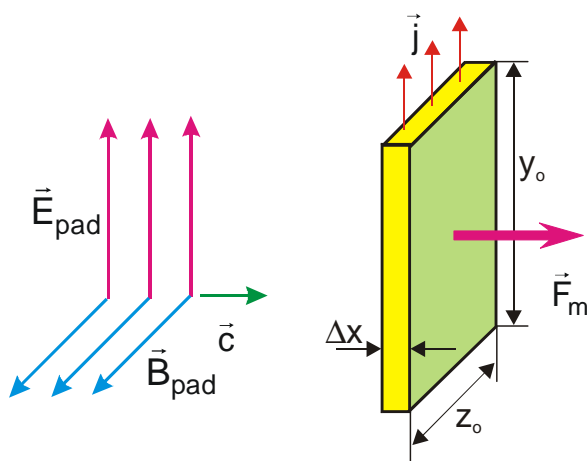
czyli

$$w = \frac{dW}{dV} = \frac{1}{2} \left(\epsilon_o E^2 + \frac{B^2}{\mu_o} \right) .$$

Otrzymujemy więc wynik zgodny z (7.14). Pierwsza składowa z prawej strony równania oznacza gęstość energii pola elektrycznego, druga zaś gęstość energii pola magnetycznego.

9.3.2. Pęd promieniowania

Wykażemy teraz, że płaska fala z rys. 9.7 przekazuje płytce o grubości Δx nie tylko energię, ale i pęd. Rozważmy prostokątny element nieskończonej płytki, którego powierzchnia wynosi $y_o z_o$ (rys. 9.8). Ponieważ $jEdt$ jest ilością ciepła Joule'a wydzielającego się w jednostce objętości w czasie dt , to ilość ciepła wydzielającego się w elemencie płytki o objętości $y_o z_o \Delta x$ wynosi



Rys. 9.8. Padająca fala wywołuje w płytce prąd $I = jz_o \Delta x$. Na ten prąd działa siła magnetyczna $\vec{F}_m = I\vec{y}_o \times \vec{B}$.

$$dW = (jEdt)(y_o z_o \Delta x) .$$

Zamieniając E na cB

$$dW = cjz_o \Delta x y_o Bdt .$$

Ponieważ prąd płynący przez rozważany element płytki wynosi $I = j(z_o \Delta x)$, więc

$$dW = cIy_o Bdt .$$

Na element prądu długości y_o , prostopadły do padającego pola magnetycznego, działa siła $\vec{F}_m = I\vec{y}_o \times \vec{B}$ w kierunku $\vec{E} \times \vec{B}$ zgodnie z kierunkiem fali padającej. Zamieniając $Iy_o B$ na F_m otrzymujemy

$$dW = cF_m dt .$$

Pęd przekazywany elementowi płytki $dp = F_m dt$, czyli $dW = cdp$, a stąd

$$dp = \frac{1}{c} dW . \quad (9.12)$$

Tak jak poprzednio, całkując po grubości płytki x otrzymujemy $p = W/c$. Wobec tego pęd przekazywany płytce przez padającą falę równy jest wielkości l/c pomnożonej przez energię rozproszoną w płytce.

Pole promieniowania posiada więc nie tylko energię ale i pęd. W dowolnym elemencie objętości pola promieniowania dV zawarta energia wynosi

$$dW = \varepsilon_0 E^2 dV ,$$

a jego pęd jest równy energii podzielonej przez c . Inaczej mówiąc, element objętości dV charakteryzuje się wektorem pędu (z uwzględnieniem związku $dW = P_S dV/c$)

$$d\vec{p} = \frac{1}{c} \left(\frac{\vec{P}_S}{c} dV \right) . \quad (9.13)$$

Energię promieniowania łatwo odczuć umieszczając rękę w strumieniu światła. Znaczna część energii przekształca się w ciepło. Jednakże pomiar pędu strumienia świetlnego jest utrudniony na skutek tego, że wartość l/c jest mała.

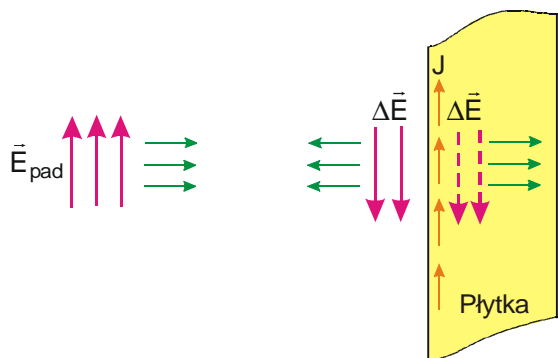
Więcej informacji o istocie promieniowania elektromagnetycznego uzyskamy w trakcie studiowania teorii i właściwości fotonu. Dowiemy się, że promieniowanie stanowią cząstki zwane fotonami. Każdy foton posiada energię $h\nu$ i pęd $h\nu/c$, gdzie h jest bardzo małą liczbą zwaną stałą Plancka.

9.3.3. Odbicie promieniowania od przewodnika

W przypadku przewodnika o wysokiej konduktywności σ , fala elektromagnetyczna nie jest pochłonięta całkowicie lecz częściowo jest odbijana. O tym łatwo się przekonać rozważając skrajny przypadek $\sigma = \infty$ (nadprzewodnik). Pole elektryczne wewnątrz nadprzewodnika zawsze przyjmuje wartość zerową (w przeciwnym przypadku nieograniczenie wzrósłby prąd). Tak więc indukowany prąd powierzchniowy okazuje się takim, że pole promieniowania $\Delta E = -E_{pad}$. Wówczas wewnątrz płytki wypadkowe pole

$$\vec{E} = \vec{E}_{pad} + \Delta \vec{E} = 0 .$$

Na lewo od płytki pole nie będzie przyjmować zerowej wartości, ponieważ jego istnienie uwarunkowane jest dwoma falami monochromatycznymi o jednakowym natężeniu, biegnącymi w przeciwnych kierunkach. W tych warunkach na lewo od płytki (rys. 9.9) powstaje fala stojąca.

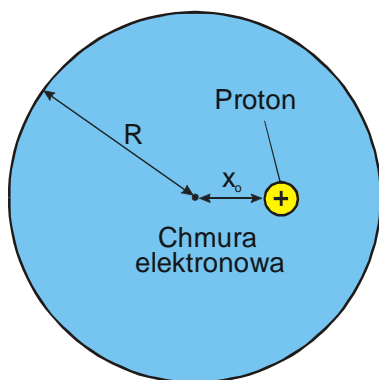


Rys. 9.9. Padająca fala wywołuje w płytce nadprzewodzącej prąd J , który promieniuje pole $\Delta\vec{E}$ równe co do wartości z \vec{E}_{pad} .

Dotychczas rozważaliśmy oddziaływanie promieniowania z materią zawierającą swobodne elektrony. Promieniowanie wywołuje pojawienie się prądu elektrycznego i efekty z nim związane. W następnym punkcie omówimy substancje, w których brak jest elektronów swobodnych. Takie substancje nazywamy dielektrykami lub izolatorami.

9.3.4. Oddziaływanie promieniowania z dielektrykiem

Zewnętrzne elektrony atomów dielektryka ulegają przemieszczaniu pod wpływem zewnętrznego pola elektrycznego. Postaramy się określić wartość takiego przemieszczenia. Zastosujemy model atomu podobny do tego, który przyjęty jest we współczesnej teorii kwantowej. Zgodnie z tym modelem, zewnętrzny elektron traktowany jest w postaci kulistego obłoku o promieniu R . Przyjmijmy, że gęstość ładunku jest stała. W stanie normalnym proton o ładunku $+e$ znajduje się w środku chmury elektronowej.



Rys. 9.10. Elektron traktowany jako jednorodna naładowana kula przesunięty na odległość x_0 względem protonu.

Założmy, że proton jest przesunięty względem środka chmury elektronowej na niewielką odległość x_0 jak to pokazano na rys. 9.10. Jeżeli teraz elektron i proton pozostawić swobodnymi, to zaczną wykonywać drgania wokół położenia równowagi z amplitudą x_0 . Siła działająca na proton zgodnie z równaniem (4.16) wynosi

$$F = eE = e \left(\frac{-e}{4\pi \epsilon_0 \epsilon_r R^3} x \right) = -k_o \frac{e^2}{R^3} x,$$

gdzie $k_o = 1/4\pi\epsilon_0\epsilon_r$. Zgodnie z III prawem Newtona, identyczna siła działa na elektron

$$m_e \frac{d^2x}{dt^2} = -\frac{k_o e^2}{R^3} x.$$

W zasadzie powinniśmy stosować masę zredukowaną $\mu = M_p m_e / (M_p + m_e)$, a nie masę elektronu m_e . Jednakże w naszym przypadku $\mu \approx m_e$. Z powyższego równania otrzymamy

$$\frac{d^2x}{dt^2} = -\left(\frac{k_oe^2}{m_eR^3}\right)x = -\omega_o^2x,$$

gdzie

$$\omega_o = \sqrt{\frac{k_oe^2}{m_eR^3}}.$$

Z rozważań tych wynika, że siła działająca na zewnętrzną chmurę elektronową wynosi

$$F_{atom} = -m\omega_o^2x, \quad (9.14)$$

gdzie ω_o oznacza częstość kołową drgań własnych elektronu atomowego.

Jeżeli na chmurę elektronową działa pole E_{pad} padającej fali, to wypadkową siłę zapiszemy w postaci

$$F_{wyp} = F_{atom} + (-e)E_{pad},$$

stąd otrzymamy następujące równanie

$$m\frac{d^2y}{dt^2} = -m\omega_o^2y - eE_{pad}.$$

Dla oznaczenia przesunięcia stosujemy wielkość y zamiast x , ponieważ będziemy rozpatrywać przypadek kiedy \vec{E}_{pad} skierowane jest wzdłuż osi y . Falę padającą w odległości x od źródła zapisujemy jako $E_{pad} = E_o \cos \omega(t - x/c)$. Wówczas powyższe równanie przyjmuje postać

$$\frac{d^2y}{dt^2} = -\omega_o^2y - \frac{eE_o}{m} \cos \omega\left(t - \frac{x}{c}\right).$$

Rozwiązaniem tego równania różniczkowego jest

$$y = -\frac{eE_o}{m(\omega_o^2 - \omega^2)} \cos \omega\left(t - \frac{x}{c}\right). \quad (9.15)$$

Można to sprawdzić podstawiając to rozwiązanie do wyjściowego równania falowego.

W powyższy sposób wykazaliśmy jak oddziałuje pojedynczy atom z falą elektromagnetyczną. Teraz rozważymy oddziaływanie z dużą ilością takich atomów zawartych w płycie ciała stałego lub w warstwie gazu.

9.3.5. Współczynnik załamania. Dyspersja

Niech na płytkę o grubości Δx składającą się z atomów poprzednio rozważanych, pada fala płaska. Zgodnie ze wzorem (9.15) pole elektryczne fali padającej E_{pad} wymusi drganie harmoniczne elektronów. Dowolny drgający elektron powinien promieniować falę elektromagnetyczną. Podobnie jak w przypadku płytki z przewodnika, powstaje fala odbita i przechodząca, lecz teraz nie ma strat na ciepło Joule'a. Cała energia zachowuje się w postaci promieniowania elektromagnetycznego, a wobec tego płytka okazuje się dla promieniowania przezroczystą. Ponadto okazuje się, że fala elektromagnetyczna propaguje się wewnątrz płytki z prędkością $u < c$. Stosunek $c/u = n$ nazywamy współczynnikiem załamania. Dla większości ciał stałych, współczynnik załamania równy jest w przybliżeniu 1,5, oznacza to, że prędkość światła w tych ciałach jest niższa o około 33%.

Jak wyjaśnić, że fala elektromagnetyczna propaguje się z prędkością $u < c$?

Z równań (9.5) i (9.7) wynika, że pola E i B opisywane są równaniami falowymi, w które wchodzi prędkość fali c . Jest tak niewątpliwie z mikroskopowego punktu widzenia. Jednakże pole wewnątrz płytki nie jest zgodne z polem promieniowania pojedynczej cząstki (lub płaskiego prądu generującego falę padającą). Pole wewnątrz płytki jest superpozycją pola fali padającej i pól promieniowania wszystkich elektronów. Każde z pól z osobna będzie propagować się z prędkością $u = c$, lecz pole wypadkowe może propagować się tak, jakby jego prędkość była mniejsza. Stosując wzór (9.15) wykażemy, że pole promieniowania każdego elektronu atomowego opóźnia się w fazie o $\pi/2$ względem pola fali padającej, która wywołuje ruch elektronów. Można oczekiwać, że w tym przypadku emitowana wypadkowa fala będzie opóźniona w fazie względem fali padającej. Okazuje się, że prędkość wypadkowej fali jest mniejsza od $u = c$, chociaż pojedyncze fale propagują się z prędkością $u = c$.

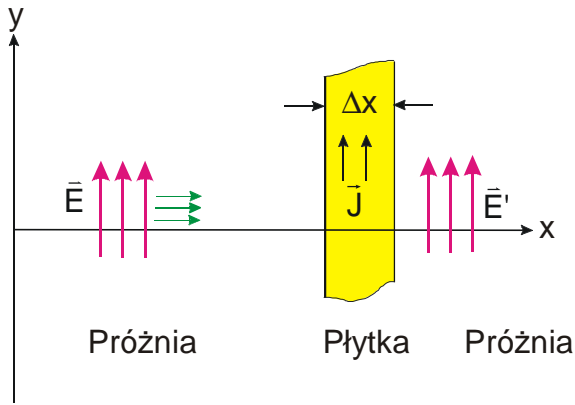
Wyprowadzimy wzór dla współczynnika załamania korzystając z rys. 9.11 i postępując w następujący sposób:

1. *Zadamy pole elektryczne fali padającej.*
2. *Obliczymy prędkość elektronów atomowych w płytce uwarunkowaną wpływem pola elektrycznego fali padającej.*
3. *Mając tę prędkość lub gęstość pola elektrycznego, obliczymy emitowane przez elektrony wtórne promieniowanie.*
4. *Zsumujemy falę padającą i fale wtórne w celu otrzymania wypadkowej emitowanej fali.*
5. *Znajdziemy związek pomiędzy fazą fali emitowanej a współczynnikiem załamania.*

Ad.1. Fala padająca

Pole elektryczne fali padającej ma postać

$$E_{pad} = E_o \cos \omega \left(t - \frac{x}{c} \right).$$



Rys. 9.11 Fala padająca \vec{E}_{pad} wytwarza w płytce prąd o gęstości j , którego promieniowanie wnosi wkład w wypadkową falę \vec{E}' .

Ad.2. Prędkość elektronu.

Różniczkując wzór (9.15) otrzymujemy wyrażenie na prędkość słabo związanych elektronów:

$$v_y = \frac{dy}{dt} = \frac{eE_o\omega}{m(\omega_o^2 - \omega^2)} \sin \omega \left(t - \frac{x}{c} \right),$$

gdzie ω_o jest częstością drgań własnych elektronów. Gęstość prądu w płytce wynosi $j = N(-e)v_y$, gdzie N jest liczbą drgających elektronów w jednostce objętości. Stosując powyżej otrzymane wyrażenie na v_y , mamy

$$j = -\frac{Ne^2\omega E_o}{m(\omega_o^2 - \omega^2)} \sin \omega \left(t - \frac{x}{c} \right).$$

Ad.3. Promieniowanie emitowane przez elektrony atomowe

Pole promieniowania emitowane przez elektrony płytki, zgodnie z (9.7), zapiszemy w postaci

$$\Delta E = -\frac{c\mu_o}{2} j \Delta x,$$

gdzie znak "–" wskazuje, że prąd j i wytwarzane przezzeń pole promieniowania charakteryzują się przeciwnymi kierunkami. Podstawiając wyrażenie dla j mamy

$$\Delta E = -\frac{c\mu_o}{2} \left[-\frac{Ne^2\omega E_o}{m(\omega_o^2 - \omega^2)} \sin \omega \left(t - \frac{x}{c} \right) \right] \Delta x.$$

Przepiszemy to wyrażenie w postaci

$$\Delta E = \Delta E_o \cos \left(\omega t - kx - \frac{\pi}{2} \right), \quad (9.16)$$

gdzie $k = \omega/c$, oraz

$$\Delta E_o = \frac{c\mu_o Ne^2\omega}{2m(\omega_o^2 - \omega^2)} E_o \Delta x. \quad (9.17)$$

Ad.4. Fala wypadkowa

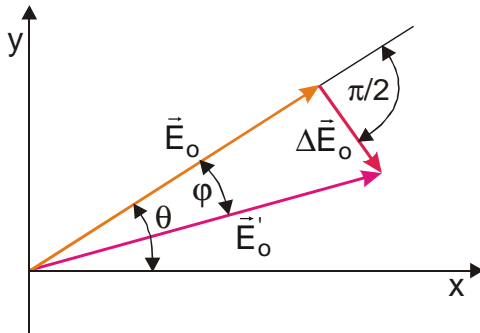
Wypadkowe pole elektryczne emitowanej fali stanowi superpozycję pola fali padającej i pola emitowanego przez elektrony atomowe

$$E' = E_{pad} + \Delta E.$$

Uwzględniając wyrażenie (9.16) otrzymamy

$$E' = E_o \cos \theta + \Delta E_o \cos\left(\theta - \frac{\pi}{2}\right),$$

gdzie $\theta = \omega(t - x/c)$. Chociaż θ rośnie z czasem, jednakże obie fale monochromatyczne zachowują stałą różnicę faz równą $\pi/2$. Stosując metodę wektorową, te dwie fale jest łatwo dodać. W danej chwili czasu t pierwsza fala przedstawia rzut wektora \vec{E}_o na oś x (rys. 9.12), druga fala rzut wektora $\Delta\vec{E}_o$ tworzącego kąt $\pi/2$ z pierwszym wektorem. Z rys. 9.12 można zauważyć, że wektor wypadkowy \vec{E}'_o przesunięty jest w fazie względem fali padającej o kąt $\varphi = \Delta E_o / E_o$ (przyjęto przy tym małe kąty zakładając, że $\Delta E_o / E_o \ll 1$).



Rys. 9.12. Wykres fazowy dla przypadku dodawania dwóch fal monochromatycznych $E_o \cos \theta$ i $\Delta E_o \cos(\theta - \pi/2)$.

Ad.5. Związek pomiędzy przesunięciem fazowym a współczynnikiem załamania

Fala padająca przechodzi przez płytkę w ciągu czasu $t = \Delta x / c$, podczas gdy fala rozchodząca się z prędkością $u = c/n$ potrzebuje więcej czasu $t' = n(\Delta x/c)$. Czoło wypadkowej fali przy przejściu przez płytkę opóźni się o $\Delta t' = (n-1)\Delta x/c$. Odpowiada to przesunięciu fazowemu

$$\varphi = \omega \Delta t' = \omega \left[(n-1) \frac{\Delta x}{c} \right].$$

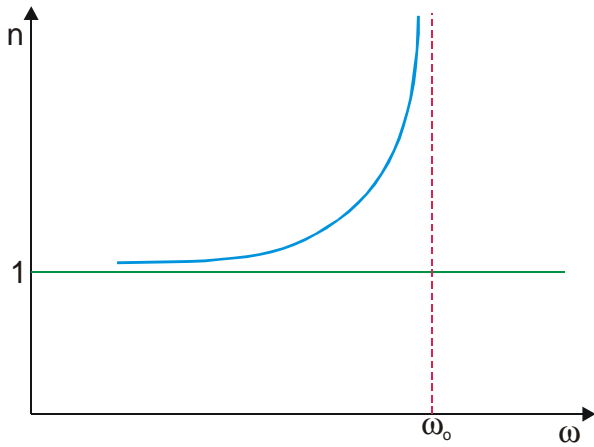
Ponieważ $\varphi = \Delta E_o / E_o$, to

$$\omega(n-1) \frac{\Delta x}{c} = \frac{\Delta E_o}{E_o}.$$

Podstawiając wyrażenie (9.17) dla ΔE_o i rozwiązując względem n , znajdujemy

$$n = 1 + \frac{Ne^2}{2\epsilon_o m(\omega_o^2 - \omega^2)}. \quad (9.18)$$

Jest to wzór określający współczynnik załamania płytki. Należy zauważyć, że n nie zależy od Δx . Nie stanowi to niespodzianki, ponieważ grubą płytkę można przedstawić w postaci zbioru cienkich płytek w których powtarza się ten sam proces. Zauważmy także, że stosowaliśmy przybliżenie zgodnie z którym pole fali padającej zmienia się słabo, tj. $\Delta E_o \ll E_o$ i stąd $(n-1) \ll 1$. W przypadku dużych n , pole E_{pad} wewnątrz płytki należy zamienić na wypadkowe pole. Komplikuje to obliczenia i nie będziemy ich tutaj przytaczać. Powyżej otrzymany wynik poprawnie określa zależność współczynnika załamania od częstości padającego promieniowania (rys. 9.13).



Rys. 9.13. Krzywa dyspersji normalnej wykreślona zgodnie ze wzorem (9.18).

Dla większości atomów $\omega_o^2 > \omega^2$, gdzie ω odnosi się do widzialnego zakresu widma. Odpowiada temu współczynnik załamania większy od 1, czyli prędkość światła mniejsza od c . Przy przejściu od zakresu czerwonego do zakresu fioletowego widma widzialnego, współczynnik załamania wzrasta i wzrasta również odchylenie promieni świetlnych przechodzących przez pryzmat, tj. ma miejsce dyspersja normalna.

Stosując równania Maxwella i uproszczony model atomów dla przewodników i dielektryków można przewidzieć wiele z obserwowanych w przyrodzie ogólnych właściwości światła i promieniowania elektromagnetycznego.

9.3.6. Pole promieniowania ładunków punktowych

Dotychczas badaliśmy pole promieniowania emitowanego przez dużą liczbę ładunków, których ruch powoduje wytworzenie prądu. Założyliśmy, że w jednostce objętości znajduje się N ładunków. Jeżeli każdy ładunek q drga według prawa $y = y_o \sin \omega t$, to gęstość prądu wynosi $j = Nq\omega y_o \cos \omega t$, a prąd w warstwie o grubości Δx wynosi

$$J = j\Delta x = (Nq\omega y_o \Delta x) \cos \omega t.$$

Wówczas stosując wyrażenie (9.7), pole promieniowania jest określone wzorem

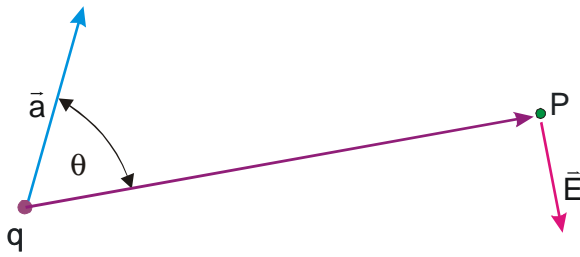
$$E_y = \frac{c\mu_o}{2} Nq\omega y_o \Delta x \cos(\omega t - kx). \quad (9.19)$$

Założmy teraz, że zamiast rozkładu ładunków mamy pojedynczy ładunek q drgający według prawa $y = y_o \sin \omega t$. Korzystając z równania Maxwella można pokazać, że w odległości r od ładunku q pole promieniowania określone jest wyrażeniem

$$E = \frac{\mu_o q \omega^2 y_o}{4\pi r} \sin \omega \left(t - \frac{r}{c} \right) \sin \theta, \quad (9.20)$$

gdzie θ jest kątem pomiędzy wektorem przyśpieszenia a wektorem wodzącym (rys. 9.14). Uwzględniając, że przyśpieszenie $a = -\omega^2 y_o \sin \omega t$, mamy

$$E = -\frac{\mu_o q}{4\pi r} a \left(t - \frac{r}{c} \right) \sin \theta. \quad (9.21)$$



Rys. 9.14. Kierunek pola promieniowania \vec{E} wytwarzanego przez ładunek punktowy q poruszający się z przyspieszeniem \vec{a} .

We wzorze tym $a(t - r/c)$ oznacza przyspieszenie z wcześniejszej chwili czasu $t - r/c$. Wektor \vec{E} skierowany jest prostopadle do wektora wodzącego \vec{r} (rys. 9.14). Wzór (9.21) jest poprawny nie tylko dla ładunku drgającego, ale również dla dowolnego poruszającego się ładunku pod warunkiem, że jego prędkość $v \ll c$. Kierunek pola \vec{B} jest prostopadły do \vec{E} i \vec{r} i tak jak poprzednio $B = E/c$.

9.4. Interferencja fal elektromagnetycznych

W poprzednim punkcie rozważaliśmy zjawiska wywołane przez pojedyncze źródło fal elektromagnetycznych, niezależnie od tego, czy tym źródłem był zmienny prąd powierzchniowy, czy też drgający ładunek elektryczny. W tym punkcie zajmiemy się badaniem zjawisk wytworzonych przez dwa lub więcej źródeł fal drgających z określonymi fazami względem siebie. Powyższe zjawisko nazywamy interferencją fal. W celu opisu interferencji dowolnego rodzaju fal stosujemy ten sam aparat matematyczny.

9.4.1. Interferencja fal promieniowanych przez dwa źródła punktowe

Rozważmy dwa dipole elektryczne S_1 i S_2 drgające w fazie w kierunku osi z (rys. 9.15). Przyjmijmy, że moment dipolowy określony jest wzorem $p = qz = qz_o \cos \omega t = p_o \cos \omega t$. Wówczas przyspieszenie ładunku dipola $a = -\omega^2 z_o \cos \omega t = -\omega^2 (p_o/q) \cos \omega t$. Podstawiając to wyrażenie do (9.21) znajdujemy pole promieniowania dipola w postaci

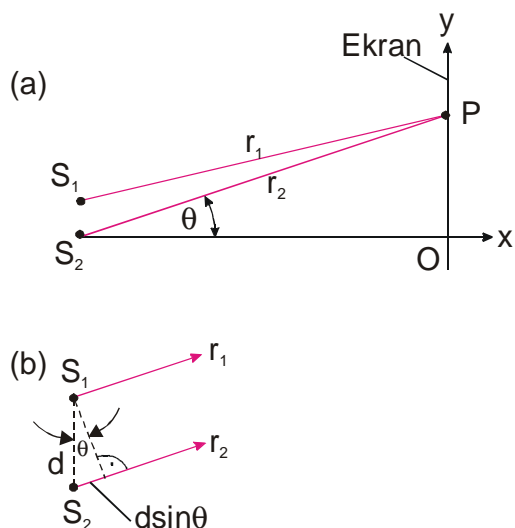
$$E = -\frac{\mu_o q}{4\pi r} \left[-\frac{\omega^2 p_o}{q} \cos \omega \left(t - \frac{r}{c} \right) \right] = \frac{\mu_o p_o \omega^2}{4\pi r} \cos \omega \left(t - \frac{r}{c} \right),$$

ponieważ $\sin \theta = 1$. Zgodnie więc z tym wyrażeniem pole elektryczne w punkcie P

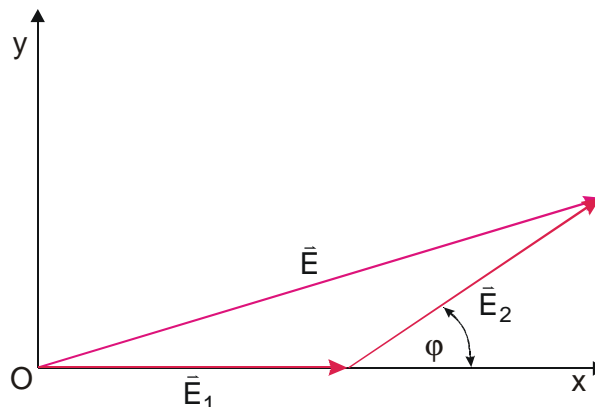
$$E = E_1 + E_2 = E_o \cos(kr_1 - \omega t) + E_o \cos(kr_2 - \omega t),$$

gdzie

$$E_o = \frac{\mu_o p_o \omega^2}{4\pi r}.$$



Rys. 9.15. (a) Dwa źródła S_1 i S_2 w odległości d od siebie. (b) Powyższe źródła w większej skali. Różnica dróg $r_2 - r_1 \approx d \sin \theta$.



Rys. 9.16. Wykres wektorowy dla przypadku dwóch źródeł o różnicy faz φ

Ponieważ obydwa źródła drgają w jednakowej fazie, więc w celu znalezienia sumy dwóch fal można zastosować metodę wektorową, co pokazano na rys. 9.16. Obydwa wektory \vec{E}_1 i \vec{E}_2 mają tę samą długość E_o . Kąt φ między wektorami równy jest różnicy faz pól \vec{E}_1 i \vec{E}_2

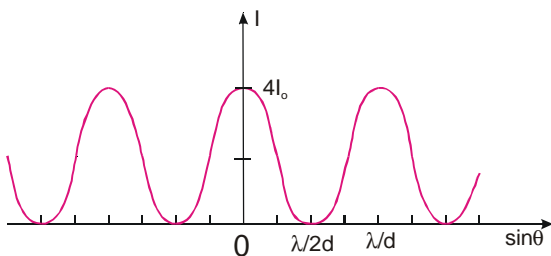
$$\varphi = (kr_2 - \omega t) - (kr_1 - \omega t) = k(r_2 - r_1).$$

Suma wektorowa \vec{E} charakteryzuje amplitudę pola wypadkowego. Stosując prawo cosinusów dla trójkąta równoramiennego (rys. 9.16) otrzymamy

$$E^2 = E_o^2 + E_o^2 + 2E_o^2 \cos \varphi = 2E_o^2(1 + \cos \varphi).$$

Natężenie fali, I , jest proporcjonalne do kwadratu amplitudy. Dlatego $I = 2I_o [1 + \cos k(r_2 - r_1)]$. Z trójkąta prostokątnego na rys. 9.15(b) widzimy, że różnica dróg $r_2 - r_1 = d \sin \theta$, jeżeli odległość od ekranu jest dostatecznie duża. Ten warunek, dla którego różnica dróg jest równa $d \sin \theta$, nazywany jest przybliżeniem Fraunhofera. W tym przypadku

$$I = 2I_o [1 + \cos(kd \sin \theta)]. \quad (9.22)$$



Rys. 9.17. Obraz interferencyjny od dwóch źródeł. Pokazano zależność intensywności od $\sin \theta$.

Na rys. 9.17 wykreślono funkcję opisaną wyrażeniem (9.22). Maksimum intensywności obserwuje się zawsze, gdy $kd \sin \theta = n2\pi$, czyli gdy

$$\sin \theta = \frac{n \lambda}{d}. \quad (9.23)$$

W tym przypadku różnica dróg, która zgodnie z (9.23) jest równa $d \sin \theta$, wynosi $n\lambda$.

Oczywiście, na ekranie oczekujemy maksimum w tym miejscu, gdzie grzbiet jednej fali jest zgodny z grzbietem drugiej fali. Może tak być tylko w przypadku, gdy różnica dróg jest równa całkowitej wielokrotności długości fali. Jeżeli różnica stanowi połowę długości fali czyli $(n + 1/2)\lambda$ to grzbiet jednej fali trafia w dolinę drugiej i będziemy mieć minimum intensywności. Miejsce geometryczne tych minimów intensywności nazywamy linią węzłów.

9.4.2. Interferencja fal od większej liczby źródeł

Założmy, że obserwator położony jest pod kątem θ względem normalnej do linii łączącej N równomiernie rozmieszczonych źródeł (rys. 9.18). Dla obserwatora różnica faz pomiędzy sąsiednimi źródłami jest równa $\varphi = k(r_2 - r_1) = kd \sin \theta$. Na rys. 9.19 przytoczono wykres fazowy dla przypadku N źródeł. Z trójkąta równoramiennego z rys. 9.19a mamy

$$\frac{E}{2} = R \sin\left(\frac{N\varphi}{2}\right).$$

Z kolei z trójkąta prostokątnego na rys. 9.19b mamy

$$\frac{E_1}{2} = R \sin\left(\frac{\varphi}{2}\right).$$

Dzieląc stronami dwa ostatnie równania

$$\frac{E}{E_1} = \frac{\sin\left(\frac{N\varphi}{2}\right)}{\sin\left|\frac{\varphi}{2}\right|}$$

i po podniesieniu do kwadratu

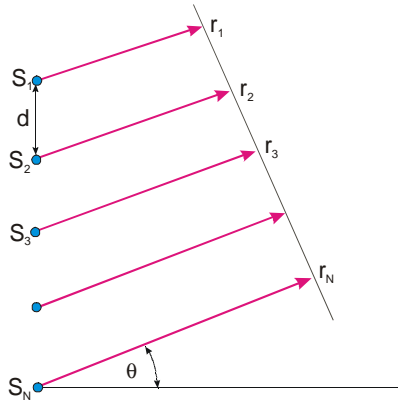
$$I = I_o \frac{\sin^2\left(\frac{N\varphi}{2}\right)}{\sin^2\left(\frac{\varphi}{2}\right)} \quad (9.24)$$

gdzie I_o jest natężeniem fali pojedynczego źródła, a $\varphi = kd \sin \theta$.

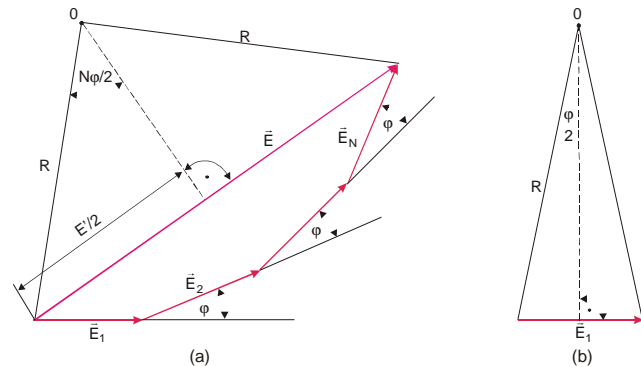
Na rys. 9.20 pokazano rozkład natężenia określony wzorem (9.24). Należy zaznaczyć, że dla $\varphi \rightarrow 0$ mamy $\sin(N\varphi/2) \rightarrow N\varphi/2$, a $\sin(\varphi/2) \rightarrow \varphi/2$, wtedy związek (9.24) można napisać w postaci

$$I \rightarrow I_o \frac{\left(\frac{N\varphi}{2}\right)^2}{\left(\frac{\varphi}{2}\right)^2} = N^2 I_o.$$

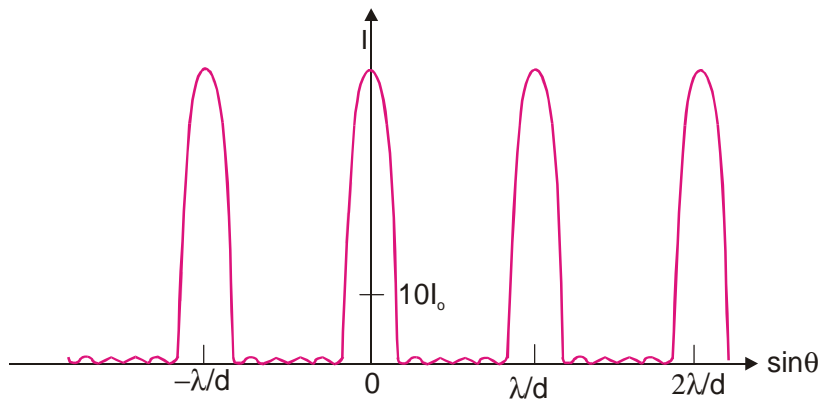
Wobec tego natężenie fali wytworzonej przez N źródeł okazuje się N^2 razy większe od natężenia fali pojedynczego źródła.



Rys. 9.18. N zsynchronizowanych źródeł we wzajemnej odległości d .



Rys. 9.19. (a) Wykres fazowy dla przypadku N źródeł przedstawionych na rys. 9.18; końce wektorów położone są na okręgu o promieniu R . (b) Wykres dla pierwszego źródła.



Rys. 9.20. Obraz interferencyjny od sześciu źródeł położonych w jednej linii.

9.4.3. Siatka dyfrakcyjna

Zarysowując równoległe linie na płaskiej płytce szklanej można otrzymać układ N szczelin. Paski szkła pomiędzy rysami będą pojedynczymi szczelinami. Jeżeli skierować na te szczeliny równoległy strumień światła monochromatycznego z jednego źródła (rys. 9.21), to otrzymamy N źródeł drgających w fazie. Dla tego przypadku, rozkład natężenia na ekranie określa wzór

$$I = I_0 \frac{\sin^2 \left[N \left(\frac{\varphi}{2} \right) \right]}{\sin^2 \left(\frac{\varphi}{2} \right)},$$

gdzie $\varphi = kd \sin \theta$. Natężenie przyjmuje wartość $I = N^2 I_0$ w tych przypadkach, kiedy mianownik przyjmuje wartość zerową, czyli kiedy

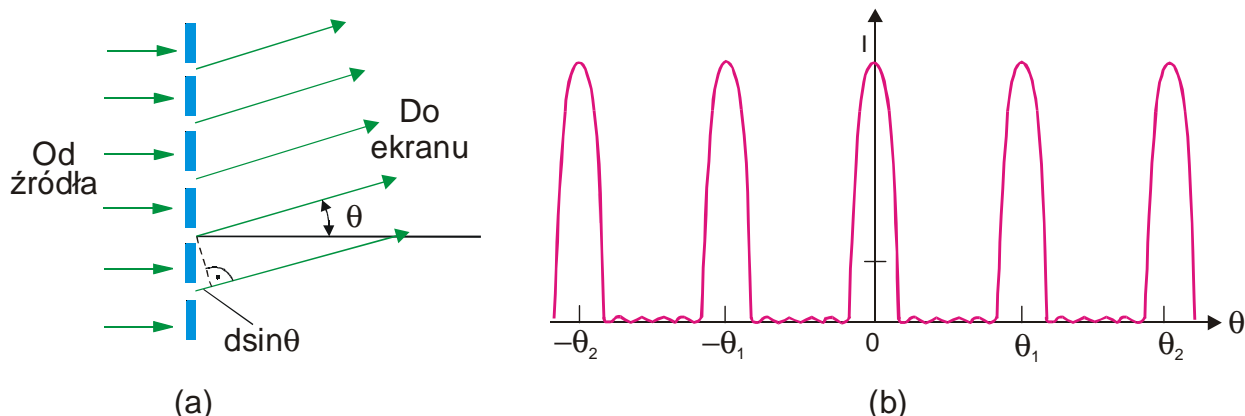
$$\varphi_n = 2\pi n,$$

lub

$$kd \sin \theta_n = 2\pi n,$$

stąd

$$\sin \theta_n = n \frac{\lambda}{d}. \quad (9.25)$$



Rys. 9.21. (a) Część siatki dyfrakcyjnej w powiększonej skali. (b) Rozkład natężenia na ekranie.

Dla pozostałych kątów θ natężenie I w przybliżeniu jest równe I_0 , tj. około N^2 razy mniejsze. W typowych siatkach dyfrakcyjnych wartość N wynosi kilka tysięcy. Równoległe promieniowanie światła za siatką można skupić za pomocą soczewki i otrzymać wyraźny obraz linii na ekranie, jak to pokazano na rys. 9.21b.

Z pomocą rys. 9.21a nietrudno otrzymać warunek (9.25). Po to, aby wszystkie równoległe promienie znajdowały się w fazie, różnica dróg dla każdej pary sąsiednich promieni musi wynosić $n\lambda$. Ponieważ ta różnica dróg wynosi $d \sin \theta$ otrzymamy

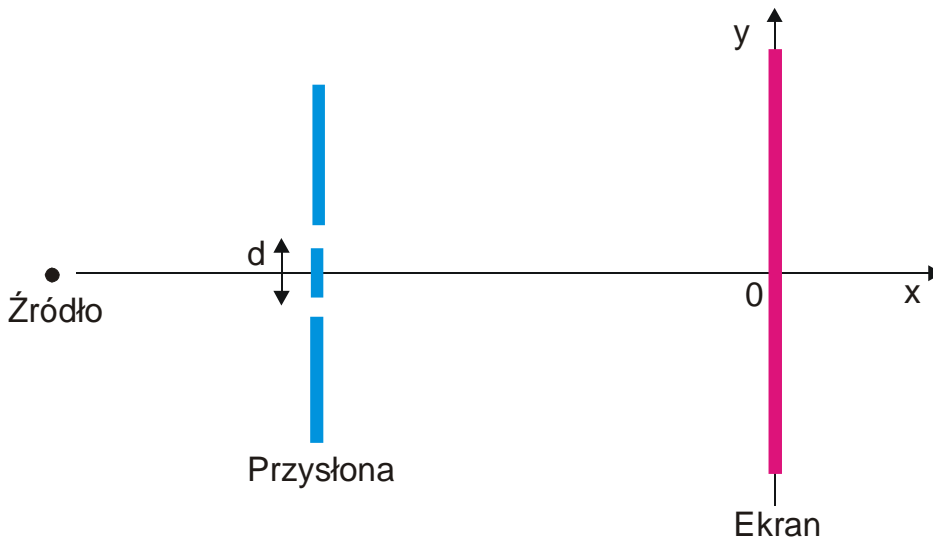
$$d \sin \theta = n\lambda, \quad \text{lub} \quad \sin \theta = \frac{n\lambda}{d}.$$

Linia spektralna odpowiadająca długości fali λ będzie obserwowana pod kątem określonym związkiem $\sin \theta = \lambda/d$. Obraz linii drugiego rzędu będzie odpowiadać $\sin \theta = 2\lambda/d$, trzeciego rzędu $\sin \theta = 3\lambda/d$, itd.

9.5. Dyfrakcja światła

Aby zaobserwować efekty interferencyjne nie jest konieczne posiadanie oddzielnych źródeł światła. Mając pojedyncze źródło światła i ekran z otworami można otrzymać dużą ilość źródeł.

Znanym doświadczeniem dotyczącym interferencji fal elektromagnetycznych jest doświadczenie z dwiema szczelinami i źródłem światła, jak to pokazano na rys. 9.22. Dwie szczeliny oświetlone są pojedynczym źródłem światła. Rozkład natężenia światła na ekranie opisany jest wzorem (9.22) i przedstawia taki obraz jakby szczeliny były zastąpione przez dwa źródła światła. Pierwszy taki eksperyment przeprowadził Thomas Young w 1803 r.



Rys. 9.22. Schemat doświadczenia dotyczącego interferencji światła z dwóch szczelin (źródłem może być laser lub pojedyncza szczelina oświetlona światłem monochromatycznym lampy); przysłona ma dwie szczeliny w odległości d od siebie.

9.5.1. Zasada Huygensa

W XVIII w. Christian Huygens na podstawie doświadczeń z falami na powierzchni wody sformułował następującą zasadę, nie dowodząc jej: *jeżeli czoło fali przechodzi przez jeden lub kilka otworów, każdy element czoła fali zachowuje się tak jakby był źródłem fali*. Na pierwszy rzut oka może to wydawać się dziwne, gdy odniesiemy to do np. siatki dyfrakcyjnej. W otworach nie mamy źródeł prądu; prądy mogą indukować się w dowolnym miejscu ekranu za wyjątkiem otworów.

Wykażemy teraz matematycznie, że pole fali padającej na ekran z otworami jest zgodne z polem promieniowania źródeł prądów wypełniających otwory. Oznaczamy pole promieniowane przez prądy indukowane w ekranie jako E_{ekr} . Wówczas na prawo od ekranu wypadkowe pole zapiszemy w postaci

$$E_{wyp} = E_{pad} + E_{ekr}, \quad (9.26)$$

gdzie E_{pad} jest polem wytwarzanym przez źródło w nieobecności ekranu. Równanie to odpowiada fizycznie sytuacji w której powstaje obserwowany obraz interferencyjny. Teraz zakryjmy otwory dodatkowymi ekranami, których rozmiary są zgodne z rozmiarami otworów. Niech E_{otw} oznacza pole promieniowane przez prądy w ekranach zakrywających otwory. Dla przypadku ekranu z ekranikami mamy

$$E_{wyp} = E_{pad} + E_{ekr} + E_{otw} = 0,$$

na prawo od ekranu, stąd

$$\begin{aligned} -E_{otw} &= E_{pad} + E_{ekr} \\ |E_{otw}|^2 &= |E_{pad} + E_{ekr}|^2 \end{aligned}$$

Widzimy więc, że prawa strona tego równania jest zgodna z polem odpowiadającym realnej sytuacji fizycznej [wzór (9.26)]; a lewa świadczy o tym, że sytuacja ta jest matematycznie równoważna rozkładowi intensywności promieniowania wywołanego źródłami prądu położonymi w otworach i emitującymi promieniowanie niezależnie od siebie. Wykazaliśmy, że jeżeli każdy element frontu falowego przechodzącego przez ekran traktować jako nowe punktowe źródło, to rozkład intensywności będzie identyczny jak w przypadku ekranu i oddzielnego źródła. Jednakże nasze rozważania należy nieznacznie skorygować ze względu na efekty graniczne (przeprowadzony przez nas dowód zakładał, że indukowane prądy mogą przecinać krawędzie otworów).

Udowodniliśmy zasadę Huygensa dla przypadku padania płaskiej fali na płaski ekran. Zasadę Huygensa można stosować dla fal o dowolnym kształcie czoła fali w nieobecności ekranu. Przy tym każdy punkt czoła fali traktowany jest jako źródło nowej fali. Takie traktowanie prowadzi do poprawnego wyniku kiedy rozważamy kształt czoła fal wtórnych, jednakże nie zawsze udaje się otrzymać poprawne wartości amplitudy fali.

9.5.2. Dyfrakcja na pojedynczej szczelinie

Równoległa wiązka światła monochromatycznego padając na pojedynczą szczelinę o szerokości a tworzy na oddalonym ekranie obraz interferencyjny pokazany na rys. 9.23. Podobna interferencja powstająca od pojedynczej szczeliny lub od krawędzi ekranu nazywana jest dyfrakcją. Promienie światła ulegają dyfrakcji (odchyleniu) od swego początkowego kierunku.

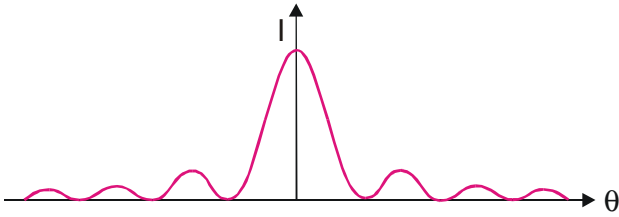
Korzystając z rys. 9.24 łatwo można określić kąt θ , przy którym obserwuje się pierwsze minimum natężenia. Zgodnie z zasadą Huygensa szczelinę możemy traktować jako zbiór pojedynczych źródeł S_1, S_2, \dots, S_N . Jeżeli różnica faz między falami emitowanymi z punktu S_1 i ze środka szczeliny wynosi π , to fale te wygaszają się wzajemnie (promienie 1 i 2 na rys. 9.24). Także wygaszą się wzajemnie fale, z których jedna emitowana jest przez źródło S_2 , a druga - przez kolejne źródło położone poniżej za promieniem 2. Analogicznie promień z S_3 i

kolejne promienie będą parami wygaszać się. Różnica dróg między promieniami 1 i 2 wynosi $(a/2)\sin\theta$. Po to aby uzyskać różnicę faz między nimi wynoszącą π , różnica dróg powinna wynosić $\lambda/2$. Kąt odpowiadający pierwszemu minimum natężenia określony jest z równania

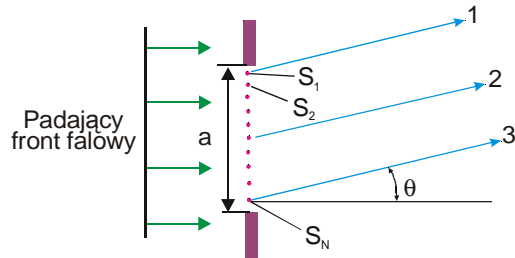
$$\frac{a}{2} \sin\theta_1 = \frac{\lambda}{2},$$

czyli

$$\sin\theta_1 = \frac{\lambda}{a}. \quad (9.27)$$



Rys. 9.23. Rozkład intensywności na oddalonym ekranie przy dyfrakcji na pojedynczej szczeliny.



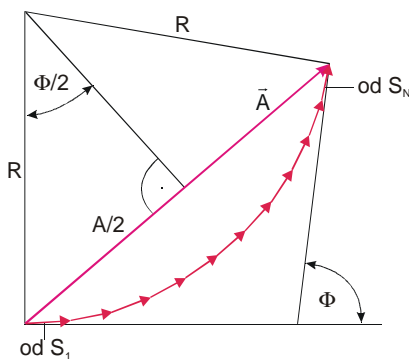
Rys. 9.24. Promieniowanie od pojedynczej szczeliny. Promienie 1 i 3 wychodzą od krawędzi, a promień 2 ze środka szczeliny.

Wartość natężenia dla dowolnego kąta θ otrzymuje się w wyniku zsumowania wkładów nieskończenie małych źródeł (rys. 9.25). Odpowiednie wektory tworzą łuk, dla którego wypadkowa różnica faz dla skrajnych promieni 1 i 3 wynosi

$$\Phi = ka \sin\theta.$$

Wypadkową amplitudę A można znaleźć z trójkąta prostokątnego

$$\sin\left(\frac{\Phi}{2}\right) = \frac{A/2}{R},$$



Rys. 9.25. Wektor \vec{A} stanowi sumę wektorową sygnałów od N źródeł przedstawionych na rys. 9.18. Φ oznacza różnicę faz pomiędzy pierwszym a ostatnim źródłem.

stąd

$$A = 2R \sin\left(\frac{\Phi}{2}\right). \quad (9.28)$$

Długość łuku równa jest A_0 . Jest to wypadkowa amplituda widziana pod kątem 0° , równa promieniowi R pomnożonemu przez kąt Φ (w radianach)

$$R\Phi = A_0.$$

Stąd znajdujemy

$$R = \frac{A_0}{\Phi}.$$

Podstawiając tę wielkość do (9.28) mamy

$$A = A_o \frac{\sin\left(\frac{\Phi}{2}\right)}{\frac{\Phi}{2}}.$$

Podnosząc do kwadratu otrzymujemy rozkład natężenia w postaci

$$I = I_o \left[\frac{\sin\left(\frac{\Phi}{2}\right)}{\frac{\Phi}{2}} \right]^2, \quad (9.29)$$

gdzie $\Phi = ka \sin \theta$.

Kolejne minima obserwuje się przy $\Phi/2 = n\pi$, lub przy

$$\frac{ka \sin \theta}{2} = n\pi,$$

czyli

$$\sin \theta_{min} = n \frac{\lambda}{a} \quad (n \geq 1)$$

Warunek ten jest zgodny z otrzymanym poprzednio.

9.6. Koherentność i niekoherentność

Dotychczas badaliśmy efekty interferencyjne spowodowane źródłami, które znajdowały się w fazie względem siebie lub z pewną stałą różnicą faz. Takie źródła nazywamy koherentnymi lub spójnymi. Koherentne źródła fal można otrzymać przyłączając dwie lub kilka anten do jednego generatora. Koherentność dwu wiązek określa ich zdolność do interferowania; wiązki spójne interferują, wiązki niespójne są tej właściwości pozbawione.

Koherentne źródła światła można otrzymać oświetlając jednym źródłem światła parę lub większą ilość szczelin lub otworów. Przed odkryciem akcji laserowej wytworzenie wiązki światła koherentnego odbywało się zawsze kosztem natężenia światła.

Koherentne wiązki światła można również otrzymać stosując półprzezroczyste zwierciadła – jak w interferometrze Michelsona. Jednakże w przypadku, kiedy obydwie ramiona interferometru mają różną długość, obraz interferencyjny może zniknąć, jeżeli różnica dróg przekracza pewną wielkość ΔL_o odpowiadającą różnicy czasu $\Delta t_o = \Delta L_o/c$. Wielkość ΔL_o nazwana jest długością koherencji, a Δt_o – czasem koherencji. Jeżeli światło źródła w interferometrze (niezależnie od tego czy stosuje się świecenie lasera, czy gazu) przeanalizować za pomocą spektrometru o wysokiej zdolności rozdzielczej, to zaobserwujemy ostrą linię o własnej

szerokości Δf w skali częstotliwości. Szerokość ta związana jest z czasem koherencji Δt_o związkiem

$$2\pi\Delta f \Delta t_o \approx 1,$$

co jest równoważne relacji (8.90).

Z mechaniki kwantowej wiadomo, że światło stanowią fotony emitowane przez różne atomy. Ciąg fal emitowanych nie jest nieskończenie długi. Wiązka światła niekoherentnego składa się z ciągu fal o skończonej długości, poprzedzielanych przypadkowymi przerwami. Docierające do punktu obserwacji dwie takie wiązki nakładają się na siebie, ale różnica faz ciągów fal obu wiązek zmienia się chaotycznie, wskutek czego interferencja nie zachodzi. Jednakże w odcinku czasu $\Delta t_o = 1/2\pi\Delta f$, gdzie Δf jest obserwowaną szerokością linii, dowolna para fotonów będzie zachowywać względem siebie stałą fazę. Fotony te zachowują się jak paczki falowe o długości $L_o = c\Delta t_o \approx c/2\pi\Delta f$.

Źródło światła monochromatycznego wysokiej jakości zachowuje się podobnie jak generator fal, którego średnia częstotliwość fal w sposób przypadkowy zmienia się w przedziale $(f_o - \Delta f/2)$ do $(f_o + \Delta f/2)$. Dwie czysto sinusoidalne fale różniące się częstotliwościami o Δf , będą koherentne przez czas $\Delta t \approx 1/2\pi\Delta f$. Najbardziej wąskim liniom widm atomowych odpowiada $\Delta t_o \approx 10^{-8}$ s. W laserze dryf częstotliwości jest mniejszy i dlatego czas koherentności jest większy. Koherentność to najważniejsza i najcenniejsza własność światła laserowego.

Koherentność światła jest definiowana również w pojęciach matematycznych. Jak wiemy drgania wektorów \vec{E} i \vec{B} o nieskończenie długim ciągu falowym opisujemy wzorem typu

$$S(t) = S_o \cos \omega_o t.$$

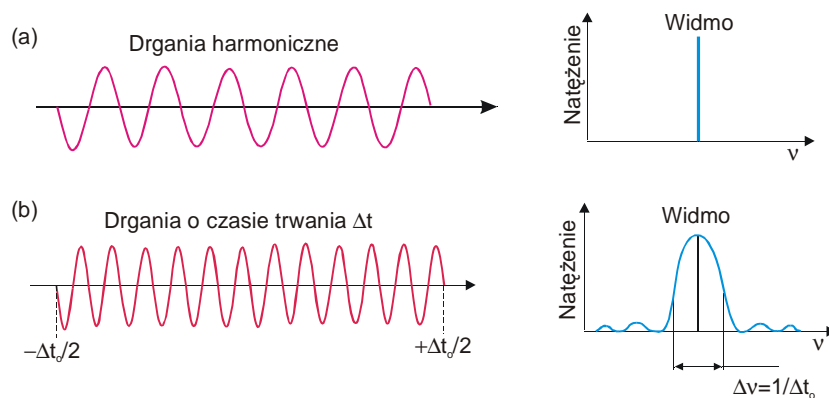
Powyższy wzór przedstawia drgania harmoniczne o widmie liniowym. Drganie, które opisuje powyższy wzór jedynie przez czas trwania Δt_o , a poza tym przedziałem czasu $S(t) = 0$, nie są harmonicznymi; nie charakteryzują się bowiem ściśle określoną częstotliwością, lecz pewnym widmem częstotliwości (patrz pkt. 9.2). Drgania te opisuje wzór

$$S(t) = \int_{-\infty}^{\infty} G(\nu) \cos 2\pi\nu t d\nu,$$

gdzie funkcja $G(\nu)$ jest amplitudą drgań o częstotliwości ν , zwana również widmem Fouriera, przy czym

$$G(\nu) = \int_{-\infty}^{\infty} G(\nu) \cos 2\pi\nu t dt = S_o \Delta t_o \frac{\sin[\pi(\nu - \nu_o)\Delta t_o]}{\pi(\nu - \nu_o)\Delta t_o}.$$

Jeżeli $\Delta\nu$ oznacza szerokość widmową linii, wówczas $\Delta\nu = 1/\Delta t_o$. Tak jak poprzednio, czas Δt_o nazywamy czasem koherencji, a długość $\Delta L_o = c\Delta t_o$ – długością koherencji.



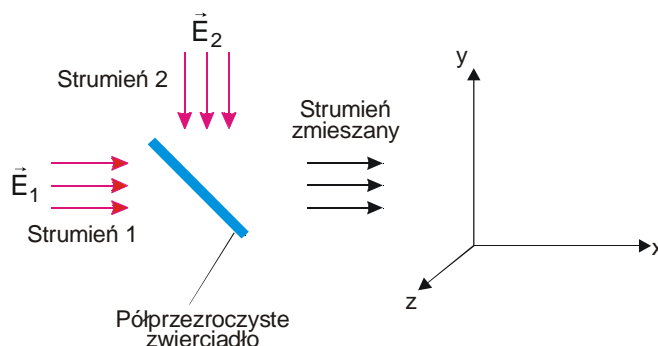
Rys. 9.26. (a) Drgania harmoniczne o widmie liniowym. (b) Drgania o czasie trwania Δt_0 i widmie Fouriera.

9.7. Polaryzacja światła

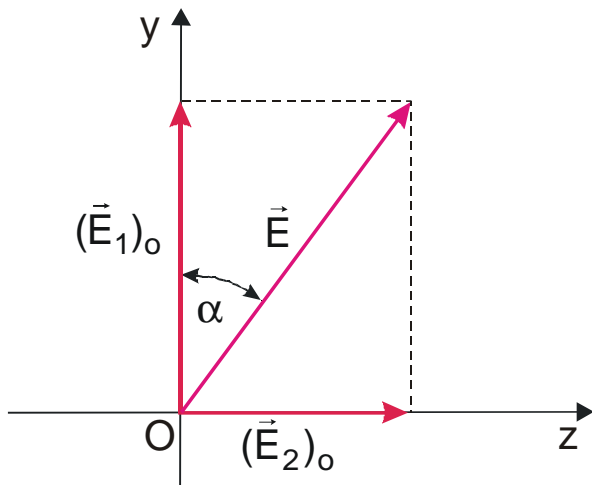
W dowolnym czasie, w każdym punkcie przestrzeni, fala elektromagnetyczna charakteryzuje się tym, że wektory \vec{E} i \vec{B} są wzajemnie prostopadłe. Za kierunek polaryzacji wybrano kierunek wektora \vec{E} . Płaszczyzną polaryzacji określa się płaszczyznę, w której leżą wektor \vec{E} i wektor kierunku propagacji fali. *Promieniowanie elektromagnetyczne, którego kierunek pola \vec{E} jest stały nazywamy płaskopolaryzowanym, lub spolaryzowanym liniowo.* W wiązce światła źródła niekoherentnego, kierunek pola elektrycznego zmienia się chaotycznie w przestrzeni pozostając jednak prostopadłym do kierunku propagacji fali. Taką wiązkę nazywamy niespolaryzowaną.

9.7.1. Polaryzacja kołowa

Przypuśćmy, że dwie koherentne wiązki światła zostały zmieszane za pomocą półprzezroczystego zwierciadła, jak to pokazano na rys. 9.27. Wiązka 1 spolaryzowana jest pionowo (wektor \vec{E}_1 położony jest w płaszczyźnie xy), a wiązka 2 spolaryzowana jest poziomo (wektor \vec{E}_2 położony jest w płaszczyźnie xz). Jaka będzie wiązka wypadkowa, jeżeli różnica faz φ pomiędzy strumieniami jest zerowa, a $E_1 = E_2$? Co będzie jeżeli $\varphi = \pi/2$?



Rys. 9.27. Dwie wiązki światła spolaryzowanego liniowo zmieszane za pomocą zwierciadła półprzezroczystego. Pole \vec{E}_2 skierowane jest na czytelnika z płaszczyzny rysunku równoległe do osi z.



Ażeby odpowiedzieć na pierwsze pytanie zapiszemy obydwie fale jako

$$E_1 = E_{10} \cos(\omega t - kx)$$

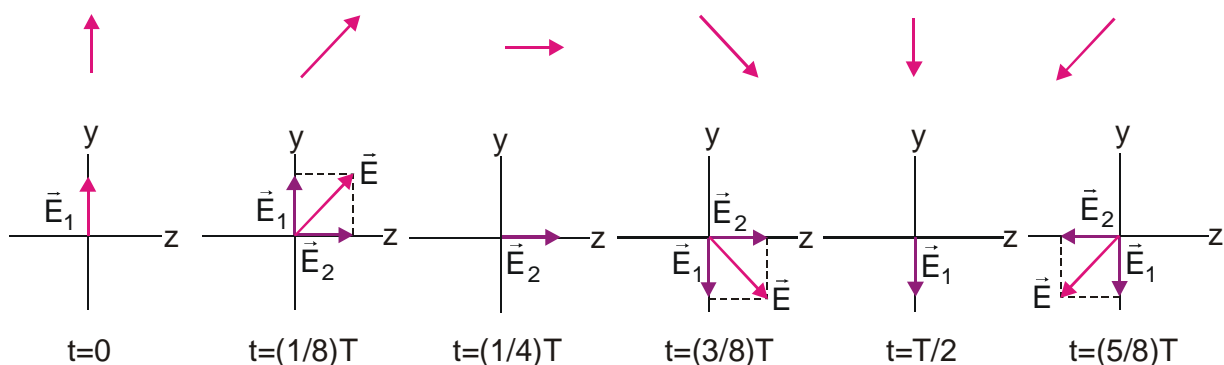
$$E_2 = E_{20} \cos(\omega t - kx)$$

przy czym kierunki pól E_{10} i E_{20} tworzą kąt prosty (rys. 9.28). Widzimy, że wektor wypadkowego pola elektrycznego zawsze położony jest w płaszczyźnie, która tworzy kąt α z pionem, przy czym $\operatorname{tg} \alpha = E_{20}/E_{10}$. Jeżeli $E_{10} = E_{20}$, to $\alpha = 45^\circ$.

Rys. 9.28. Rzut na płaszczyznę yz pól pokazanych na rys. 9.27.

Tak więc wypadkowa wiązka światła jest płaskospolaryzowana, przy czym płaszczyzna polaryzacji tworzy kąt 45° z pionem. Jeżeli dwie wiązki przesunięte są w fazie o $\pi/2$, to przy $x = 0$ mamy $E_y = E_{10} \cos \omega t$ i $E_z = E_{20} \cos(\omega t - \pi/2)$.

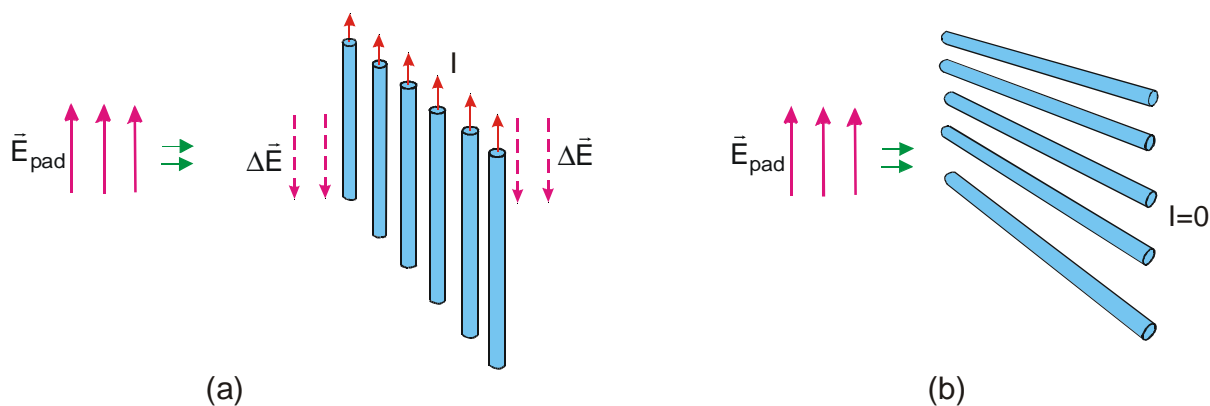
Na rys. 9.29 pokazano położenie wektorów \vec{E}_1 i \vec{E}_2 w kolejnych chwilach czasu. Jak widać z rysunku, wypadkowy wektor \vec{E} jest stały co do wartości i obraca się zgodnie ze wskazówką zegara wokół osi x wykonując jeden obrót w okresie drgań T . Taką polaryzację fali nazywamy lewą polaryzacją kołową. Jeżeli wektor \vec{E} obraca się w kierunku przeciwnym do wskazówek zegara (kiedy patrzymy w kierunku propagacji wiązki), to taką polaryzację nazywamy prawą polaryzacją kołową. Pod pewnymi względami światło spolaryzowane kołowo zachowuje się jak niespolaryzowane. Jednakże, jeżeli mieszać wiązki o jednakowym natężeniu, z których jedna jest spolaryzowana kołowo w lewo a druga – w prawo, to w rezultacie otrzymamy wiązkę płaskospolaryzowaną.



Rys. 9.29. Pole wypadkowe w płaszczyźnie yz w kolejnych chwilach czasu dla przypadku kiedy wiązka 2 opóźnia się w fazie w stosunku do wiązki 1 o kąt $\pi/2$. Wypadkowe pole \vec{E} obraca się zgodnie ze wskazówką zegara. E_y i E_z mają jednakowe amplitudy.

9.7.2. Polaryzatory

Większość źródeł światła emituje fale niekoherentne, niespolaryzowane. Wiązkę światła niespolaryzowanego można spolaryzować, jeżeli przepuścić ją przez polaryzator. Ekran wykonany z cienkich równoległych drucików jest pięknym przykładem polaryzatora dla fal milimetrowych (mikrofal); pokazano to na rys. 9.30. Jeżeli wiązka promieniowania mikrofalowego spolaryzowana jest pionowo i druciki także ułożone są pionowo (rys. 9.30a), to w każdym druciku indukuje się prąd I . Jak wyjaśniono w pkt. 9.3.3, indukowany prąd emituje pole $\Delta\vec{E} = -\vec{E}_{pad}$. Dlatego na prawo od polaryzatora, wypadkowe pole $\vec{E} = \vec{E}_{pad} + \Delta\vec{E} = 0$. Tak więc przy takiej orientacji, polaryzator zachowuje się analogicznie do idealnego zwierciadła, które nie przepuszcza wiązki. Jeżeli natomiast druciki są prostopadłe do \vec{E}_{pad} to pionowe prądy nie indukują się, nie powstaje dodatkowe promieniowanie i padająca fala przechodzi bez strat.



Rys. 9.30. Fala elektromagnetyczna z polaryzacją pionową padająca na ekran równoległych drucików (a) Ekran z pionowymi drucikami odbija falę. (b) Ekran z poziomymi drucikami nie odbija fali; fala przechodzi przez ekran bez osłabienia.

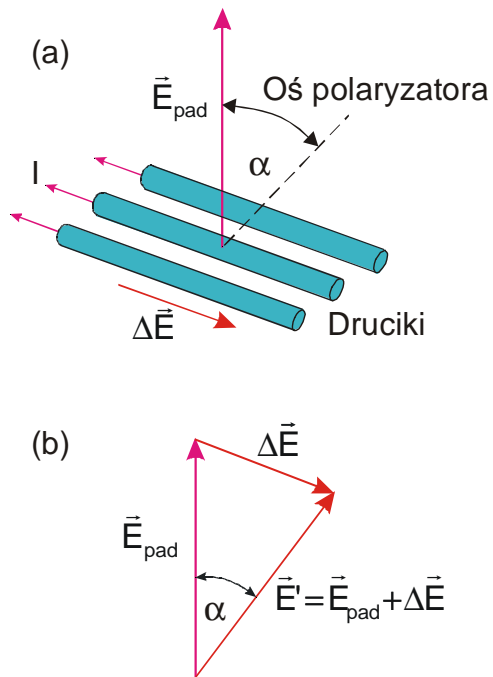
Za oś polaryzatora przyjmujemy linię prostopadłą do linii, w której kierunku położone są druciki (rys. 9.31). Z rys. 9.31 widać, że jeżeli oś polaryzatora tworzy kąt α z kierunkiem \vec{E}_{pad} , to polaryzator będzie promieniować pole $\Delta\vec{E}$ pod kątem prostym do osi. Ponieważ $\Delta\vec{E}$ kompensuje składową \vec{E}_{pad} w tym kierunku, więc wypadkowe pole \vec{E}' będzie przedstawiać składową pola \vec{E}_{pad} równoległą do osi, a więc

$$E' = E_{pad} \cos \alpha,$$

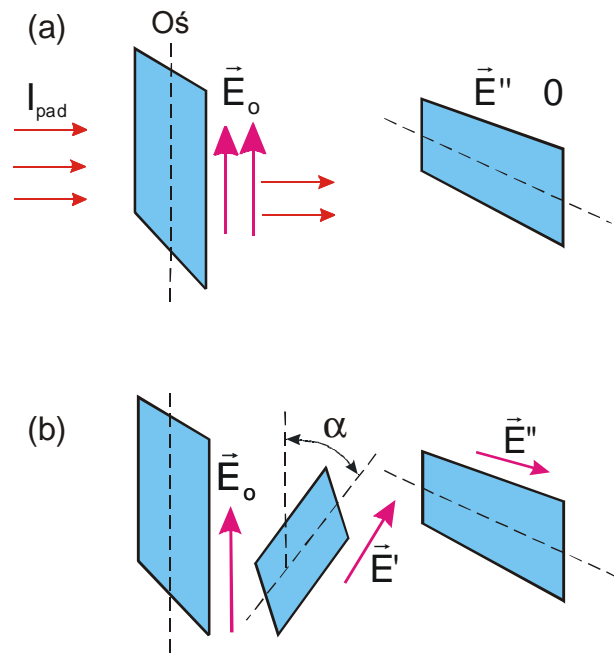
czyli

$$I' = I_{pad} \cos^2 \alpha. \quad (9.30)$$

W przypadku idealnego polaryzatora natężenie jest proporcjonalne do $\cos^2 \alpha$. Polaryzator przepuszcza maksimum natężenia w przypadku kiedy jego oś skierowana jest wzdłuż płaszczyzny polaryzacji. Dowolne promieniowanie po przejściu przez polaryzator jest płaskopolaryzowane w kierunku osi polaryzatora.



Rys. 9.31. (a) Widok przekroju poprzecznego wiązki (wiązka wnika w płaszczyznę rysunku); wektor pionowej polaryzacji wiązki tworzy kąt α z osią polaryzatora stanowiącego druciki promieniujące pole $\Delta \vec{E}$. (b) Wypadkowe pole \vec{E}' za drucikami.



Rys. 9.32. (a) Dwa wzajemnie prostopadłe polaroidy całkowicie wygaszają światło. (b) Światło przechodzi jeżeli pomiędzy nimi umieścić trzeci polaroid.

Na podobnej zasadzie oparte jest działanie polaroidowego filtra świetlnego. Polaroid wykonany jest ze specjalnie zorientowanej masy plastycznej, która składa się przede wszystkim z długich równoległych łańcuchów molekuł, wzdłuż których może płynąć prąd elektryczny. Wobec tego, polaroid stanowi mikroskopową odmianę polaryzatora z równoległymi drucikami. Oś filtra polaroidowego jest prostopadła do łańcuchów molekularnych. W przypadku światła niespolaryzowanego, składowe pola \vec{E} równoległe do łańcuchów molekularnych ulegają pochłonięciu. Po przejściu przez filtr polaroidowy pozostają jedynie te składowe, które są równoległe do osi polaroidu. Jeżeli za pierwszym polaroidem umieścić drugi, w taki sposób, aby ich osie były wzajemnie prostopadłe, jak pokazano na rys. 9.32a, to wiązka ulega całkowitemu pochłonięciu i z drugiego polaroidu światło nie wychodzi. Jeżeli teraz między dwoma skrzyżowanymi polaroidami umiejscowić trzeci polaroid, jak to pokazano na rys. 9.32b, to światło ponownie pojawi się. Ażeby to wyjaśnić założmy, że na środkowy polaroid pada światło, którego natężenie jest równe $I_o = I_{pad} / 2$. Za drugim polaroidem światło będzie spolaryzowane pod kątem α i będzie charakteryzowało się natężeniem $I' = I_o \cos^2 \alpha$. Oś ostatniego polaroidu tworzy kąt $(\pi/2 - \alpha)$ z płaszczyzną polaryzacji światła. Tak więc

$$I'' = I' \cos^2 \left(\frac{\pi}{2} - \alpha \right) = (I_o \cos^2 \alpha) \cos^2 \left(\frac{\pi}{2} - \alpha \right) = \frac{I_o}{4} \sin^2 2\alpha.$$

Wyrażenie to ma maksymalną wartość dla $\alpha = \pi/4$, przy czym w przypadku idealnych polaroidów końcowe natężenie wynosi $I_{pad}/8$.

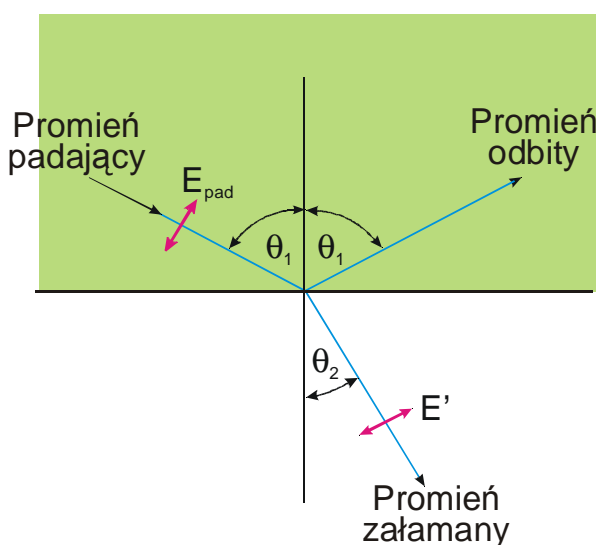
9.7.3. Polaryzacja przez odbicie

Niespolaryzowane światło słoneczne staje się spolaryzowane przy odbiciu. Polaroidowe szkła zmniejszają natężenie światła słonecznego odbitego od powierzchni wody, piasku, szkła itp. Jeżeli oś polaroidu jest prostopadła do płaszczyzny polaryzacji światła odbitego to odbicie gaśnie.

Jeżeli wiązka światła pada pod kątem θ_1 na powierzchnię ośrodka nieprzewodzącego, to część światła odbija się pod tym samym kątem, a część ulega załamaniu zgodnie z prawem Snelliusa (patrz pkt. 9.9.2) pod kątem θ_2 , tak że

$$\frac{\sin \theta_1}{\sin \theta_2} = n.$$

Odbite światło może być emitowane tylko dzięki drganiom atomów nieprzewodzącego ośrodka. Dzięki istnieniu we wzorze (9.21) czynnika $\sin \theta$, elektrony nie emitują promieniowania w kierunku swego ruchu. Jeżeli padające światło jest spolaryzowane jak pokazano na rys. 9.33, to elektrony będą drgały w kierunku \vec{E}' . W tym przypadku światło nie będzie odbijane ponieważ odbijany promień skierowany byłby w kierunku ruchu elektronów. Jednakże jeżeli padające światło spolaryzowane jest prostopadle do płaszczyzny rysunku, to odbicie jest dopuszczalne.



Rys. 9.33. Zmiana polaryzacji przy odbiciu. Promienie odbite i załamane są wzajemnie prostopadłe, tj. kierunek pola E' jest zgodny z kierunkiem promienia odbitego.

Na rys. 9.33 mamy $\theta_1 + \theta_2 = \pi/2$. Podstawiając do prawa Snelliusa $\theta_2 = (\pi/2) - \theta_1$ otrzymamy

$$\frac{\sin \theta_1}{\sin \left(\frac{\pi}{2} - \theta_1 \right)} = n,$$

czyli

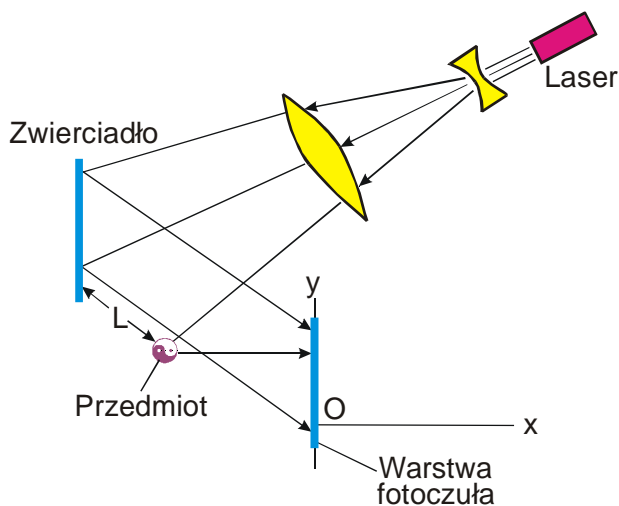
$$\operatorname{tg} \theta_1 = n. \quad (9.31)$$

Widzimy, że jeżeli światło niespolaryzowane pada pod kątem Brewstera, to światło odbite jest spolaryzowane prostopadle do płaszczyzny rysunku. Ten warunek powstania polaryzacji przy odbiciu nazywamy prawem Brewstera.

9.8. Holografia

Holografia zaczęła rozwijać się w latach sześćdziesiątych z chwilą pojawienia się koherentnych źródeł światła stanowiących lasery. W swym zwykłym kształcie hologram wygląda podobnie do negatywu fotograficznego. Jednakże negatyw ten posiada jedną znakomitą właściwość. Chociaż sam jest płaski, to oglądany w świetle monochromatycznym, daje pełny trójwymiarowy obraz realnego przedmiotu. Obraz zarejestrowany na negatywie nie jest podobny do przedmiotu widzianego przez obserwatora w przestrzeni. Negatyw jest bardziej podobny do obrazu wielu nałożonych na siebie odcisków palców. Pomimo tego, że żadna część negatywu nie jest podobna do pozostałych części, to jednak gdy ją umieścimy w wiązce światła monochromatycznego to odtworzymy trójwymiarowy obraz, ale już z gorszą rozdzielczością optyczną. Ponadto, przy ekspozycji negatywu jak i przy odtwarzaniu obrazu trójwymiarowego nie stosuje się soczewki.

Zasada holografii stanowi pogładową ilustrację falowej natury światła i tego w czym tkwi istotna różnica pomiędzy światłem koherentnym a niekoherentnym. Ażeby otrzymać negatyw holograficzny (hologram), przedmiot oświetla się wiązką światła koherentnego, zwykle laserowego. Przy czym na warstwę fotoczułą (błone) pada światło odbite od przedmiotu, a także część wiązki światła odbitego od zwierciadła (rys. 9.34). Jeżeli światło padające na hologram jest koherentne, to na warstwie fotoczułej powstanie obraz interferencyjny. Konieczna długość koherentności wynosi $2L$, gdzie L jest odległością pomiędzy przedmiotem a zwierciadłem.



Rys. 9.34. Sposób otrzymywania hologramu. Na błonę fotoczułą pada światło laserowe odbite od przedmiotu jak i wiązka odbita od zwierciadła.

Na to pytanie można odpowiedzieć stosując następujący uproszczony opis matematyczny zasady holografii. Załóżmy, że warstwa fotoczuła hologramu położona jest w płaszczyźnie yz . Wówczas amplitudę fali odbitej przez przedmiot w płaszczyźnie yz możemy napisać w postaci

$$E = a(y, z) \cos[\omega t + \Phi(y, z)]. \quad (9.32)$$

Jeżeli wywołać warstwę fotoczułą i oświetlić analogiczną wiązką światła, to odtwarzane jest początkowe czoło fali w miejscu położenia warstwy w chwili ekspozycji. Z zasady Huygensa wynika, że odtworzone czoło fali będzie przemieszczać się w kierunku oka obserwatora dokładnie tak samo jak to zachodziłoby z czołem fali wyjściowej. Lecz w jaki sposób za pomocą obrazu na błonie udaje się odtworzyć czoło fali z prawidłowymi wartościami amplitud i faz wzdłuż całej jego powierzchni, jeżeli błona jest jedynie czuła na natężenie światła?

Jeżeli wywołamy błonę to możemy odtworzyć powyższe zachowanie amplitudy tak, że w oku powstaje wrażenie takie, że widzimy wyjściowy przedmiot. Przypuśćmy teraz, że mając taki rozkład amplitudy fali na błonie, oświetlamy ją płaską falą lasera. Wówczas rozkład pola elektrycznego w płaszczyźnie błony ma postać

$$E_{\text{wyp}} = E_o \cos \omega t + a \cos(\omega t + \Phi),$$

gdzie $a = a(y, z)$ i $\Phi = \Phi(y, z)$. Ponieważ natężenie jest proporcjonalne do kwadratu E_{wyp} , mamy

$$I = E_o^2 \cos^2 \omega t + 2E_o a \cos \omega t \cos(\omega t + \Phi) + a^2 \cos^2(\omega t + \Phi).$$

Uśredniony w czasie kwadrat cosinusa daje 1/2. Tak więc średnia wartość natężenia wynosi

$$\langle I \rangle = \frac{I_o}{2} + E_o a \langle [\cos \Phi + \cos(2\omega t + \Phi)] \rangle + \frac{a^2}{2}.$$

Zastosowano tutaj znaną matematyczną tożsamość trygonometryczną $\cos \alpha \cos \beta = (1/2) \cos(\beta - \alpha) + (1/2) \cos(\beta + \alpha)$. Średnia wartość $\cos(2\omega t + \Phi)$ jest równa zeru, dlatego

$$\langle I \rangle = K_1 + E_o a(y, z) \cos \Phi(y, z), \quad (9.33)$$

gdzie $K_1 = (I_o + a^2)/2$.

Należy zaznaczyć, że w przypadku stosowania źródła światła o długości koherencji przewyższającej $2L$, udaje się zachować informację o rozkładzie fazy $\Phi(y, z)$ na błonie. Poczernienie błony jest proporcjonalne do $\langle I \rangle$. Jeżeli skierować na negatyw wiązkę lasera o natężeniu $I' \cos^2 \omega t$, to za negatywem otrzymamy

$$I = I' \cos^2 \omega t [1 - K_2 (K_1 + E_o a \cos \Phi)].$$

Odpowiednie pole elektryczne proporcjonalne jest do pierwiastka kwadratowego z wyrażenia, więc

$$E = E' \cos \omega t [1 - (K_1 K_2 + K_2 E_o a \cos \Phi)]^{1/2} \approx K_3 \cos \omega t + 2K_4 a \cos \omega t \cos \Phi,$$

gdzie

$$K_3 = 1 - \frac{K_1 K_2}{2} \quad \text{i} \quad K_4 = \frac{K_2 E_o}{4}.$$

Stosując ponownie związek trygonometryczny $2 \cos \alpha \cos \beta = \cos(\alpha - \beta) + \cos(\alpha + \beta)$, mamy

$$E = K_3 \cos \omega t + K_4 a(y, z) \cos[\omega t + \Phi(y, z)] + K_4 a(y, z) \cos[\omega t - \Phi(y, z)].$$

$$\left(\begin{array}{c} \text{światło bezpośrednio} \\ \text{z lasera} \end{array} \right) + K_4 \left(\begin{array}{c} \text{światło odbite} \\ \text{od przedmiotu} \end{array} \right) + K_4 \left(\begin{array}{c} \text{światło od przedmiotu} \\ \text{z odwróconą fazą} \end{array} \right)$$

Wkład pierwszego członu rejestrowany jest przez oko obserwatora jako światło laserowe; drugiego członu – w postaci światła odbitego od przedmiotu jakby przedmiot faktycznie znajdował się w swym pierwotnym położeniu; trzeci człon przejawia się w postaci jeszcze jednego wtórnego (a zarazem zbytecznego) obrazu rzeczywistego.

Chociaż holografię odkryto w 1949 r., pozostała nauką samą w sobie do początku lat sześćdziesiątych. Dopiero po konstrukcji pierwszych laserów znalazła szerokie zastosowanie. Obecnie holografia stanowi jeden z głównych kierunków badań optycznych.

9.9. Optyka geometryczna

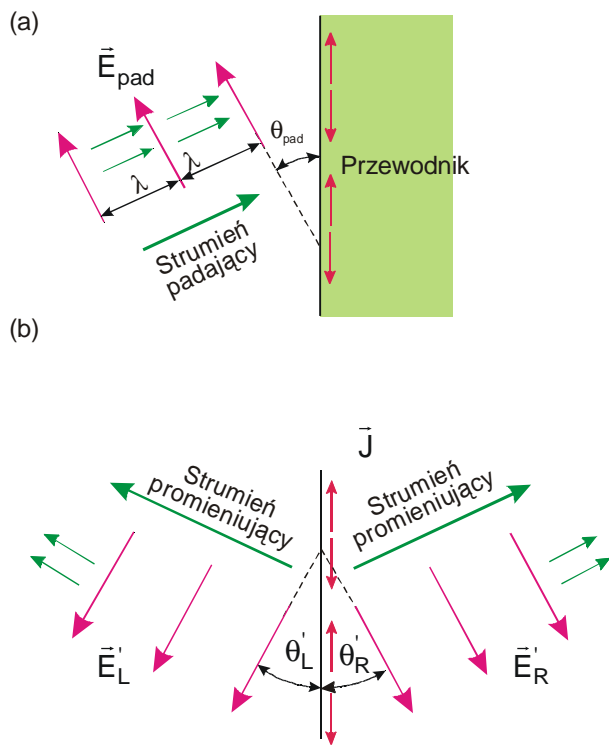
Z ogólnego punktu widzenia przez badanie zjawisk falowych rozumie się również badanie tego, co nazywamy optyką fizyczną. Właśnie dotychczas zajmowaliśmy się tym działem fizyki. Obecnie w dużym skrócie przedstawimy główne zagadnienia optyki geometrycznej.

Długość fali świetlnej jest na tyle mała w porównaniu z rozmiarami większości przyrządów optycznych, że efekty interferencyjne nie ujawniają się. Fale świetlne rozprzestrzeniają się wzdłuż linii prostych prostopadłych do czoła fali. Dowolna taka prosta wzdłuż kierunku propagacji fal świetlnych nazywana jest promieniem świetlnym. Wykażemy, że promienie świetlne podlegają prawu odbicia (od zwierciadeł) i załamania (w przezroczystych ośrodkach np. soczewkach). Stosując te dwa prawa i zwykłe zasady geometrii euklidesowej można zbudować opis matematyczny lub obraz geometryczny propagacji promieni świetlnych. Taki opis matematyczny promieni świetlnych stanowi oddzielny dział fizyki i nosi nazwę optyki geometrycznej. Ponieważ w danym przypadku nowymi zasadami fizycznymi są jedynie prawo odbicia i załamania światła, więc rozpatrzmy je szczegółowo. Pozostałe problemy omówimy w skrócie.

9.9.1. Prawo odbicia

Zgodnie z prawem odbicia, kąt pod którym promień świetlny pada na odbijającą powierzchnię jest równy kątowi odbicia. Kąt padania określony jest jako kąt pomiędzy wiązką padającą i prostopadłą do odbijającej powierzchni. Wobec tego kąt pomiędzy czołem fali i powierzchnią odbijającą także jest równy kątowi padania θ_{pad} .

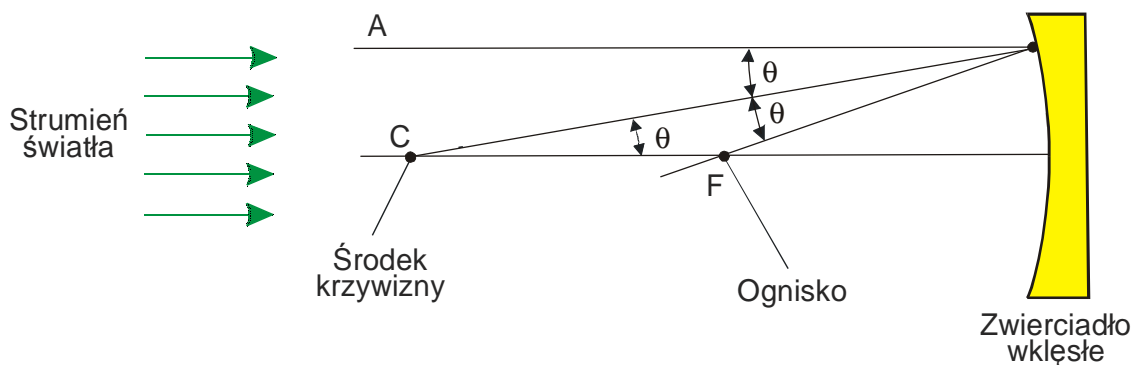
Jak widać z rys. 9.35a, padające pole elektryczne indukuje prąd powierzchniowy $J(y,t)$ który zmienia się z y według prawa sinusoidalnego. W przypadku przewodnika, prąd powierzchniowy okazuje się być takim, że pole wewnątrz przewodnika zawsze jest równe zero. Oznacza to, że promieniowane przez prąd pole na prawą stronę dokładnie kompensuje E_{pad} . Tak więc, jak widać z rys. 9.35b $E'_R = -E_{pad}$ i $\theta'_R = \theta_{pad}$.



Rys. 9.35. (a) Trzy kolejne położenia czoła fali padającej pod kątem padania θ_{pad} ; w przewodniku indukuje się prąd powierzchniowy $J(y)$, którego maksima odpowiadają przecięciu czoła fali z powierzchnią przewodnika. (b) Pole promieniowania spowodowane jedynie prądem $J(y)$.

Warunki symetrii wymagają aby $E'_L = E'_R$ i $\theta'_L = \theta'_R$ [obraz fizyczny nie zmieni się, jeżeli na rys. 9.35b zamienić miejscami prawą stronę na lewą]. Udowodniliśmy więc, że w przypadku powierzchni przewodzącej amplituda fali odbitej zachowuje się lecz jej składowa wzdłuż powierzchni zmienia swój kierunek na przeciwny.

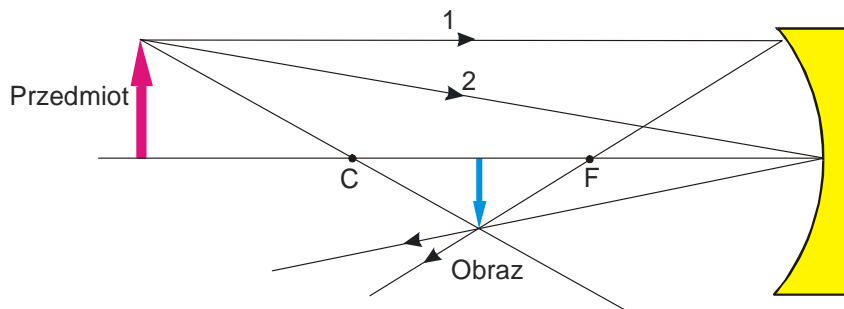
Jako przykład zastosowania prawa odbicia pokażemy w jaki sposób zwierciadło wklęsłe spełnia funkcję soczewki skupiającej. Wiadomo dobrze, że soczewka prosta lub szkła powiększające skupiają równoległą wiązkę promieni słonecznych w punkcie nazywanym ogniskiem. W podobny sposób zachowuje się zwierciadło wklęsłe. Jak widać z rys. 9.36, ogniskowa zwierciadła wklęsłego równa jest połowie jego promienia krzywizny. Na tym rysunku z wiązki promieni równoległych wybrano promień AP .



Rys. 9.36. Równoległa wiązka światła padająca na zwierciadło wklęsłe o promieniu CP .

Niech θ oznacza kąt pomiędzy tym promieniem a normalną do zwierciadła CP . Zauważymy, że odcinek CP jest promieniem krzywizny zwierciadła. Zgodnie z prawem odbicia kąt APC jest równy kątowi FPC ; wówczas trójkąt FPC jest równoramienny. Tak więc boki CF i FP są sobie równe i długość dowolnego z nich równa jest w przybliżeniu połowie odległości pomiędzy punktami C i P , czyli równa promieniowi krzywizny. Na rys. 9.37 pokazano jak można graficznie zbudować obraz przedmiotu (w danym przypadku strzałki) jeżeli znane jest położenie ogniska F . Przeprowadzamy z końca strzałki promień 1 równoległe do osi zwierciadła, a

promień 2 do środka zwierciadła. Obraz końca strzałki położony jest w punkcie przecięcia tych promieni. Wszystkie pozostałe promienie wychodzące z końca strzałki także przejdą przez ten sam punkt obrazu lub w jego pobliżu.



Rys. 9.37. Tworzenie obrazu przez zwierciadło wklęsłe. Za pomocą promieni 1 i 2 określa się graficzne położenie obrazu.

Zwierciadło wklęsłe można stosować do otrzymania obrazu oddalonego przedmiotu. Obraz można powiększyć za pomocą krótkoogniskowej soczewki. Takie urządzenie nazwane jest teleskopem, a jego powiększenie równe jest stosunkowi ogniskowych. W teleskopach astronomicznych błony fotoczule umieszczone są bezpośrednio w ognisku dużego zwierciadła wklęsłego. Ten szeroko rozpowszechniony typ teleskopu był odkryty przez Newtona i nazywany jest teleskopem odbiciowym.

9.9.2. Prawo załamania

Prawo załamania potwierdza, że przy przejściu z jednego ośrodka przezroczystego do drugiego, promień świetlny zmienia swój kierunek. Na przykład, jeżeli promień świetlny pada z powietrza na wodę lub szkło, to odchyła się w stronę normalnej do powierzchni rozdzielającej ośrodki. Rozważmy rys. 9.38. Pokażemy, że na granicy zmienia się kierunek propagacji fali. Na rysunku pokazano dwa kolejne położenia czoła falowego AB i $A'B'$. Niech λ_1 oznacza długość fali świetlnej w ośrodku 1, a λ_2 – w ośrodku 2; wówczas

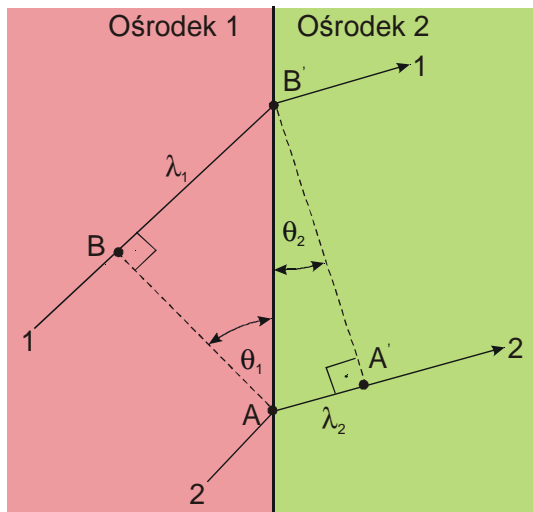
$$\lambda_1 = \frac{u_1}{f} \quad ; \quad \lambda_2 = \frac{u_2}{f}.$$

Z trójkąta prostokątnego ABB' znajdujemy

$$\sin \theta_1 = \frac{\lambda_1}{AB'},$$

a z trójkąta prostokątnego $A'AB$

$$\sin \theta_2 = \frac{\lambda_2}{AB'}.$$



Rys. 9.38. Dwa kolejne położenia czoła falowego, kiedy fala przechodzi przez powierzchnię rozdziału szkło – powietrze.

Podzielimy pierwszy związek przez drugi

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\lambda_1}{\lambda_2} = \frac{u_1}{u_2} = \frac{c/u_2}{c/u_1}.$$

W punkcie 9.3.5 pokazaliśmy, że c/u to współczynnik załamania ośrodka. Stąd

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1}, \quad (9.34)$$

gdzie n_1 i n_2 współczynniki załamania odpowiednio ośrodka 1 i 2. Jest to prawo Snelliusa.

9.9.3. Soczewki

Za pomocą prawa Snelliusa można obliczać właściwości optyczne soczewek. Podobnie jak zwierciadło wklęsłe, soczewka skupiająca odchyła wiązkę promieni równoległych i skupia je w odległości F od soczewki. Odległość F nazywamy ogniskową soczewki.

Położenie obrazu można określić graficznie w sposób pokazany na rys. 9.39. Prowadzimy promień 1 równoległe do poziomej osi. Soczewka odchyła go w ten sposób, że przechodzi przez ognisko F . Przez środek soczewki prowadzimy promień 2. Przecięcie tych dwóch promieni daje punkt, w którym położony jest obraz.

Za pomocą rys. 9.39 można otrzymać ilościowy związek pomiędzy odległością przedmiotu s i odległością obrazu s' . Zauważmy, że trójkąt ABO jest podobny do trójkąta $A'B'O$. Z podobieństwa trójkątów mamy

$$\frac{A'B'}{AB} = \frac{s'}{s}. \quad (9.35)$$

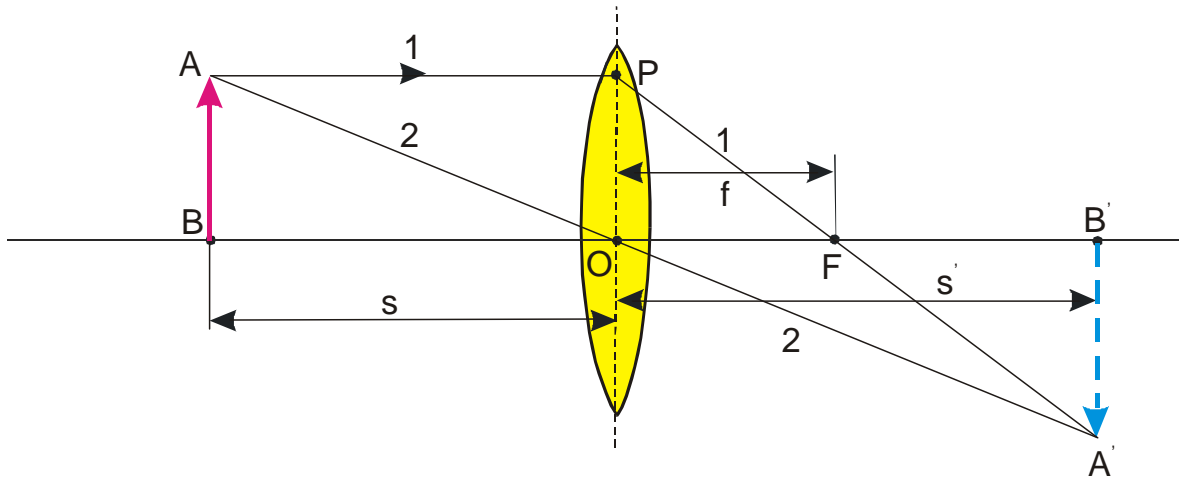
Oprócz tego trójkąt POF podobny jest do trójkąta $A'B'F$, tak że

$$\frac{A'B'}{PO} = \frac{s'-f}{f}. \quad (9.36)$$

Ponieważ $PO = AB$, lewe strony we wzorach (9.35) i (9.36) są równe. Przyrównując między sobą prawe strony, otrzymujemy

$$\frac{1}{f} = \frac{1}{s} + \frac{1}{s'}. \quad (9.37)$$

Związek ten między odległościami od przedmiotu i obrazu nazywany jest wzorem cienkiej soczewki.



Rys. 9.39. Przedmiot AB położony jest w odległości s od soczewki o ogniskowej f . Obraz $A'B'$ położony jest w odległości s' od soczewki.

Koniecznym jest ustalić w jakich przypadkach należy traktować wielkość s , s' i f jako dodatnie, a w jakich jako ujemne. Zwykle przy rozwiązywaniu jakichkolwiek zagadnień, elementy optyczne umieszcza się w taki sposób ażeby światło biegło przez soczewkę z lewa na prawo. Wówczas wielkość s' traktuje się jako dodatnią, jeżeli obraz położony jest na prawo od soczewki i ujemną – jeżeli obraz położony jest na lewo od niej. W przypadku soczewki rozpraszającej wielkość f jest ujemna. Wielkość s będzie ujemna, jeżeli promienie wychodzące z soczewki schodzą się w urojony przedmiot (może to być urojony obraz wytworzony na lewo od soczewki).

ROZDZIAŁ 10

FALOWA NATURA MATERII

Zjawiska fizyczne i opisujące je prawa przyjęto dzielić na klasyczne i współczesne. Fizyka współczesna opiera się o falową naturę materii i zawiera nową fundamentalną stałą odkrytą przez Plancka w 1900 r. Przekonanie wśród fizyków o falowej naturze materii ugruntowało się w latach dwudziestych XX w.

Zgodnie z takim określeniem fizyki współczesnej, to wszystko o czym dotychczas mówiliśmy jest fizyką klasyczną. W tym rozdziale pokażemy, że wszystkie cząstki elementarne charakteryzują się właściwościami falowymi, które wpływają na ich zachowanie szczególnie w niedużych odległościach. Nie wyjaśniając uprzednio falowych właściwości materii, niemożliwe byłoby zrozumienie budowy atomów i molekuł, właściwości cząstek elementarnych oraz takich działów fizyki jak: fizyka ciała stałego, fizyka jądrowa, czy też astrofizyka. Wynikające z falowej natury materii podstawowe założenia i formalizm matematyczny stanowią przedmiot badań mechaniki kwantowej.

Zdumiewający przewrót w naszych poglądach na czas i przestrzeń spowodował Einstein. Jednakże zobaczymy, że falowa natura materii, dualizm korpuskularno-falowy i ich konsekwencje, okażą się bardziej zdumiewające i przeczące zdrowemu rozsądkowi niż einsteinowski postulat o stałości prędkości światła we wszystkich inercjalnych układach odniesienia. Falowa natura materii jakościowo przejawia się w tym, że każdą cząstkę cechują właściwości falowe i odwrotnie, dowolna fala charakteryzuje się właściwościami cząstek. Jako pierwszy przykład demonstrujący tę właściwość fal, rozważymy promieniowanie termiczne.

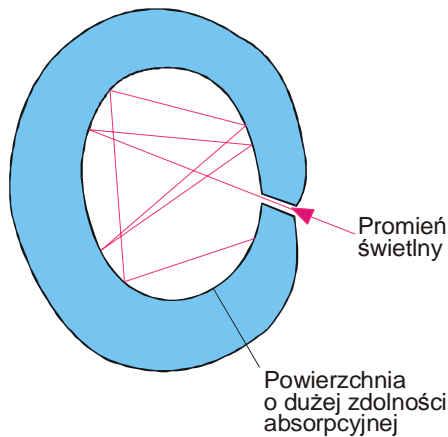
10.1. Promieniowanie termiczne

Jednym z podstawowych zjawisk fizycznych, którego nie można wyjaśnić na gruncie fizyki klasycznej, jest promieniowanie termiczne ciał (zwane też promieniowaniem cieplnym lub temperaturowym). Usunięcie sprzeczności między wynikami doświadczalnymi a teorią dało początek teorii kwantów.

Promieniowaniem termicznym nazywamy promieniowanie wysyłane przez ciała ogrzane do pewnej temperatury. Emitowane fale mają długości fal z szerokiego zakresu widma, ale zwykle główna część przypada na podczerwień. Promieniowanie termiczne jest wynikiem przyśpieszeń jakich doznają ładunki elektryczne atomów i cząstek. Zatem promieniowanie to powstaje kosztem ich ruchu cieplnego.

Promieniowanie termiczne charakteryzujemy wprowadzając pojęcie *zdolności emisyjnej ciała* $e(\nu, T)$ zdefiniowanej tak, że $e(\nu, T)d\nu$ jest energią promieniowania wysyłanego w jednostce czasu z jednostki powierzchni o temperaturze T , w postaci fal elektromagnetycznych o częstotliwościach zawartych w przedziale od ν do $\nu + d\nu$.

Widmo promieniowania termicznego wysyłanego przez ciało zależy w pewnym stopniu od składu tego ciała. Wprowadzono pojęcie ciała emitującego promieniowanie o widmie mającym charakter uniwersalny. Ciało takie nazywamy *ciałem doskonale czarnym*, ponieważ całkowicie absorbuje promieniowanie termiczne nań padające. Przykładem ciała, które ma właściwości zbliżone do ciała doskonale czarnego, jest sadza. Bardzo dobrym modelem ciała doskonale czarnego jest nieprzezroczyste ciało zawierające wnękę z bardzo małym otworem wejściowym (rys. 10.1). Ponadto zakładamy, że ścianka wnęki w dużym stopniu absorbuje padające promieniowanie. Powierzchnia otworu zachowuje się tak, jak powierzchnia ciała doskonale czarnego.



Rys. 10.1. Model ciała doskonale czarnego.

Ogólnie biorąc dla dowolnego ciała a i r zależą od częstotliwości padającego promieniowania i temperatury powierzchni, tak że $a = a(\nu, T)$ i $r = r(\nu, T)$. Między wielkościami a i r zachodzi związek

$$a(\nu, T) + r(\nu, T) = 1. \quad (10.1)$$

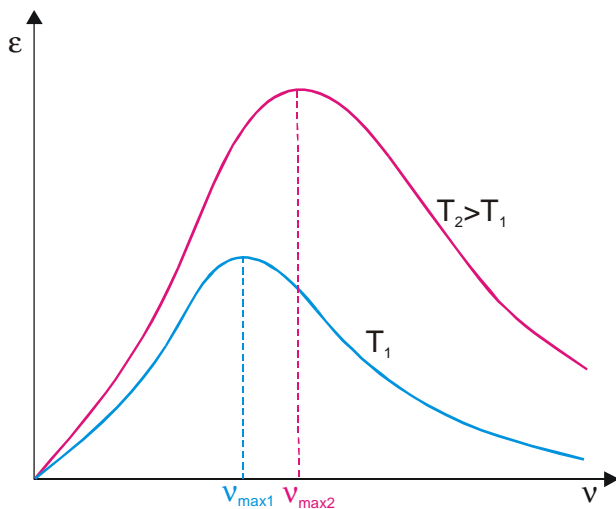
Dla ciała doskonale czarnego, niezależnie od częstotliwości promieniowania i temperatury powierzchni $a = 1$ i $r = 0$.

Podstawowym prawem odnoszącym się do promieniowania termicznego ciał jest prawo Kirchhoffa, które mówi, że *stosunek zdolności emisyjnej do zdolności absorpcyjnej jest dla wszystkich powierzchni jednakowy*

$$\frac{e(\nu, T)}{a(\nu, T)} = \varepsilon(\nu, T). \quad (10.2)$$

Funkcja $\varepsilon(\nu, T)$ jest pewną funkcją uniwersalną. Jej sens fizyczny jest natychmiast jasny, gdy przyjąć $a(\nu, T) = 1$. Wówczas $\varepsilon(\nu, T) = e(\nu, T)$, tzn., że funkcja $\varepsilon(\nu, T)$ jest zdolnością emisyjną ciała doskonale czarnego.

Aby uzasadnić prawo Kirchhoffa wyobraźmy sobie zbiornik (ciało z wnęką) pozostający w temperaturze T . Jeżeli wyciąć w tym zbiorniku otwór o jednostkowej powierzchni, to w ciągu jednostki czasu otworem zbiornika wypromieniuje ilość energii równa $\varepsilon(\nu, T)d\nu$. Jeżeli otwór nie jest wycięty, to pada na jednostkową powierzchnię w czasie jednostki czasu taka sama ilość energii. Część tej energii zostanie zaabsorbowana $a(\nu, T)\varepsilon(\nu, T)d\nu$. W stanie równowagi taką samą ilość energii ścianka musi wypromieniować. Zatem $e(\nu, T) = a(\nu, T)\varepsilon(\nu, T)$, co jest właśnie prawem Kirchhoffa.



Rys. 10.2. Zależność zdolności emisyjnej ciała doskonale czarnego od częstotliwości dla dwóch temperatur T_1 i T_2 .

Z powyższego widać jednoznacznie, że zdolność emisyjną dowolnej powierzchni uzyskujemy mnożąc jej zdolność absorpcyjną $a(\nu, T)$ przez zdolność emisyjną ciała doskonale czarnego $\varepsilon(\nu, T)$. Ponieważ zawsze $a \leq 1$, więc $e(\nu, T) \leq \varepsilon(\nu, T)$, tzn. zdolność emisyjna każdej powierzchni nie jest większa od zdolności emisyjnej ciała doskonale czarnego.

Pomiary zdolności emisyjnej ciała doskonale czarnego doprowadziły do wyników zaprezentowanych na rys. 10.2, gdzie przedstawiono $\varepsilon(\nu, T)$ jako funkcję częstotliwości dla dwóch różnych temperatur T_1 i T_2 ($T_2 > T_1$).

Całkowita energia E wypromieniowana przez jednostkową powierzchnię w czasie jednostki czasu jest równa polu powierzchni ograniczonej każdą z krzywych. Energia ta rośnie z temperaturą według prawa

$$E = \sigma T^4. \quad (10.3)$$

Wzór powyższy nazywa się prawem Stefana-Boltzmana, gdyż był po raz pierwszy wyznaczony przez Stefana, a teoretycznie wyprowadzony przez Boltzmana. Stała σ jest stałą Stefana-Boltzmana a jej wartość wynosi $5,67 \times 10^{-8} \text{ Wm}^{-2}\text{K}^{-4}$.

Funkcja $\varepsilon(\nu, T)$ ma maksimum, które zależy od temperatury. Im wyższa temperatura, tym maksimum przypada dla wyższej częstotliwości. Między ν_{max} a T zachodzi prosta zależność

$$\nu_{max} = const \times T. \quad (10.4)$$

Prawo to nazywa się prawem Wiena.

Pod koniec XIX w. przeprowadzono bardzo staranne pomiary promieniowania termicznego ciała doskonale czarnego. Okazało się jednak, że próby wyprowadzenia prawa opisującego to widmo oparte na zasadach fizyki klasycznej, prowadzą do absurdalnych wyników. Np. Rayleigh i Jeans stosując prawa klasycznej elektrodynamiki dla promieniowania zrównoważonego (w którym promieniowanie emitowane przez drgające elektrony atomowe stanowiące oscylatory, jest pochłaniane przez inne atomy) otrzymali wzór

$$\varepsilon(\nu, T) = \frac{2\pi\nu^2}{c^2} kT. \quad (10.5)$$

Zauważmy, że z tego wzoru wynika, iż gęstość energii jest proporcjonalna do ν^2 i przy $\nu \rightarrow \infty$ staje się nieskończona. Jest to oczywiście sprzeczne z eksperymentem. Jedynie w zakresie niskich częstotliwości zgodność jest dobra.

Próbując usunąć rozbieżności między teorią a doświadczeniem, M. Planck w 1900 r. wysunął hipotezę, że *elektryczny oscylator harmoniczny stanowiący model elementarnego źródła promieniowania, w procesie emisji promieniowania może tracić energię tylko porcjami, czyli kwantami ΔE , o wartości proporcjonalnej do częstości ν jego drgań własnych*. Czyli

$$\Delta E = h\nu, \quad (10.6)$$

gdzie współczynnik proporcjonalności h nosi nazwę stałej Plancka i wynosi $6,626 \times 10^{-34}$ Js. Wymiarem h jest *działanie* = (energia) \times (czas) = (długość) \times (pęd) = (moment pędu). W związku z tym stałą Plancka nazywa się też elementarnym kwantem działania. Uogólniając swoje rozważania Planck zapostulował, że energia oscylatora może przyjmować wartości

$$E_n = nh\nu, \quad n = 0, 1, 2, \dots \quad (10.7)$$

gdzie n jest liczbą kwantową. Jeżeli teraz przyjąć, że rozkład oscylatorów po możliwych dyskretnych stanach energii jest określony rozkładem Boltzmann'a (patrz pkt. 15.4) to prawdopodobieństwo znajdowania się oscylatorów w stanie o energii $nh\nu$ w temperaturze T jest równe $p_n = c \exp(-nh\nu/kT)$, gdzie c jest stałym współczynnikiem określonym z warunku unormowania $\sum_n p_n = 1$. Wówczas średnia energia oscylatora wynosi

$$\langle E \rangle = \sum_{n=0}^{\infty} p_n nh\nu = \frac{\sum_{n=0}^{\infty} nh\nu \exp(-nh\nu/kT)}{\sum_{n=0}^{\infty} \exp(-nh\nu/kT)}. \quad (10.8)$$

Oznaczając $\exp(-h\nu/kT)$ przez x i rozpisując wzór (10.8) otrzymamy

$$\langle E \rangle = \frac{h\nu x (1 + 2x + 3x^2 + \dots)}{1 + x + x^2 + x^3 + \dots}$$

Zauważmy, że

$$1 + 2x + 3x^2 + \dots = \frac{1}{(1-x)^2},$$

natomiast

$$1 + x + x^2 + x^3 + \dots = \frac{1}{1-x}. \quad (x < 1)$$

Zatem

$$\langle E \rangle = h\nu \frac{\exp(-h\nu/kT) \frac{1}{(1-x)^2}}{\frac{1}{1-x}} = \frac{h\nu}{\exp(h\nu/kT) - 1}.$$

Uwzględniając, że w stanie równowagi termicznej wydatek energii promieniowania oscylatorów jest całkowicie kompensowany przez pochłanianie padającego na nie promieniowania, Planck pokazał, że

$$\varepsilon(\nu, T) = \frac{2\pi\nu^2}{c^2} \langle E \rangle.$$

Wobec tego

$$\varepsilon(\nu, T) = \frac{2\pi h \nu^3}{c^2} \frac{1}{\exp(h\nu/kT) - 1}. \quad (10.9)$$

Wzór (10.9) jest słynnym wzorem Plancka na zdolność emisyjną ciała doskonale czarnego. Wzór ten określa rozkład widmowy promieniowania ciała, który jest w bardzo dobrej zgodności z doświadczeniem.

Znając $\varepsilon(\nu, T)$ możemy wyliczyć całkowitą energię emitowaną w jednostce czasu z jednostkowej powierzchni ciała doskonale czarnego poprzez scałkowanie wyrażenia $\varepsilon(\nu, T)$.

Otrzymujemy w ten sposób całkowitą zdolność emisyjną E ciała doskonale czarnego

$$E = \int_0^{\infty} \varepsilon(\nu, T) d\nu = \frac{2\pi h}{c^2} \int_0^{\infty} \frac{\nu^3 d\nu}{\exp(h\nu/kT) - 1}.$$

Wprowadzając pomocniczą zmienną $x = h\nu/kT$ dostajemy

$$E = \frac{2\pi k^4}{c^2 h^3} T^4 \int_0^{\infty} \frac{x^3 dx}{e^x - 1}.$$

Występująca tu całka wynosi $\pi^4/15$. Ostatecznie więc mamy

$$E = \frac{2\pi^5 k^4}{15c^2 h^3} T^4 = \sigma T^4.$$

Uzyskaliśmy teoretycznie wzór Stefana-Botzmanna, przy czym stała Stefana-Botzmanna jest równa

$$\sigma = \frac{2\pi^5 k^4}{15c^2 h^3}.$$

Interesujące jest jeszcze zachowanie się funkcji $\varepsilon(\nu, T)$ dla małych częstotliwości. Występującą w mianowniku funkcję wykładniczą można rozwinąć w szereg, zostawiając dwa pierwsze wyrazy. Dostajemy wówczas

$$\exp(h\nu/kT) - 1 \approx 1 + h\nu/kT - 1 = h\nu/kT.$$

Wzór (10.9) dla niskich częstotliwości promieniowania przybiera więc postać

$$\varepsilon(\nu, T) = \frac{2\pi}{c^2} kT\nu^2.$$

Jest to właśnie wzór Reyleigha-Jeansa.

Oznaczając podobnie jak poprzednio $x = h\nu/kT$ i narzucając warunek istnienia maksimum ($d\varepsilon/dx = 0$), mamy

$$3e^{-x} = 3 - x.$$

Pierwiastek tego równania wynosi około 2,822, stąd wynika

$$\nu_{max} = \frac{x_{max}k}{h} T = 2,822 \frac{k}{h} T = 5,877 \times 10^{10} \text{ s}^{-1} \text{ K}^{-1} T.$$

Otrzymaliśmy zatem prawo przesunięć Wiena. Również to prawo uzyskane z prawa Plancka jest w bardzo dobrej zgodności z doświadczeniem. Widzimy więc, że postulat Plancka o tym, że energia nie może być wypromieniowana w sposób ciągły, doprowadził do teoretycznego wyjaśnienia promieniowania ciała doskonale czarnego.

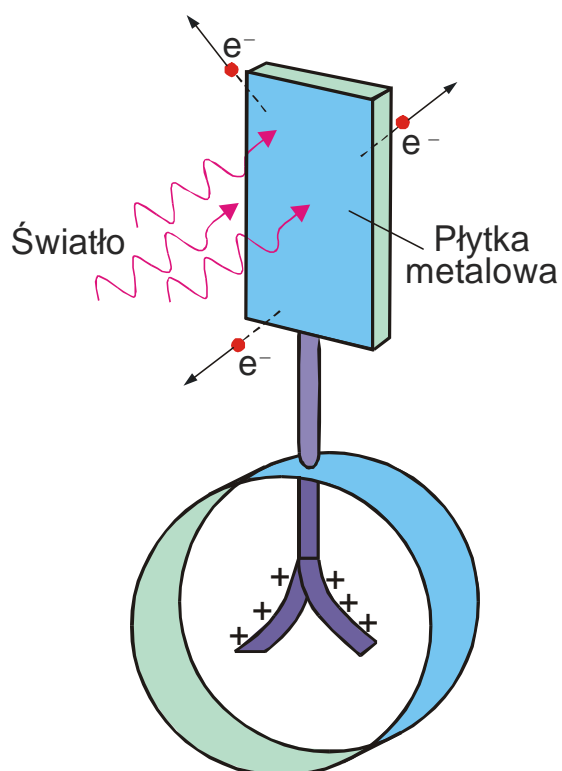
Porcje energii promienistej emitowanej przez ciało wynoszą $h\nu$. Porcje te zostały nazwane kwantami lub fotonami. Hipoteza Plancka dała początek fizyce kwantowej, a stała h występuje obecnie we wielu równaniach fizyki atomowej, jądrowej i ciała stałego.

10.2. Fotoefekt

W końcu XIX w. odkryto elektron. Wkrótce potem zauważono, że elektrony uciekają z niektórych powierzchni metalicznych, kiedy na powierzchnię pada światło (rys. 10.3). Od czasu eksperymentów dyfrakcyjnych przeprowadzonych przez Younga na dwóch szczelinach nie było wątpliwości, że światło jest falą. Taki pogląd pozwalał wyjaśnić fotoefekt. Amplitudę drgań

swobodnego elektronu w zmiennym polu elektrycznym $E = E_o \cos \omega t$ zgodnie z wyrażeniem (9.15) zapiszemy w postaci $A = eE_o / m\omega^2$. Dlatego można było oczekiwać, że elektron znajdujący się w pobliżu powierzchni opuści metal gdy amplituda A przekroczy pewną krytyczną wartość.

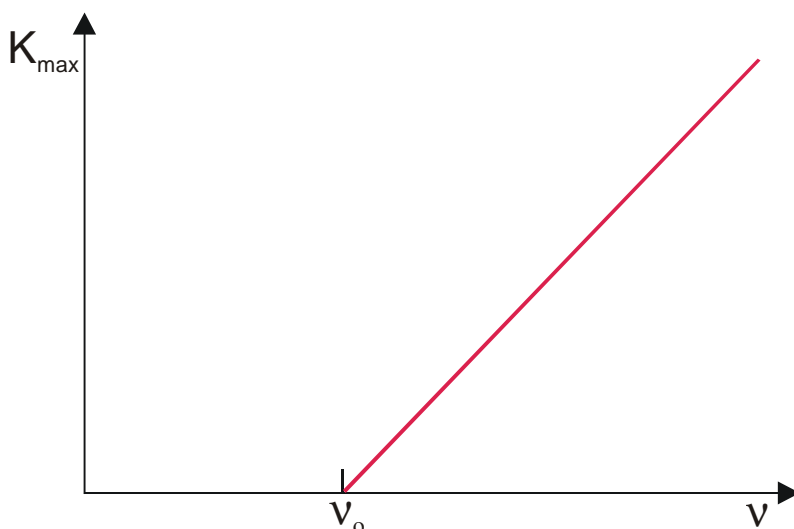
Z falowej teorii światła wynikają następujące wnioski: elektron nie opuści metalu dopóki E_o nie przekroczy określonej wartości krytycznej, energia emitowanych elektronów wzrasta proporcjonalnie do E_o^2 , jeżeli wielkość E_o (a także natężenie) zachować stałą a częstotliwość światła zwiększać, to liczba emitowanych elektronów powinna zmniejszyć się. Jednakże wyniki eksperymentalne obaliły powyższe przewidywania:



- progowego natężenia nie zaobserwowano, liczba uciekających elektronów okazała się ściśle proporcjonalna do E_o^2 przy dowolnie małej intensywności padającego promieniowania,
- energia elektronów okazała się niezależna od wielkości E_o ,
- zauważono zależność energii elektronów od częstotliwości; okazało się, że istnieje progowa częstotliwość ν_o , powyżej której energia emitowanych elektronów rośnie liniowo ze wzrostem częstotliwości – faktycznie energia kinetyczna elektronów zmienia się w przedziale od zera do maksymalnej wartości K_{max} . Na rys. 10.4 pokazano otrzymaną eksperymentalną zależność K_{max} od częstości ν .

Rys. 10.3. Neutralny elektroskop połączony z płytką metalową. Przy oświetleniu płytki przez światło wybijane są fotoelektrony i listki elektroskopu ładowane są dodatnio.

W 1905 r. A. Einstein podał poprawne wyjaśnienie fotoefektu. Zaproponował śmiałą w tym czasie ideę, że światło stanowi zbiór kwantów z których każdy posiada energię $h\nu$. Założył również, że te kwanty światła (obecnie zwane fotonami) zachowują się podobnie do cząstek materialnych i że przy zderzeniu foton może być pochłonięty, a cała jego energia przekazana jest elektronowi. Wówczas nawet Planckowi wydawało się to dziwne. Jak wobec tego pogodzić prawo interferencji fal z tym, że jednocześnie światło składa się z cząstek? Przecież np. w doświadczeniu z dwoma szczelinami cząstka przechodziłaby bądź przez jedną bądź przez drugą szczelinę, co wyklucza utworzenie obrazu interferencyjnego.



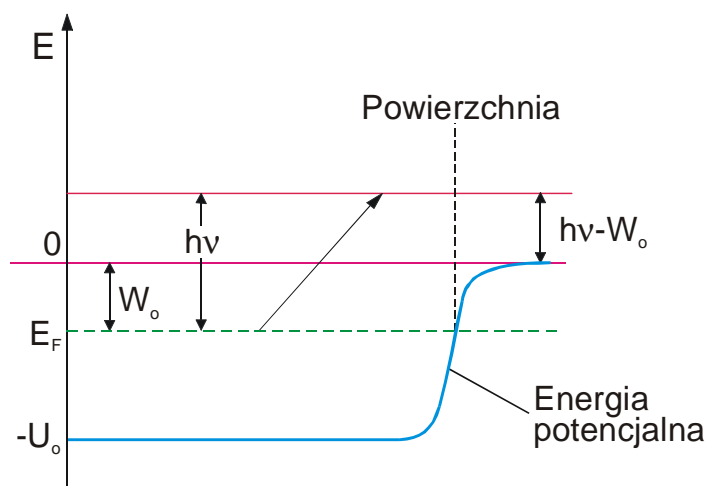
Rys. 10.4. Zależność maksymalnej energii kinetycznej elektronów wybitych z metalu od częstotliwości światła.

Teoria Einsteina wyjaśnia fakty eksperymentalne. Załóżmy, że aby elektron mógł opuścić metal konieczna jest strata energii W_o . Wówczas przy pochłonięciu fotonu o energii $h\nu$, energia elektronu opuszczającego powierzchnię wynosi $h\nu - W_o$. Jest to możliwa maksymalna energia kinetyczna

$$K_{max} = h\nu - W_o. \quad (10.10)$$

Powyższy związek jest zgodny z przebiegiem prostej eksperymentalnej przedstawionej na rys. 10.4. Einstein przewidział, że nachylenie prostej powinno być określone stałą Plancka h . Wielkość W_o nazywana jest pracą wyjścia i zależy od rodzaju metalu. Jeżeli elektron jest początkowo w stanie spoczynku, to przedostając się do metalu przyjmując energię kinetyczną U_o . Inaczej mówiąc, układ elektron w metalu można przedstawić w postaci jamy potencjału o głębokości U_o , jak schematycznie pokazano na rys. 10.5. Wewnątrz metalu zewnętrzne elektrony atomu są swobodne (tj. nie są związane z określonymi atomami), a ich energia kinetyczna może się zmieniać od zera do E_F . Wielkość E_F nazwana jest energią Fermiego. Jeżeli elektronowi o energii Fermiego dostarczyć dodatkowo energii W_o , to jego energia $K = E_F + W_o$ zaledwie wystarcza na to ażeby opuścić metal. Innymi słowy, kiedy elektron opuści metal, jego energia staje się równa $K = 0$. Z rys. 10.5 widzimy, że $W_o + E_F = U_o$, czyli $W_o = U_o - E_F$.

W zjawisku fotoefektu po pochłonięciu fotonu o energii $h\nu$, elektron z poziomu o energii E_F przechodzi na wyższy poziom energetyczny. Przy tym energia elektronu na zewnątrz metalu okazuje się równa $h\nu - W_o$. Jest to maksymalna możliwa energia, którą może posiadać wybity elektron $K_{max} = h\nu - W_o$. Jeżeli elektron znajduje się na niższym poziomie (poniżej linii przerywanej) i pochłania foton o tej energii, to energia na zewnątrz metalu będzie mniejsza od K_{max} .



Rys. 10.5. Jama potencjału w której znajdują się elektrony metalu. Elektron o energii E_F pochłania foton i przechodzi na wyższy poziom energetyczny.

10.3. Efekt Comptona

Stosując klasyczną elektrodynamikę w pkt. 3.5.4 wykazaliśmy, że światło przenoszące energię E posiada pęd $p = E/c$. Tak więc kwant świetlny o energii $E = h\nu$ powinien charakteryzować się pędem $p = h\nu/c$. Jeżeli zamienić ν/c na $1/\lambda$, to

$$p = \frac{h}{\lambda}. \quad (10.11)$$

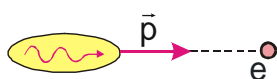
Einstein przewidział, że kwanty świetlne (fotony) będą zachowywać się podobnie do cząstek elementarnych o pędzie $p = h/\lambda$. W przypadku fotoefektu ten minimalny pęd przekazywany jest całej próbce metalu i wybitemu z niego elektronowi. Pęd przyjęty przez metal w takich warunkach jest zbyt mały i nie można go zmierzyć, jednakże przy zderzeniu fotonu ze swobodnym elektronem wielkość przekazywanego pędu można zmierzyć. Proces ten – rozpraszanie fotonu na elektronie swobodnym, nazywany jest efektem Comptona. Po raz pierwszy proces ten był eksperymentalnie potwierdzony przez A. Comptona w 1923 r.

Wprowadzimy teraz związek łączący długość fali rozproszonego fotonu z kątem rozpraszania i długością fali fotonu przed zderzeniem. Niech foton o energii pc i pędzie \vec{p} zderza się z nieruchomym elektronem o energii spoczynkowej mc^2 . Po zderzeniu pęd fotonu będzie równy \vec{p}' i skierowany pod kątem θ , jak to pokazano na rys. 10.6.

Pęd elektronu odrzutu będzie równy \vec{p}'_e , a całkowita energia relatywistyczna E'_e . Stosujemy mechanikę relatywistyczną, ponieważ prędkość elektronu może osiągnąć prędkości bliskie prędkości światła. Zgodnie z prawem zachowania energii, całkowita energia przed zderzeniem równa jest całkowitej energii po zderzeniu, wobec tego

$$pc + mc^2 = p'c + E'_e,$$

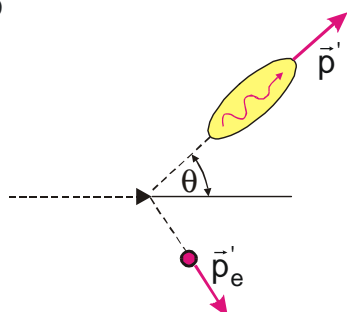
Do



czyli

$$(p - p' + mc)^2 = (E'_e / c)^2. \quad (10.12)$$

Po



Prawo zachowania pędu daje

$$\vec{p} - \vec{p}' = \vec{p}'_e.$$

Podnosząc obie strony do kwadratu

$$p^2 - 2\vec{p}\vec{p}' + p'^2 = p_e'^2$$

i odejmując ostatnie równanie od (10.12) mamy

Rys. 10.6. Efekt Comptona. Zderzenie fotonu ze swobodnym elektronem.

$$m^2 c^2 - 2pp' + 2pmc - 2p' mc + 2pp' \cos \theta = \frac{E_e'^2}{c^2} - p_e'^2.$$

Na podstawie (3.49) widzimy, że prawą stronę można zamienić na $m^2 c^2$

$$m^2 c^2 - 2p'(p + mc - p \cos \theta) + 2pmc = m^2 c^2,$$

stąd znajdujemy

$$p' = \frac{p}{1 + \frac{p}{mc}(1 - \cos \theta)}.$$

Wykorzystując fakt, że $p = h / \lambda$ otrzymujemy

$$\frac{1}{\lambda'} = \frac{1}{\lambda + \frac{h}{mc}(1 - \cos \theta)},$$

czyli

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos \theta). \quad (10.13)$$

W eksperymencie Compton stosował promieniowanie rentgenowskie o znanej długości fali i zauważył, że długość fali fotonów zwiększa się zgodnie z przewidywaniem według wzoru (10.13).

Promieniowanie termiczne, fotoefekt, efekt Comptona i wiele innych eksperymentów z udziałem światła i atomów potwierdziły, że światło faktycznie zachowuje się jakby składało się z cząstek o energii $h\nu$ i pędzie h/λ .

10.4. Dualizm korpuskularno-falowy

Jeżeli byśmy w pierwszych eksperymentach ze światłem zaobserwowali efekt Comptona i fotoefekt, to bylibyśmy przekonani o tym, że światło jest strumieniem fotonów, które zachowują się jak wszystkie "przyzwoite" cząstki. Przy takim układzie rzeczy zaobserwowanie obrazu interferencyjnego od dwóch szczelin wywołałoby zdumienie. Faktycznie; jak cząstki mogą wykazywać właściwości klasycznych fal? Przecież cząstka może przejść tylko przez jedną lub drugą szczelinę.

Paradoks ten stał się bardziej znaczącym w 1927 r. kiedy to amerykańscy fizycy C. Davisson i L. Germer odkryli właściwości falowe elektronu. W rzeczywistości trzy lata wcześniej Louis de Broglie w swej rozprawie doktorskiej założył, że związek (10.11) słuszny jest nie tylko dla fotonów, lecz w ogóle dla wszystkich cząstek. Czyli

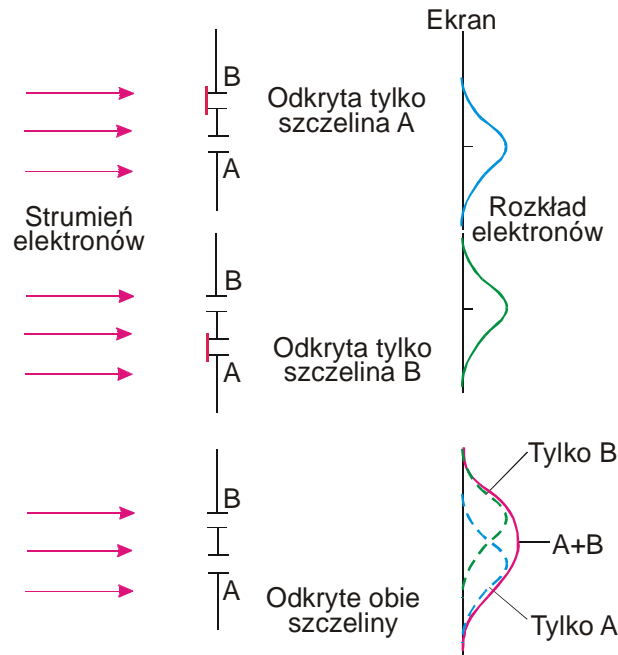
$$p = \frac{h}{\lambda} \quad \text{i} \quad E = h\nu. \quad (10.14)$$

De Broglie założył, że wiązka cząstek dowolnego rodzaju będzie tworzyć obraz interferencyjny na odpowiedniej podwójnej szczelinie charakterystyczny dla doświadczenia Younga.

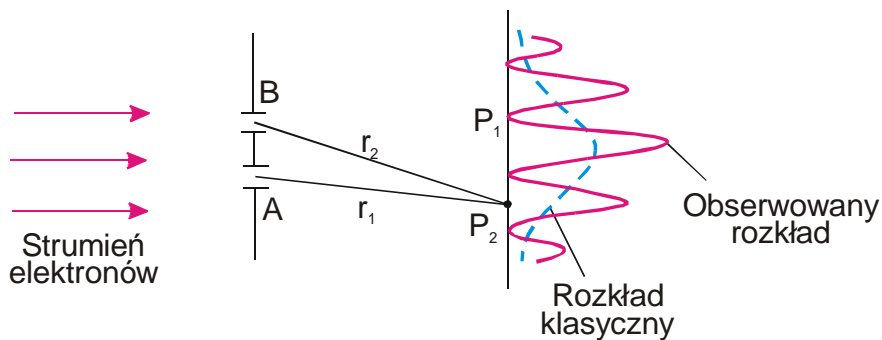
W tych czasach hipoteza de Broglie'a wydawała się wariacką, bodajże niewłaściwą do ubiegania się o stopień doktora. Zaledwie po upływie trzech lat nauka przeżyła wstrząs – eksperyment potwierdził hipotezę. Wstrząs był spowodowany tym, że wydawało się niemożliwe aby takie cząstki jak elektrony zachowywały się jednocześnie jak cząstki i fale.

W przypadku fotonów paradoks można byłoby usunąć zakładając, że pojedynczy foton przed przejściem przez dwie szczeliny zdolny jest rozszczepić się a następnie interferować ze sobą. Jednakże w przypadku elektronów, w przyrodzie nigdy nie zaobserwowano połowy lub części elektronu. Niezależnie od tego czy detektor znajduje się za szczeliną A lub B (rys. 10.7) elektron zawsze wykrywany jest w całości. Z tego punktu widzenia dochodzimy do wniosku, że pojedynczy elektron może przejść tylko przez jedną z dwóch szczelin na rys. 10.7 Tak więc rozkład elektronów na ekranie powinien być sumą rozkładów dla każdej szczeliny oddzielnie. Chociaż logika wyводу wydaje się być nieskazitelną, rozkład charakterystyczny dla A+B nie ma miejsca! Zamiast tego obserwujemy klasyczny obraz interferencyjny dla dwóch szczelin przedstawiony na rys. 10.8. Nie zachodzi więc zaprzeczenie czystej logiki? Przecież wszystko to wygląda jakby $100 + 100 = 0$. Załóżmy, że w punkcie P_1 na rys. 10.8 znajduje się licznik Geigera rejestrujący w każdej sekundzie 100 elektronów kiedy otwarta jest dowolna ze szczelin A lub B. Przy tym, gdy otwarte są obie szczeliny jednocześnie, licznik przestaje rejestrować elektrony. Oznacza to, że w punkcie P_1 przypada minimum interferencyjne ($r_2 - r_1 = \lambda/2$). Jeżeli początkowo otworzyć tylko szczelinę A, a później stopniowo otwierać szczelinę B, to zgodnie ze zdrowym rozsądkiem możemy oczekiwać że prędkość zliczeń w miarę odkrywania szczeliny B będzie stopniowo wzrastać od 100 do 200 zliczeń na sekundę. Zamiast tego obserwujemy zmniejszanie prędkości zliczeń od 100 do zera. Ponadto, jeżeli licznik Geigera umieścić w

punkcie P_2 , to w miarę otwierania szczeliny B prędkość zliczeń będzie stopniowo wzrastać od 100 do 400 zliczeń na sekundę, kiedy druga szczelina jest całkowicie otwarta. Wobec tego $100 + 100 = 400$.



Rys. 10.7. Rozkład intensywności elektronów zgodnie z fizyką klasyczną.



Rys. 10.8. Rozkład intensywności elektronów zgodnie z teorią kwantową.

Jedyny sposób wyjaśnienia tych paradoksalnych wyników polega na stworzeniu nowego formalizmu matematycznego pozwalającego opisać falowe właściwości cząstek materialnych na poziomie mikroświata, a zatem także poprawnie przewidującego obserwowane zjawiska interferencyjne. Formalizm ten musi być wewnętrznie spójny.

10.5. Funkcja falowa

Formalizm matematyczny za pomocą którego usuwa się opisane powyżej paradoksy, przypisuje każdej cząstce materialnej funkcję falową $\Psi(x, y, z, t)$ będącą funkcją współrzędnych i czasu. Pamiętajmy, że natężenie jest proporcjonalne do kwadratu amplitudy funkcji falowej.

Powróćmy na chwilę do eksperymentu z elektronami. Elektrony przepuszczone przez szczelinę padały w określone punkty na ekranie, z tym, że tam gdzie obserwowaliśmy maksima, elektrony padały częściej. Okazuje się jednak, że nie możemy z góry przewidzieć, w którym miejscu dany elektron padnie na ekran. Znajdując natomiast rozkład natężenia w obrazie dyfrakcyjnym można określić prawdopodobieństwo, że elektron padnie w określonym miejscu ekranu. Zatem kwadrat amplitudy funkcji falowej jest proporcjonalny do gęstości prawdopodobieństwa znalezienia elektronu w danym elemencie obszaru. Ponieważ funkcja falowa Ψ jest na ogół funkcją zespoloną, to kwadrat amplitudy tej funkcji wynosi

$$|\Psi|^2 = \Psi \cdot \Psi^*,$$

gdzie Ψ^* jest funkcją sprzężoną z Ψ . Tak więc $|\Psi|^2 dx dy dz$ jest proporcjonalne do prawdopodobieństwa znalezienia elektronu w elemencie objętości $dx dy dz$. Ponieważ suma prawdopodobieństw znalezienia elektronu w poszczególnych elementach objętości rozciągnięta na całą przestrzeń musi być równa jeden, zatem można tak określić funkcję falową, aby był spełniony warunek

$$\int_V |\Psi|^2 dV = 1.$$

Jest to warunek unormowania funkcji falowej. Wówczas $|\Psi|^2$ jest równe gęstości prawdopodobieństwa znalezienia elektronu.

Formalnie funkcja falowa Ψ charakteryzuje się właściwościami klasycznych fal, lecz nie reprezentuje takich wielkości jak np. wychylenie cząstki z położenia równowagi.

Jeżeli zdarzenie może przebiegać w kilku wzajemnie wykluczających się sposobach (jak powiedzmy, przy przechodzeniu cząstek przez jedną ze szczelin A i B) to funkcja falowa takiego zdarzenia przedstawia sumę funkcji falowych każdego ze sposobów

$$\Psi = \Psi_1 + \Psi_2.$$

To twierdzenie (zasada superpozycji) jest identyczne z zasadą dodawania amplitud fal w optyce. W rozważanym wyżej przykładzie, Ψ_1 opisuje falę przechodzącą przez szczelinę A , a Ψ_2 – falę przechodzącą przez szczelinę B . Na ekranie obie funkcje falowe pokrywają się i dają klasyczny obraz interferencyjny z dwóch szczelin, przy czym n -te maksimum określone jest wyrażeniem $\sin \theta_n = n\lambda / d$ (patrz wzór (9.23)).

Być może, że ten formalizm stanowiący podstawę falowej lub kwantowej mechaniki wprowadza u czytelnika niepokój, że bardziej głębokie zrozumienie zjawisk umknęło naszej uwadze. Jednakże, to nie jest tak.

Przedstawiony formalizm pociąga za sobą szereg niepewności wymagających dalszej interpretacji fizycznej. Przypuśćmy, że mamy tak słaby strumień elektronów, że wysyłamy pojedyncze elektrony. I w tym przypadku po dostatecznie długiej ekspozycji otrzymuje się

charakterystyczny obraz dyfrakcyjny. A więc falowy charakter jest cechą pojedynczych cząstek. Zgodnie z prezentacją falową każdemu elektronowi odpowiada paczka falowa dzieląc się jednakowo pomiędzy dwie szczeliny. Jednakże umieszczając licznik Geigera za szczeliną A, komorę Wilsona lub inny detektor, zauważymy, że przez szczelinę nigdy nie przechodzi połówka elektronu. Na tym polega atomizm (zasada niepodzielności), według którego intensywność fali za szczeliną A charakteryzuje prawdopodobieństwo znalezienia całkowitego elektronu w tym miejscu. Ponadto jeżeli detektor umieścić za szczeliną A, to obraz interferencyjny wygładza się i otrzymuje się klasyczny wynik, przekształcając obraz interferencyjny (rys. 10.8) w klasyczny (rys. 10.7). Wielu fizyków, włączając Einsteina, próbowało wymyślić takie doświadczenie w rezultacie którego można byłoby, nie naruszając obrazu interferencyjnego, ustalić przez którą szczelinę przeszła dana cząstka; jednakże wszystkie te próby były nieudane.

Wobec tego co przedstawiają fale odpowiadające elektronowi? Na to pytanie należy odpowiedzieć tak jak w przypadku fotonów. Fale elektromagnetyczne propagują się swobodnie w pustej przestrzeni. W odróżnieniu od fal mechanicznych w tym przypadku nie istnieje ośrodek przenoszący drgania. Funkcja falowa Ψ nie stanowi bezpośrednio obserwowanej wielkości i w tym sensie nie wykonuje ruchu drgającego. Fale klasyczne i fale odpowiadające cząstkom podlegają równaniom matematycznym tego samego typu. Lecz w przypadku klasycznym amplituda fali jest bezpośrednio obserwowana, a dla funkcji falowej Ψ – nie.

10.6. Dyfrakcja elektronów

Eksperyment z dyfrakcją elektronów na dwóch szczelinach jest bardziej złożony, ponieważ charakterystyczna długość fali elektronów jest dużo mniejsza od długości fal świetlnych zakresu widzialnego. Obliczymy długość fali elektronu przyśpieszanego napięciem $V = 1000 \text{ V}$, tzn. o energii kinetycznej $K = 1000 \text{ eV} = 1.6 \times 10^{-16} \text{ J}$. Wówczas

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mK}}$$

i po podstawieniu danych liczbowych otrzymujemy $4 \times 10^{-11} \text{ m}$. Jest to więc wielkość rzędu promienia atomu. Ponieważ λ jest także porównywalne z odległością między atomami w ciele stałym, nasuwa się wniosek, że wiązka elektronów odbijając się od płaszczyzn krystalograficznych powinna wykazać analogiczne efekty jak w przypadku promieni rentgenowskich. Uporządkowany szereg atomów na powierzchni metalu działa podobnie do szczelin cienkiej siatki dyfrakcyjnej. Właśnie w powyższy sposób Davisson i Germer badali rozpraszanie powolnych elektronów na płytce niklowej. Na rys. 10.9 pokazano schematycznie urządzenie do obserwacji dyfrakcji elektronów od powierzchni kryształu. Jako detektor cząstek można stosować ekran luminescencyjny. Znając wartość kąta θ przy którym obserwuje się

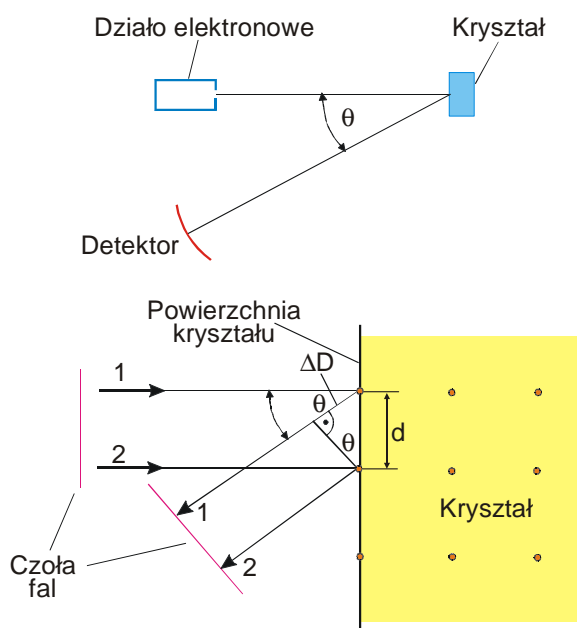
maksimum intensywności, można określić stałą Plancka. Jak widać z rys. 10.9b, $\Delta D = d \sin \theta$ w pierwszym maksimum intensywności powinno być równe długości fali h/p . Wobec tego

$$\frac{h}{p} = d \sin \theta,$$

stąd

$$h = pd \sin \theta.$$

Należy zauważyć, że w tym eksperymencie, jak i w kilku innych mających wyjątkowo duże znaczenie dla rozwoju fizyki, wielkiego odkrycia dokonano przypadkowo. Davisson i Germer nie postawili sobie za cel badanie dyfrakcji elektronów. W 1926 r. Davisson przedstawił na konferencji w Anglii pewne wyniki badań rozpraszania elektronów na powierzchni niklu. Uczni europejscy zwrócili mu uwagę, że wyniki te można lepiej zinterpretować dyfrakcją elektronów niż klasycznym rozpraszaniem które badał. Po upływie kilku miesięcy Davisson i Germer otrzymali nowe wyniki jednoznacznie potwierdzające falową naturę elektronów, które pozwoliły określić wartość stałej Plancka z dokładnością do 1%. Wkrótce po pojawieniu się w 1924 hipotezy de Broglie'a, angielski fizyk Thompson przystąpił do systematycznego badania dyfrakcji elektronów na cienkich foliach metalowych. Jednakże dopiero w 1928 r. otrzymał transmisyjne efekty dyfrakcyjne wiązki elektronów przechodzących przez złotą folię o grubości 10^{-5} m. Widzimy, że staranne badania i przemyślane podejście okazało się mniej szczęśliwe od "przypadkowości" Davissona i Germera. Jednakże doświadczenie tych dwóch amerykańskich uczonych jest dobrym przykładem istoty podejścia naukowego. Jeżeli eksperymentator, nawet przypadkowo, zauważy niezrozumiały dla niego efekt, to należy dokładnie zbadać go, dopóki nie osiągnie się pełnej jasności.



Rys. 10.9. (a) Przyrząd do obserwowania dyfrakcji elektronów od powierzchni kryształu.
(b) Część kryształu silnie powiększona.

Obecnie szczegółowo badane są obrazy dyfrakcyjne wytwarzane nie tylko przez elektrony, protony, ale również przez całe atomy. Falowa natura materii jest wszechstronnie sprawdzona i żadnych odchyłeń od przewidywań teorii nie udało się dotychczas zaobserwować.

ROZDZIAŁ 11

MECHANIKA KWANTOWA

11.1. Paczki falowe

U podstaw mechaniki kwantowej leży związek de Broglie'a $p = h/\lambda$. Jednakże pęd przyjęto wyrażać nie poprzez długość fali λ , ale przez liczbę falową $k = 2\pi/\lambda$

$$p = \frac{h}{2\pi} \frac{2\pi}{\lambda} = \frac{h}{2\pi} k.$$

Wielkość $h/2\pi$ spotykana jest bardzo często, dlatego wprowadzono specjalne oznaczenie \hbar

$$\hbar = \frac{h}{2\pi},$$

$$p = \hbar k. \quad (11.1)$$

Rozważmy cząstkę poruszającą się wzdłuż osi x , której długość fali jest równa λ_o . Liczba falowa cząstki $k_o = 2\pi/\lambda_o$. Czy można funkcję falową przyjąć w postaci $\Psi = A \cos(k_o x - \omega t)$? W tym przypadku gęstość rozkładu prawdopodobieństwa ma postać $|\Psi|^2 \sim \cos^2(k_o x - \omega t)$, tj. w dowolnej chwili czasu t na osi x znajdowałyby się punkty, w których niemożliwe byłoby zaobserwowanie cząstki, podczas gdy w rzeczywistości można ją z jednakowym prawdopodobieństwem znaleźć w dowolnym punkcie na osi x . Aby usunąć tę sprzeczność, funkcję falową należy przyjąć w postaci $\Psi = A e^{i(k_o x - \omega t)}$. Wówczas

$$|\Psi|^2 = \Psi^* \Psi = \left(A e^{-i(k_o x - \omega t)} \right) \left(A e^{i(k_o x - \omega t)} \right) = A^2.$$

Widzimy, że zastosowanie zespolonej funkcji falowej rozwiązuje wskazane powyżej trudność i daje równomierny rozkład prawdopodobieństwa na osi x .

Ze wzoru Eulera wynika, że urojoną i rzeczywistą część funkcji Ψ stanowią fale monochromatyczne

$$\text{Re} \Psi = A \cos(k_o x - \omega t),$$

$$\text{Im} \Psi = A \sin(k_o x - \omega t).$$

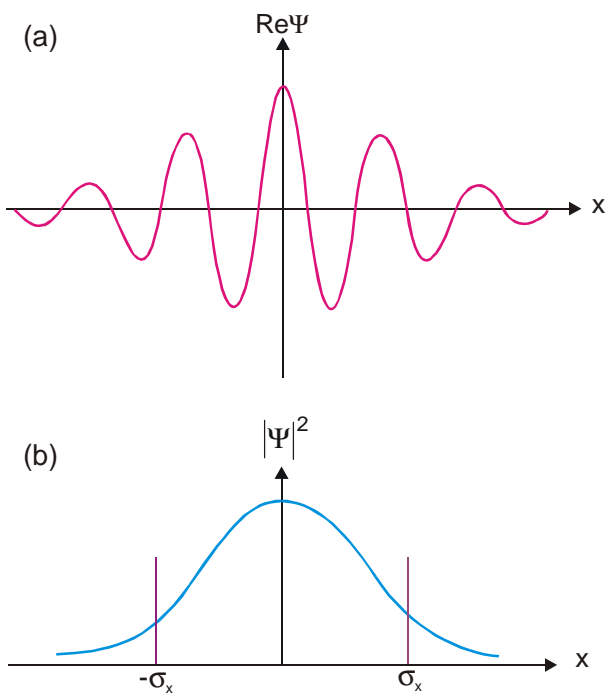
Udowodniliśmy, że jeżeli pęd cząstki posiada określoną wartość, to cząstkę można znaleźć z jednakowym prawdopodobieństwem w dowolnym punkcie przestrzeni. Inaczej mówiąc, jeżeli

pęd cząstki jest dokładnie znany, to nic nie wiemy o jej miejscu położenia. Jednakże w większości sytuacji fizycznych wiadomo, że cząstka znajduje się w określonym obszarze przestrzeni. Rozważmy na przykład, następującą funkcję falową w chwili czasu $t = 0$

$$\Psi(x,0) = A \exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp(ik_0 x). \quad (11.2)$$

Na rys. 11.1a przytoczono rzeczywistą część tej funkcji, a na rys. 11.1b pokazano odpowiednio rozkład prawdopodobieństwa

$$|\Psi|^2 = A^2 \exp\left(-\frac{x^2}{2\sigma_x^2}\right).$$



Należy zauważyć, że w ponad 50% przypadków cząstkę można zaobserwować w przedziale od $x = -\sigma_x$ do $x = \sigma_x$. Funkcja $\exp(-x^2/2\sigma_x^2)$ przedstawia znany rozkład Gaussa, gdzie σ_x jest odchyleniem średniokwadratowym, które będziemy nazywać nieokreślonością wielkości x i oznaczać przez Δx . Taka zlokalizowana funkcja nazywana jest paczką falową. Chociaż przedstawiona na rys. 11.1a fala nie jest czysto monochromatyczną, podobną paczkę falową można przedstawić w postaci fal monochromatycznych. W celu ilustracji tego faktu, rozpatrzmy paczkę falową w chwili czasu $t = 0$ i dobierzemy stosowną superpozycję fal monochromatycznych typu $\exp(ikx)$. W tym celu koniecznym jest znalezienie współczynników B_n w następującym wyrażeniu:

$$\Psi = \exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp(ik_0 x) = \sum B_n \exp(ik_0 x).$$

Rys. 11.1. Paczka falowa w postaci rozkładu Gaussa: (a) zależność rzeczywistej części funkcji falowej od x ; (b) zależność kwadratu modułu funkcji falowej (lub gęstości prawdopodobieństwa) od x .

W wyrażeniu tym zawarta jest nieskończona liczba fal monochromatycznych. Dlatego przejdziemy od sumowania do całkowania

$$\Psi = \exp\left(-\frac{x^2}{4\sigma_x^2}\right) \exp(ik_0 x) = \int B(k) \exp(ik_0 x) dx. \quad (11.3)$$

Posłużymy się następującą tożsamością matematyczną:

$$\exp\left(-\frac{x^2}{4\sigma_x^2}\right)\exp(ik_0x) = \frac{\sigma_x}{\sqrt{\pi}} \int \exp\left[-\sigma_x^2(k-k_0)^2\right]\exp(ikx)dx. \quad (11.4)$$

Porównując wzory (11.3) i (11.4), mamy

$$B(k) = \frac{\sigma_x}{\sqrt{\pi}} \exp\left[-\sigma_x^2(k-k_0)^2\right].$$

W końcu zamieniając k na p/\hbar mamy

$$B(p) = \frac{\sigma_x}{\sqrt{\pi}} \exp\left[-\frac{(p-p_0)^2}{\left(\frac{\hbar}{2\sigma_x}\right)^2}\right].$$

11.2. Zasada nieoznaczoności

Na rys. 11.2 przedstawiono rozkład pędu dla przypadku dwóch paczek falowych o różnej szerokości. Należy zauważyć, że czym węższa jest przestrzennie paczka falowa tym szerszy rozkład po pędzie. Ponieważ prawdopodobieństwo znalezienia cząstki w stanie opisanym funkcją falową $B(k)\exp(ikx)$ jest proporcjonalne do kwadratu amplitudy, prawdopodobieństwo różnych wartości pędu określone jest funkcją:

$$|B(p)|^2 = \frac{\sigma_x^2}{\pi} \exp\left[-\frac{(p-p_0)^2}{2\left(\frac{\hbar}{2\sigma_x}\right)^2}\right]. \quad (11.5)$$

Widzimy, że wyrażenie $|B(p)|^2$ jest także rozkładem Gaussa dla p i można je napisać w postaci:

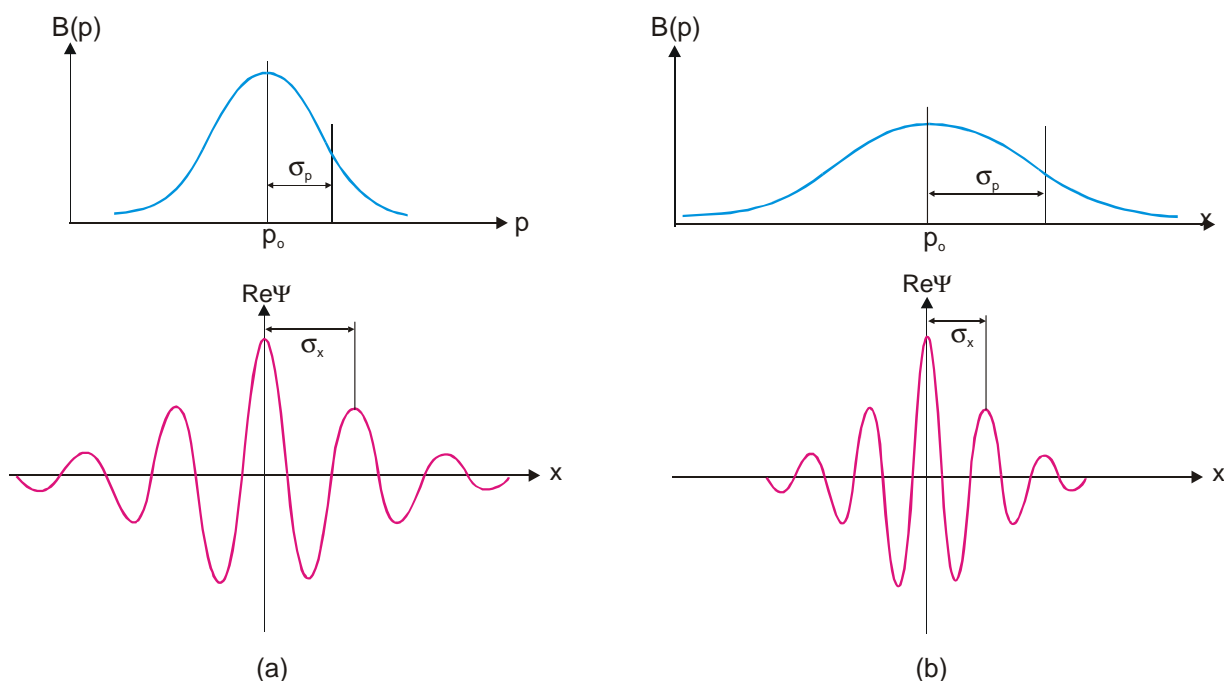
$$|B(p)|^2 = \frac{\sigma_x^2}{\pi} \exp\left[-\frac{(p-p_0)^2}{2\sigma_p^2}\right], \quad (11.6)$$

gdzie σ_p jest odchyleniem średniokwadratowym czyli "nieoznaczonością" wielkości p . Porównując wyrażenia (11.5) i (11.6) otrzymujemy

$$\sigma_p = \frac{\hbar}{2\sigma_x}, \quad \sigma_p\sigma_x = \frac{\hbar}{2}. \quad (11.7)$$

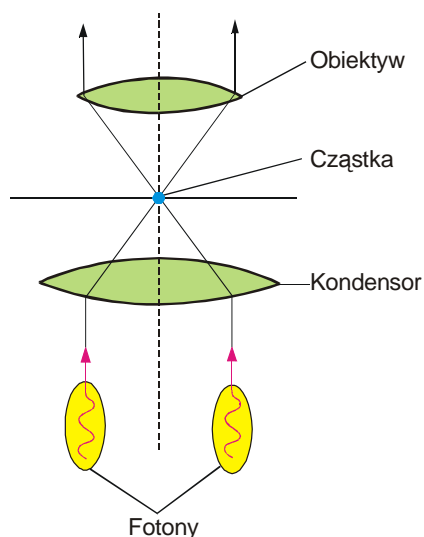
Wobec tego, w przypadku funkcji falowej w postaci rozkładu Gaussa iloczyn szerokości paczki falowej przez szerokość funkcji rozkładu po pędzie jest równy $\hbar/2$. W ogólnym przypadku mamy

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (11.8)$$



Rys. 11.2. Funkcja rozkładu $B(p)$ względem pędu (u góry) i odpowiadająca jej paczka falowa (poniżej). Szerokość paczki falowej na rys. (a) przewyższa dwa razy szerokość na rys. (b). Zauważamy, że w obydwu przypadkach iloczyn $\sigma_x \sigma_p$ jest jednakowy.

Zasada nieoznaczoności potwierdza, że jeżeli cząstka zlokalizowana jest w przestrzeni z odchyleniem średniokwadratowym Δx , to jej pęd nie ma określonej wartości, lecz charakteryzuje się rozkładem $|B(p)|^2$ o "szerokości" Δp . Fizycznie to oznacza, że niemożliwe jest jednoczesne dokładne określenie wartości współrzędnej i pędu cząstki.



Rys. 11.3. Oddziaływanie fotonów z cząstką w mikroskopie.

Jeżeli wiadomo, że cząstka jest w spoczynku, to nieokreśloność jej pędu $\Delta p = 0$. Można byłoby pomyśleć, że za pomocą mikroskopu uda się określić położenie cząstki, a tym samym obalić zasadę nieoznaczoności. Jednakże w najlepszym przypadku, mikroskop pozwala określić położenie cząstki z dokładnością do długości fali stosowanego źródła. Wobec tego $\Delta x \approx \lambda$. Ponieważ $\Delta p = 0$ to iloczyn $\Delta x \Delta p$ także powinien być równy zero i zasada nieoznaczoności jest naruszona! Czy to jest prawdziwe rozumowanie?

Rozważmy to zagadnienie z pozycji mechaniki kwantowej. Posługujemy się światłem, a teoria kwantowa twierdzi, że światło składa się z fotonów o pędzie $p = h/\lambda$. Ażeby zaobserwować cząstkę, to na niej powinien ulec rozproszeniu lub pochłonięciu (w skrajnym

przypadku) jeden z fotonów wiązki światła zebranej soczewką skupiającą (rys. 11.3). Wobec tego cząstce będzie przekazany pęd h/λ . Stąd w chwili obserwacji położenia cząstki z dokładnością $\Delta x \approx \lambda$, nieokreśloność jej pędu jest $\Delta p \geq h/\lambda$. Mnożąc te nieoznaczoności otrzymujemy

$$\Delta x \Delta p \geq \lambda \frac{h}{\lambda} = h,$$

co jest zgodne z (11.8). Przykład ten ilustruje wewnętrzną spójność mechaniki kwantowej. Fizycy razem z matematykami usilnie poszukiwali sprzeczności w podobnych zagadnieniach, lecz nie udało się im ich zaobserwować.

11.3. Właściwości paczek falowych

Wcześniej wyjaśnialiśmy, że paczka falowa propaguje się nie z prędkością fali $u = \omega/k$, ale z prędkością grupową $v_g = d\omega/dk$. Zgodnie ze związkami de Broglie'a $\hbar\omega = E$ i $\hbar k = p$ dla wszystkich cząstek. Zamieńmy w wyrażeniu $E = p^2/2m$ wielkość E na $\hbar\omega$, a p na $\hbar k$; wówczas

$$\hbar\omega = \frac{(\hbar k)^2}{2m}.$$

Różniczkując to wyrażenie po k

$$\hbar \frac{d\omega}{dk} = \frac{\hbar^2 k}{m},$$

$$\frac{d\omega}{dk} = \frac{\hbar k}{m} = \frac{p}{m} = v,$$

czyli

$$v_g = v.$$

Wobec tego przedstawienie zlokalizowanej cząstki w postaci paczki falowej prowadzi do wiernego wyniku klasycznego. Paczka falowa przemieszcza się z prędkością równą prędkości cząstki.

Rozpatrzmy teraz dwie cząstki z których jedna posiada prędkość v_g , a druga – prędkość $v_g + \Delta v_g$. W chwili czasu $t = 0$ ich współrzędne są zgodne, a po upływie pewnego czasu t cząstki rozejdą się na odległość

$$\Delta x = (\Delta v_g) t. \tag{11.9}$$

Udowodnimy, że pojedynczej paczce falowej właściwy jest rozrzut wartości prędkości grupowej Δv_g , który zgodnie z (11.9) powinien prowadzić do zwiększenia szerokości Δx . Oceńmy teraz wielkość Δv_g . Mamy

$$\Delta v_g = \frac{dv_g}{dp} \Delta p.$$

Stosując wynik poprzednio otrzymany, możemy w miejsce v_g napisać v

$$\Delta v_g = \frac{dv}{dp} \Delta p \approx \frac{1}{m} \Delta p. \quad (11.10)$$

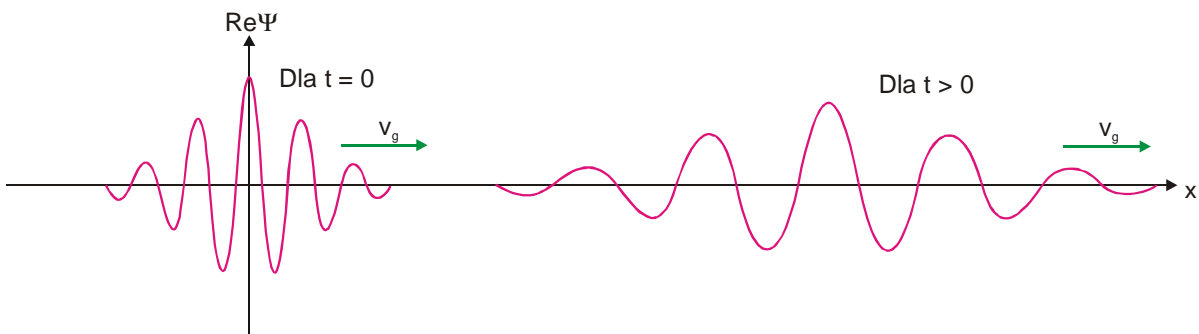
Początkowa wartość Δp jest ograniczona, zgodnie z zasadą nieoznaczoności, wielkością $\hbar/\Delta x_o$, gdzie Δx_o oznacza nieokreśloność położenia początkowego, czyli szerokość wyjściowej paczki falowej. Podstawiając tę wielkość do (11.10) otrzymujemy

$$\Delta v_g \approx \frac{1}{m} \left(\frac{\hbar}{\Delta x_o} \right).$$

Podstawienie ostatniego wyrażenia do (11.9) daje

$$\Delta x \approx \frac{\hbar}{m\Delta x_o} t.$$

Widzimy więc, że szerokość paczki falowej rośnie proporcjonalnie do t . Wkrótce zobaczymy, że podobnego "rozpływania się" paczki falowej można uniknąć umieszczając cząstkę w jamie potencjału. Na rys. 11.4 pokazano jak deformuje się paczka falowa z upływem czasu.



Rys. 11.4. Paczka falowa w dwóch kolejnych chwilach czasu. Paczka porusza się w prawo z prędkością grupową zgodnie z prędkością cząstki.

Aby otrzymać ilościowe wyobrażenie o prędkości rozpływania się paczki falowej w przypadku cząstki swobodnej, rozpatrzmy swobodny elektron zlokalizowany w chwili początkowej w obszarze $\Delta x_o = 10^{-10}$ m (typowy rozmiar atomu). Po upływie sekundy będziemy mieć $\Delta x = \hbar t / (m\Delta x_o) \approx 1100$ km. Widzimy, że po jednej sekundzie chmura elektronowa w swych

rozmiarach okaże się większa od szerokości Polski. Chociaż teoria kwantowa pozwala ściśle określić zachowanie funkcji falowej w przyszłości, jeżeli jest ona znana w chwili początkowej, nie ma jednak istotnego znaczenia ponieważ funkcja falowa bardzo szybko rozplywa się po całej przestrzeni.

Mechanika kwantowa pozwala wyjść z jednego kłopotu o charakterze filozoficznym z którym spotkała się fizyka klasyczna. W okresie dominacji fizyki klasycznej uważano, że jeżeli w pewnej chwili czasu t_0 znane byłyby dokładnie wielkości współrzędnych i prędkości wszystkich cząstek we Wszechświecie, to stosując ściśle prawa fizyczne w zasadzie można byłoby całkowicie opisać obraz przeszłości i przyszłości. Przy tym Wszechświat wyobrażano sobie jako gigantyczny mechanizm. Opierając się na podobnych argumentach, pewni filozofowie mogliby dojść do wniosku, że wszelkie działania człowieka (przecież człowiek również składa się z protonów, neutronów i elektronów) są w pełni zdeterminowane. Wiadomo jednak, że podobnych obliczeń przyszłości i przeszłości nigdy nie uda się wykonać z powodu ogromnej liczby cząstek we Wszechświecie. Wszystkie temu podobne argumenty powodowały niepokój tych, którzy chcieliby wierzyć w swobodną wolę.

Z zasady nieoznaczoności wynika, że istnieją bardziej fundamentalne przeszkody aby można byłoby wykonać takie obliczenie i wobec tego determinizm klasyczny obecnie "nie ciąży" nad fizykami. Jednakże to nie oznacza, że mamy prawo powoływać się na mechanikę kwantową jako na dowód istnienia swobodnej woli.

11.4. Cząstka w studni potencjału

Rozpatrzmy cząstkę zamkniętą w jednowymiarowej studni potencjału o idealnie odbijających ściankach, pomiędzy którymi odległość wynosi L . Na prawo od ścianki w punkcie $x = 0$ (rys. 11.5) zachodzi nałożenie dwóch fal rozchodzących się w przeciwnych kierunkach. W tym przypadku

$$\Psi(x,t) = Be^{ikx-i\omega t} - Be^{-ikx-i\omega t} = B(e^{ikx} - e^{-ikx})e^{-i\omega t}.$$

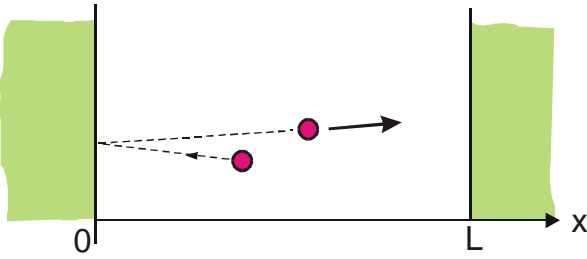
Wybraliśmy znak minus ze względu na to, że Ψ powinno przyjmować zerową wartość przy $x = 0$. Stosując znany wzór

$$\sin(kx) = \frac{e^{ikx} - e^{-ikx}}{2i},$$

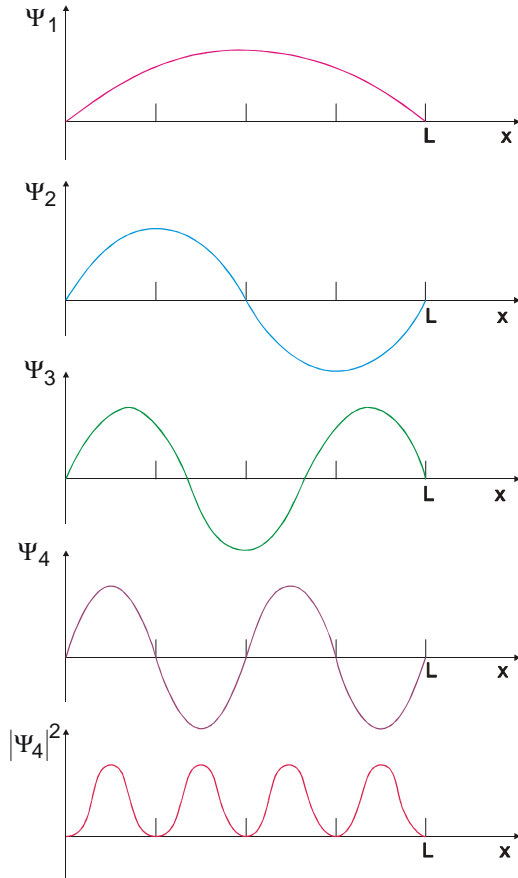
napiszemy $\Psi(x,t)$ w postaci

$$\Psi(x,t) = 2iBe^{-i\omega t} \sin(kx) = Ae^{-i\omega t} \sin(kx), \quad (11.11)$$

gdzie $A = 2Bi$. Funkcja $\Psi(x)$ powinna przyjmować wartość zerową przy $x = L$ i $x = 0$. Podstawiając do (11.11) w miejsce x wielkość L , otrzymamy



Rys. 11.5. Cząstka odbija się od lewej ścianki studni o długości L .



Rys. 11.6. Pierwsze cztery fale stojące odpowiadające cząstce w studni; na najniższym rysunku pokazano gęstość prawdopodobieństwa cząstki w stanie $n = 4$.

Należy zauważyć, że najniższa możliwa energia $\pi^2\hbar^2/2mL^2$ odpowiada $n = 1$, a odpowiadająca jej funkcja falowa przedstawia połowę sinusoidy. Energię odpowiadającą $n = 1$ nazywamy energią stanu podstawowego. W mechanice kwantowej cząstka w studni nie może posiadać energii mniejszej niż $\pi^2\hbar^2/2mL^2$, wskutek tego, że Ψ w studni nie może być funkcją zerową. W fizyce klasycznej cząstka może mieć zerową energię.

Ażeby mieć wyobrażenie o skali energii, rozważmy elektron zamknięty w studni o rozmiarach typowych dla atomu – 10^{-10} m. W tym przypadku $E_n = (37,2n^2)$ eV. Na rys. 11.7

$$\sin(kL) = 0.$$

Równość ta jest spełniona kiedy $kL = n\pi$ gdzie n jest liczbą całkowitą. Widzimy, że dozwolone są tylko takie wartości liczby falowej k_n , które spełniają równanie

$$k_n = \frac{n\pi}{L}. \quad (11.12)$$

Tym samym zażądaliśmy, aby w studni ułożyła się całkowita liczba półfal, co jest zgodne z warunkiem powstania fali stojącej na strunie:

$$L = n \frac{\lambda}{2}.$$

Na rys. 11.6 przedstawiono funkcje falowe $\Psi_n(x) = A \sin(n\pi/L)x$ dla $n = 1, 2, 3, 4$. Odpowiednie wartości pędu zapiszemy w postaci

$$p_n = \hbar k_n,$$

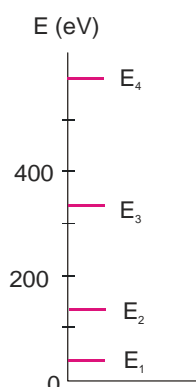
czyli z uwzględnieniem (11.12)

$$p_n = n \frac{\pi\hbar}{L}. \quad (11.13)$$

Tym pędem odpowiadają wartości energii kinetycznej

$$E_n = \frac{p_n^2}{2m} = n^2 \frac{\pi^2\hbar^2}{2mL^2}. \quad (11.14)$$

przedstawiono cztery najniższe poziomy energii. Energia E_1 porównywalna jest co do wartości z energią kinetyczną elektronu w atomie wodoru.



Rys. 11.7. Cztery najniższe poziomy energetyczne elektronu znajdującego się w studni o szerokości 10^{-10} m.

Elektrony na wyższych poziomach energetycznych mogą emitować fotony i przechodzić na niższe poziomy. Ponieważ energia elektronu w studni może przyjmować jedynie określone dyskretne wartości, to energia (lub długości fal) emitowanych fotonów przez elektrony także przyjmuje dyskretne wartości. Takie "widmo" charakteryzujące się dyskretnymi wartościami energii emitowanych fotonów nazywamy liniowym.

11.5. Równanie Schrödingera

Dotychczas mieliśmy do czynienia z cząstkami swobodnymi, które charakteryzowały się określonym pędem, a stąd i określoną energią. W bardziej ogólnym przypadku na cząstkę mogą działać siły zewnętrzne scharakteryzowane energią potencjalną oddziaływania $U(x)$. Przy tym, ponieważ całkowita energia

$$E = \frac{p^2}{2m} + U(x), \quad (11.15)$$

pozostaje stała (stany stacjonarne), wzrostowi energii potencjalnej U ze wzrostem x towarzyszyć będzie zmniejszenie pędu \vec{p} z odpowiednim zwiększeniem długości fali. Wobec tego, funkcji falowej powinna odpowiadać zmieniająca się długość fali. Na rys. 11.8b pokazana jest funkcja falowa, której długość fali zwiększa się ze wzrostem x . Dokładną postać funkcji falowej $\Psi(x)$ ze zmieniającą się długością fali można znaleźć rozwiązując równanie różniczkowe zwane równaniem Schrödingera. Znajdziemy to równanie dla przypadku kiedy $U(x)$ można aproksymować funkcją schodkową przedstawioną na rys. 11.8c.

W ogólnym przypadku funkcja falowa cząstki w studni ma postać

$$\Psi = A \sin(kx + \varphi),$$

gdzie pęd p można otrzymać ze związku $E = p^2/2m + U_1$

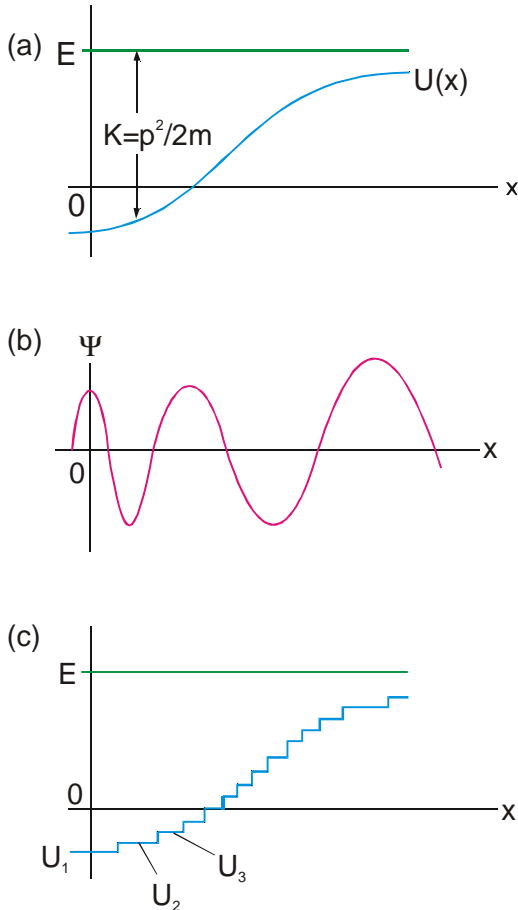
$$p = \sqrt{2m(E - U_1)},$$

$$k = \sqrt{\frac{2m}{\hbar^2}(E - U_1)}.$$

Wobec tego druga pochodna Ψ ma postać

$$\frac{d^2\Psi}{dx^2} = -k^2 A \sin(kx + \varphi) = -k^2\Psi,$$

$$\frac{d^2\Psi}{dx^2} = -\frac{2m}{\hbar^2}[E - U_1]\Psi.$$



Rys. 11.8. (a) Ze wzrostem x wzrasta energia potencjalna, a k zmniejsza się; (b) odpowiadająca funkcja falowa $\Psi(x)$, której długość fali wzrasta z x ; (c) aproksymacja funkcją schodkową funkcji $U(x)$ przedstawionej na rys. (a).

Ponieważ kwadrat modułu funkcji falowej określa gęstość prawdopodobieństwa znalezienia cząstki, więc tylko te rozwiązania równania Schrödingera mają sens fizyczny i są funkcjami falowymi, które tak jak to prawdopodobieństwo są: jednoznaczne, ciągłe wraz z pierwszymi pochodnymi oraz nie rosną nieograniczenie w nieskończoność. Jeżeli cząstka zamknięta jest w studni potencjalnej, to prawdopodobieństwo jej znalezienia na zewnątrz jest zerowe; tak więc w tym przypadku warunek graniczny znalezienia cząstki przy dużych wartościach $|x|$ jest zerowy.

Temu warunkowi granicznemu odpowiadają jedynie określone wartości E (będziemy je oznaczać przez E_n) i odpowiadające im funkcje falowe Ψ_n . Wartości energii E_n nazywamy wartościami własnymi, a odpowiadające im funkcje falowe Ψ_n – funkcjami własnymi.

Równanie to jest słuszne dla obszaru U_1 . Ponieważ to równanie jest również słuszne dla U_2, U_3, \dots, U_j , a dowolną funkcję $U(x)$ można przedstawić w postaci doboru małych "schodków", to U_j można zamienić na $U(x)$.

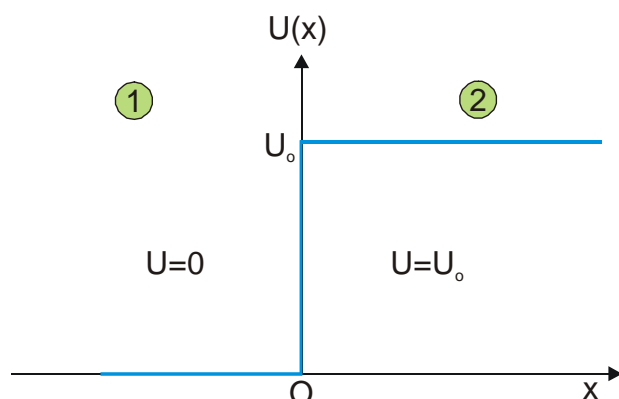
$$\frac{d^2\Psi}{dx^2} = -\frac{2m}{\hbar^2}[E - U(x)]\Psi. \quad (11.16)$$

Jest to znane stacjonarne, jednowymiarowe równanie Schrödingera. Jest ono słuszne w układach nierelatywistycznych pod warunkiem, że rozkład prawdopodobieństwa nie zmienia się w czasie; inaczej mówiąc, jest ono słuszne w przypadkach kiedy funkcje mają postać fal stojących. Istnieje także niestacjonarne, czyli zależne od czasu, równanie Schrödingera stosowane przy rozwiązywaniu zadań, w których paczka falowa zmienia się w czasie.

Rozważymy kilka przypadków cząstki w studni potencjału. W celu znalezienia stanów stacjonarnych (fal stojących) stosujemy stacjonarne równanie Schrödingera.

11.6. Bariera potencjału. Efekt tunelowy

Przedyskutujmy teraz jednowymiarowy ruch cząstek w obszarze, w którym energia potencjalna zmienia się skokowo. Na początek rozpatrzmy skok potencjału przedstawiony na rys. 11.9



Rys. 11.9. Skok potencjału.

$$U(x) = \begin{cases} 0 & \text{dla } x < 0 \\ U_0 & \text{dla } x > 0 \end{cases}$$

W praktyce nigdy nie ma dokładnie prostokątnego skoku potencjału. Jednakże model ten jest dobrym przybliżeniem wielu sytuacji fizycznych, np. skoku potencjału istniejącego na powierzchni metalu (patrz rys. 10.5).

Niech cząstka porusza się z lewa na prawo wzdłuż osi x i założmy, że $E > U_0$. Według praw mechaniki klasycznej, w punkcie $x = 0$ na cząstkę będzie działać siła opóźniająca $F = -dU/dx$, i cząstka będzie poruszała się ze zmniejszoną prędkością. Zobaczymy, że w przypadku kwantowym jest inaczej. Istnieje pewne prawdopodobieństwo odbicia od progu.

Dla $x < 0$ równanie Schrödingera i jego rozwiązanie mają postać

$$\frac{d^2\Psi_1}{dx^2} + \frac{2m}{\hbar^2} E \Psi_1 = 0,$$

$$\Psi_1(x) = A_1 e^{ik_1 x} + B_1 e^{-ik_1 x}, \quad (11.17)$$

gdzie

$$k_1 = \sqrt{\frac{2m}{\hbar^2} E}.$$

Natomiast dla $x > 0$ mamy

$$\frac{d^2\Psi_2}{dx^2} + \frac{2m}{\hbar^2} (E - U_0) \Psi_2 = 0,$$

$$\Psi_2(x) = A_2 e^{ik_2 x} + B_2 e^{-ik_2 x}, \quad (11.18)$$

gdzie

$$k_2 = \sqrt{\frac{2m}{\hbar^2} (E - U_0)}.$$

W funkcji Ψ_1 [wzór (11.17)] człon $A_1 e^{ik_1 x}$ przedstawia falę propagującą się w kierunku dodatnich wartości osi x (falę padającą), a człon $B_1 e^{-ik_1 x}$ – falę propagującą się w kierunku ujemnych wartości osi x (falę odbitą). Ponieważ dla $x > 0$ nie ma skoku potencjału, nie ma więc fizycznych powodów pojawienia się fali odbitej i dlatego $B_2 = 0$. Z warunku ciągłości funkcji falowej Ψ i jej pochodnej $d\Psi/dx$ w punkcie $x = 0$, mamy

$$\begin{aligned} \Psi_1(0) &= \Psi_2(0) & ; & & A_1 + B_1 &= A_2 \\ \left(\frac{d\Psi_1}{dx} \right)_{|x=0} &= \left(\frac{d\Psi_2}{dx} \right)_{|x=0} & : & & k_1(A_1 - B_1) &= k_2 A_2 \end{aligned}$$

Wyrażając B_1 i A_2 za pomocą A_1 i podstawiając do wzorów (11.17) i (11.18) otrzymujemy

$$\begin{aligned} \Psi_1 &= A_1 e^{ik_1 x} + A_1 \frac{k_1 - k_2}{k_1 + k_2} e^{-ik_1 x}, \\ \Psi_2 &= A_1 \frac{2k_1}{k_1 + k_2} e^{ik_2 x}. \end{aligned}$$

Możemy teraz obliczyć współczynnik transmisji T , który zdefiniujemy jako stosunek gęstości strumienia cząstek przechodzących do gęstości strumienia cząstek padających. Klasycznie gęstość strumienia cząstek jest to liczba cząstek przechodzących w jednostce czasu przez jednostkową powierzchnię prostopadłą do kierunku wiązki i jest równa iloczynowi gęstości cząstek przez ich prędkość. Ponieważ w ujęciu kwantowym odpowiednikiem gęstości cząstek jest gęstość prawdopodobieństwa $|\Psi|^2$, dlatego współczynnik transmisji T wyniesie

$$T = \frac{v_2 |A_2|^2}{v_1 |A_1|^2}, \quad (11.19)$$

gdzie v_1 i v_2 są prędkościami cząstki w obszarze 1 i 2.

Ponieważ

$$v_1 = \frac{p_1}{m} = \frac{\hbar k_1}{m} \quad ; \quad v_2 = \frac{p_2}{m} = \frac{\hbar k_2}{m},$$

więc ostatecznie

$$T = \frac{k_2}{k_1} \left(\frac{2k_1}{k_1 + k_2} \right)^2 = \frac{4k_1 k_2}{(k_1 + k_2)^2} = \frac{4 \sqrt{\frac{E - U_o}{E}}}{\left(1 + \sqrt{\frac{E - U_o}{E}} \right)^2}. \quad (11.20)$$

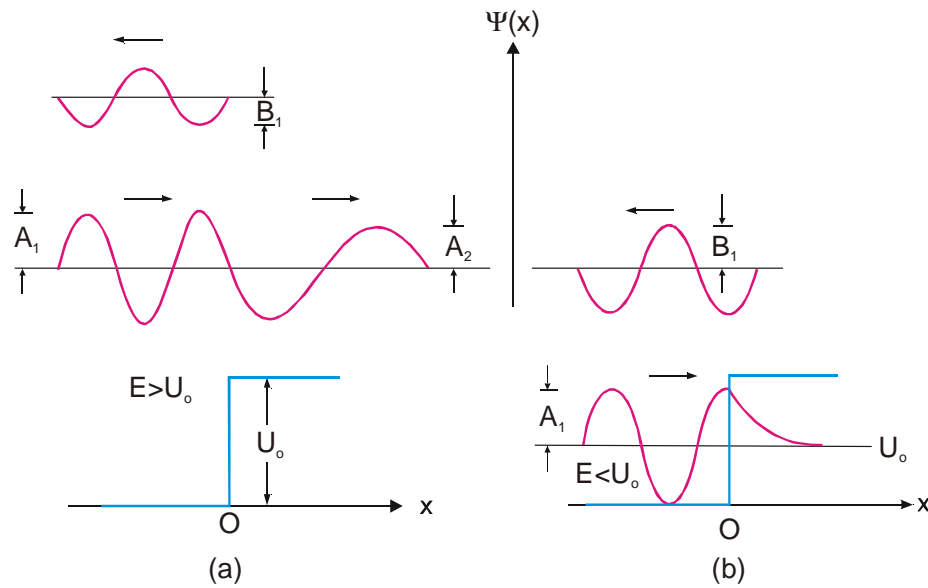
Podobnie obliczamy współczynnik odbicia R

$$R = \frac{v_1 B_1^2}{v_2 A_1^2} = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2} = \frac{\left(1 - \sqrt{\frac{E - U_o}{E}}\right)^2}{\left(1 + \sqrt{\frac{E - U_o}{E}}\right)^2}.$$

Oczywiście musi zachodzić związek $R + T = 1$, co jak łatwo sprawdzić jest w naszym przypadku spełnione.

Jeżeli uwzględnimy falowy charakter cząstek, to wynik ten nie powinien być zaskoczeniem. Gdy fale świetlne padają na granicę między dwoma przezroczystymi ośrodkami, to część światła odbija się, a reszta przechodzi do drugiego ośrodka.

Odbicie cząstek powoduje, że np. w zjawisku fotoemisji elektrony mogą zostać zawrócone do metalu, nawet jeśli energia jest większa od pracy wyjścia. Może to prowadzić do zmniejszenia czułości fotokomórek, zwłaszcza dla światła o częstotliwości niewiele większej od częstotliwości progowej.



Rys. 11.10. Funkcje falowe dla cząstek poruszających się w obszarze w którym występuje skok potencjału: (a) $E > U_o$, (b) $E < U_o$.

Rozpatrzmy teraz przypadek, gdy $E < U_o$. W obszarze pierwszym ($x < 0$) funkcja falowa jest taka sama jak poprzednio. Natomiast w obszarze drugim ($x > 0$) równanie Schrödingera i funkcja falowa mają postać:

$$\frac{d^2 \Psi_2}{dx^2} - \frac{2m}{\hbar^2} (U_o - E) \Psi_2 = 0,$$

$$\Psi_2(x) = A_3 e^{\lambda x} + B_3 e^{-\lambda x}, \quad (11.21)$$

gdzie

$$\chi = \sqrt{\frac{2m}{\hbar^2}(U_o - E)}.$$

Ponieważ Ψ_2 nie może wzrastać nieograniczenie, należy przyjąć, że $A_3 = 0$.

Korzystając ponownie z warunków ciągłości funkcji falowej i jej pochodnych w punkcie $x = 0$, otrzymujemy

$$B_1 = -\frac{k_1 - i\chi}{k_1 + i\chi} A_1,$$

$$B_3 = \frac{2k_1}{k_1 + i\chi} A_1.$$

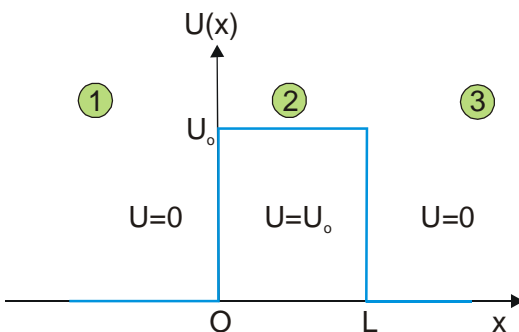
Współczynnik odbicia

$$R = \frac{B_1 B_1^*}{A_1 A_1^*} = 1.$$

Zgodnie ze wzorem (11.21) fala wchodząca do obszaru drugiego ($x > 0$) jest wykładniczo tłumiona i gęstość prawdopodobieństwa jest proporcjonalna do $\exp(-2\chi x)$. Na głębokości $x_o = 1/2\chi$ gęstość prawdopodobieństwa maleje e razy; jest to tzw. efektywna głębokość przenikania cząstek przez barierę. Na taką odległość oddalają się np. elektrony od powierzchni metalu, jeżeli ich energia kinetyczna jest mniejsza o 0,01 eV od skoku energii potencjalnej na powierzchni metalu. Według mechaniki klasycznej cząstka o energii mniejszej od wysokości bariery nie może znaleźć się w obszarze tej bariery. Uzupełnieniem przeprowadzonych rozważań jest rys. 11.10.

Rozpatrzmy teraz przypadek, kiedy cząstki padają na barierę o skończonej grubości (rys. 11.11). Energia potencjalna zmienia się zgodnie z zależnością

$$U(x) = \begin{cases} 0 & \text{dla } x < 0 \\ U_o & \text{dla } 0 < x < l \\ 0 & \text{dla } x > l \end{cases}$$



Rys. 11.11. Bariera potencjału o skończonej szerokości.

Przeprowadzone powyżej rozważania pozwalają przypuszczać, że jeżeli $E < U_o$, to nastąpi przenikanie cząstek przez barierę, natomiast dla $E > U_o$ nastąpi odbicie. Przenikanie cząstek przez barierę o skończonej grubości, gdy $E < U_o$, nosi nazwę efektu tunelowego. Rozpatrzmy właśnie taki przypadek.

Równanie Schrödingera w obszarze 1 i 3 ma postać

$$\frac{d^2\Psi}{dx^2} + \frac{2m}{\hbar^2} E\Psi = 0,$$

natomiast w obszarze drugim

$$\frac{d^2\Psi}{dx^2} - \frac{2m}{\hbar^2} (U_o - E)\Psi = 0.$$

Rozwiązaniami w poszczególnych obszarach są funkcje

$$\Psi_1 = A_1 e^{ikx} + B_1 e^{-ikx} \quad \text{w obszarze 1}$$

$$\Psi_2 = A_2 e^{\chi x} + B_2 e^{-\chi x} \quad \text{w obszarze 2}$$

$$\Psi_3 = A_3 e^{ikx} \quad \text{w obszarze 3}$$

gdzie

$$k = \sqrt{\frac{2m}{\hbar^2} E} \quad ; \quad \chi = \sqrt{\frac{2m}{\hbar^2} (U_o - E)}.$$

Korzystając z warunku ciągłości Ψ i $d\Psi/dx$ otrzymujemy dla $x = 0$

$$A_1 + B_1 = A_2 + B_2,$$

$$ik(A_1 - B_1) = \chi(A_2 - B_2),$$

natomiast dla $x = l$

$$A_2 e^{\chi l} + B_2 e^{-\chi l} = A_3 e^{ikl},$$

$$\chi(A_2 e^{\chi l} - B_2 e^{-\chi l}) = ikA_3 e^{ikl}.$$

Rozwiązując powyższy układ równań możemy wyznaczyć A_3 przez A_1 . Elementarne obliczenia prowadzą do wzoru

$$\frac{A_3}{A_1} = \frac{4i\chi k e^{ikl}}{(k + i\chi)^2 e^{\chi l} - (k - i\chi)^2 e^{-\chi l}}, \quad (11.22)$$

Możemy teraz wyliczyć współczynnik transmisji

$$T = \frac{v_3}{v_1} \frac{A_3 A_3^*}{A_1 A_1^*} = \frac{A_3 A_3^*}{A_1 A_1^*},$$

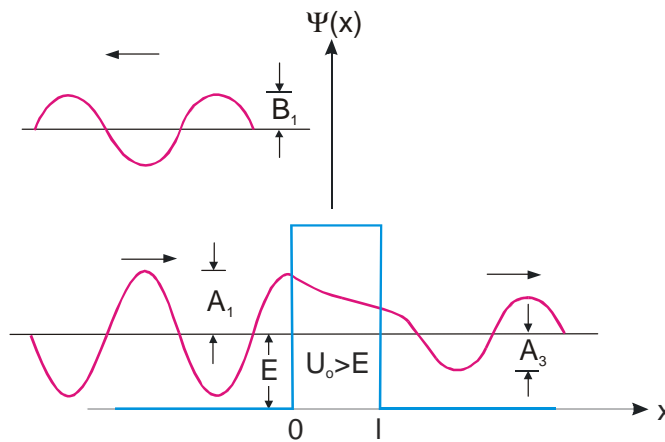
ponieważ prędkości w obszarze 1 i 3 są jednakowe. Uwzględniając wzór (11.22) otrzymujemy

$$T = \frac{16k^2 \chi^2}{(k^2 + \chi^2)^2 (e^{2\chi l} + e^{-2\chi l} - 2) + 16k^2 \chi^2}. \quad (11.23)$$

Bardzo często spełniony jest warunek $\chi l \gg 1$. Wówczas z dobrym przybliżeniem

$$T = \frac{16k^2 \chi^2}{(k^2 + \chi^2)^2} e^{-2\chi l} = 16 \frac{E}{U_o} \left(1 - \frac{E}{U_o}\right) \exp\left[-\frac{2l}{\hbar} \sqrt{2m(U_o - E)}\right]. \quad (11.24)$$

Ze wzoru tego wynika, że prawdopodobieństwo przenikania bardzo szybko maleje wraz ze wzrostem szerokości bariery.



Rys. 11.12. Funkcja falowa dla cząstek o energii $E < U_o$ padających z lewej strony na barierę potencjału o skończonej szerokości.

Dla przykładu rozważmy wiązkę elektronów o energii $E = 8 \text{ eV}$ padającą na barierę o wysokości $U_o = 10 \text{ eV}$ i szerokości $l = 2 \times 10^{-10} \text{ m}$. Wówczas ze wzoru (11.24) otrzymamy $T = 0.12$ i stąd $R = 0.88$. Przy szerokości bariery $l = 5 \times 10^{-10} \text{ m}$ współczynnik transmisji wynosi już tylko 0.01.

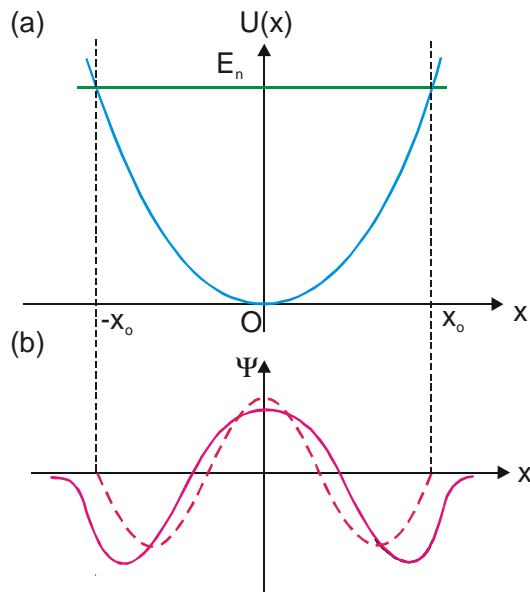
Na rys.11.12 przedstawiono funkcje falowe w poszczególnych obszarach.

11.7. Oscylator harmoniczny

Przypadek oscylatora harmonicznego jest bardzo ważny w fizyce, gdyż wiele układów można z dobrym przybliżeniem traktować właśnie jak oscylatory harmoniczne. Z modelu oscylatora korzystaliśmy już przy omawianiu promieniowania termicznego ciał.

Przypomnijmy, że w przypadku klasycznym oscylatora harmonicznego, energia potencjalna $U(x) = kx^2/2 = m\omega_{kl}^2 x^2/2$. Cząstka wykonuje drgania harmoniczne o częstości kołowej $\omega_{kl} = \sqrt{k/m}$, a jej energia może przyjmować dowolne wartości, w tym również wartość zerową (indeks "kl" oznacza "klasyczny"). Jak się przekonamy, w mechanice kwantowej istnienie

warunków granicznych przy dużych $|x|$ prowadzi do tego, że dozwolony jest jedynie zbiór energii $E_n = (n - 1/2)\hbar\omega_{kl}$, gdzie n jest dodatnią liczbą całkowitą.



Rys. 11.13: (a) Energia potencjalna oscylatora harmonicznego. (b) Fala stojąca trzeciego rzędu (linia ciągła – rozwiązanie dokładne, linia przerywana – rozwiązanie przybliżone).

Przed bezpośrednim zastosowaniem równania Schrödingera, przeprowadzimy przybliżone obliczenia. Zilustrują one podstawowe metody obliczeń kwantowomechanicznych. Spróbujmy rozwiązać to zadanie analogicznie do zadania z cząstką w studni. Na rys. 11.13 widać, że długość studni wynosi $L = 2x_0$, gdzie x_0 jest maksymalnym wychyleniem czyli amplitudą drgań. W przypadku fali stojącej n -tego rzędu na odcinku o długości L , nałoży się n półfal

$$\frac{n\lambda_n}{2} = 2x_0.$$

lub

$$\lambda_n = \frac{4x_0}{n}.$$

Średnia wartość pędu

$$\langle p \rangle = \frac{h}{\lambda} = \frac{h}{\frac{4x_0}{n}} = n \frac{h}{4x_0}.$$

Średnia energia kinetyczna

$$\langle K \rangle = \frac{\langle p \rangle^2}{2m} = \frac{n^2 h^2}{32m x_0^2}.$$

Całkowita energia E jest dwa razy większa

$$E = \frac{n^2 h^2}{16m x_0^2}$$

i jest także równa maksymalnej energii potencjalnej

$$E = \frac{1}{2} m \omega_{kl}^2 x_0^2.$$

Mnożąc dwa ostatnie równania, otrzymujemy

$$E^2 = \frac{n^2 h^2 \omega_{kl}^2}{32},$$

lub

$$E = \frac{\pi}{2\sqrt{2}} n\hbar\omega_{kl}.$$

W granicach dokładności naszych "obliczeń" można przyjąć, że $\pi/2\sqrt{2} \approx 1$ i wówczas

$$E_n \approx n\hbar\omega_{kl},$$

gdzie $n = 1, 2, 3, \dots$

Przeprowadzimy teraz dokładne obliczenia stosując równanie Schrödingera. Podstawiając do równania Schrödingera $U = (1/2)m\omega_{kl}^2 x^2$, otrzymujemy

$$\frac{d^2\Psi}{dx^2} = -\frac{2m}{\hbar^2} \left(E - \frac{1}{2} m\omega_{kl}^2 x^2 \right) \Psi. \quad (11.25)$$

Ponieważ mamy znaleźć rozwiązanie ze zmieniającą się długością fali, spróbujemy zastosować funkcję Gaussa. Jeżeli odgadujemy rozwiązanie, sprawdzamy czy jest ono spełnione. Takie postępowanie jest ogólnie przyjętym sposobem rozwiązywania problemów. Wobec tego przyjmujemy funkcję falową w postaci

$$\Psi(x) = e^{-ax^2}.$$

Obliczając drugą pochodną

$$\frac{d^2\Psi}{dx^2} = -2ae^{-ax^2} + 4a^2 x^2 e^{-ax^2}$$

i podstawiając do równania (11.25), mamy

$$\left(-2a + 4a^2 x^2 \right) e^{-ax^2} = -\frac{2m}{\hbar^2} \left(E - \frac{1}{2} m\omega_{kl}^2 x^2 \right) e^{-ax^2},$$

$$-2a + (4a^2)x^2 = -\frac{2mE}{\hbar^2} + \left(\frac{m^2\omega_{kl}^2}{\hbar^2} \right) x^2.$$

Przyrównując współczynniki przy x^2 , mamy

$$4a^2 = \frac{m^2\omega_{kl}^2}{\hbar^2},$$

$$a = \frac{m\omega_{kl}}{2\hbar}.$$

Z porównania członów wolnych wynika

$$-2a = -\frac{2mE}{\hbar^2},$$

$$E = \frac{\hbar^2 a}{m} = \frac{\hbar^2}{m} \frac{m\omega_{kl}}{2\hbar} = \frac{\hbar\omega_{kl}}{2}.$$

Widzimy, że funkcja Gaussa jest rozwiązaniem pod warunkiem, że

$$E_1 = \frac{1}{2} \hbar\omega_{kl}.$$

W tym przypadku

$$\Psi_1(x) = e^{-(m\omega_{kl}/2\hbar)x^2}.$$

Przez podstawienie do równania (11.25) można przekonać się, że rozwiązaniem odpowiadającym fali stojącej drugiego rzędu będzie

$$\Psi_2(x) = xe^{-(m\omega_{kl}/2\hbar)x^2}.$$

Funkcja ta jest rozwiązaniem tylko wtedy gdy

$$E_2 = \frac{3}{2} \hbar\omega_{kl}.$$

Należy zaznaczyć, że odległość energetyczna pomiędzy dwoma sąsiednimi poziomami wynosi $E_2 - E_1 = \hbar\omega_{kl}$. Odnosi się to również do bardziej wysokich poziomów energetycznych opisanych wyrażeniem

$$E_n = \left(n - \frac{1}{2}\right) \hbar\omega_{kl}, \quad \text{gdzie } n = 1, 2, 3, \dots \quad (11.26)$$

Widzimy więc, że oscylator harmoniczny emituje fotony, których częstotliwość jest zgodna z częstotliwością oscylatora klasycznego w przypadku kiedy przejścia zachodzą pomiędzy sąsiednimi poziomami.

Jeżeli mamy zbiór oscylatorów harmonicznych, jak to jest np. w ciele stałym, to wymiana energii między nimi może odbywać tylko kwantami o energii $\hbar\omega_{kl}$. Widmo wartości energii jest dyskretne. Zauważmy ponadto, że oscylator kwantowy nie może mieć energii równej zero. Najmniejszej energii równej $\hbar\omega_{kl}/2$ nie może przekazać na zewnątrz; oscylator harmoniczny posiada ją nawet w temperaturze zera bezwzględnego. Jest to konsekwencją zasady nieoznaczoności, gdyby bowiem energia była zerowa i drgania nie występowały, to byłby możliwy jednoczesny pomiar położenia i pędu.

ROZDZIAŁ 12

ATOM WODORU

W 1911 r. E. Rutheford wykonując słynne eksperymenty nad rozpraszaniem cząstek α na bardzo cienkiej folii złota stwierdził, że atom składa się z bardzo małego, dodatnio naładowanego jądra oraz z chmury ujemnie naładowanych elektronów. Rozmiary jądra są rzędu 10^{-14} m, rozmiary atomu – rzędu 10^{-10} m. Masa jądra jest tysiące razy większa od masy elektronu (masa protonu = 1836 mas elektronu swobodnego). Aby w wyniku działania sił elektrostatycznych elektron nie spadł na jądro, należy przyjąć, że krąży on po orbicie kołowej. Według mechaniki klasycznej energia elektronu może przyjmować dowolne wartości. Tymczasem analiza widmowa promieniowania emitowanego przez wodór wykazuje, że poziomy energetyczne elektronu muszą mieć nieciągłą strukturę. Również założenie o ruchu elektronu po orbicie kołowej jest sprzeczne z elektrodynamiką klasyczną. Wiadomo bowiem, że ładunek poruszający się ruchem niejednostajnym promieniuje energię. Elektron więc powinien tracić energię i poruszając się po spirali spaść na jądro.

12.1. Postulaty Bohra

Pierwszą próbą teoretycznego opisu nieciągłości poziomów energetycznych elektronu w atomie była teoria atomu Bohra. Bohr sądził, że możliwe orbity elektronów są analogiczne do klasycznych kołowych orbit planet i usiłował znaleźć zasadę, która dopuszczałaby jedynie określone wartości energii lub promieni orbit. Oparł swój model na następujących postulatach:

- I. *Elektrony poruszają się w atomach nie promieniując energii, po takich orbitach kołowych, że moment pędu elektronu jest równy całkowitej wartości stałej \hbar*

$$mvr = n\hbar \quad n = 1, 2, 3, \dots \quad (12.1)$$

- II. *Przejścia elektronu z orbity o energii E_n na orbitę, gdzie energia wynosi E_m , towarzyszy emisja lub absorpcja fotonu o częstotliwości ν określonej wzorem*

$$|E_n - E_m| = h\nu \quad (12.2)$$

Zwróćmy uwagę, że model Bohra nie tłumaczy, dlaczego elektron krążący po orbicie nie promieniuje energii. Jest to przyjmowane jako postulat, aby wyjaśnić fakty eksperymentalne.

Z pierwszego postulatu Bohra otrzymujemy wyrażenie na dopuszczalne poziomy energetyczne elektronu w atomie wodoru. Ponieważ $v = n\hbar/mr$, więc energia kinetyczna

$$K = \frac{mv^2}{2} = \frac{1}{2} m \left(\frac{n\hbar}{mr} \right)^2 = \frac{e^2}{8\pi\epsilon_0 r} \quad (12.3)$$

Ostatnie wyrażenie otrzymano uwzględniając równość siły elektrostatycznej z siłą odśrodkową. Z wyrażenia (12.3) wyliczymy dozwolone orbity

$$r_n = \frac{4\pi\epsilon_0 \hbar^2}{me^2} n^2 = r_0 n^2 \quad (12.4)$$

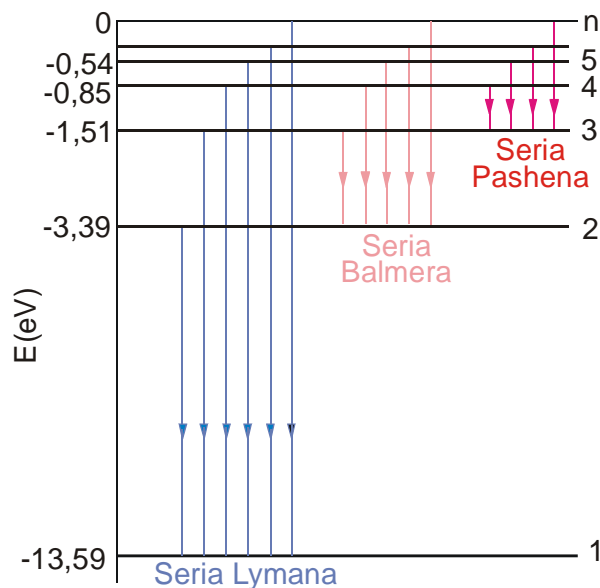
gdzie $r_0 = 4\pi\epsilon_0 \hbar^2 / me^2 = 5,29 \times 10^{-11}$ m jest promieniem Bohra. Promienie orbit są więc skwantowane i wynoszą: $r_0, 4r_0, 9r_0, \dots$, itd.

Natomiast dopuszczalne wartości energii wynoszą

$$E = K + U = \frac{e^2}{8\pi\epsilon_0 r} - \frac{e^2}{4\pi\epsilon_0 r} = -\frac{e^2}{8\pi\epsilon_0 r}$$

stąd

$$E_n = -\frac{me^4}{32\pi^2 \epsilon_0^2 \hbar^2} \frac{1}{n^2} = -13,59 \frac{1}{n^2} \text{ eV} \quad (12.5)$$



Rys. 12.1. Schemat poziomów energetycznych atomu wodoru.

Aby elektron przesunąć do nieskończoności, gdzie miałby energię zerową, trzeba dostarczyć energię 13,59 eV przyjmując, że elektron znajduje się w stanie podstawowym $n = 1$. Jest to energia jonizacji atomu, tzn. energia potrzebna do tego, aby elektron wyrwać z atomu. Energię potrzebną do przeniesienia elektronu ze stanu podstawowego do stanu wyższego (wzbudzonego) nazywamy energią wzbudzenia. Na rys. 12.1 przedstawiono schemat dozwolonych poziomów energetycznych dla atomu wodoru. Zauważmy, że gdy n rośnie, poziomy leżą coraz bliżej siebie i różnica między nimi jest

bardzo mała. Kiedy elektron odrywa się od jądra ($r \rightarrow \infty$), staje się elektronem swobodnym i może przyjmować dowolną wartość energii większą od zera.

Chociaż teoria Bohra jest przestarzała, jednak jest bardzo prosta. Jej znaczenie historyczne jest duże i z tego powodu ją rozważaliśmy. Symbole teorii bohrowskiej stosowane są do chwili obecnej. Współczesny model atomu był zaproponowany w 1926 r. wkrótce po sformułowaniu

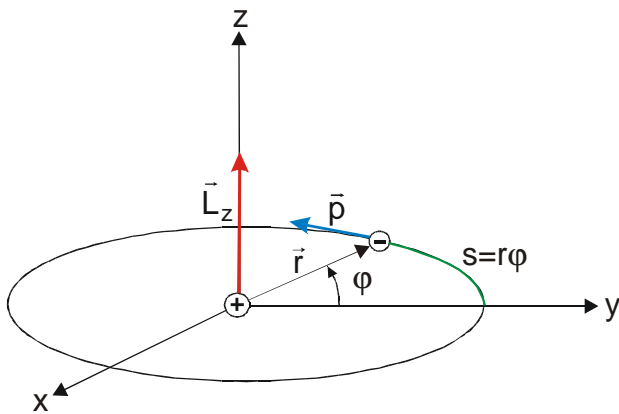
równania Schrödingera. Szczegółowe omówienie teorii atomu w ujęciu kwantowym jest zagadnieniem skomplikowanym. Z tego powodu nasze podejście będzie raczej opisowe. Położymy nacisk na interpretację fizyczną wyników, nie będziemy natomiast zajmować się trudnymi obliczeniami matematycznymi. Nasze rozważania rozpoczniemy od omówienia orbitalnego momentu pędu elektronu w atomie.

12.2. Orbitalny moment pędu

Aby wyjaśnić widmo liniowe wodoru, Bohr przyjął pewne założenia, które na gruncie mechaniki klasycznej nie mają teoretycznego uzasadnienia. Można bowiem zapytać:

- dlaczego moment pędu elektronu jest skwantowany?
- dlaczego elektron nie emituje promieniowania i nie spada na jądro?

Odpowiemy na te pytania poniżej.



Rys. 12.2. Paczka falowa poruszająca się po okręgu o promieniu r .

Założmy, że elektron (paczka falowa) porusza się po orbicie kołowej o promieniu r . Wówczas moment pędu elektronu względem osi z wynosi

$$L_z = r \hbar k$$

gdzie uwzględniliśmy, że $p = \hbar k$. Ponieważ droga jaką przebył elektron na orbicie wynosi $s = r d\varphi$, to funkcja falowa opisująca stan elektronu może być zapisana w postaci

$$\Psi(r) = \Psi_o e^{iks} = \Psi_o e^{ikr\varphi}$$

Po wykonaniu jednego obiegu po orbicie elektron znajduje się w tym samym miejscu

$$\Psi(\varphi) = \Psi(\varphi + 2\pi)$$

lub

$$\Psi_o e^{ikr\varphi} = \Psi_o e^{ikr(\varphi + 2\pi)}$$

Równanie to będzie spełnione jeżeli

$$e^{ikr2\pi} = 1 \tag{12.6}$$

Ponieważ $\exp(2\pi ikr) = \cos(2\pi kr) + i \sin(2\pi kr)$, więc z równania (12.6) wynika, że iloczyn kr musi być liczbą całkowitą, którą oznaczymy przez m_l

$$kr = m_l \qquad m_l = 0, \pm 1, \pm 2, \dots$$

Mnożąc obustronnie przez \hbar otrzymujemy

$$\hbar k r = m_l \hbar$$

czyli

$$L_z = m_l \hbar \quad (12.7)$$

Wykazaliśmy w ten sposób, że orbitalny moment pędu względem osi z (lub innego ustalonego kierunku) jest skwantowany i może przyjmować wartości $0, \pm \hbar, \pm 2\hbar, \dots$, itd. Istnieje ścisły dowód tego twierdzenia, jednakże wychodzi to poza ramy tego skryptu. Ponieważ składowa momentu pędu nie może być większa od całkowitego momentu pędu, więc przy ustalonej wartości całkowitego momentu pędu istnieje pewna maksymalna wartość m_l . Oznaczmy ją przez l . Między l i m_l zachodzi warunek $|m_l| \leq l$. Z powyższego wynika, że liczba m_l może przybierać $2l+1$ wartości: $-l, -l+1, \dots, 0, 1, \dots, l-1, l$.

Rozpatrzmy teraz zbiór elektronów, które mają ten sam orbitalny moment pędu L , ale różne wartości L_z . Obliczmy średnią wartość kwadratu momentu pędu wyrażając ją przez składowe

$$\langle L^2 \rangle = \langle L_x^2 \rangle + \langle L_y^2 \rangle + \langle L_z^2 \rangle$$

Ponieważ żaden kierunek w przestrzeni nie jest wyróżniony, więc $\langle L_x^2 \rangle = \langle L_y^2 \rangle = \langle L_z^2 \rangle$. Ponadto $\langle L^2 \rangle = L^2$. Wynika to z faktu, że L^2 jest wielkością skalarną i wszystkie elektrony mają jednakowy moment pędu. Ostatecznie więc otrzymujemy:

$$L^2 = 3 \langle L_z^2 \rangle = 3 \frac{\sum_{m_l=-l}^l m_l^2 \hbar^2}{2l+1} = 3 \frac{\hbar^2}{2l+1} \sum_{m_l=-l}^l m_l^2$$

Ponieważ

$$\sum_{m_l=-l}^l m_l^2 = 2(1^2 + 2^2 + \dots + l^2) = \frac{l(l+1)(2l+1)}{6}$$

więc

$$L^2 = 3 \frac{\hbar^2}{2l+1} 2 \frac{l(l+1)(2l+1)}{6} = l(l+1)\hbar^2$$

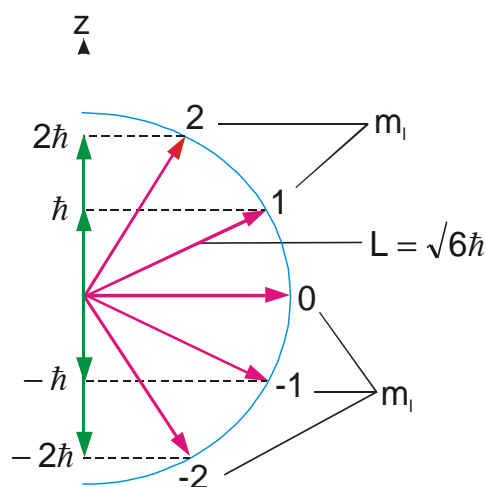
lub

$$L = \sqrt{l(l+1)}\hbar \quad (12.8)$$

gdzie $l = 0, 1, 2, \dots$ itd.

Wykazaliśmy więc, że orbitalny moment pędu i jego rzut na ustalony kierunek są skwantowane. Ilość możliwych ustaleń wektora L wynosi $2l+1$, przy czym kąty między L_z i L spełniają warunek

$$\cos \theta = \frac{L_z}{L} = \frac{m_l}{\sqrt{l(l+1)}} \quad (12.9)$$



Rys. 12.3. Przestrzenne kwantowanie orbitalnego momentu pędu. W przypadku przedstawionym na rysunku $l = 2$ i $L = \sqrt{6}\hbar$. Rzuty wektora L na wyróżniony kierunek wynoszą: $2\hbar, \hbar, 0, -\hbar, -2\hbar$.

Zapamiętajmy także, że funkcja falowa związana z magnetyczną liczbą falową ma postać

$$\Psi = \Psi_o e^{im_l \varphi}$$

przy czym stałą Ψ_o wyznaczymy z warunku normalizacji

$$\int_0^{2\pi} \Psi \Psi^* d\psi = |\Psi_o|^2 \int_0^{2\pi} e^{im_l \varphi} e^{-im_l \varphi} d\varphi = |\Psi_o|^2 2\pi = 1$$

Zatem $\Psi_o = (\sqrt{2\pi})^{-1}$. Ostatecznie więc

$$\Psi = \frac{1}{\sqrt{2\pi}} e^{im_l \varphi} \quad (12.10)$$

12.3. Równanie Schrödingera dla atomu wodoru

Pokażemy teraz jak korzystając z równania Schrödingera można wyznaczyć poziomy energetyczne elektronu w atomie wodoru.

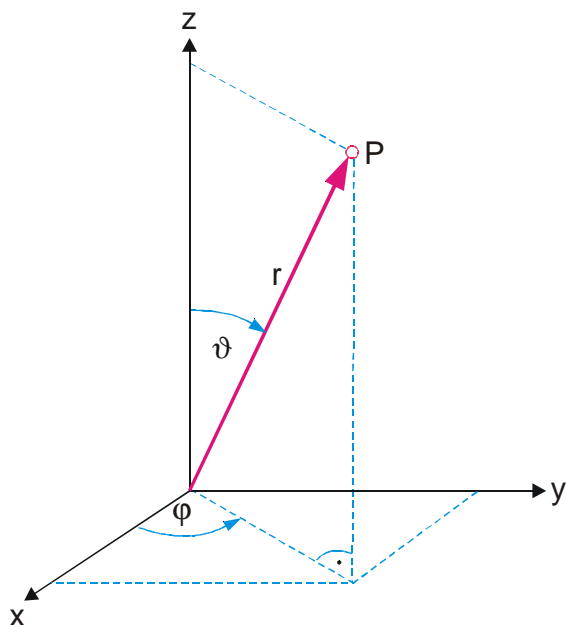
Energia potencjalna oddziaływania elektron-jądro ma postać

$$U(r) = -\frac{e^2}{4\pi\epsilon_0 r}$$

Na rys. 12.3 przedstawiono przykładowo możliwe ustawienia wektora L dla przypadku gdy $l = 2$. Mówimy często, że wektor orbitalnego momentu pędu jest skwantowany w przestrzeni. Aby mówić o pewnej składowej wektora L , musimy wyróżnić określony kierunek w przestrzeni. Jest to najczęściej kierunek pola magnetycznego (zewnętrznego lub własnego, wytworzonego przez atom). Przyjmiemy dalej, że ten wyróżniony kierunek pokrywa się z osią z .

Liczbę l określającą wartość orbitalnego momentu pędu nazywamy orbitalną (azymutalną) liczbą kwantową, a liczbę m_l – magnetyczną liczbą kwantową.

Ponieważ potencjał ma symetrię sferyczną, wygodniej jest wprowadzić sferyczny układ współrzędnych. W układzie tym położenie określone jest przez promień wodzący r , kąt biegunowy ϑ i kąt azymutalny φ , co przedstawiono schematycznie na rys. 12.4.



Rys. 12.4. Kartezjańskie (x, y, z) i sferyczne (r, ϑ, φ) współrzędne punktu P .

Współrzędne sferyczne związane są ze współrzędnymi kartezjańskimi następującymi związkami

$$x = r \sin \vartheta \cos \varphi$$

$$y = r \sin \vartheta \sin \varphi$$

$$z = r \cos \vartheta$$

Przypomnijmy, że jądro znajduje się w początku układu współrzędnych i jest nieruchome (jego masa jest 1836 razy większa od masy elektronu).

Równanie Schrödingera dla przypadku trójwymiarowego

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2} = -\frac{2m}{\hbar^2} (E - U) \Psi \quad (12.11)$$

należy przekształcić do sferycznego układu współrzędnych.

Przejsie z równania (12.11) ze współzrędnymi x, y, z do układu sferycznego nie jest trudne lecz należy przeprowadzić żmudne przekształcenia. Dlatego napiszemy końcowy wynik

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial \Psi}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2 \Psi}{\partial \varphi^2} = -\frac{2m}{\hbar^2} (E - U) \Psi \quad (12.12)$$

Wobec tego równanie Schrödingera dla atomu wodoru posiada następującą postać

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Psi}{\partial r} \right) + \frac{1}{r^2} \left[\frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial \Psi}{\partial \vartheta} \right) + \frac{1}{\sin^2 \vartheta} \frac{\partial^2 \Psi}{\partial \varphi^2} \right] = -\frac{2m}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) \Psi \quad (12.13)$$

Równanie to jest skomplikowane i spróbujemy je rozwiązać, a właściwie odgadnąć rozwiązanie tylko w pewnych przypadkach.

W ogólnym przypadku funkcja falowa Ψ zależy od trzech zmiennych: r, ϑ, φ . Można założyć jednak, że istnieją pewne sytuacje fizyczne, że Ψ jest tylko funkcją r . Wówczas żaden kierunek w przestrzeni nie jest wyróżniony, a więc orbitalny moment pędu musi być równy zero. Stan ten nazywamy stanem s . Dla takiego stanu równanie Schrödingera ma postać

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Psi}{dr} \right) + \frac{2m}{\hbar^2} \left(E + \frac{e^2}{4\pi\epsilon_0 r} \right) \Psi = 0 \quad (12.14)$$

Funkcja Ψ musi spełniać pewne warunki: musimy żądać aby dla $r \rightarrow \infty, \Psi \rightarrow 0$. Ponadto, prawdopodobieństwo znalezienia elektronu musi osiągnąć maksimum w pewnej odległości od jądra i jednocześnie być równe zero dla $r = 0$, gdyż tam znajduje się jądro. Najprostsza funkcja spełniająca te warunki ma postać

$$\Psi(r) = \Psi_0 e^{-r/r_0} \quad (12.15)$$

gdzie r_0 jest pewną stałą mającą wymiar długości. Prawdopodobieństwo znalezienia elektronu w elemencie objętości $dV = 4\pi r^2 dr$ (objętość warstwy kulistej o promieniu r i grubości dr) wynosi:

$$p = \Psi^2(r) 4\pi r^2 dr = 4\pi r^2 e^{-2r/r_0} \Psi_0^2 dr$$

Widać, że $p(0) = 0$ i $p(\infty) = 0$, natomiast różniczkując wyrażenie $r^2 e^{-2r/r_0}$ przekonamy się, że p osiąga maksimum dla $r = r_0$. Zatem przyjęta funkcja spełnia nasze warunki. Podstawiając wyrażenie (12.15) do równania (12.14) i wykonując różniczkowanie otrzymujemy

$$\left(\frac{1}{r_0^2} - \frac{2}{r_0 r} + \frac{2m}{\hbar^2} E + \frac{me^2}{2\pi\epsilon_0 \hbar^2} \frac{1}{r} \right) \Psi_0 e^{-r/r_0} = 0$$

Ponieważ $\Psi_0 e^{-r/r_0} \neq 0$, więc

$$\frac{1}{r_0^2} - \frac{2}{r_0 r} + \frac{2m}{\hbar^2} E + \frac{me^2}{2\pi\epsilon_0 \hbar^2} \frac{1}{r} = 0$$

Równanie to będzie spełnione tylko wówczas, jeżeli niezależne wyrazy zawierające $1/r$ i wyrazy wolne będą równały się zero, co daje

$$-\frac{2}{r_0} + \frac{me^2}{2\pi\epsilon_0 \hbar^2} = 0$$

$$\frac{1}{r_0^2} + \frac{2mE}{\hbar^2} = 0$$

Z równań tych wyliczymy r_0 i E

$$r_0 = \frac{4\pi\epsilon_0 \hbar^2}{me^2} = 5,29 \times 10^{-11} \text{ m} \quad (12.16)$$

$$E = \frac{\hbar^2}{2mr_0^2} = -\frac{me^4}{32\pi^2 \epsilon_0^2 \hbar^2} = -13,59 \text{ eV} \quad (12.17)$$

W ten sposób stwierdzamy, że przyjęta przez nas funkcja będzie rozwiązaniem równania Schrödingera, jeżeli r_0 i E będą określone wzorami (12.16) i (12.17). Przypomnijmy w tym miejscu, że takimi samymi wzorami wyrażają się promień orbity i najniższy poziom

energetyczny elektronu w atomie wodoru wg modelu Bohra. Teraz jednak wzory na r i E otrzymaliśmy bez żadnych założeń, rozwiązując jedynie równanie Schrödingera. Zwróćmy też uwagę, że r_o będziemy interpretować jako tę odległość od jądra, przy której prawdopodobieństwo znalezienia elektronu osiągnie wartość maksymalną. Wracając do modelu Bohra należy zaznaczyć, że postulat (12.1) różni się od tego o czym dyskutowaliśmy powyżej z dwu powodów. Po pierwsze, że przyjęcie klasycznej orbity traci sens szczególnie dla elektronu, którego stan scharakteryzowany jest falą stojącą. Po drugie, teraz wiemy, że moment pędu równy jest nie $n\hbar$ lecz $l\hbar$ (jak się dalej przekonamy jest mniejszy od bohrowskiej wartości). Wobec tego fakt, że teoria Bohra prawidłowo opisuje widmo poziomów energetycznych atomu wodoru, jest szczęśliwym zbiegiem okoliczności. To, że prawdziwa odpowiedź może okazać się wynikiem nieprawidłowych rozważań, nie jest czymś niezwykłym.

Postępując w sposób analogiczny jak powyżej, można wykazać że funkcja

$$\Psi = \Psi_o \left(1 - \frac{r}{2r_o} \right) e^{-r/r_o}$$

jest rozwiązaniem równania Schrödingera, jeżeli

$$E = -\frac{1}{4} \frac{me^4}{32\pi^2 \epsilon_o^2 \hbar^2} \quad (12.18)$$

Uzyskane wyrażenia na wartości energii elektronu (12.17) i (12.18) sugerują, że poziomy energetyczne elektronu w atomie wodoru można wyrazić w postaci

$$E_n = -\frac{me^4}{32\pi^2 \epsilon_o^2 \hbar^2} \frac{1}{n^2}$$

gdzie $n = 1, 2, 3, \dots$ (całkowita liczba dodatnia) zwana główną liczbą kwantową. Wniosek ten całkowicie zgadza się z przewidywaniami teorii Bohra.

Omówimy teraz jakościowo nowy przypadek, gdy funkcja falowa zależy nie tylko od r , ale także od ϑ i φ . Postępowanie w tym przypadku jest następujące. Funkcję falową $\Psi(r, \vartheta, \varphi)$ przedstawiamy w postaci iloczynu trzech funkcji z których każda zależy od jednej zmiennej

$$\Psi(r, \vartheta, \varphi) = R(r)\Theta(\vartheta)\Phi(\varphi) \quad (12.19)$$

Podstawiając tak przedstawioną funkcję falową do równania Schrödingera otrzymujemy wówczas trzy równania, z których każde opisuje zachowanie się funkcji falowej w zależności od r, ϑ, φ . Równania te nazywają się odpowiednio radialne, biegunowe, azymutalne. Poniżej w sposób skrótowy to przedstawimy.

Podstawiając funkcję (12.19) do równania (12.12), otrzymujemy

$$-\frac{\hbar^2}{2m} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R\Theta\Phi}{\partial r} \right) + \frac{1}{r^2 \sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial R\Theta\Phi}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \frac{\partial^2 R\Theta\Phi}{\partial \varphi^2} \right] +$$

$$+U(r)R\Theta\Phi = ER\Theta\Phi \quad (12.20)$$

Obliczając odpowiednie pochodne, a następnie mnożąc ostatnie równanie przez wyrażenie $r^2 \sin^2 \theta$ i dzieląc przez $R\Theta\Phi$, otrzymamy po uporządkowaniu

$$-\frac{\sin^2 \theta}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) - \frac{\sin \theta}{\Theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right) - \frac{2m}{\hbar^2} r^2 \sin^2 (E - U) = \frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial^2 \varphi} \quad (12.21)$$

Prawa strona tego równania zależy tylko od φ , a lewa tylko od r i θ . Stąd też równanie to będzie spełnione, jeżeli obie strony będą równe jednej i tej samej stałej, którą oznaczamy przez $-m_l^2$. Taki wybór stałej separacji wynika z ogólnej teorii rozwiązywania tego typu równań. Zatem równanie dla Φ przyjmie postać

$$\frac{d^2 \Phi}{d\varphi^2} = -m_l^2 \Phi \quad (12.22)$$

Jeżeli teraz do pozostałej części przedostatniego równania w miejsce prawej strony podstawimy $-m_l^2$, to dzieląc równanie przez $\sin^2 \theta$ i porządkując, otrzymamy

$$\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right) + \frac{2mr^2}{\hbar^2} (E - U) = \frac{m^2}{\sin^2 \theta} - \frac{1}{\Theta \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Theta}{\partial \theta} \right). \quad (12.23)$$

Również w tym równaniu po obu stronach występują różne zmienne. Będzie więc ono spełnione tylko wówczas, gdy obie jego strony będą tożsamościowo równe stałej, którą w tym przypadku oznaczmy przez $l(l+1)$ (znow powód takiego uzasadnienia jak powyżej)

$$\frac{m^2 \Theta}{\sin^2 \theta} - \frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = l(l+1)\Theta \quad (12.24)$$

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{2m}{\hbar^2} (E - U)R = l(l+1) \frac{R}{r^2} \quad (12.25)$$

W ten sposób udało się rozseparować równanie Schrödingera trzech zmiennych na trzy równania różniczkowe jednej zmiennej. Najprostsze z nich jest równanie na Φ ; znamy jego rozwiązanie opisane równaniem (12.10).

Z równań (12.22), (12.24) i (12.25) można wyznaczyć możliwe wartości energii, momentu pędu i jednej jego składowej. Wielkości te, jak wiemy, są określone przez podanie n , l , m_l . Okazuje się przy tym, że dla danego n , orbitalna liczba kwantowa może przyjmować wartości: $l = 0, 1, 2, \dots, n-1$. W ten sposób możemy stwierdzić, że trzy liczby kwantowe opisujące stan elektronu w atomie wodoru, są związane ze sobą w następujący sposób:

- główna liczba kwantowa $n = 1, 2, 3, \dots$
- orbitalna liczba kwantowa $l = 0, 1, 2, \dots, n-1$
- magnetyczna liczba kwantowa $m_l = -l, -l+1, \dots, 0, \dots, l-1, l$

Oznacza to, że przy określonych n , a więc określonej wartości energii, liczba możliwych wartości l i m_l , czyli liczba niezależnych rozwiązań równania Schrödingera będzie wynosiła

$$\sum_{l=0}^{n-1} (2l+1) = n^2 \quad (12.26)$$

Jeżeli elektron jest w takim stanie, że jednej wartości energii odpowiada kilka niezależnych rozwiązań równania Schrödingera, to mówimy, że stan taki jest n^2 -krotnie zwyrodniały.

Dla zadanych wartości liczb kwantowych n , l , m_l ; funkcja falowa ma określoną postać którą oznaczamy symbolem Ψ_{n,l,m_l} i nazywamy orbitalem atomowym. Każdy orbital oznaczony jest tym samym symbolem co odpowiadający mu stan elektronowy. Istnieją ogólnie przyjęte symbole liczb kwantowych n i l . Tworzy się je z cyfry oznaczającej główną liczbę kwantową i litery przyporządkowanej liczbie l w następujący sposób

$$l = 0, 1, 2, 3, 4$$

$$s, p, d, f, g$$

Dla stanu podstawowego atomu wodoru $n = 1$, a zatem $l = m_l = 0$. Stan ten oznaczamy symbolem $1s$.

Dla najniższego stanu wzbudzonego $n = 2$, a zatem $l = 0$ lub $l = 1$. Dla $l = 0$ (stan $2s$) musi być również $m_l = 0$, natomiast dla $l = 1$ (stany $2p$) możemy mieć $m_l = -1, 0$ lub 1 . W celu rozróżnienia stanów $2p$ o różnej wartości liczby m_l stosuje się niekiedy dodatkowy indeks podający wartości tej liczby.

W tabeli poniżej zestawiono stany kwantowe atomu wodoru dla pewnych wartości głównej liczby kwantowej.

Tabela 12.1. Funkcje falowe atomu wodoru

Stan kwantowy	n	l	m_l	Funkcje falowe
$1s$	1	0	0	$\Psi_{100} = \frac{1}{\sqrt{\pi}} \left(\frac{1}{r_0} \right)^{3/2} e^{-r/r_0}$
$2s$	2	0	0	$\Psi_{200} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{r_0} \right)^{3/2} \left(2 - \frac{r}{r_0} \right) e^{-r/2r_0}$
$2p_0$	2	1	0	$\Psi_{210} = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{r_0} \right)^{3/2} \frac{r}{r_0} e^{-r/2r_0} \cos \vartheta$
$2p_1$	2	1	1	$\Psi_{211} = \frac{1}{8\sqrt{\pi}} \left(\frac{1}{r_0} \right)^{3/2} \frac{r}{r_0} e^{-r/2r_0} \sin \vartheta e^{i\varphi}$
$2p_{-1}$	2	1	-1	$\Psi_{21-1} = \frac{1}{8\sqrt{\pi}} \left(\frac{1}{r_0} \right)^{3/2} \frac{r}{r_0} e^{-r/2r_0} \sin \vartheta e^{-i\varphi}$

Dokładna analiza funkcji falowych wykazuje, że część radialna funkcji zależy od n i l , a część kątowa – od l i m_l . Tak więc możemy zapisać, że

$$\Psi_{nlm_l} = R_{nl}(r)\Theta_{lm_l}(\vartheta)\Phi_{m_l}(\varphi) \quad (12.27)$$

Stałe współczynniki występujące w funkcjach zamieszczonych w tabeli 12.1 są wynikiem ich unormowania. Jako przykład wyznaczania stałych normujących rozpatrzmy funkcję Ψ_{100} w postaci Ae^{-r/r_0} . Musi być spełniony warunek

$$\int |\Psi_{100}|^2 dV = \int A^2 e^{-2r/r_0} dV = 1$$

Element objętości we współrzędnych sferycznych wynosi

$$dV = r^2 \sin \vartheta dr d\vartheta d\varphi$$

Uwzględniając to w powyższym wzorze otrzymujemy

$$A^2 \int_0^\infty r^2 e^{-2r/r_0} dr \int_0^\pi \sin \vartheta d\vartheta \int_0^{2\pi} d\varphi = 1$$

Całkowanie względem ϑ i φ jest elementarne, natomiast

$$\int_0^\infty r^2 e^{-2r/r_0} dr = \frac{2}{(2/r_0)^3}$$

zatem

$$A^2 \frac{2}{(2/r_0)^3} 4\pi = 1$$

Stąd wynika, że

$$A = \frac{1}{\sqrt{\pi}} \left(\frac{1}{r_0} \right)^{3/2}$$

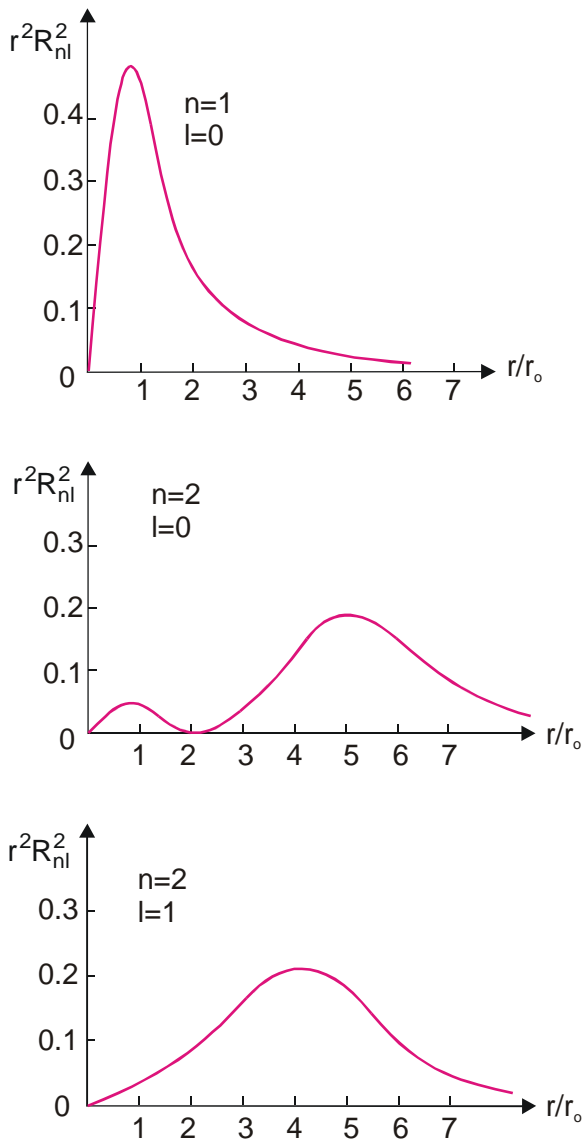
Podobnie wyznaczamy stałe w innych przypadkach, choć obliczenia są bardziej złożone.

Znajomość funkcji falowych pozwala obliczyć prawdopodobieństwo znalezienia elektronu w określonym elemencie objętości. Wygodnie jest rozpatrywać prawdopodobieństwo radialne (zależne od r) i prawdopodobieństwo kątowe (zależne od ϑ i φ).

Prawdopodobieństwo radialne jest zdefiniowane jako prawdopodobieństwo znalezienia elektronu w warstwie kulistej o promieniu r i grubości dr . Ponieważ objętość warstwy kulistej jest proporcjonalna do r^2 , zatem

$$p(r)dr = R_{nl}^2 r^2 dr$$

Stąd $p(r) = R_{nl}^2 r^2$ jest gęstością prawdopodobieństwa.

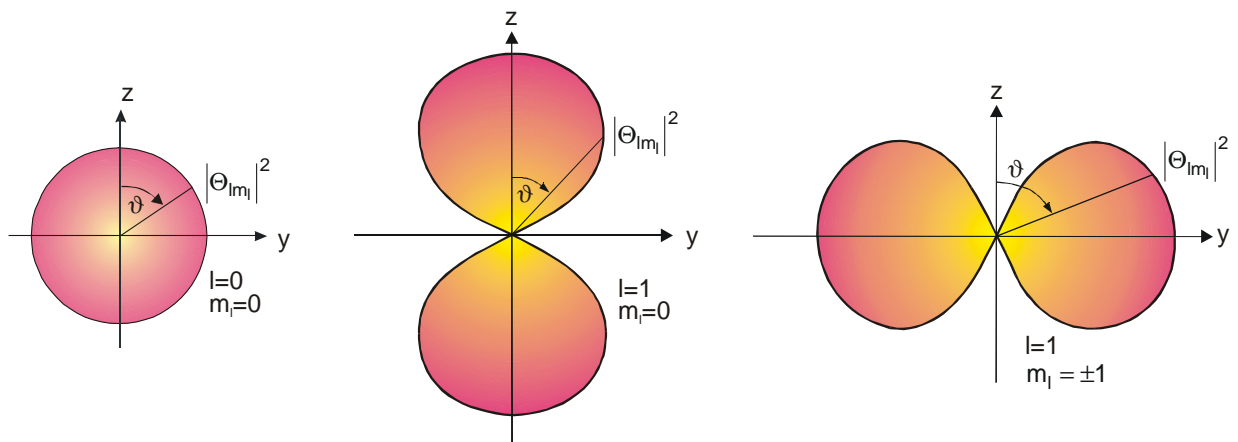


Rys. 12.5. Radialne gęstości prawdopodobieństwa dla atomu wodoru gdy $n = 1$ i 2 .

Na rys. 12.5 przedstawiono zależność $p(r)$ dla $n = 1$ i 2 w zależności od r/r_0 . W stanie $1s$ gęstość prawdopodobieństwa osiąga maksimum dla $r = r_0$. W stanie $2s$ istnieją dwa maksima: jedno dla $r = r_0$, drugie dla $r \cong 5r_0$, przy czym to maksimum jest znacznie większe od pierwszego. W stanie $2p$ gęstość prawdopodobieństwa osiąga maksimum w pobliżu wartości $r = 4r_0$.

Analizując część kątową funkcji falowych możemy wyznaczyć kątową gęstość prawdopodobieństwa $p(\vartheta, \varphi)$. Zauważmy przede wszystkim, że dla wszystkich stanów zależność funkcji falowych od kąta φ jest postaci $e^{im\varphi}$. Zatem gęstość prawdopodobieństwa nie będzie zależała od φ , gdyż $e^{im\varphi} e^{-im\varphi} = 1$.

Na rys. 12.6 przedstawiono zależność $p(\vartheta)$ od kąta ϑ we współrzędnych biegunowych. Miarą prawdopodobieństwa znalezienia elektronu w danym kierunku jest odległość między początkiem układu współrzędnych i punktem przecięcia prostej poprowadzonej pod danym kątem, z wykresem funkcji $\Theta_{lm_l}^2$.



Rys. 12.6. Wykresy biegunowe kierunkowej zależności gęstości prawdopodobieństwa dla atomu wodoru w przypadku $l = 0$ i $l = 1$.

W stanie s ($l = 0$) gęstość prawdopodobieństwa jest stała. Chmura elektronowa wykazuje symetrię kulistą. W stanie $l = 1$ i $m = 0$, $p(\vartheta) \sim \cos^2 \vartheta$; dla $l = 1$ i $m = \pm 1$, $p(\vartheta) \sim \sin^2 \vartheta$. W każdym przypadku gęstość prawdopodobieństwa wykazuje symetrię obrotową względem osi z .

Funkcja całkowitej gęstości prawdopodobieństwa znalezienia elektronu w odległości r od jądra pod kątem ϑ jest równa $p(r)p(\vartheta)$.

12.4. Orbitalny moment magnetyczny

Z ruchem orbitalnym elektronu związany jest moment magnetyczny elektronu. Moment magnetyczny zamkniętego, płaskiego obwodu z prądem wynosi

$$\vec{\mu} = I \cdot \vec{S}$$

gdzie \vec{S} jest wektorem, którego długość jest równa liczbowo powierzchni obwodu, a kierunek wektora jest prostopadły do płaszczyzny obwodu. Cząstka o ładunku q poruszająca się po orbicie kołowej z prędkością v , wytwarza prąd o natężeniu równym iloczynowi q przez częstotliwość obrotów $v/2\pi r$. Wartość orbitalnego momentu magnetycznego μ_l wynosi więc

$$\mu_l = \frac{qv}{2\pi r} \pi r^2 = \frac{qvr}{2}$$

W zapisie wektorowym

$$\mu_l = \frac{q}{2} \vec{r} \times \vec{v} = \frac{q}{2m} \vec{r} \times \vec{p} = \frac{q}{2m} \vec{L}$$

gdzie $\vec{L} = \vec{r} \times \vec{p}$ jest orbitalnym momentem pędu. Stosunek μ/L nazywamy stosunkiem giromagnetycznym. Dla elektronu

$$\mu_l = -\frac{e}{2m} L = -\frac{e}{2m} \sqrt{l(l+1)} \hbar = -\mu_B \sqrt{l(l+1)} \quad (12.28)$$

gdzie

$$\mu_B = \frac{e\hbar}{2m} = 9,27 \times 10^{-24} \text{ Am}^2 \quad (12.29)$$

nazywamy magnetonem Bohra.

Magneton Bohra jest jednostką atomowego momentu magnetycznego. Znak minus we wzorze (12.28) oznacza, że μ_l jest skierowany przeciwnie niż \vec{L} , co spowodowane jest ujemnym ładunkiem elektronu.

Jeżeli cząstkę o momencie magnetycznym umieścimy w zewnętrznym polu magnetycznym o indukcji B , to działa na nią moment siły

$$\vec{T} = \vec{\mu}_l \times \vec{B} \quad (12.30)$$

który dąży do ustawienia wektora momentu magnetycznego wzdłuż kierunku pola.

Energia potencjalna momentu magnetycznego w zewnętrznym polu magnetycznym wynosi natomiast

$$U = \vec{\mu}_l \cdot \vec{B} = -\mu_l B \cos \theta \quad (12.31)$$

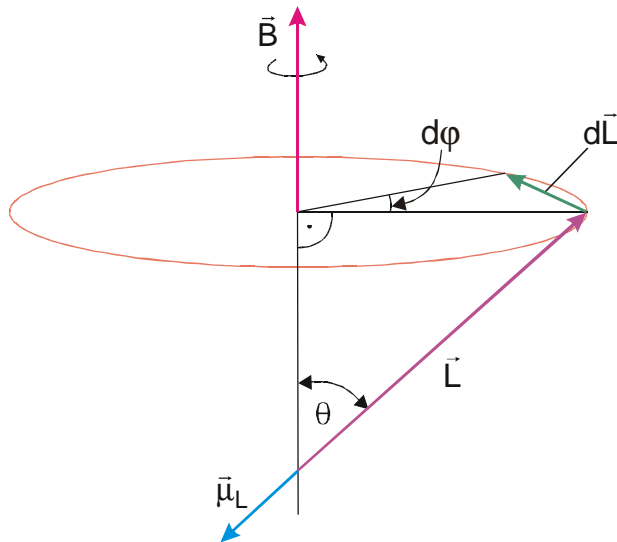
gdzie θ jest kątem między $\vec{\mu}_l$ i \vec{B} .

Wiemy już, że orbitalny moment pędu i jego rzut na dowolny, ale ustalony kierunek są skwantowane. Kąt między L i L_z jest zawsze różny od zera. Zewnętrzne pole magnetyczne wyróżnia pewien kierunek w przestrzeni. Oznacza to, że wektor \vec{L} , a więc i wektor $\vec{\mu}_l$, nigdy nie mogą ustawić się równolegle czy też antyrównolegle do zewnętrznego pola. Składowa orbitalnego momentu magnetycznego w kierunku zewnętrznego pola wynosi

$$\mu_{lz} = \mu_l \cos \theta = -\frac{e\hbar}{2m} \sqrt{l(l+1)} \cos \theta = -\frac{e\hbar}{2m} m_l \quad (12.32)$$

Widzimy teraz, dlaczego m_l nazywa się magnetyczną liczbą kwantową. Określa ona bowiem rzut momentu magnetycznego na kierunek zewnętrznego pola magnetycznego. Jak już wspomnieliśmy powyżej, kąt θ między wektorami $\vec{\mu}_l$ i \vec{B} jest zawsze różny od zera, więc w zewnętrznym polu magnetycznym na moment magnetyczny zawsze będzie działać moment siły $\vec{T} = -(e/2m)\vec{L} \times \vec{B}$ (rys. 12.7). Moment ten powoduje zmianę momentu pędu $d\vec{L}$ zgodnie z prawem Newtona $d\vec{L}/dt = \vec{T}$. Zmiana $d\vec{L}$ powoduje precesję \vec{L} o kąt $\omega_L dt$, gdzie ω_L jest prędkością kątową precesji zwanej prędkością Larmona. Z rys. 12.7 widać, że

$$dL = L \sin \theta \omega_L dt$$



czyli

$$\frac{dL}{dt} = L \omega_L \sin \theta = T = \frac{e}{2m} LB \sin \theta$$

Tak więc

$$\omega_L = \frac{eB}{2m} \quad (12.33)$$

i nazywana jest częstością Larmona.

Powróćmy jeszcze raz do wyrażenia na energię potencjalną momentu magnetycznego w polu magnetycznym (12.31)

$$U = -\mu_l B \cos \theta = \frac{e\hbar}{2m} B m_l = \mu_B m_l B \quad (12.34)$$

Rys. 12.7. Oddziaływanie momentu magnetycznego z zewnętrznym polem magnetycznym.

Oznacza to, że pole magnetyczne będzie zmieniało poziomy energetyczne elektronu. Ponieważ dla ustalonego n i l istnieje $2l+1$ możliwych wartości m_l , więc pierwotny poziom energetyczny zostanie rozszczepiony na $2l+1$ podpoziomów. Jeżeli np. elektron w atomie wodoru znajduje się w stanie $2p$ ($l = 1$) w polu magnetycznym o indukcji $B = 1\text{T}$ to po podstawieniu wartości liczbowych otrzymujemy

$$U = \begin{cases} 9,27 \times 10^{-24} \text{ J} & m_l = 1 \\ 0 & m_l = 0 \\ -9,27 \times 10^{-24} \text{ J} & m_l = -1 \end{cases}$$

Z powyższego wynika, że rozszczepienie poziomów energetycznych nie jest duże, ale może być z powodzeniem mierzone eksperymentalnie. Jedną z możliwych metod pomiaru polega na obserwacji zmian długości linii widmowych emitowanych przez atomy wzbudzone znajdujące się w zewnętrznym polu magnetycznym. Pojedyncza linia widmowa może ulec rozszczepieniu na trzy blisko siebie położone linie, przy czym odstęp energetyczny między dwoma kolejnymi liniami wynosi $\mu_B B$.

Dociekliwy czytelnik może zapytać, dlaczego stwierdziliśmy, że linie widmowe mogą rozszczepiać się tylko na trzy blisko położone siebie linie, a nie na $2l+1$ linii, skoro tyle jest podpoziomów energetycznych. Wiąże się to z tzw. regułami wyboru dla liczb kwantowych l i m_l , ale problemem tym nie będziemy się zajmować.

Wspomniane tu zjawisko rozszczepienia linii widmowych w zewnętrznym polu magnetycznym nosi nazwę zjawiska Zeemana. Zjawisko to potwierdza skwantowanie orbitalnego momentu pędu w atomie.

12.5. Spin elektronu

12.5.1. Spinowy moment pędu i spinowy moment magnetyczny

Precyzyjne pomiary linii widmowych pozwalają stwierdzić, że nawet bez obecności zewnętrznego pola magnetycznego, składają się one z kilku położonych blisko siebie linii. Takie rozszczepienie pojedynczej linii widmowej nazywane jest strukturą subtelną linii widmowych. Pewnym przypadkiem struktury subtelnej jest oczywiście zjawisko Zeemana.

Okazuje się, że elektron poza omówionym powyżej orbitalnym momentem pędu, ma własny moment pędu, zwany spinowym lub krótko spinem. Jego istnienie nie jest związane z ruchem elektronu lecz jest jego naturalną (wewnętrzną) właściwością. Jego istnienie zaproponowali w 1925 roku Goudsmit i Uhlenbeck w celu wyjaśnienia struktury subtelnej linii widmowych.

Spinowy moment pędu elektronu wynosi

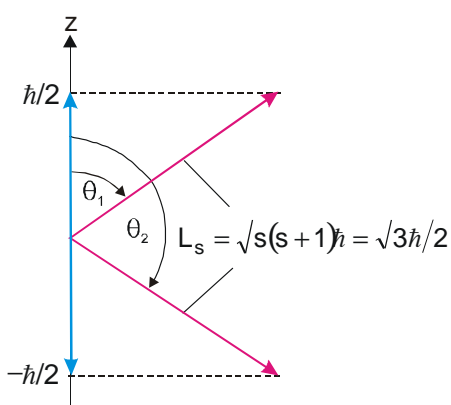
$$L_s = \sqrt{s(s+1)}\hbar \quad (12.35)$$

gdzie spinowa liczba kwantowa s ma tylko jedną wartość $1/2$. Wobec tego spin elektronu ma wartość $L_s = \sqrt{3}\hbar/2$ jest podstawową właściwością elektronu, taką jak masa i ładunek.

Rzut spinowego momentu pędu na wyróżniony kierunek, np. określony zewnętrznym polem magnetycznym, jest skwantowany i wynosi

$$L_{sz} = m_s \hbar \quad (12.36)$$

gdzie m_s przyjmuje dwie wartości, $1/2$ lub $-1/2$; i nosi nazwę magnetycznej spinowej liczby kwantowej. Liczbie kwantowej $m_s = 1/2$ odpowiada spin skierowany w górę, a $m_s = -1/2$ odpowiada spin skierowany w dół (rys. 12.8). Możliwe kąty między wektorem L_s a wyróżnionym kierunkiem w przestrzeni wynoszą $54,7^\circ$ i $125,3^\circ$.



Rys. 12.8. Spinowy moment pędu elektronu i jego rzut na określony kierunek w przestrzeni.

Stan elektronu w atomie będziemy więc opisywać za pomocą czterech liczb kwantowych: n , l , m_l i m_s . Z poprzednich rozważań pamiętamy, że liczba niezależnych stanów związanych z główną liczbą kwantową wynosi n^2 . Po uwzględnieniu spinu każdy stan "rozszczenia się" na dwa stany różniące się od siebie rzutem spinu. Wobec tego całkowita liczba stanów związanych z daną liczbą kwantową n wynosi $2n^2$.

Ze spinowym momentem pędu L_s związany jest spinowy moment magnetyczny μ_s , który wynosi

$$\vec{\mu}_s = -\frac{e}{m} \vec{L}_s = -\frac{e\hbar}{m} \sqrt{s(s+1)} \quad (12.37)$$

Rzut spinowego momentu magnetycznego na wyróżniony kierunek wynosi

$$\mu_{sz} = -\frac{e\hbar}{m} \sqrt{s(s+1)} \cos \theta = \pm \frac{e\hbar}{2m} = \pm \mu_B \quad (12.38)$$

a więc μ_{sz} może przyjmować dwie wartości różniące się znakiem, co do modułu równe magnetonowi Bohra.

Warto też zauważyć, że stosunek giromagnetyczny dla momentu spinowego jest dwukrotnie większy niż dla momentu orbitalnego. Fakt ten nazywamy anomalią magnetomechaniczną.

12.5.2. Całkowity moment pędu i całkowity moment magnetyczny

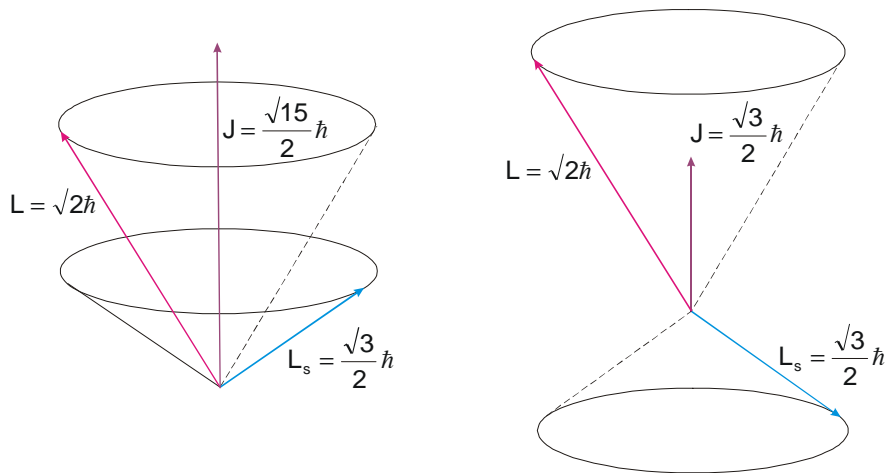
Na podstawie naszych dotychczasowych wiadomości możemy powiedzieć, że elektron poruszający się po orbicie wokół jądra ma całkowity moment pędu \vec{J} równy sumie wektorów \vec{L} i \vec{L}_s . Zapiszemy zatem

$$\vec{J} = \vec{L} + \vec{L}_s \quad (12.39)$$

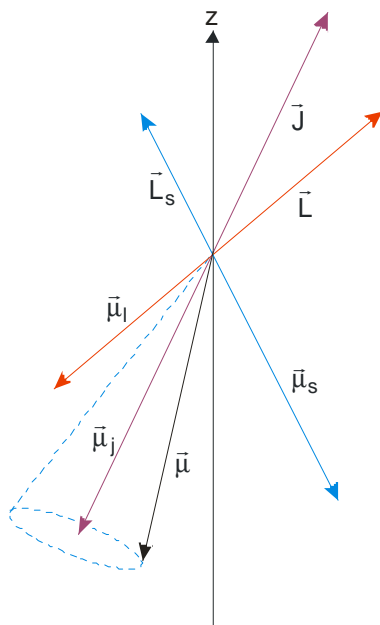
Okazuje się, że całkowity moment pędu, podobnie jak \vec{L} i \vec{L}_s , jest skwantowany i wynosi

$$J = \sqrt{j(j+1)}\hbar \quad (12.40)$$

gdzie j jest liczbą kwantową całkowitego momentu pędu wynoszącą $j = l + s$ lub $j = l - s$. Ponieważ $s = 1/2$, więc $j = l - 1/2$ lub $j = l + 1/2$. Na rys. 12.9 pokazano przykładowo dodawanie wektorów \vec{L} i \vec{L}_s przyjmując, że $l = 1$. Zauważmy, że ponieważ wektory \vec{L} , \vec{L}_s i \vec{J} są skwantowane, więc istnieją tylko pewne dozwolone orientacje tych wektorów względem siebie.



Rys. 12.9. Możliwe rezultaty dodawania orbitalnego i spinowego momentu pędu elektronu dla $l = 1$ i $s = 1/2$.



Rys 12.10. Wypadkowy moment magnetyczny elektronu.

Możliwe wartości liczbowe łatwo wytłumaczyć jeśli rozważyć rzut wektorów \vec{L} i \vec{L}_s na określony kierunek, np. oś z . Maksymalna wartość $L_z = l\hbar$, maksymalna wartość $L_{sz} = \hbar/2$, dlatego maksymalna wartość J_z może wynosić $(l+1/2)\hbar$, a minimalna wartość $(l-1/2)\hbar$. Odpowiednio więc j może być równe $l+1/2$ lub $l-1/2$.

Warto dodać, że często stan elektronu w atomie charakteryzujemy przez podanie liczb n , l , oraz j . Przyjęto taki zapis, że n podajemy jako liczbę, l – odpowiednią literę, zaś wartość j jako indeks przy literze. Na przykład zapis $3d_{3/2}$ oznacza, że $n = 3$, $l = 2$ i $j = 3/2$. Podobnie jak w przypadku momentu pędu \vec{L} i

\vec{L}_s , składowa \vec{J} w określonym kierunku jest skwantowana i wynosi

$$J_z = m_j \hbar \quad (12.41)$$

gdzie $m_j = -j, -j+1, \dots, j-1, j$; tzn. przyjmuje łącznie $2j+1$ wartości różniące się o 1 i zawarte w przedziale $-j \leq m_j \leq j$.

Również całkowity moment magnetyczny $\vec{\mu}$ jest sumą momentu magnetycznego orbitalnego i spiralnego

$$\vec{\mu} = \vec{\mu}_l + \vec{\mu}_s \quad (12.42)$$

Jednakże na skutek tego, że $\vec{\mu}_s/L_s = 2\mu_l/L$, kierunek wektora $\vec{\mu}$ nie pokrywa się z kierunkiem wektora \vec{J} (rys. 12.10). Jeżeli na elektron nie działa żaden moment sił zewnętrznych, to zgodnie z zasadą zachowania momentu pędu, wartość i kierunek wektora \vec{J} nie zmieniają się. Oznacza to, że wektor momentu magnetycznego wykonuje ruch precesyjny wokół wektora \vec{J} . Dlatego na zewnątrz występuje jedynie składowa μ_j momentu magnetycznego $\vec{\mu}$ w kierunku wektora \vec{J} . Rzutując $\vec{\mu}_l$ i $\vec{\mu}_s$ na kierunek wektora \vec{J} otrzymujemy

$$\mu_j = -\frac{e\hbar}{2m} \sqrt{l(l+1)} \cos(\vec{L}, \vec{J}) - \frac{e\hbar}{m} \sqrt{s(s+1)} \cos(\vec{L}_s, \vec{J}) \quad (12.43)$$

Wyznaczając z twierdzenia cosinusów kąty między wektorami \vec{L} i \vec{J} oraz \vec{L}_s i \vec{J} mamy

$$\cos(\vec{L}, \vec{J}) = \frac{j(j+1) + l(l+1) - s(s+1)}{2\sqrt{j(j+1)}\sqrt{l(l+1)}}$$

$$\cos(\vec{L}_s, \vec{J}) = \frac{j(j+1) + s(s+1) - l(l+1)}{2\sqrt{j(j+1)}\sqrt{s(s+1)}}$$

Podstawiając powyższe do wzoru (12.43) otrzymujemy

$$\mu_j = -\frac{e\hbar}{2m} g \sqrt{j(j+1)} = -\mu_B g \sqrt{j(j+1)} \quad (12.44)$$

gdzie

$$g = \frac{3j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}$$

jest tzw. czynnikiem Landego. Czynnikiem Landego określa stosunek całkowitego momentu magnetycznego do momentu pędu w tych przypadkach, w których moment pędu jest częściowo spinowy i częściowo orbitalny. Gdy moment magnetyczny jest tylko spinowy, to $g = 2$, natomiast dla momentu orbitalnego $g = 1$.

Ponieważ z momentem magnetycznym związane jest pole magnetyczne, więc na orbitalny i spinowy moment magnetyczny działać będą momenty sił, co spowoduje precesję wektorów $\vec{\mu}_l$ i

$\vec{\mu}_s$, a więc również \vec{L} i \vec{L}_s wokół \vec{J} . Natomiast jeżeli atom znajdzie się w zewnętrznym polu magnetycznym, to wówczas wektory \vec{J} i $\vec{\mu}_j$ wykonują precesję wokół kierunku wektora \vec{B} . Zgodnie z regułami kwantowania przestrzennego, $\vec{\mu}_j$ i \vec{J} nie mogą być równoległe do \vec{B} . Składowe wektora $\vec{\mu}_j$ w kierunku zewnętrznego pola magnetycznego wynoszą

$$\mu_{jz} = -\mu_B g m_j \quad (12.45)$$

Moment magnetyczny $\vec{\mu}_j$ w zewnętrznym polu magnetycznym uzyskuje energię potencjalną

$$U = -\vec{\mu}_j \cdot \vec{B} = \mu_B g B \sqrt{j(j+1)} \cos \theta$$

Ponieważ $\sqrt{j(j+1)} \cos \theta = m_j$, więc

$$U = \mu_B g m_j B \quad (12.46)$$

W polach magnetycznych o indukcyjności $B = 1\text{T}$ wartość energii potencjalnej momentu magnetycznego wynosi około 10^{-4} eV . Jest to wartość bardzo mała w porównaniu z energią ruchów termicznych w temperaturze pokojowej ($kT = 0.026\text{ eV}$ dla $T = 300\text{K}$).

Znajomość zachowania się atomu w zewnętrznym polu magnetycznym jest bardzo ważna przy opisie właściwości magnetycznych materii.

12.5.3. Oddziaływanie spin-orbita

Omówimy teraz oddziaływanie między momentem magnetycznym orbitalnym a momentem magnetycznym spinowym. Oddziaływanie to nazywamy krótko oddziaływaniem spin-orbita.

Jeżeli elektron o ładunku e porusza się po orbicie kołowej wokół jądra z prędkością v , to w układzie odniesienia związanym z elektronem, jądro porusza się z prędkością $-v$. Z prawa Biota-Savarta-Laplace'a znajdujemy wartość indukcji pola magnetycznego w miejscu gdzie znajduje się elektron

$$\vec{B} = \frac{\mu_o}{4\pi} \frac{Id\vec{l} \times \vec{r}}{r^3} = -\frac{\mu_o}{4\pi} \frac{e\vec{v} \times \vec{r}}{r^3} = \frac{\mu_o e}{4\pi m r^3} \vec{L} \quad (12.47)$$

gdzie podstawiliśmy: $Id\vec{l} = -e\vec{v}$, $\vec{r} \times m\vec{v} = -(m\vec{v} \times \vec{r}) = -\vec{L}$. Obliczymy wartość pola magnetycznego przyjmując $n = 2$ i $l = 1$ (dla $n = 1$, $l = 0$ i również $L = 0$). Wówczas $L = \sqrt{l(l+1)}\hbar = \sqrt{2}\hbar$, $r = n^2 r_o = 2,12 \times 10^{-10}\text{ m}$ i podstawiając stałe fizyczne występujące w wyrażeniu (12.47) otrzymujemy $B = 0,27\text{ T}$. Jest to stosunkowo silne pole magnetyczne uzyskiwane w typowych elektromagnesach z rdzeniem żelaznym

Spinowy moment magnetyczny uzyskuje zatem energię:

$$U = -\vec{\mu}_s \cdot \vec{B} = -\left(-\frac{e}{m} \vec{L}_s\right) \frac{\mu_o e}{4\pi m r^3} \vec{L} = \frac{\mu_o e^2}{4\pi m^2 r^3} \vec{L} \cdot \vec{L}_s \quad (12.48)$$

Wartości iloczynu skalarnego $\vec{L} \cdot \vec{L}_s$ obliczamy korzystając z oczywistej zależności

$$\vec{J} \cdot \vec{J} = J^2 = (\vec{L} + \vec{L}_s)(\vec{L} + \vec{L}_s) = L^2 + L_s^2 + 2\vec{L}\vec{L}_s$$

gdzie \vec{J} , \vec{L} , \vec{L}_s są wektorami odpowiednio całkowitego, orbitalnego i spinowego momentu pędu.

Wobec tego

$$\vec{L} \cdot \vec{L}_s = \frac{1}{2}(J^2 - L^2 - L_s^2) = \frac{\hbar^2}{2}[j(j+1) - l(l+1) - s(s+1)]$$

Ostatecznie więc wyrażenie (12.48) zapiszemy w postaci

$$U = \frac{\mu_o e^2 \hbar^2}{8\pi m^2 r^3} [j(j+1) - l(l+1) - s(s+1)] \quad (12.49)$$

Z powyższego wynika, że największa wartość energii oddziaływania spin-orbita, a więc i największe rozszczepienie poziomów energetycznych będzie na orbitach najbliższych jądra. Ze względów wspomnianych wcześniej należy wykluczyć stany o $l = 0$.

Całkowita energia elektronu zależy więc nie tylko od głównej, ale również od orbitalnej liczby kwantowej. Wielkość oddziaływania spin-orbita nie jest duża, np. dla $n = 2$, $l = 1$, $j = 3/2$ i $r = 2^2 r_o$ ($r_o = 0,53 \times 10^{-10}$ m – promień Bohra) otrzymujemy $U = 3,2 \times 10^{-5}$ eV. Jednak dla atomów wieloelektronowych zależność energii od orbitalnej liczby kwantowej jest znacznie większa.

Dla ustalonego l energia oddziaływania spin-orbita przyjmuje duże wartości, gdyż j może być równe $l + 1/2$ lub $l - 1/2$. Wielkość tego rozszczepienia wynosi

$$U_{l+1/2} - U_{l-1/2} = \frac{\mu_o e^2 \hbar^2}{8\pi m^2 r^3} (2l + 1) \quad (12.50)$$

przy czym $U_{l+1/2} > 0$, natomiast $U_{l-1/2} < 0$.

Nasze rozważania dotyczące oddziaływania spin-orbita mają charakter przybliżony, ale wyprowadzone zależności dobrze opisują omawiane efekty.

Z faktu, że energia elektronu w atomie zależy od ustawienia momentu magnetycznego spinowego względem momentu orbitalnego wynika, że poziom energetyczny określony główną liczbą kwantową n rozszczepia się na kilka podpoziomów w zależności od wartości liczb kwantowych j , l i s . Mówimy wtedy, że linie widmowe mają strukturę subtelną.

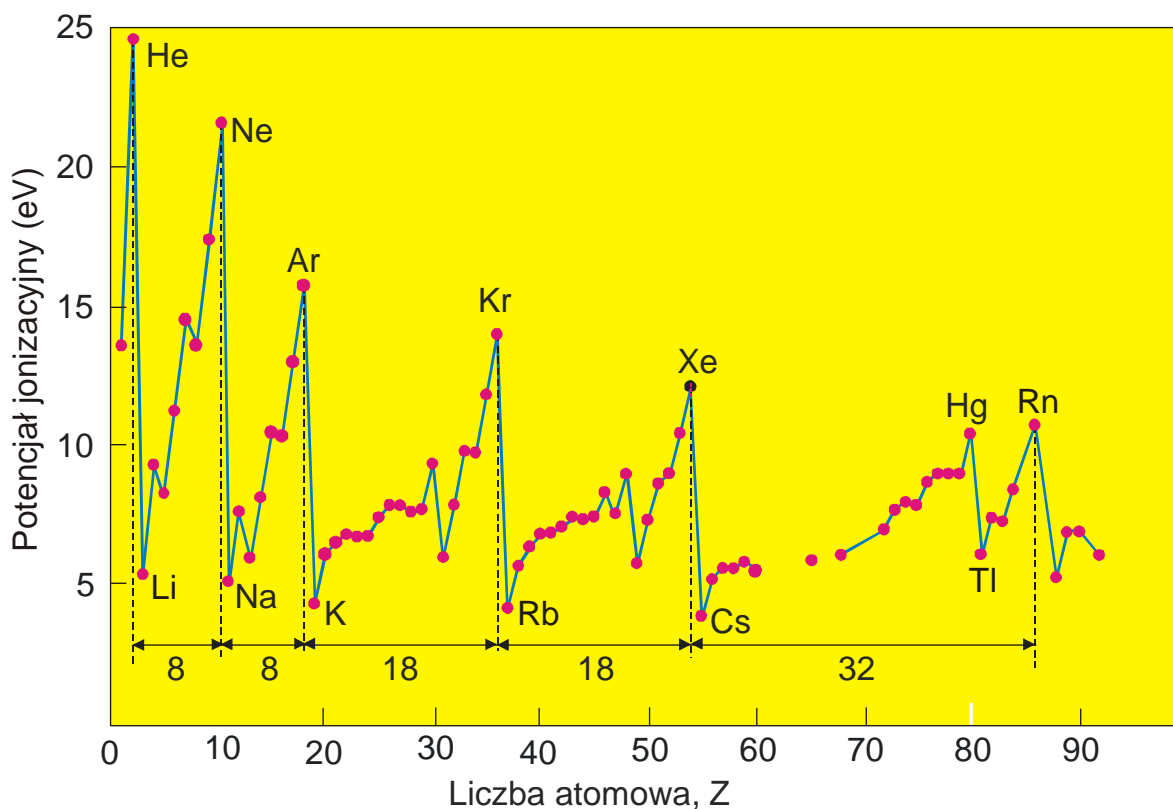
ROZDZIAŁ 13

FIZYKA ATOMOWA

Nasze dotychczasowe rozważania dotyczyły głównie atomu wodoru, gdzie elektron oddziałuje tylko z jądrem, a pole elektrostatyczne jest polem kulombowskim. Omówimy teraz atomy wieloelektronowe. Zastanowimy się najpierw, jak są rozmieszczone elektrony w atomach o liczbie atomowej $Z > 1$. Należałoby oczekiwać, że wszystkie elektrony w atomie znajdującym się w stanie podstawowym będą na najniższym poziomie magnetycznym. Tymczasem fakty doświadczalne wskazują, że elektrony rozmieszczają się na różnych poziomach.

13.1. Zakaz Pauliego

Jak wynika z układu okresowego pierwiastków, okresy zmian właściwości fizycznych i chemicznych pierwiastków tworzą kolejność liczb 2, 8, 8, 18, 18, 32 (rys. 13.1). W roku 1925 W. Pauli przedstawił prostą regułę która automatycznie wyjaśnia istnienie grup z 2, 8, 18 i 32 pierwiastkami.



Rys. 13.1. Zależność potencjału jonizacyjnego elektronów od liczby atomowej Z.

Pauli zapostulował, że *jeden orbital elektronowy mogą zajmować nie więcej niż dwa elektrony*. Tak więc w stanie z $n = 1$ mogą znajdować się dwa elektrony. Liczbie kwantowej $n = 2$ odpowiadają 4 orbitale: $(n, l, m_l) = (2, 0, 0), (2, 1, 1), (2, 1, 0), (2, 1, -1)$. Wobec tego w stanie z $n = 2$ może znajdować się 8 elektronów. W ten sposób otrzymaliśmy liczby 2 i 8. Liczbę 18 można otrzymać dodając 5 orbitali z $l = 2$ i 4 orbitale z $l = 0$. Te dziewięć orbitali może być zajęte przez 18 elektronów.

Widzimy więc, że liczby 2, 8 i 18 są prostą konsekwencją zakazu Pauliego, a także zasady kwantowomechanicznej zgodnej z którą $-l \leq m_l \leq l$ i $0 \leq l \leq n - 1$. W chwili pojawienia się zasada Pauliego była nowym postulatem, którego w tym czasie nie można było wyprowadzić z ogólnych założeń. Wkrótce po odkryciu spinu, Pauli i Dirac opracowali relatywistyczną teorię cząstek o spinie $1/2$ i stwierdzili, że warunek relatywistycznej inwariantności prowadzi do funkcji falowych elektronu, które automatycznie spełniają zakaz Pauliego. Oznaczało to, że zasada Pauliego nie była "wzięta z sufitu". Dla cząstki o spinie $1/2$ rzuty jej spinu na oś z mogą przyjmować jedną z dwóch wartości $(1/2)\hbar$ lub $-(1/2)\hbar$. Zasada zakazu potwierdza, że w danym stanie określonym czterema liczbami kwantowymi (n, l, m_l, s) może znajdować się nie więcej niż jeden elektron. Jest to równorzędne sformułowaniu zasady Pauliego, że na danym orbitalu mogą znajdować się nie więcej niż dwa elektrony.

Podsumowując to możemy stwierdzić, że na n -tej orbicie liczba elektronów może być co najwyżej równa

$$N = 2 \sum_{l=0}^{n-1} 2l + 1 = 2n^2,$$

gdzie czynnik 2 uwzględnia dwie możliwe orientacje spinu. Możemy już teraz powiedzieć, że rozmieszczenie elektronów w atomie będzie takie, aby układ miał minimalną energię, ale jednocześnie nie był naruszony zakaz Pauliego.

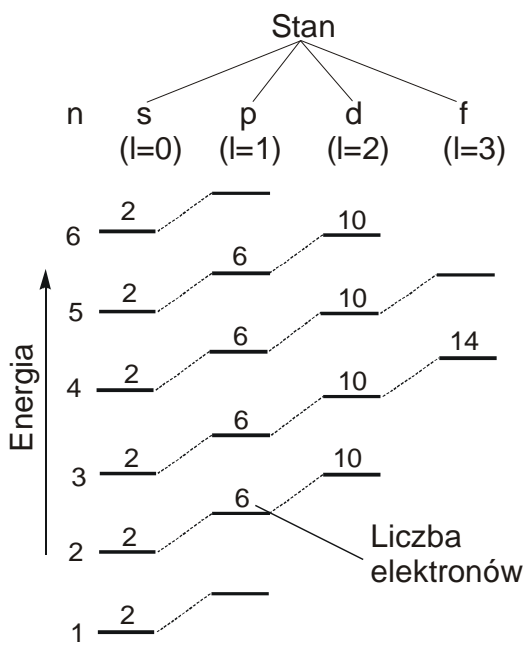
13.2. Atomy wieloelektronowe

Aby wyznaczyć wartości energii elektronów w atomach wieloelektronowych należałoby rozwiązać równanie Schrödingera dla układu złożonego z jądra i wielu elektronów. Jest to zadanie praktycznie niewykonalne. Jedynie dla atomu helu znaleziono dokładne rozwiązanie drogą obliczeń numerycznych. Dla atomów wieloelektronowych stosujemy na ogół przybliżone metody opisu.

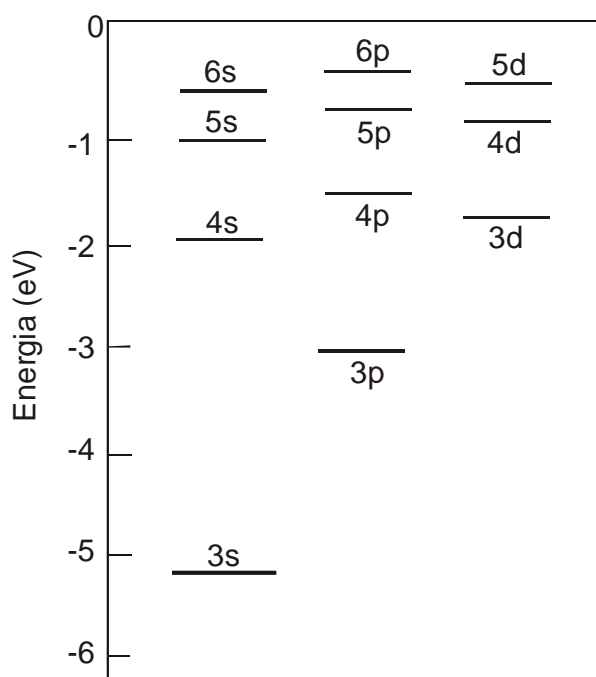
W atomach wodoropodobnych, tzn. w atomach $(Z-1)$ -krotnie zjonizowanych, mamy sytuację podobną do atomu wodoru. Promienie orbit i poziomy energetyczne opisujemy wzorami

$$r = \frac{4\pi\epsilon_0\hbar^2}{mZe^2} n^2, \quad E = -\frac{mZ^2e^4}{32\pi^2\epsilon_0^2\hbar^2} \frac{1}{n^2}.$$

Dla atomów wieloelektronowych najprościej jest przyjąć, że każdy elektron porusza się w polu elektrycznym o symetrii kulistej, która jest wynikiem nałożenia się pola kulombowskiego jądra z polem elektronów otaczających to jądro. Tak więc blisko jądra, pole jest w przybliżeniu polem kulombowskim o potencjale $Ze/4\pi\epsilon_0 r$. Powoduje to, że energia wiązania elektronów na pierwszej orbicie bardzo silnie zależy od Z . Okazuje się, że dla ciężkich atomów jest ona proporcjonalna do $(Z-1)^2$, a więc jest tylko nieznacznie mniejsza niż w atomie wodoropodobnym. W miarę oddalania się od jądra, ekranujący wpływ elektronów wzrasta i pole przestaje być kulombowskie. Można przyjąć, że jądro ma pewien ładunek efektywny eZ_{ef} , przy czym $Z_{ef} < Z$. Na orbitach najbardziej zewnętrznych wpływ ekranowania jest tak duży, że energia wiązania nie zależy praktycznie od Z i jest w przybliżeniu równa energii wiązania elektronu w atomie wodoru.



Rys. 13.2. Schematyczne przedstawienie wpływu orbitalnej liczby kwantowej l na wartości energii wiązania elektronu na danej orbicie. Stany o większej wartości l mają wyżej położone poziomy energetyczne.



Rys. 13.3. Niektóre poziomy energetyczne sodu.

W atomach wieloelektronowych energia elektronu w atomie zależy znacznie bardziej od wartości orbitalnej liczby kwantowej niż w atomie wodoru. Dla przykładu rozpatrzmy drugą orbitę. W stanie $2s$ rozkład gęstości prawdopodobieństwa znalezienia elektronu wykazuje dwa maksima: pierwsze w odległości około $0.8r_0$ od jądra, drugie w odległości $5r_0$ od jądra (rys. 12.5). Natomiast w stanie $2p$ gęstość prawdopodobieństwa blisko jądra jest mała i występuje tylko jedno maksimum w odległości $4r_0$ od jądra. Oznacza to, że w stanie $2s$ elektron będzie często przebywał stosunkowo blisko jądra i w efekcie będzie bardziej "odczuwał" ładunek jądra niż w stanie $2p$. Stan $2s$ będzie stanem silniej związanym. Im większa wartość głównej liczby kwantowej, tym zależność energii wiązania elektronu od orbitalnej liczby kwantowej jest

większa (rys. 13.2). Okazuje się również, że podpoziomy charakteryzujące się dużym l mogą mieć większe wartości energii niż podpoziomy o małym l , ale o większej wartości głównej liczby kwantowej. Jako przykład, na rys. 13.3 pokazano poziomy energetyczne atomu sodu gdzie stan $3d$ leży wyżej niż stan $4s$, a stan $4f$ wyżej niż $5p$.

13.3. Układ okresowy pierwiastków

Układ okresowy pierwiastków (patrz Tabele 13.1 i 13.2) przedstawia takie uporządkowanie pierwiastków według ich liczby atomowej, które odzwierciedla periodyczność zmian właściwości fizykochemicznych pierwiastków. Pierwiastki o podobnych właściwościach tworzą grupy, które w układzie okresowym występują jako kolumny pionowe. Wiersze układu okresowego nazywamy okresami.

Właściwości pierwiastków uwarunkowane są konfiguracją elektronów w atomach. Przez konfigurację elektronów rozumiemy określony układ obsadzeń powłok i podpowłok przez elektrony. Układ ten zapisujemy w sposób następujący: nl^k , gdzie n , l – odpowiednio główna i orbitalna liczba kwantowa, k – liczba elektronów na danej podpowłoce.

Omówimy teraz pobieżnie układ okresowy pierwiastków.

Wodór ($Z = 1$)

Dyskutowany szczegółowo poprzednio. Jeden elektron w stanie $1s$. Energia jonizacji wynosi 13,6 eV

Hel ($Z = 2$)

Dwa elektrony o przeciwnie ustawionych spinach w atomie $1s$. Energia jonizacji wynosi 24,6 eV. Zauważmy, że energia ta nie wzrasta czterokrotnie w stosunku do energii jonizacji atomu wodoru na skutek oddziaływania między elektronami. Duża energia jonizacji powoduje, że hel jest chemicznie obojętny. Hel kończy pierwszy okres układu okresowego pierwiastków.

Lit ($Z = 3$) – Neon ($Z = 10$)

W kolejności trzeci elektron zaczyna wypełnienie drugiej powłoki. Lit zaczyna drugi okres układu okresowego pierwiastków. Konfiguracja elektronów jest następująca: $1s^2 2s^1$. Występuje silne ekranowanie jądra przez elektrony w stanie $1s$ i dlatego elektron $2s$ jest słabo związany. Energia jonizacji Li wynosi 5,4 eV. Natomiast na oderwanie drugiego elektronu potrzeba aż 75,6 eV. Lit jest więc zawsze jednowartościowy. W atomie berylu ($Z = 4$) czwarty elektron wypełnia powłokę $2s$. Od boru ($Z = 5$) do neonu ($Z = 10$) elektrony obsadzają powłokę $2p$. W miarę wypełniania tej podpowłoki wzrasta energia jonizacji atomu, gdyż wzrasta Z , a jądro jest w każdym z tych przypadków ekranowane przez dwa elektrony w stanie $1s$. Dla boru energia jonizacji wynosi 8,3 eV, a dla neonu – 21,6 eV. Atom neonu o konfiguracji elektronów $1s^2 2s^2 2p^6$ posiada całkowicie zapełnioną drugą powłokę i zamyka drugi okres.

Tabela 13.1. Konfiguracje elektronowe atomów

Z	Symbol	Konfiguracja elektronowa	Energia jon. (eV)	Z	Symbol	Konfiguracja elektronowa	Energia jon. (eV)
1	H	$1s$	13.595	55	Cs	$[Xe] 6s$	3.893
2	He	$1s^2$	24.581	56	Ba	$6s^2$	5.210
3	Li	$[He] 2s$	5.390	57	La	$5d6s^2$	5.61
4	Be	$2s^2$	9.320	58	Ce	$4f^36s^2$	6.54
5	B	$2s^22p$	8.296	59	Pr	$4f^36s^2$	5.48
6	C	$2s^22p^2$	11.256	60	Nd	$4f^46s^2$	5.51
7	N	$2s^22p^3$	14.545	61	Pm	$4f^56s^2$	
8	O	$2s^22p^4$	13.614	62	Fm	$4f^66s^2$	5.6
9	F	$2s^22p^5$	17.418	63	Eu	$4f^76s^2$	5.67
10	Ne	$2s^22p^6$	21.559	64	Gd	$4f^75d^2$	6.16
11	Na	$[Ne] 3s$	5.138	65	Tb	$4f^96s^2$	6.74
12	Mg	$3s^2$	7.644	66	Dy	$4f^{10}6s^2$	6.82
13	Al	$3s^2p$	5.984	67	Ho	$4f^{11}6s^2$	
14	Si	$3s^2p^2$	8.149	68	Er	$4f^{12}6s^2$	
15	P	$3s^2p^3$	10.484	69	Tm	$4f^{13}6s^2$	
16	S	$3s^2p^4$	10.357	70	Yb	$4f^{14}6s^2$	6.22
17	Cl	$3s^2p^5$	13.01	71	Lu	$4f^{14}5d^26s^2$	6.15
18	Ar	$3s^2p^6$	15.755	72	Hf	$4f^{14}5d^26s^2$	7.0
19	K	$[Ar] 4s$	4.339	73	Ta	$4f^{14}5d^36s^2$	7.88
20	Ca	$4s^2$	6.111	74	W	$4f^{14}5d^46s^2$	7.98
21	Sc	$3d^24s^2$	6.54	75	Re	$4f^{14}5d^56s^2$	7.87
22	Ti	$3d^24s^2$	6.83	76	Os	$4f^{14}5d^66s^2$	8.7
23	V	$3d^34s^2$	6.74	77	Ir	$4f^{14}5d^76s^2$	9.2
24	Cr	$3d^54s$	6.76	78	Pt	$4f^{14}5d^86s^2$	8.88
25	Mn	$3d^54s^2$	7.432	79	Au	$[Xe, 4f^{14}5d^{10}]6s$	9.22
26	Fe	$3d^64s^2$	7.87	80	Hg	$6s^2$	10.434
27	Co	$3d^74s^2$	7.86	81	Tl	$6s^26p$	6.106
28	Ni	$3d^84s^2$	7.633	82	Pb	$6s^26p^2$	7.415
29	Cu	$3d^{10}4s$	7.724	83	Bi	$6s^26p^3$	7.287
30	Zn	$3d^{10}4s^2$	9.391	84	Po	$6s^26p^4$	8.43
31	Ga	$3d^{10}4s^24p$	6.00	85	At	$6s^26p^5$	
32	Ge	$3d^{10}4s^24p^2$	7.88	86	Rn	$6s^26p^6$	10.745
33	As	$3d^{10}4s^24p^3$	9.81	87	Fr	$[Rn] 7s$	5.277
34	Se	$3d^{10}4s^24p^4$	9.75	88	Ra	$7s^2$	6.9
35	Br	$3d^{10}4s^24p^5$	11.84	89	Ac	$6d7s^2$	
36	Kr	$3d^{10}4s^24p^6$	13.996	90	Th	$6d^27s^2$	
37	Rb	$[Kr] 5s$	4.176	91	Pa	$5f^26d7s^2$	
38	Sr	$5s^2$	5.692	92	U	$5f^36d7s^2$	4.0
39	Y	$4d^55s^2$	6.377	93	Np	$5f^46d7s^2$	
40	Zr	$4d^25s^2$	6.835	94	Pu	$5f^56d7s^2$	
41	Nb	$4d^45s$	6.881	95	Am	$5f^77s^2$	
42	Mo	$4d^55s$	7.10	96	Cm	$5f^76d7s^2$	
43	Tc	$4d^55s^2$	7.228	97	Bk	$5f^86d7s^2$	
44	Ru	$4d^75s$	7.365	98	Cf	$5f^{10}7s^2$	
45	Rh	$4d^85s$	7.461	99	Es	$5f^{11}7s^2$	
46	Pd	$4d^{10}$	8.33	100	Fm	$5f^{12}7s$	
47	Ag	$4d^{10}5s$	7.574	101	Mv	$5f^{13}7s^2$	
48	Cd	$4d^{10}5s^2$	8.991	102	No	$5f^{14}7s^2$	
49	In	$4d^{10}5s^25p$	5.785	103	Lw	$5f^{14}6d7s^2$	
50	Sn	$4d^{10}5s^25p^2$	7.342	104	Ku	$5f^{14}6d^27s^2$	
51	Sb	$4d^{10}5s^25p^3$	8.639				
52	Te	$4d^{10}5s^25p^4$	9.01				
53	I	$4d^{10}5s^25p^5$	10.454				
54	Xe	$4d^{10}5s^25p^6$	12.127				

Tabela 13.2. Układ okresowy pierwiastków

0																																				
1a	Niemetale										Półmetale							Niemetale																		
1	H 1,008	2	He 4,0026	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18																	
2	Li 6,939	Be 9,012	B 10,81	C 12,01	N 14,007	O 15,999	F 18,99	Ne 20,18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36										
3	Na 22,989	Mg 24,31	Al 26,98	Si 28,09	P 30,97	S 32,06	Cl 35,45	Ar 39,94	K 39,102	Ca 40,08	Sc 44,956	Ti 47,88	V 50,94	Cr 51,99	Mn 54,94	Fe 55,85	Co 58,93	Ni 58,71	Cu 63,54	Zn 65,37	Ga 69,72	Ge 72,59	As 74,92	Se 78,96	Br 79,91	Kr 83,80										
4	K 39,102	Ca 40,08	Sc 44,956	Ti 47,88	V 50,94	Cr 51,99	Mn 54,94	Fe 55,85	Co 58,93	Ni 58,71	Cu 63,54	Zn 65,37	Ga 69,72	Ge 72,59	As 74,92	Se 78,96	Br 79,91	Kr 83,80	Rb 85,47	Sr 87,62	Y 88,905	Zr 91,22	Nb 92,90	Mo 95,94	Tc 98,906	Ru 101,0	Rh 102,90	Pd 106,4	Ag 107,87	Cd 112,40	In 114,82	Sn 118,7	Sb 121,75	Te 127,6	I 126,9	Xe 131,7
5	Rb 85,47	Sr 87,62	Y 88,905	Zr 91,22	Nb 92,90	Mo 95,94	Tc 98,906	Ru 101,0	Rh 102,90	Pd 106,4	Ag 107,87	Cd 112,40	In 114,82	Sn 118,7	Sb 121,75	Te 127,6	I 126,9	Xe 131,7	Cs 132,90	Ba 137,34	La 138,91	Ce 140,12	Pr 140,91	Nd 144,24	Pm (147)	Sm 150,35	Eu 151,96	Gd 157,25	Tb 158,92	Dy 162,5	Ho 164,93	Er 167,26	Tm (169)	Yb 173,04	Lu 174,97	
6	Cs 132,90	Ba 137,34	La 138,91	Ce 140,12	Pr 140,91	Nd 144,24	Pm (147)	Sm 150,35	Eu 151,96	Gd 157,25	Tb 158,92	Dy 162,5	Ho 164,93	Er 167,26	Tm (169)	Yb 173,04	Lu 174,97	Rn (222)	Fr (223)	Ra (226)	Ac (227)	Th 232,038	Pa (231)	U 238,03	Np (237)	Pu (244)	Am (243)	Cm (247)	Bk (247)	Cf (251)	Es (254)	Fm (253)	Md (258)	No (259)	Lw (261)	
7	Fr (223)	Ra (226)	Ac (227)	Th 232,038	Pa (231)	U 238,03	Np (237)	Pu (244)	Am (243)	Cm (247)	Bk (247)	Cf (251)	Es (254)	Fm (253)	Md (258)	No (259)	Lw (261)	Rn (222)	Fr (223)	Ra (226)	Ac (227)	Th 232,038	Pa (231)	U 238,03	Np (237)	Pu (244)	Am (243)	Cm (247)	Bk (247)	Cf (251)	Es (254)	Fm (253)	Md (258)	No (259)	Lw (261)	
Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	Metale	Metale	Lantanowce	Aktynowce	

Sód ($Z = 11$) – Argon ($Z = 18$)

Trzeci okres rozpoczyna atom sodu Na o konfiguracji elektronów $1s^2 2s^2 2p^6 3s^1$, a kończy atom argonu o całkowicie wypełnionej podpowłoce $3p$ ($1s^2 2s^2 2p^6 3s^2 3p^6$). Energia jonizacji wzrasta od 5,1 eV dla Na do 15,8 eV dla Ar. Sód wykazuje dużą aktywność chemiczną, argon jest chemicznie obojętny.

Potas ($Z = 19$) i dalej

Czwarty okres rozpoczyna potas ($Z = 19$). Dziewiętnasty elektron w atomie potasu nie wypełnia podpowłoki $3d$, ale zajmuje stan $4s$. Jest to bowiem stan o niższej energii niż stan $3d$. Zatem konfiguracja elektronów w atomie potasu jest następująca: $1s^2 2s^2 2p^6 3s^2 3p^6 4s^1$. W atomie wapnia zostaje wypełniona podpowłoka $4s^2$, a od skandiu ($Z = 21$) zaczyna się wypełnianie powłoki $3d$. Pierwiastki od skandiu do niklu ($Z = 28$) zwane są pierwszą grupą przejściową. Pierwiastki te mają bardzo podobne właściwości chemiczne i prawie jednakowe energie jonizacji: od 6,5 eV dla Sc do 7,6 eV dla Ni. Dzieje się tak dlatego, że wypełniona podpowłoka $4s$ ekranuje od wpływów zewnętrznych wypełnioną podpowłokę $2d$. Warto zapamiętać, że żelazo, kobalt i nikiel mają niezapełnioną podpowłokę $3d$, co określa właściwości magnetyczne tych pierwiastków. Miedź ($Z = 29$) ma już nieco inne właściwości chemiczne, ponieważ atom Cu ma już tylko jeden elektron na podpowłoce $4s$.

Nie będziemy dalej analizować konfiguracji elektronów w atomach. Poznaliśmy już ogólne zasady obsadzania podpoziomów przez elektrony i wynikające stąd właściwości pierwiastków.

13.4. Promieniowanie atomów wzbudzonych

Po wprowadzeniu równania Schrödingera, mechanikę kwantową zaczęto stosować w teorii oddziaływań elektromagnetycznych, w wyniku czego powstała elektrodynamika kwantowa. Szczegółowe rozważania tej teorii wychodzą poza ramy skryptu. Dla nas istotny jest fakt ustanowiony przez elektrodynamikę kwantową, że naładowana cząstka może pochłonąć lub emitować pojedyncze fotony, że teoria pozwala dokładnie obliczyć odpowiednie prawdopodobieństwa pochłonięcia lub emisji fotonu.

13.4.1. Widma optyczne

Atomy po zaabsorbowaniu pewnej porcji energii przechodzą do stanu wzbudzonego. Jeżeli dostarczona energia wynosi kilka lub kilkanaście elektronowoltów, to możliwe są jedynie przejścia elektronów walencyjnych na wyższe poziomy energetyczne. Wiemy bowiem, że energie wiązania tych elektronów wynoszą kilka-kilkanaście elektronowoltów. Takie energie mogą atomy uzyskać poprzez:

- ogrzewanie, przy którym energia jest przekazywana w wyniku zderzeń w ruchu cieplnym,

- wyładowania elektryczne,
- oświetlenie promieniowaniem widzialnym i nadfioletowym,
- reakcje chemiczne.

Wzbudzone atomy samorzutnie przechodzą do stanu niższego, a każdemu przejściu towarzyszy emisja kwantu promieniowania o częstości ν_{nm} równej

$$\nu_{nm} = \frac{E_n - E_m}{h},$$

gdzie E_n i E_m – energia elektronu odpowiednio na wyższym i niższym poziomie energetycznym. Linie emisyjne promieniowania uwarunkowane tymi przejściami energetycznymi, nazywamy widmem optycznym, gdyż widmo to leży w obszarze widzialnym, bliskim nadfioletu i bliskiej podczerwieni. Układ linii jest na ogół bardzo złożony, ale charakterystyczny dla danego pierwiastka. Fakt ten wykorzystuje się w analizie widmowej do określania jakościowego i ilościowego składu danej substancji.

Obserwując widma można stwierdzić, że nie zachodzą przejścia między wszystkimi poziomami energetycznymi. Okazuje się, że możliwe są tylko przejścia między takimi poziomami, których liczby kwantowe l i j spełniają tzw. reguły wyboru

$$\Delta l = \pm 1,$$

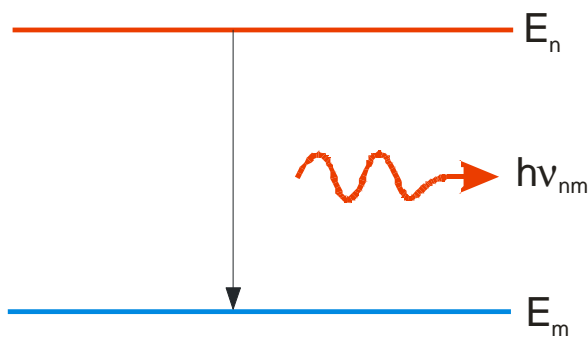
$$\Delta j = 0, \pm 1.$$

Zatem przejścia zachodzą między poziomami, których liczby kwantowe l różnią się o jeden, a liczby kwantowe j są jednakowe lub różnią się o jeden. Mówiąc ściślej, reguły powyższe nie zabraniają w sposób bezwzględny innych przejść, lecz mówią tylko, że przejścia te są mało prawdopodobne.

Jeżeli atom jest wzbudzony do wyższego stanu z którego może powrócić do stanu niżej położonego na drodze takiego właśnie mało prawdopodobnego przejścia, to może on pozostać w tym stanie przez bardzo długi okres czasu. Poziom taki nazywamy metatrwałym (metastabilnym). O ile typowy czas życia atomu w stanie wzbudzonym wynosi około 10^{-8} s, to poziomy metatrwały charakteryzują się czasami życia rzędu 10^{-2} s, tzn. ich czasy życia są miliony razy dłuższe.

13.4.2. Promieniowanie spontaniczne i wymuszone

Wiemy już, że atom w stanie wzbudzonym przechodzi samorzutnie do stanu energetycznie niższego, emitując przy tym kwant promieniowania. Tego rodzaju emisję nazywamy emisją spontaniczną (samorzutną). Przedstawiono ją schematycznie na rys. 13.4. Omówimy nieco bliżej ten problem.



Rys. 13.4. Emisja spontaniczna.

Niech na wyższym wzbudzonym poziomie energetycznym będzie w chwili początkowej $N_n(0)$ atomów. Liczba przejść spontanicznych w czasie dt będzie proporcjonalna do $N_n(t) dt$. Należy także oczekiwać, że szybkość przejść będzie różna dla różnych atomów. Możemy więc napisać, że na skutek emisji spontanicznej z poziomu n -tego ubędzie dN_n atomów, przy czym

$$dN_n = -A_{nm} N_n dt, \quad (13.1)$$

gdzie A_{nm} jest stałą charakteryzującą szybkość przejść między poziomem n -tym i m -tym. Jest to współczynnik emisji spontanicznej. Znak minus występuje dlatego, że na skutek emisji liczba atomów na poziomie n -tym zmniejsza się. Pisząc wzór (13.1) w postaci

$$\frac{dN_n}{N_n} = -A_{nm} dt$$

i całkując otrzymujemy:

$$\ln N_n = -A_{nm} t + C.$$

Dla $t = 0$, $N_n = N_n(0)$, a więc $C = \ln N_n(0)$. Zatem

$$N_n(t) = N_n(0) e^{-A_{nm} t}.$$

Aby bliżej określić stałą A_{nm} , obliczamy średni czas życia atomów w stanie wzbudzonym. Liczba atomów dN_n , które przeżyły w stanie wzbudzonym czas t , a następnie w przedziale czasu $(t, t + dt)$ przeszły do niższego stanu, wynosi

$$dN_n = N_n(t) A_{nm} dt = N_n(0) e^{-A_{nm} t} A_{nm} dt.$$

Ponieważ każdy z tych atomów był w stanie wzbudzonym przez czas t , wobec tego czas życia dN atomów wynosi

$$t dN_n = N_n(0) e^{-A_{nm} t} A_{nm} t dt.$$

Całkowity czas życia wszystkich atomów jest równy

$$T = \int_0^{\infty} N_n(0) e^{-A_{nm} t} A_{nm} t dt.$$

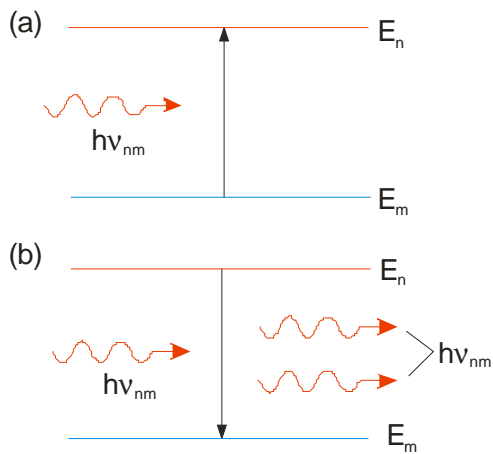
A zatem średni czas życia atomów na n -tym poziomie (ze względu na przejście na poziom m -ty) wynosi

$$\tau_{nm} = \frac{T}{N_n(0)} = A_{nm} \int_0^{\infty} e^{-A_{nm}t} dt = \frac{1}{A_{nm}}. \quad (13.2)$$

Tak więc współczynnik spontanicznego przejścia, określający szybkość przejść dla emisji spontanicznej, jest równy odwrotności średniego czasu życia atomów w stanie wzbudzonym i określa prawdopodobieństwo przejścia spontanicznego w jednostce czasu do stanu o niższej energii.

Teoretyczne wyliczenia średniego czasu życia atomu w stanie wzbudzonym dają wartość około 10^{-8} s, tzn. stała emisji spontanicznej A_{nm} wynosi około 10^8 s.

Oprócz emisji spontanicznej występować może także emisja wymuszona, co pokazano schematycznie na rys. 13.5. W procesie absorpcji wymuszonej foton promieniowania o częstotliwości ν_{nm} powoduje przejście atomu z poziomu niższego na poziom wyższy, przy czym foton zostaje zaabsorbowany przez atom.



Rys. 13.5. Absorpcja (a) i emisja (b) wymuszona.

W procesie emisji wymuszonej padający foton powoduje przejście wzbudzonego atomu do stanu energetycznie niższego. W wyniku tego przejścia wzrasta liczba fotonów, gdyż obok fotonu padającego pojawia się drugi foton, emitowany przez atom. Fakt, że foton promieniowania może spowodować przejście z poziomu niższego na wyższy i odwrotnie ma swoją analogię klasyczną. Jeżeli na oscylator harmoniczny działa periodycznie siła zewnętrzna, to oscylator wykonuje drgania wymuszone o częstotliwości siły zewnętrznej. Jeżeli

częstotliwość ta jest równa częstotliwości rezonansowej oscylatora, to wówczas opóźnienie fazowe drgań oscylatora w stosunku do siły wymuszającej wynosi 90° i siła ta przekazuje energię oscylatorowi. Gdyby jednak oscylator wyprzedzał w fazie siłę zewnętrzną to wówczas oscylator przekazuje energię na zewnątrz. Traktując atom jako oscylator, a siłę zewnętrzną jako pole energetyczne promieniowania, otrzymujemy klasyczne wyjaśnienie emisji i absorpcji wymuszonej.

Możemy przyjąć, że częstotliwość przejść wymuszonych między poziomami n i m będzie proporcjonalna do gęstości widmowej promieniowania $u(\nu_{nm})$. Gęstość energii promieniowania jest proporcjonalna do zdolności emisyjnej. Rozważania których nie będziemy przeprowadzać prowadzą do wniosku, że $\varepsilon(\nu, T) = (1/4)cu(\nu, T)$. Korzystając ze wzoru (10.9) otrzymamy

$$u(\nu_{nm}) = \frac{8\pi h \nu_{nm}^3}{c^3} \frac{1}{\exp(h\nu_{nm}/kT) - 1}. \quad (13.3)$$

Analogicznie do wzoru (13.1) napiszemy, że ilość przejść wymuszonych z poziomu n -tego na poziom m -ty w czasie dt będzie wynosiła

$$dN_n = B_{nm}u(\nu_{nm})N_n dt ,$$

natomiast liczba przejść z poziomu m -tego na n -ty

$$dN_m = B_{mn}u(\nu_{nm})N_m dt ,$$

gdzie B_{nm} jest współczynnikiem emisji wymuszonej, a B_{mn} – współczynnikiem absorpcji wymuszonej. W wyrażeniach powyższych opuściliśmy znak minus, gdyż interesuje nas liczba przejść, a nie zmiana liczby atomów.

Współczynniki A_{nm} , B_{nm} i B_{mn} charakteryzujące procesy emisji spontanicznej, emisji i absorpcji wymuszonej nazywamy także współczynnikami Einsteina.

Jeżeli układ składający się z N_n atomów na poziomie n -tym i N_m atomów na poziomie m -tym jest w określonej temperaturze T w równowadze z promieniowaniem elektromagnetycznym o gęstości widmowej $u(\nu_{nm})$, to liczba przejść z poziomu wyższego na niższy musi być taka sama jak z niższego na wyższy. Możemy więc zapisać

$$A_{nm}N_n dt + B_{nm}u(\nu_{nm})N_n dt = B_{mn}u(\nu_{nm})N_m dt . \quad (13.4)$$

Liczby atomów na n -tym i m -tym poziomie możemy wyrazić przez funkcję rozkładu Boltzmann'a (patrz pkt. 15.4). Mamy więc

$$\frac{N_n}{N_m} = \frac{\exp(-E_n/kT)}{\exp(-E_m/kT)} = \exp[-(E_n - E_m)/kT] = \exp(-h\nu_{nm}/kT). \quad (13.5)$$

Po uwzględnieniu tego wzoru w (13.4), otrzymujemy

$$A_{nm} + B_{nm}u(\nu_{nm}) = B_{mn}u(\nu_{nm})e^{h\nu_{nm}/kT}. \quad (13.6)$$

Z ostatniego wzoru obliczymy $u(\nu_{nm})$

$$u(\nu_{nm}) = \frac{\frac{A_{nm}}{B_{nm}}}{\frac{B_{mn}}{B_{nm}} \exp(h\nu_{nm}/kT) - 1}.$$

Jeżeli porównać powyższe wyrażenie z (13.3), to otrzymujemy

$$\frac{B_{nm}}{B_{mn}} = 1, \quad (13.7)$$

$$\frac{A_{nm}}{B_{nm}} = \frac{8\pi h \nu_{nm}^3}{c^3}. \quad (13.8)$$

Z wyrażenia (13.7) wynika, że współczynniki emisji i absorpcji wymuszonej są równe. Ponieważ jednak w stanie równowagi termodynamicznej $N_n \ll N_m$, to absorpcja wymuszona zdecydowanie przewyższa emisję wymuszoną. Oznacza to, że promieniowanie elektromagnetyczne jest absorbowane w układzie. Gdyby jednak doprowadzić do sytuacji w której $N_n > N_m$, to wówczas promieniowanie będzie w układzie wzmacniane. Fakt ten jest wykorzystywany w laserach. Musimy jednak pamiętać, że warunek $N_n > N_m$ nie może być nigdy spełniony dla atomów będących w stanie równowagi termodynamicznej z otoczeniem.

Powróćmy jeszcze do przypadku, kiedy stosunek N_n / N_m jest dany równaniem (13.5). Wówczas na podstawie zależności (13.6) możemy wyliczyć stosunek prawdopodobieństwa A_{nm} emisji spontanicznej do prawdopodobieństwa $B_{nm}u(\nu_{nm})$ emisji wymuszonej.

$$\frac{A_{nm}}{B_{nm}u(\nu_{nm})} = \exp(h\nu_{nm} / kT) - 1.$$

Dla promieniowania w obszarze widzialnym $\nu_{nm} = 5 \times 10^{14}$ Hz, w temperaturze 300 K otrzymujemy

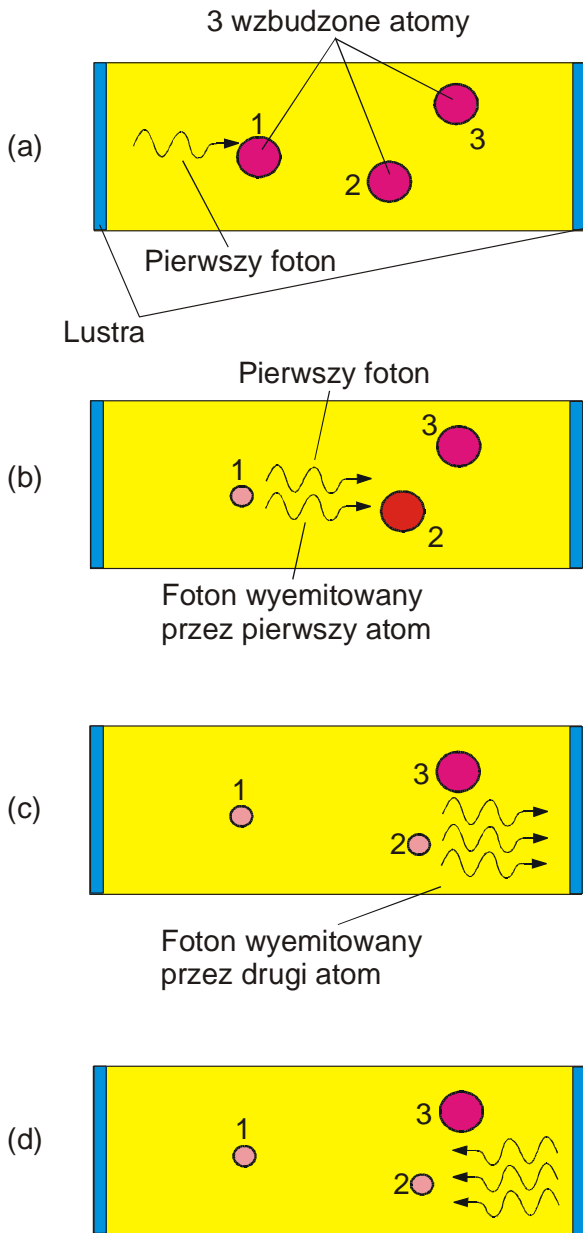
$$\frac{A_{nm}}{B_{nm}u(\nu_{nm})} \gg 1,$$

tzn. emisja spontaniczna jest zdecydowanie bardziej prawdopodobna niż emisja wymuszona. Mówiąc inaczej, promieniowanie emitowane przez klasyczne źródła światła jest promieniowaniem spontanicznym.

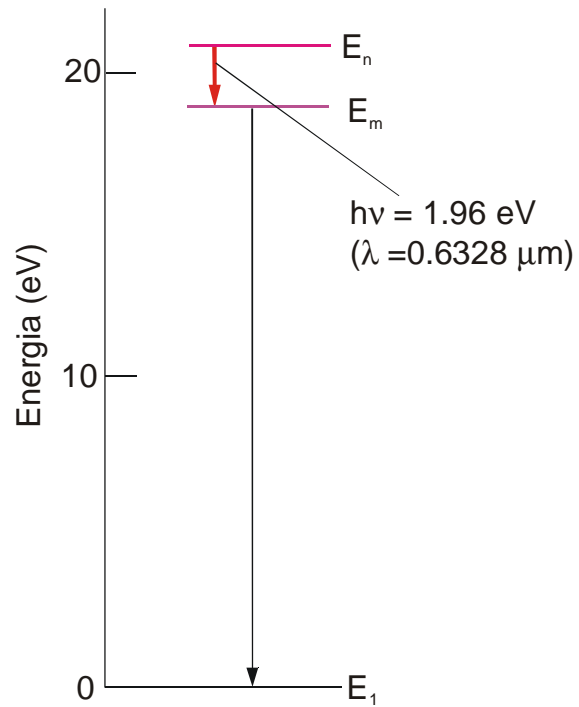
13.4.3. Laser

Założmy, że mamy próbkę zawierającą atomy z których większość znajduje się w stanie wzbudzonym E_n . Taki stan obsadzeń można osiągnąć np. w wyniku naświetlania substancji światłem (jest to tzw. pompowanie optyczne), w wyniku zderzeń ze strumieniem elektronów i zderzeń z innymi wzbudzonymi atomami. Prześledzimy teraz foton, który był emitowany jako pierwszy. Przechodząc przez ośrodek foton ten stymuluje wymuszone promieniowanie napotkanych atomów. Jeśli próbkę umieścimy pomiędzy dwoma zwierciadłami, to proces ma charakter reakcji łańcuchowej i zachodzi tak długo, dopóki nie wypromieniują wszystkie atomy (rys. 13.6). Przy tym wszystkie fotony są w fazie jeden z drugim. Jeżeli jedno ze zwierciadeł jest częściowo odbijające (półprzezroczyste), to wychodzące na zewnątrz światło jest koherentne. Stanowi ono ciągłą monochromatyczną falę elektromagnetyczną. Oczywiście, aby proces emisji wymuszonej przeważał nad procesem pochłaniania, większość atomów powinna znajdować się w stanie wzbudzonym. Na rys. 13.7 przytoczono schemat poziomów lasera helowo-neonowego, w którym w stanie E_n znajduje się więcej atomów niż w stanie E_m . W laserze helowo-neonowym atomy neonu przechodzą do wzbudzonego stanu E_n na skutek zderzeń ze wzbudzonymi atomami helu. Przejście do stanu E_m uwarunkowane jest emisją wymuszoną.

Następnie na skutek zderzeń ze ściankami, atomy neonu szybko przechodzą do stanu podstawowego.



Rys. 13.6. Promieniowanie wymuszone w laserze: (a) pierwszy foton pada na wzbudzony atom, (b) atom emituje foton, który pada na drugi wzbudzony atom, (d) trzy odbite fotony.

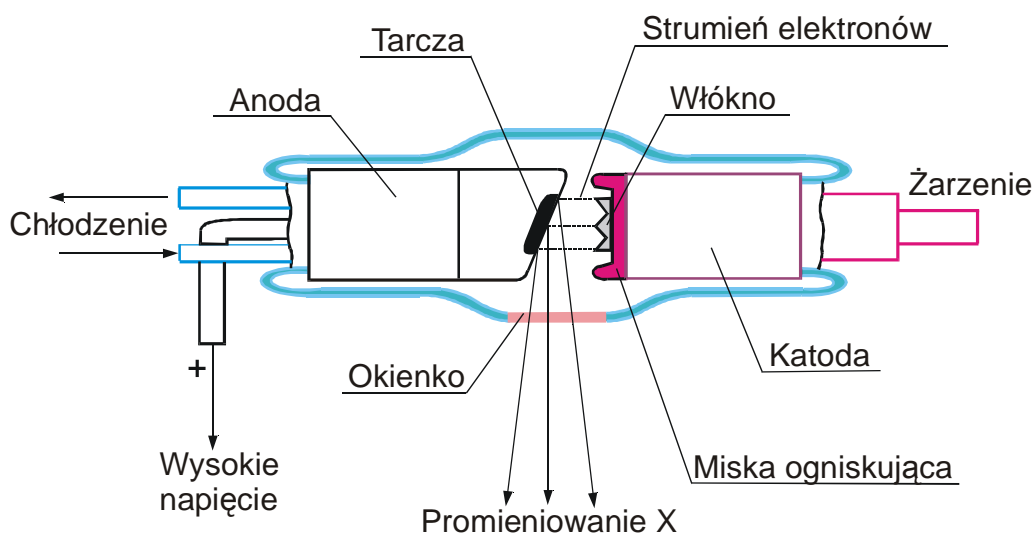


Rys. 13.7. Trzy poziomy energetyczne atomu neonu.

Do czasu opracowania lasera w 1960 r. wszystkie dostępne człowiekowi źródła promieniowania w podczerwieni, zakresie widzialnym i ultrafioletowym były niekoherentne, tj. fazy pojedynczych fotonów miały przypadkowe wartości względem siebie. Z chwilą pojawienia się pierwszego lasera osiągnięto znaczne sukcesy w ich technologii co pozwoliło budować lasery impulsowe jak i lasery ciągłego działania na zakres długości fal promieniowania podczerwonego, widzialnego i ultrafioletowego.

13.5. Promieniowanie rentgenowskie

Omówiliśmy powyżej powstawanie i niektóre właściwości widm optycznych. Przypomnijmy, że powstają one w wyniku dostarczania do atomu energii od kilku do kilkunastu elektronowoltów. W takim przypadku ulegają wzbudzeniu tylko elektrony z powłoki zewnętrznej. Natomiast elektrony z powłok wewnętrznych mają energię wiązania 10^4 – 10^5 eV. Dostarczając do atomu takiej energii można wybić elektron z tych powłok. Atom pozostaje wówczas w stanie bardzo wysokiego wzbudzenia, gdyż ubył silnie związany elektron. Atom powraca do stanu podstawowego w wyniku przejść elektronów z orbit wyższych i w konsekwencji emisji fotonów o bardzo wysokiej częstotliwości. Otrzymamy wówczas widmo liniowe, które nazywamy liniowym widmem rentgenowskim (odkrył je Röntgen w 1895 r.) Długości fal promieniowania rentgenowskiego wynoszą zwykle 10^{-9} – 10^{-11} m. Z doświadczeń wynika, że oprócz widma liniowego występuje również widmo ciągłe promieniowania rentgenowskiego.



Rys.13.8. Schemat lampy rentgenowskiej

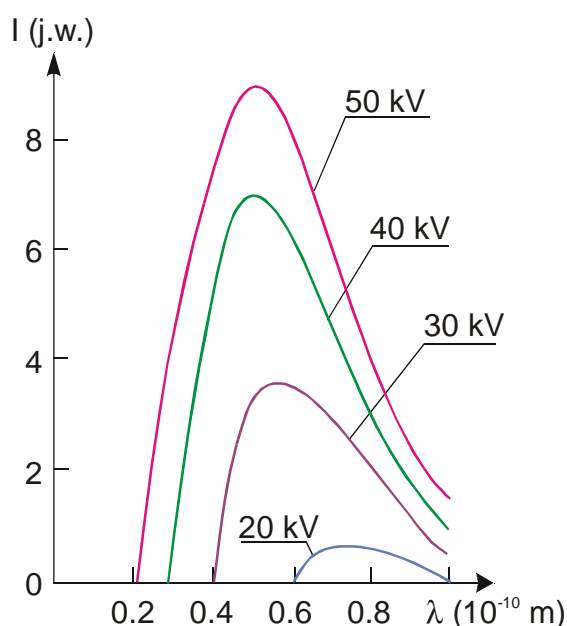
Promieniowanie rentgenowskie zwane także promieniowaniem X (gdyż początkowo nie była znana jego natura) wytwarzane jest w lampie rentgenowskiej, której schemat przedstawiono na rys. 13.8. Elektrony emitowane przez żarzoną katodę są przyspieszane przez różnicę potencjałów rzędu 10^5 V i uderzają w anodę, gdzie są hamowane. Hamowanie to jest wynikiem oddziaływania kulombowskiego padających elektronów z jądrami atomów materiału anody. Elektron poruszający się z opóźnieniem a emituje falę elektromagnetyczną o mocy proporcjonalnej do a^2 . Ponieważ $a = F/m \sim Z/m$, więc moc promieniowania jest proporcjonalna do Z^2/m^2 . Dlatego do wytwarzania promieniowania rentgenowskiego stosuje się elektrony (o małej masie m) oraz anody o dużym Z . Jeśli całkowita energia elektronu zmienia się na skutek hamowania w anodzie na energię promieniowania, to wówczas energia kwantu promieniowania będzie maksymalna i wyrazi się wzorem

$$h\nu_{max} = h \frac{c}{\lambda_{min}} = eV,$$

gdzie V jest napięciem między katodą i anodą.

Energia elektronu jest również w bardzo dużym stopniu tracona w wyniku kolejnych zderzeń z atomami anody. Na skutek tego powstają kwanty promieniowania o częstotliwości mniejszej od częstotliwości maksymalnej. Ponieważ w wyniku zderzeń elektrony mogą tracić różne ilości energii, dlatego widmo promieniowania wytwarzane przez hamowane elektrony będzie widmem ciągłym (rys. 13.9), odciętym od strony krótkich fal. Maksimum natężenia tego promieniowania odpowiada najbardziej prawdopodobnemu stopniowi przemiany energii elektronu w energię promieniowania. Okazuje się, że długość fali λ_{Imax} dla której występuje maksimum natężenia

wynosi



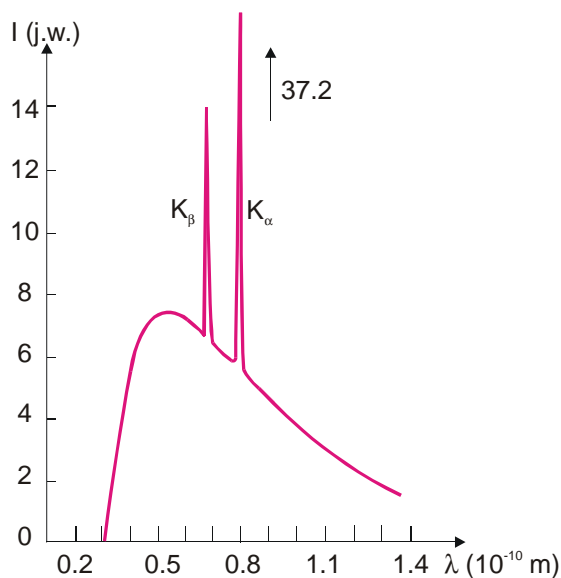
$$\lambda_{Imax} = 1.5\lambda_{min}.$$

Wraz ze wzrostem V rośnie więc energia fotonów. Wzrasta także liczba fotonów emitowanych w jednostce czasu.

Jak już wspomniano wcześniej, oprócz widma ciągłego, anoda rentgenowska emituje również widmo liniowe, zwane także charakterystycznym, gdyż zależy ono od materiału anody. Widmo charakterystyczne składa się z szeregu linii i występuje na tle widma ciągłego. Na rys. 13.10 przedstawiono widmo promieniowania rentgenowskiego emitowanego z anody molibdenowej.

Rys. 13.9. Zależność natężenia widma ciągłego od długości fali i napięcia przyłożonego do lampy rentgenowskiej.

Mechanizm powstawania rentgenowskiego widma liniowego jest następujący. Jeśli elektron z wiązki padającej przechodzi dostatecznie blisko elektronów znajdujących się na powłokach wewnętrznych atomu materiału anody, to występuje bardzo silne oddziaływanie kulombowskie między tymi elektronami. W wyniku tego elektron z powłoki wewnętrznej może zostać wybity i przejść na leżące wyżej poziomy energetyczne lub opuścić atom. Wówczas atom jest w stanie wzbudzonym i elektrony z wyższych powłok przechodzą na niższą, opróżnioną powłokę. Atom emituje kwant promieniowania o energii równej różnicy energii odpowiednich poziomów energetycznych. Ponieważ powłoki o $n = 1, 2, 3, 4$ itd., przyjęto oznaczać w spektroskopii rentgenowskiej odpowiednio przez K, L, M, N , itd., rozróżnia się serie K, L, M , itp. Seria K powstaje, jeżeli elektron przechodzi na opróżnione miejsce w powłoce K . Przy symbolu serii podaje się jeszcze indeksy $\alpha, \beta, \gamma, \dots$, dla zaznaczenia powłok z których następuje przeskok (z L na K – linia K_{α} , z M na K – linia K_{β}).



Rys. 13.10. Widmo rentgenowskie ciągłe i charakterystyczne molibdenu.

Na podstawie wcześniejszych rozważań wiemy, że energia wiązania elektronu w atomie wieloelektronowym wynosi

$$E = -\frac{me^4}{32\pi^2 \epsilon_0^2 \hbar^2} \frac{1}{n^2} Z_{ef}^2.$$

Na pierwszej powłoce $Z_{ef} = Z - 1$, na drugiej $Z_{ef} = Z - 7.4$. Wobec tego energia kwantów promieniowania dla serii K wyraża się wzorem

$$h\nu = \frac{me^4}{32\pi^2 \epsilon_0^2 \hbar^2} \left(1 - \frac{1}{n^2}\right) (Z-1)^2 \quad n = 2, 3, 4.$$

natomiast dla serii L

$$h\nu = \frac{me^4}{32\pi^2 \epsilon_0^2 \hbar^2} \left(\frac{1}{2^2} - \frac{1}{n^2}\right) (Z-7.4)^2 \quad n = 3, 4, 5...$$

Ogólnie można powiedzieć, że dla promieniowania rentgenowskiego

$$\sqrt{\nu} \sim (Z - b),$$

gdzie b jest stałą dla danej serii. Prawo to nosi nazwę prawa Moseleya i zostało sformułowane na długo przed powstaniem mechaniki kwantowej.

Promieniowanie rentgenowskie przechodząc przez warstwę materii ulega osłabieniu w wyniku rozproszenia i absorpcji. Zmiana natężenia promieniowania dI po przejściu przez warstwę o grubości dx wynosi

$$dI = -\alpha I dx, \quad (13.9)$$

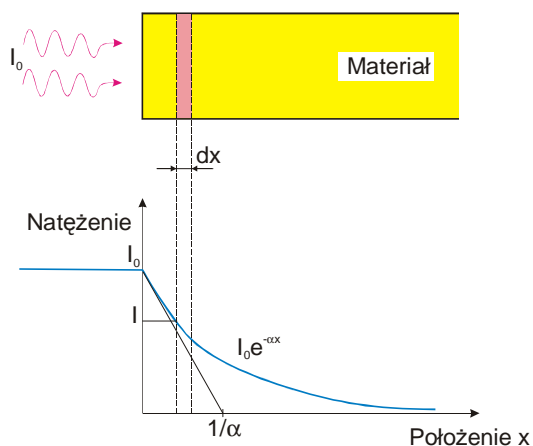
gdzie α jest liniowym współczynnikiem pochłaniania.

Zakładając, że natężenie promieniowania padającego na warstwę o grubości x wynosi I_0 (patrz rys. 13.11), z równania (13.9) otrzymujemy natężenie promieniowania I po przejściu przez warstwę o grubości x

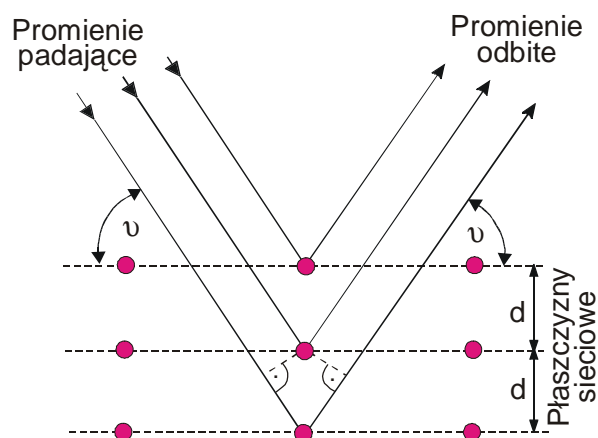
$$I(x) = I_0 e^{-\alpha x}. \quad (13.10)$$

Zależność absorpcji promieni rentgenowskich od rodzaju ciała absorbującego umożliwia ich zastosowanie w diagnostyce medycznej.

Ponieważ długości fal promieniowania rentgenowskiego są porównywalne z odległościami atomów w ciałach stałych, można je z powodzeniem również stosować do określania struktury krystalicznej. Dziedziną nauki tym się zajmującą jest rentgenografia.



Rys. 13.11. Absorpcja promieniowania rentgenowskiego



Rys. 13.12. Dyfrakcja wiązki promieni rentgenowskich na kryształach.

Wiązka promieni padających ulegając dyfrakcji zostaje częściowo odbita przez płaszczyzny sieciowe. Wzmocnienie interferencyjne promieni odbitych od kolejnych płaszczyzn zachodzi wówczas, gdy różnice dróg tych promieni są równe całkowitej wielokrotności długości fali w kryształach. Na podstawie rys. 13.12 można stwierdzić, że zachodzi to wówczas, gdy jest spełniony tzw. warunek Wulfa-Bragga wyrażony wzorem

$$2d \sin \vartheta = n\lambda, \quad (13.11)$$

gdzie n – liczba całkowita; d – odległość między płaszczyznami sieciowymi kryształu. Ponieważ $\sin \vartheta$ nie może być większy od jedności, natomiast n jest większe od jedności, zatem odbicie zachodzi dla fal o długości większej niż $2d$.

ROZDZIAŁ 14

SIEĆ KRYSTALICZNA CIAŁ STAŁYCH

W poprzednim rozdziale zapoznaliśmy się z budową powłok elektronowych atomu. Atomy, dzięki występującym między nimi siłom przyciągania, mogą łączyć się w cząsteczki chemiczne (molekuły) albo w zespoły zawierające olbrzymie ilości atomów, które w zależności od stanu skupienia zaliczamy do ciał stałych lub cieczy. Przedmiotem naszych rozważań w dalszych częściach skryptu będą właściwości elektronowe ciał stałych, a w szczególności półprzewodników, dlatego w rozdziale tym zaznajomimy się z budową tych ciał.

Ciało stałe definiuje się jako ciało odznaczające się stabilnością kryształu. W zależności od stopnia uporządkowania struktury wewnętrznej ciała stałe dzielimy na krystaliczne i amorficzne, czyli bezpostaciowe.

Jeśli położenia atomów wykazują okresowe uporządkowanie w przestrzeni (uporządkowanie dalekiego zasięgu) to takie ciało stałe nazywamy ciałem krystalicznym lub kryształem. W przypadku gdy uporządkowanie występuje w całej objętości kryształu, to taki kryształ nazywamy monokryształem. Jeżeli natomiast prawidłowe uporządkowanie występuje tylko wewnątrz pewnych obszarów zwanych ziarnami krystalicznymi, wówczas mówimy o polikryształach. Zwykle większość występujących w przyrodzie i otrzymanych sztucznie ciał krystalicznych stanowią polikryształy.

Jedną z cech charakterystycznych kryształu jest anizotropia jego właściwości polegająca na tym, że kryształ wykazuje w różnych kierunkach różne właściwości fizyczne.

W odróżnieniu od kryształów, ciecze i ciała amorficzne (których typowym przykładem jest szkło) nie mają prawidłowej struktury krystalicznej i nie wykazują anizotropii właściwości. W ciałach tych występuje tzw. uporządkowanie bliskiego zasięgu.

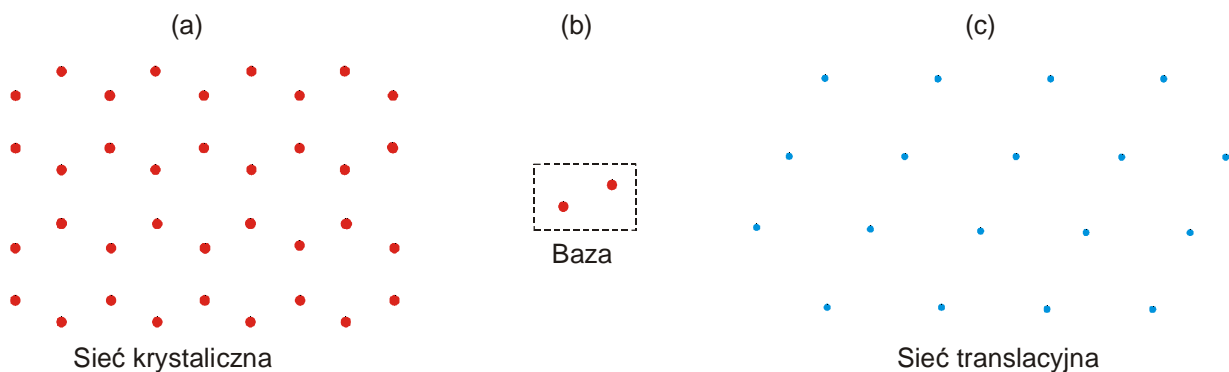
Ciała krystaliczne charakteryzują się ponadto skokową zmianą właściwości przy przejściach z jednej fazy do drugiej. W ciałach amorficznych właściwie nie istnieje granica między fazami (ciekłą i stałą). Z tego powodu ciała amorficzne są często traktowane jako przechłodzone ciecze.

W tym miejscu należy zaznaczyć, że niektóre ciecze również charakteryzują się uporządkowaniem cząsteczek. Związana z tym jest anizotropia ich różnych właściwości fizycznych, mechanicznych, optycznych, elektrycznych, magnetycznych i wielu innych. Takie ciecze nazywamy ciekłymi kryształami.

14.1. Sieć krystaliczna i układy krystalograficzne

Idealny nieskończony kryształ jest strukturą utworzoną przez regularne powtarzanie w przestrzeni identycznych elementów. Ten charakterystyczny dla danego ciała układ atomów nazywamy jego siecią krystaliczną.

Najmniejszy pojedynczy element sieci, powtarzający się periodycznie w trzech wymiarach przestrzeni, nazywamy bazą tej sieci. W najprostszym przypadku baza jest utworzona z jednego atomu danego pierwiastka. Na ogół zawiera ona 2 lub więcej atomów, przy czym mogą to być atomy tego samego lub różnych pierwiastków. Przykładowo, w kryształach diamentu bazę stanowią dwa atomy węgla, a w kryształach soli kuchennej jeden atom sodu i jeden atom chloru. Grupy atomów stanowiące bazę są ustawione w kryształach wg określonej sieci przestrzennej, którą nazywamy siecią translacyjną lub siecią punktową. Jest ona utworzona z punktów nazywanych węzłami sieci (rys. 14.1). W celu ułatwienia zobrazowania obrano tu sieć dwuwymiarową. Można jednak z łatwością rozszerzyć rozważania na sieć trójwymiarową.



Rys. 14.1. Dwuwymiarowa sieć krystaliczna (a), jej baza (b) i sieć translacyjna (c).

Symetria translacyjna sieci wyraża się tym, że położenie dowolnego węzła w stosunku do węzła przyjętego za punkt odniesienia, może być określone za pomocą wektora translacji

$$\vec{R} = m\vec{a} + n\vec{b} + p\vec{c},$$

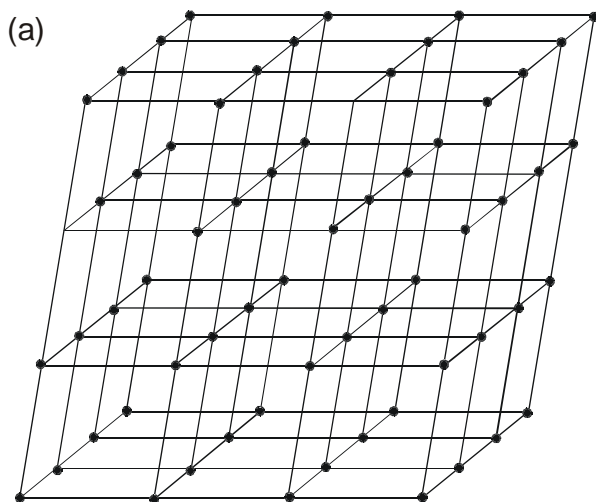
gdzie: m, n, p – liczby całkowite, zaś $\vec{a}, \vec{b}, \vec{c}$ – odpowiednio dobrane wektory, nazywane elementarnymi wektorami translacji. Długości tych wektorów, tzn. liczby a, b, c nazywamy stałymi sieci.

Równoległoscian zbudowany na elementarnych wektorach translacji nazywamy komórką elementarną prostą lub prymitywną. Jak widać z rys. 14.2, sieć przestrzenną kryształu stanowi nieskończony zbiór jednakowych komórek ułożonych równoległe względem siebie i wypełniających całą przestrzeń.

W komórce prostej węzły sieci przestrzennej znajdują się tylko w narożach, a zatem na jedną komórkę przypada jeden węzeł. W niektórych rodzajach sieci wygodniej jest posługiwać się komórkami o większych rozmiarach, ale za to o niższym stopniu symetrii niż komórka prosta.

Są to tzw. komórki złożone, które oprócz węzłów w narożach zawierają dodatkowe węzły. Sieci takie nazywamy sieciami centrowanymi.

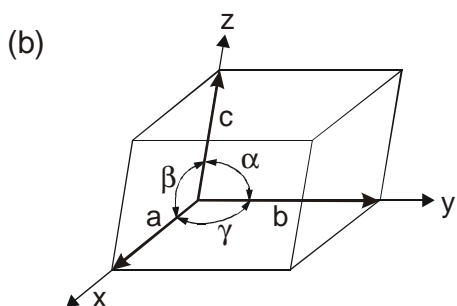
Komórka elementarna jest scharakteryzowana przez tzw. parametry sieci, tzn. długości jej krawędzi a, b, c oraz kąty α, β, γ zawarte między tymi krawędziami. Kierunki wyznaczone przez wektory $\vec{a}, \vec{b}, \vec{c}$ nazywamy osiami krystalograficznymi.



W zależności od stosunku długości i wzajemnej orientacji krawędzi komórki elementarnej, istnieje 14 typów sieci krystalograficznych, tzw. sieci Bravais'go.

W zależności od stopnia symetrii, sieci Bravais'go dzieli się na siedem układów krystalograficznych: regularny, heksagonalny, tetragonalny, trygonalny (romboedryczny), rombowy, jednoskośny i trójskośny (rys. 14.3).

W układzie regularnym, komórką elementarną jest sześcian. Występują trzy rodzaje sieci: prosta, centrowana przestrzennie i centrowana powierzchniowo.



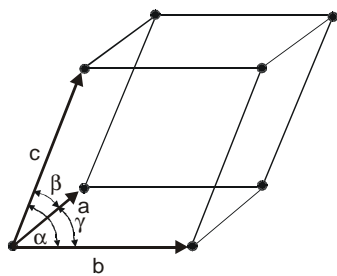
Elementarna komórka układu heksagonalnego ma postać prostego graniastoslupa, którego podstawą jest romb o kątach 60° i 120° . Trzy takie komórki tworzą graniastoslup heksagonalny, który bardziej oddaje symetrię układu niż komórka prosta.

Rys. 14.2. Sieć przestrzenna (a) i komórka elementarna (b).

W układzie tetragonalnym komórką elementarną jest prostopadłościan o podstawie kwadratowej. Występują dwa rodzaje sieci: prosta i centrowana przestrzennie. W układzie trygonalnym komórka elementarna ma postać romboedru (w związku z tym układ ten nazywa się romboedrycznym). Kąty α, β, γ są jednakowe, różne od 90° i mniejsze od 120° .

W układzie jednoskośnym komórka elementarna ma postać pochyłego równoległoscianu, którego dwie pary ścian są prostokątami, a dwie ściany są równoległobokami. W układzie tym występują dwie sieci: prosta i z centrowaną podstawą.

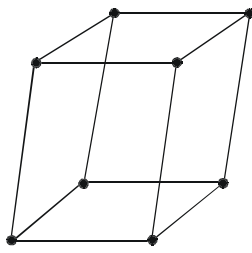
Komórka elementarna układu trójskośnego jest równoległoscianem, którego krawędzie a, b i c i kąty α, β, γ są różne.



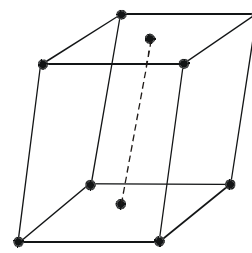
Trójskośny

$$\alpha \neq \beta \neq \gamma$$

$$a \neq b \neq c$$



Prymitywna



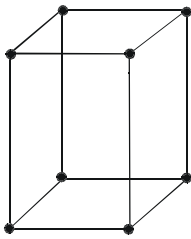
O centrowanej podstawie

Jednoskośny

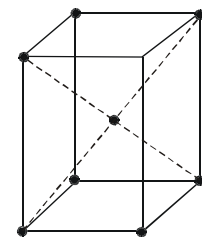
$$\alpha = \beta = 90^\circ$$

$$\gamma = 90^\circ$$

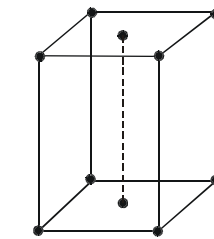
$$a \neq b \neq c$$



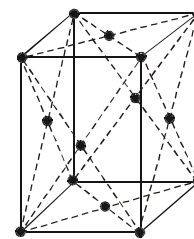
Prymitywna



Centrowana przestrzennie



O centrowanej podstawie

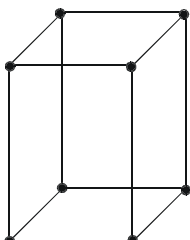


Centrowana płasko

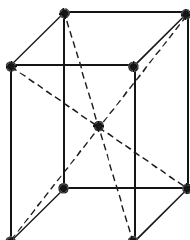
Rombowy

$$\alpha = \beta = 90^\circ$$

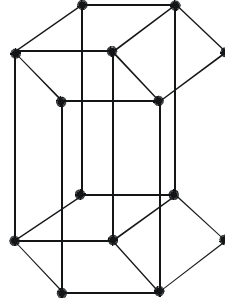
$$a \neq b \neq c$$



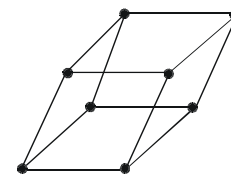
Prymitywna



Centrowana przestrzennie



Heksagonalny



Trygonalny (romboedryczny)

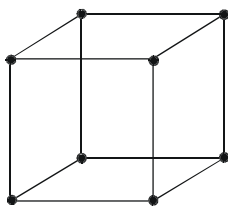
$$\alpha = \beta = \gamma \neq 90^\circ$$

$$a = b = c$$

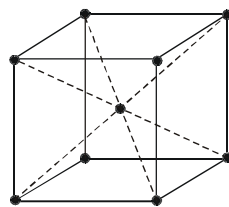
Tetragonalny

$$\alpha = \beta = \gamma = 90^\circ$$

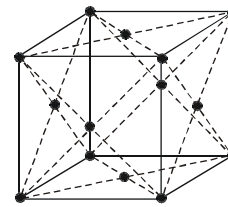
$$a = b \neq c$$



Prymitywna



Centrowana przestrzennie



Centrowana płasko

Regularny

$$\alpha = \beta = \gamma = 90^\circ$$

$$a = b = c$$

Rys. 14.3. Układy krystalograficzne.

Oprócz parametrów sieci a, b i c ; α, β i γ strukturę krystaliczną charakteryzują:

- liczba najbliższych sąsiadów, zwana liczbą koordynacyjną,
- odległość między najbliższymi atomami w kryształach,
- liczba atomów w komórce elementarnej,
- współczynnik upakowania, czyli stosunek objętości kryształu zajętej przez atomy traktowane jako kulki do całkowitej objętości kryształu.

14.2. Oznaczenie węzłów, kierunków i płaszczyzn w kryształach

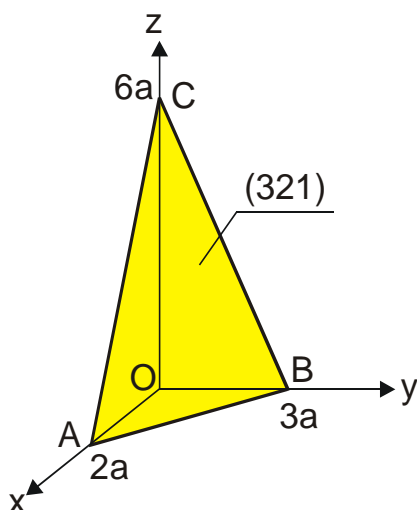
Wskaźniki węzłów

Położenie dowolnego węzła sieci, względem obranego początku układu współrzędnych, jest określone poprzez podanie trzech jego współrzędnych x, y, z , które można wyrazić następująco

$$x = ma, \quad y = nb, \quad z = pc,$$

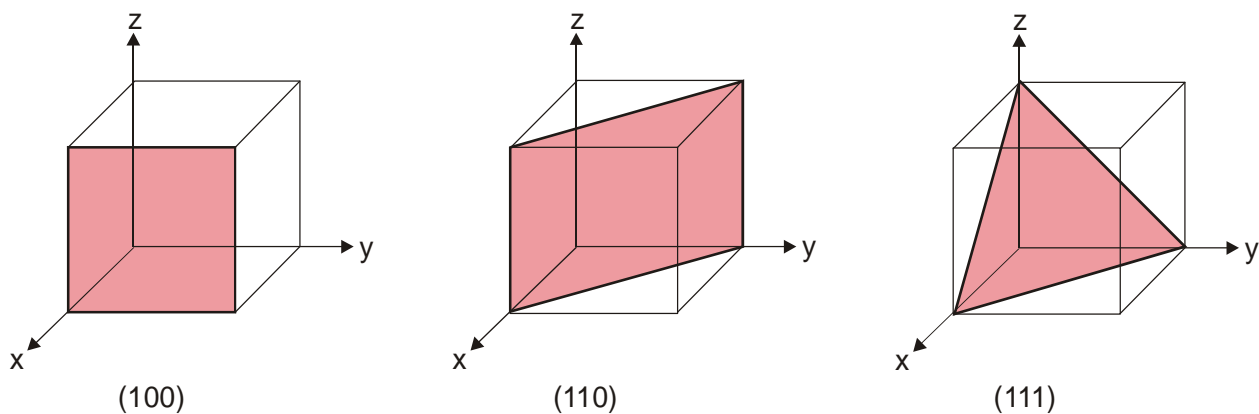
gdzie a, b, c – długości krawędzi komórki elementarnej; m, n, p – liczby całkowite. Jeśli za jednostkę długości przyjąć stałe a, b, c to współrzędnymi węzła będą liczby m, n, p . Liczby te nazywają się wskaźnikami węzła i są zapisane następująco: mnp . Dla wskaźników ujemnych znak minus umieszcza się nad wskaźnikiem. Jeżeli węzeł znajduje się wewnątrz komórki, to liczby m, n i p będą liczbami ułamkowymi.

Wskaźniki kierunków



Rys. 14.4. Płaszczyzna (321).

Wskaźniki kierunków w kryształach stanowią zbiór najmniejszych liczb u, v, w , które mają się do siebie tak, jak rzuty wektora równoległego do danego kierunku na osie krystalograficzne. Wskaźniki zapisuje się w nawiasach kwadratowych $[u, v, w]$. Symbol prostej przechodzącej przez dwa punkty leżące na tej prostej otrzymujemy w ten sposób, że odejmujemy od siebie współrzędne tych punktów i otrzymane liczby sprowadzamy do liczb całkowitych. Kierunki równoległe mają wskaźniki identyczne.



Rys. 14.5. Przykłady symboli płaszczyzn sieciowych w sieci regularnej.

Wskaźniki płaszczyzn

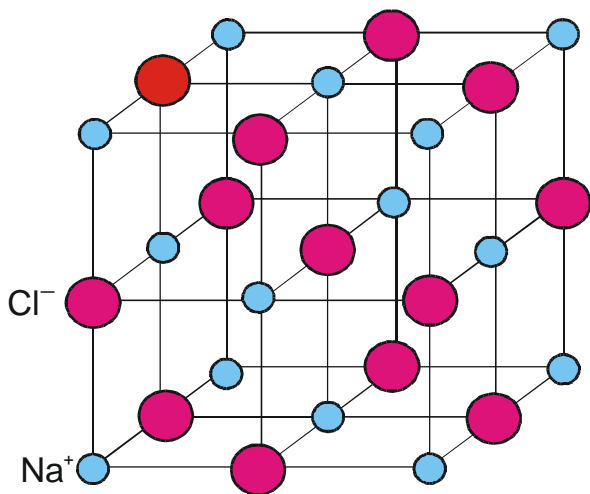
Położenie płaszczyzn jest określone przez podanie trzech odcinków A , B , C , które płaszczyzna odcina na osiach sieci. Wskaźniki płaszczyzn znajdujemy w następujący sposób. Odcinki A , B , C wyrażamy w jednostkach osiowych i piszemy odwrotności tych odcinków $1/A$, $1/B$, $1/C$. Otrzymane ułamki sprowadzamy do najmniejszego wspólnego mianownika D . Liczby całkowite

$$h = \frac{D}{A}, \quad k = \frac{D}{B}, \quad l = \frac{D}{C},$$

są wskaźnikami płaszczyzny i zapisujemy je w postaci (hkl) .

14.3. Proste struktury krystaliczne

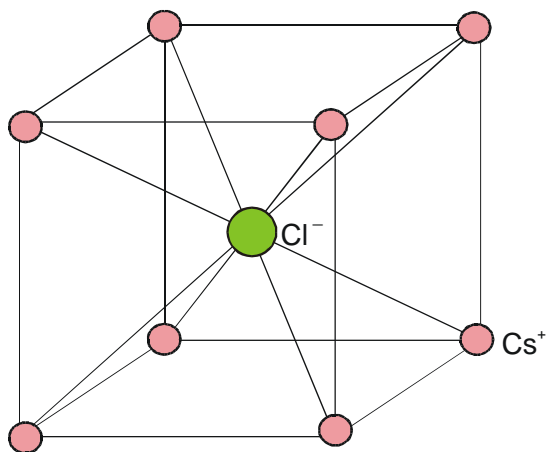
Omówimy teraz proste struktury krystaliczne.



Rys. 14.6. Struktura NaCl.

Struktura chlorku sodu

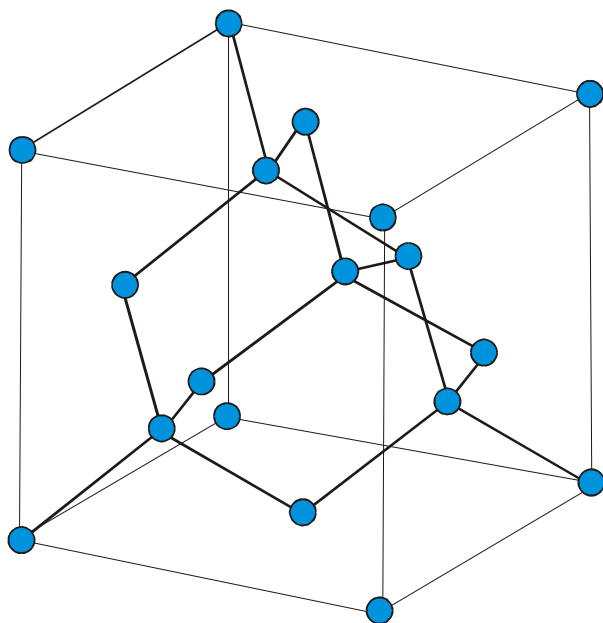
Na rys. 14.6 przedstawiono strukturę chlorku sodu. W elementarnej komórce znajdują się cztery jony Na i cztery jony Cl, czyli cztery cząsteczki NaCl. Każdy jon jednego rodzaju jest otoczony sześcioma jonami drugiego rodzaju, a w związku z czym liczba koordynacyjna równa się 6.



Rys. 14.7. Struktura CsCl.

Struktura chlorku cezu

Strukturę chlorku cezu przedstawiono na rys. 14.7. Jest to struktura regularna przestrzennie centrowana. W komórce elementarnej znajduje się jeden jon Cs i jeden jon Cl. Każdy jon jest otoczony ośmioma jonami znaku przeciwnego, a więc liczba koordynacyjna równa się osiem.



Rys. 14.8. Struktura diamentu.

Struktura diamentu

Struktura diamentu jest strukturą regularną, będącą kombinacją dwóch sieci regularnych powierzchniowo centrowanych, przesuniętych względem siebie o jedną czwartą głównej przekątnej. Rys. 14.8 przedstawia strukturę diamentu. Każdy atom ma czterech najbliższych sąsiadów. W komórce elementarnej znajduje się osiem atomów, a odległość między nimi wynosi $a\sqrt{3}/4$, gdzie a jest długością krawędzi komórki elementarnej. We wszystkich węzłach sieci znajdują się atomy tego samego rodzaju. W strukturze tej krystalizuje między innymi diament, krzem i german.

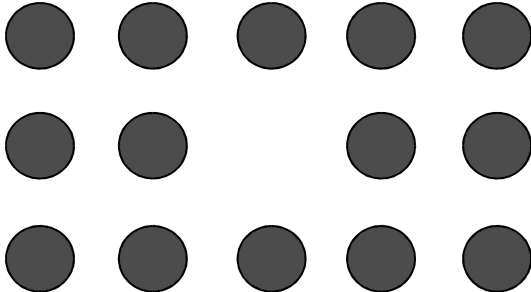
Jeżeli dwie przecinające się sieci regularne centrowane powierzchniowo obsadzone są przez atomy różnych pierwiastków, powstaje nowa struktura zwana strukturą siarczku cynku. W strukturze tej krystalizują między innymi GaAs, CuCl oraz ZnS od którego pochodzi nazwa struktury.

14.4. Defekty sieci krystalicznej

W rzeczywistych kryształach występują odstępstwa od idealnej struktury krystalicznej. Oznacza to, że prawidłowość rozmieszczenia atomów jest w pewien sposób zaburzona. Odchylenia od idealnej struktury nazywamy niedoskonałościami lub defektami.

Spośród różnych defektów sieci, najważniejsze są tzw. defekty punktowe. Defektami punktowymi mogą być: puste węzły sieci (luki), atomy międzywęzłowe oraz domieszki chemiczne.

Najprostszym defektem jest luka w sieci, zwana też defektem Schottky'ego, która jest wywołana brakiem atomu lub jonu w węźle sieci (rys. 14.9).



Rys. 14.9. Luka w sieci (defekt Schottky'ego).

Defekt Schottky'ego można sobie wyobrazić jako przemieszczenie atomu z węzła sieci leżącego w głębi kryształu do węzła leżącego na jego powierzchni. Istnienie defektów Schottky'ego powoduje zmniejszenie gęstości kryształu (wzrasta objętość przy niezmięnionej masie).

Luki występują we wszystkich kryształach. Energia E_v potrzebna do wytworzenia luki wynosi zwykle około 1 eV. Ponieważ energia ta jest znacznie większa od średniej energii drgań cieplnych atomów, wnioskujemy, że tylko niewielka ilość atomów na skutek drgań będzie mogła wyrwać się z sieci i zająć miejsce na powierzchni kryształu. Można wykazać, że w temperaturze T koncentracja luk wyraża się wzorem

$$n_v = n \exp(-E_v / kT),$$

gdzie n jest koncentracją atomów, E_v – energią aktywacji potrzebną do wytworzenia luki. Ze wzrostem temperatury liczba luk bardzo silnie wzrasta i w pobliżu temperatury topnienia wynosi około 0.1%. W temperaturze pokojowej ilość luk wynosi zaledwie $10^{-10}\%$.

Innym rodzajem defektów punktowych są *atomy międzywęzłowe*, czyli *dotatkowe atomy nie zajmujące węzłów sieci*. Atom może też zająć położenie międzywęzłowe przeskakując z węzła sieci i pozostawiając za sobą lukę. Oba rodzaje tych defektów przedstawiono na rys. 14.10. Obliczenia wskazują, że równowagowa koncentracja atomów międzywęzłowych n_i może być określona za pomocą następującej zależności

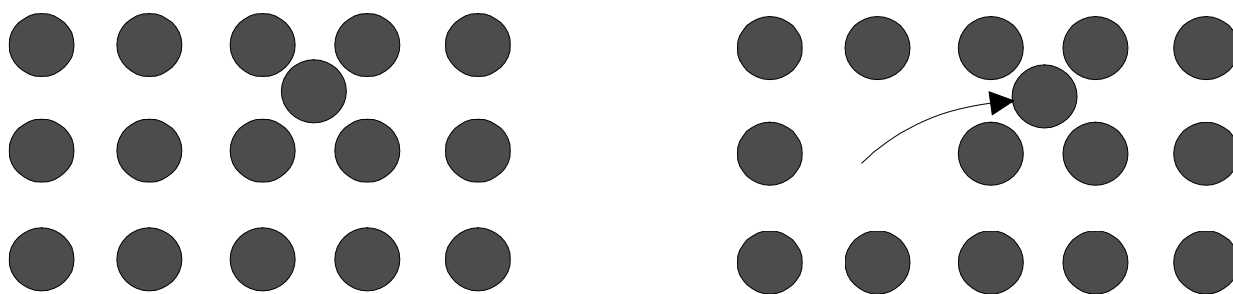
$$n_i = kn \exp(-E_i / kT),$$

gdzie E_i jest energią aktywacji potrzebną do wytworzenia atomu międzywęzłowego, n – liczbą węzłów w jednostce objętości, k – liczbą całkowitą (na ogół bliską 1) określającą liczbę równoważnych położeń międzywęzłowych wokół atomu w węźle.

Energia E_i jest większa od energii E_v , dlatego zwykle koncentracja atomów międzywęzłowych jest znacznie mniejsza niż koncentracja luk.

Luki i atomy międzywęzłowe nie pozostają trwale w miejscach gdzie zostały wytworzone, lecz przemieszczają się w kryształach na skutek dyfuzji. Dyfuzja jest procesem samorzutnym prowadzącym do wyrównania rozkładu koncentracji defektów punktowych wskutek ich

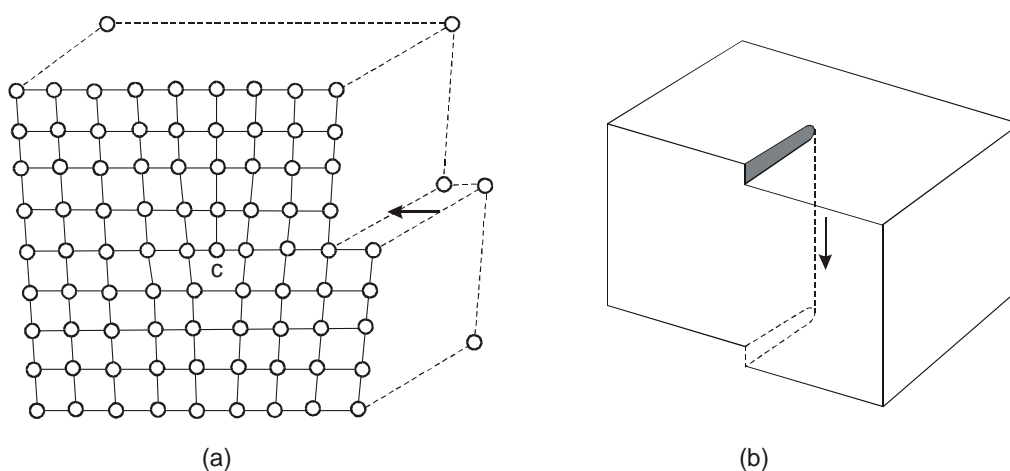
nieuporządkowanych ruchów cieplnych. Ruchliwość atomów międzywęzłowych jest na ogół większa niż ruchliwość luk.



Rys. 14.10. Atomy międzywęzłowe.

Oprócz drgań mogą istnieć jeszcze inne powody występowania defektów w kryształach. Wówczas ich ilość przewyższa wartości równowagowe dla danej temperatury określone powyższymi wzorami. Inne sposoby wytwarzania defektów to: gwałtowne ochłodzenie kryształu, obróbka plastyczna, bombardowanie kryształu wysokoenergetycznymi naładowanymi cząstkami.

Oprócz defektów punktowych dużą rolę w fizyce ciała stałego (szczególnie półprzewodników) odgrywają defekty liniowe, zwane dyslokacjami. Wśród nich wyróżniamy dyslokacje krawędziowe i śrubowe.



Rys. 14.11. Dyslokacja krawędziowa (a) i śrubowa (b)

Dyslokacja krawędziowa jest najprostszym rodzajem dyslokacji polegającym na tym, że w pewnej płaszczyźnie sieciowej ciągłość sieci urywa się wzdłuż linii (krawędzi) oznaczonej na rys. 14.11a symbolem c . To zaburzenie struktury można sobie wyobrazić jako wynik "wsunięcia" półpłaszczyzny sieciowej do jednej części kryształu. Strzałka wskazuje kierunek, w którym górna część sieci krystalicznej została przesunięta w stosunku do dolnej.

Dyslokacja śrubowa oznacza granicę między przesuniętą a nieprzesuniętą częścią kryształu. Granica ta biegnie równoległe do kierunku poślizgu; a nie prostopadle, jak to zachodzi w

przypadku dyslokacji krawędziowej. Można sobie wyobrazić, że dyslokacja śrubowa powstaje przy częściowym nacięciu kryształu nożem i przesunięciu części naciętej o jeden odstęp atomowy równoległe do krawędzi cięcia. Dyslokacja śrubowa przekształca kolejne płaszczyzny atomowe w powierzchnie śrubowe, z czego wywodzi się nazwa tej dyslokacji.

Bardzo poważnym defektem strukturalnym jest powierzchnia kryształu na której raptownie urywa się periodyczna struktura sieci krystalicznej. Ponadto powierzchnia kryształu jest narażona na oddziaływanie zewnętrznego otoczenia. Na powierzchni gromadzi się (zostają zaabsorbowane) znaczne ilości obcych atomów.

Defekty sieci krystalicznej, a szczególnie obce atomy, w zasadniczy sposób wpływają na właściwości fizyczne, zwłaszcza elektryczne i optyczne ciał stałych. Występują one w każdym, nawet najbardziej czystym kryształ. Współczesne sposoby oczyszczania nie pozwalają na uzyskanie substancji zawierających mniej niż $10^{-9}\%$ domieszek, tzn. około 10^{17} atomów domieszkowych w 1 m^3 .

Najbardziej czułe na obce domieszki są półprzewodniki, w których wprowadzenie nawet bardzo małych ilości pewnych pierwiastków (np. 1 atom na miliard atomów własnych) może spowodować wzrost przewodności elektrycznej o kilka rzędów wielkości. W celu więc osiągnięcia określonych właściwości fizycznych półprzewodników stosuje się domieszkowanie kryształów.

Inne rodzaje defektów, które stanowią odstępstwo od idealnej struktury, takie jak drgania cieplne atomów czy też defekty elektronowe, będą przedmiotem naszych rozważań w dalszych rozdziałach.

14.5. Drgania sieci krystalicznej

Rozpatrując sieć krystaliczną zakładaliśmy, że atomy są nieruchome, utrzymywane przez oddziaływania międzyatomowe. W rzeczywistości atomy wykonują drgania wokół swoich położenia równowagi. Drgania te występują w każdej temperaturze, nawet w temperaturze zera bezwzględnego jako tzw. drgania zerowe. Amplituda drgań cieplnych nie przekracza 10^{-11} m .

Z drganiami atomów związanych jest wiele ważnych zagadnień, jak np. rozszerzalność cieplna i przewodnictwo cieplne ciał stałych. Drgania cieplne atomów są najważniejszą przyczyną występowania oporu elektrycznego.

Ponieważ amplituda drgań cieplnych jest znacznie mniejsza od odległości międzyatomowych można w pewnym przybliżeniu przyjąć, że oddziaływanie między atomami jest harmoniczne, tzn. że siła oddziaływania jest proporcjonalna do wychylenia.

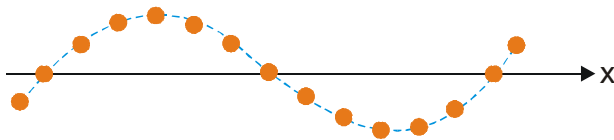
Drgania każdego atomu naruszają równowagę sąsiednich atomów, w wyniku czego rozciągają się one na całą sieć w postaci fal sprężystych rozchodzących się w kryształach we wszystkich możliwych kierunkach. Taki ruch atomów w postaci fali sprężystej nosi nazwę drgań normalnych. Fale te nazywamy też falami sieciowymi. Widmo częstotliwości drgań sieci

obejmuje zakres od zera do ponad 10^{13} Hz. Fale sieciowe o najmniejszych częstotliwościach są to dobrze nam znane fale akustyczne.

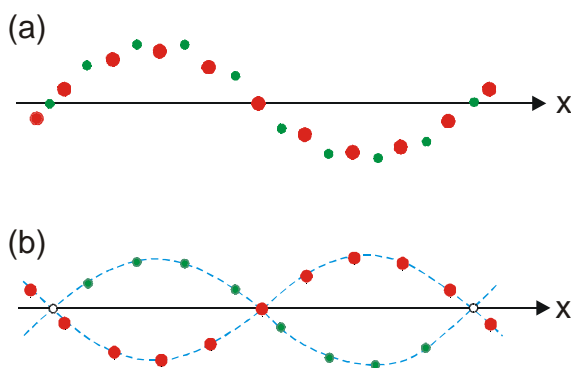
Zamiast rozpatrywać drgania 3-wymiarowej sieci krystalicznej, posłużymy się prostym jednowymiarowym modelem w postaci periodycznego łańcucha złożonego z oddziaływujących na siebie atomów.

Jeden z możliwych sposobów drgań takiego łańcucha przedstawiono na rys. 14.12. W łańcuchu zbudowanym nie z jednego, lecz z dwóch rodzajów atomów, mogą występować następujące dwa typy drgań:

- drgania akustyczne (rys. 14.13a), w których oba rodzaje atomów drgają w tych samych fazach, oraz
- drgania optyczne (rys. 14.13b), w których atomy różnych rodzajów drgają w przeciwnych fazach. Drgania tego typu mogą być wywołane zmiennym polem elektrycznym, np. falą elektromagnetyczną o częstotliwości zwykle zakresu podczerwieni – tym tłumaczy się nazwa "drgania optyczne".



Rys. 14.12 Drgania jednowymiarowego modelu sieci krystalicznej zbudowanej z jednakowych atomów.

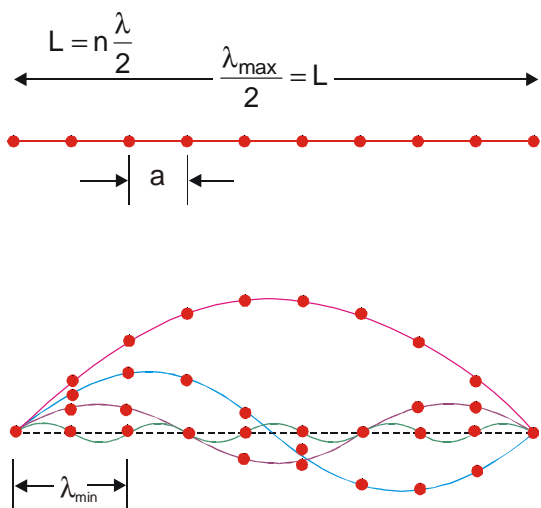


Rys. 14.13. Drgania jednowymiarowego modelu sieci krystalicznej zbudowanej z dwu rodzajów atomów: (a) drgania akustyczne, (b) drgania optyczne.

Drgania akustyczne obejmują zakres mniejszych częstotliwości (od zera do pewnej częstotliwości granicznej) i mniejszych wartości energii, natomiast drgania optyczne obejmują zakres tylko największych częstotliwości i są związane ze stosunkowo wysokoenergetycznymi zjawiskami elektromagnetycznymi.

Rozważmy bliżej drgania akustyczne. Ponieważ drgania każdego atomu można rozłożyć na dwa drgania poprzeczne i jedno podłużne, liczba drgań które mają powstać w kryształach wynosi $3N$, gdzie N jest całkowitą liczbą atomów. Mówiąc inaczej, w ciele stałym propagują się trzy fale akustyczne: jedna podłużna i dwie poprzeczne. Fale te dochodzą do powierzchni kryształu, odbijają się od niej i interferują z falami nadbiegającymi tworząc układ fal stojących. Oznacza to, że na całej długości kryształu L musi odłożyć się całkowita liczba połówek długości fal

$$L = n \frac{\lambda}{2}, \quad (14.1)$$



Rys. 14.14. Przykłady możliwych drgań jednowymiarowej sieci krystalicznej.

gdzie v jest prędkością fazową fali sprężystej. Przyjmując $a = 3 \times 10^{-10}$ m, $v = 6 \times 10^3$ m/s, otrzymujemy $v_{max} = 10^{13}$ Hz.

Największa długość fali wynosi $\lambda_{max} = 2L$, a odpowiadająca jej częstotliwość ν_{min} wynosi

$$\nu_{min} = \frac{v}{\lambda_{max}} = \frac{v}{2L}.$$

Przyjmując $L = 0,1$ m, $v = 5 \times 10^3$ m/s otrzymujemy $\nu_{min} = 2,5 \times 10^5$ Hz. Ponieważ częstotliwość ta jest bardzo mała, można przyjąć, że $\nu_{min} = 0$.

Ze wzoru (14.1) wynika, że liczba drgań normalnych n o długości fal równej lub większej od λ_n wynosi

$$n = \frac{2L}{\lambda_n}.$$

Analogicznie liczba fal stojących w kryształach trójwymiarowym o objętości $V = L^3$ (np. sześcian o krawędzi L), o długości równej lub większej od λ powinna wynosić

$$n = \left(\frac{2L}{\lambda} \right)^3 = 8 \frac{V}{\lambda^3}.$$

Z dokładnych obliczeń wynika, że

$$n = 4\pi \frac{V}{\lambda^3} = 4\pi \frac{v^3}{v^3} V.$$

Różniczkując to wyrażenie względem ν otrzymujemy

$$dn = 3 \frac{4\pi v^2 V}{v^3} dv = g(\nu) d\nu. \quad (14.3)$$

Kilka drgań mogących powstać w przypadku jednowymiarowego kryształu przedstawiono na rys. 14.14.

Najkrótsza długość fali, która może się propagować w kryształach, jest równa podwójnej odległości między atomami

$$\lambda_{min} = 2a,$$

gdzie a jest odległością między atomami. Fali tej odpowiada maksymalna częstotliwość

$$\nu_{max} = \frac{v}{\lambda_{min}} = \frac{v}{2a}, \quad (14.2)$$

Wzór (14.3) określa liczbę drgań normalnych zawartych w przedziale częstotliwości od ν do $\nu+d\nu$. Funkcja $g(\nu)$ nosi nazwę funkcji rozkładu drgań normalnych. Ponieważ w danej objętości V zawierającej N atomów jest $3N$ drgań normalnych (modów), funkcja $g(\nu)$ musi spełniać następujący warunek unormowania:

$$\int_0^{\nu_0} g(\nu) d\nu = 3N, \quad (14.4)$$

gdzie ν_0 jest częstotliwością maksymalną, zwaną też częstotliwością Debye'a. Podstawiając do wzoru (14.4) wyrażenie $g(\nu)d\nu$ i całkując otrzymamy

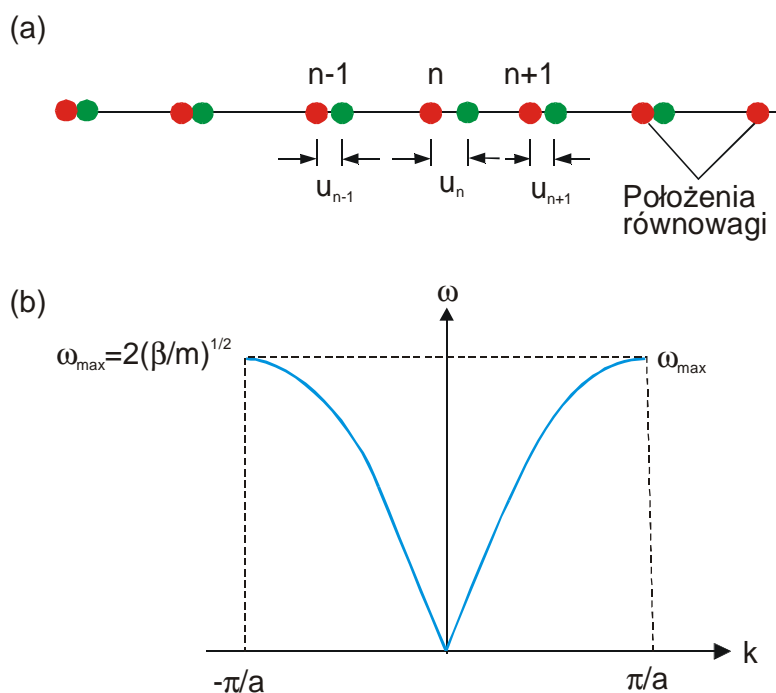
$$\frac{4\pi\nu_0^3 V}{\nu^3} = 3N,$$

skąd

$$\nu_0 = \left(\frac{3N\nu^3}{4\pi V} \right)^{1/3} = \nu \left(\frac{3n}{4\pi} \right)^{1/3}, \quad (14.5)$$

gdzie $n = N/V$ jest koncentracją atomów. Częstotliwość ν_0 ma identyczną interpretację co częstotliwość wyrażona wzorem (14.2). Dla $n = 5 \times 10^{28} \text{ m}^{-3}$ i $\nu = 5 \times 10^3 \text{ m/s}$, $\nu_0 = 1,1 \times 10^{13} \text{ Hz}$.

Przeanalizujemy teraz drgania nieskończenie długiej, jednowymiarowej sieci krystalicznej składającej się z jednakowych atomów mających masy m i znajdujących się w odległości a jeden od drugiego (rys. 14.15).



Rys. 14.15. Drgania nieskończenie długiej sieci jednowymiarowej (a)

i zależność ω od k dla tych drgań (b).

Na każdy atom działają dwie siły pochodzące od obu jego sąsiadów. Jeśli przez u_n oznaczyć wychylenie n -tego atomu z położenia równowagi, to w przybliżeniu harmonicznym równanie ruchu tego atomu ma postać:

$$m \frac{d^2 u_n}{dt^2} = -\beta(u_n - u_{n-1}) - \beta(u_n - u_{n+1}) = \beta(u_{n+1} + u_{n-1} - 2u_n). \quad (14.6)$$

Rozwiązanie równania (14.6) przewidujemy w postaci fali płaskiej propagującej się wzdłuż sieci. Przesunięcie n -tego atomu może więc być zapisane w postaci

$$u_n = A e^{i(\omega t - k n a)}. \quad (14.7)$$

Podstawiając wyrażenie (14.7) do (14.6) otrzymujemy

$$\omega^2 = \frac{\beta}{m} [2 - (e^{i k a} + e^{-i k a})] = 2 \frac{\beta}{m} [1 - \cos k a] = 4 \frac{\beta}{m} \sin^2 \left(\frac{k a}{2} \right).$$

Częstość ω wynosi zatem

$$\omega = 2 \sqrt{\frac{\beta}{m}} \left| \sin \frac{k a}{2} \right|. \quad (14.8)$$

Zależność dyspersyjną ω od k przedstawiono na rys. 14.15b. Ponieważ najkrótsza długość fali wynosi $2a$, to wartość wektora falowego $k = 2\pi/\lambda$ może zmienić się w przedziale

$$-\frac{\pi}{a} \leq k \leq \frac{\pi}{a}. \quad (14.9)$$

Rozpatrujemy tutaj zarówno dodatnie jak i ujemne wartości k , ponieważ fala może propagować się w jedną jak i w drugą stronę.

Z wyrażenia (14.8) wynika, że fale akustyczne wykazują dyspersję, tzn. zależność prędkości fal od częstotliwości

$$v = \frac{\omega}{k} = 2 \sqrt{\frac{\beta}{m}} \frac{\left| \sin \frac{k a}{2} \right|}{k}. \quad (14.10)$$

Jedynie w przypadku gdy $ka \ll 1$, możemy funkcję *sinus* przybliżyć przez jej argument i wówczas otrzymujemy

$$v = \sqrt{\frac{\beta}{m}} a. \quad (14.11)$$

Biorąc pod uwagę, że $a \approx 3 \times 10^{-10}$ m, $v = 5 \times 10^3$ m/s, warunek $ka \ll 1$ jest spełniony do częstotliwości rzędu 10^{11} Hz. Oznacza to, że do tej częstotliwości fale nie wykazują dyspersji. Przy częstotliwościach wysokich prędkość fali jest zależna od częstotliwości.

Omawiając kwantowy oscylator harmoniczny stwierdziliśmy, że jego energia jest skwantowana i możliwe wartości energii określa wzór

$$E_n = \left(n + \frac{1}{2} \right) h\nu \quad n = 0, 1, 2, \dots \quad (14.12)$$

Wiemy również, że wyrażenia klasyczne muszą być zastąpione przez kwantowe, jeśli wartości wielkości fizycznych mających wymiar stałej Plancka (energia×czas) są rzędu tej stałej. Oszacujemy zatem iloczyn ET , gdzie E – energia oscylatora, T – okres drgań

$$ET = m \frac{\omega^2 A^2}{2} T = \pi m \omega A^2.$$

Podstawiając wartości liczbowe: $m = 10^{-25}$ kg, $\omega = 10^{13}$ Hz, $A = 10^{-11}$ m otrzymujemy $ET = 3 \times 10^{-34}$ Js. Wartość ta jest rzędu stałej Plancka, co oznacza, że energia drgań atomów jest skwantowana i wyraża się wzorem (14.12). Minimalna porcja (kwant energii) takich drgań (podobnie jak w przypadku fotonów) wynosi

$$E = h\nu = \hbar\omega. \quad (14.13)$$

Kwant ten nosi nazwę fononu. Każdy atom wymieniający energię z atomami sąsiednimi czyni to kwantami $h\nu$. Fale sprężyste możemy więc traktować jako zbiór fononów, które przemieszczają się w kryształach z prędkością tej fali. Kryształ z drganiami sieci możemy traktować jako obszar wypełniony fononami. Fonon posiada nie tylko energię ale i pęd

$$\vec{p} = \hbar\vec{k}. \quad (14.14)$$

Jednak fonony nie są cząstkami w zwykłym tego słowa znaczeniu. Są one korpuskularnym opisem fal sprężystych w ciele stałym i są ściśle związane z siecią krystaliczną. Istnieją tylko wtedy, gdy istnieje sieć. Fonony nie mogą opuścić kryształu, dlatego nazywamy je quasi-cząstkami. Należy podkreślić, że korpuskularny charakter fal sprężystych przejawia się nie tylko w przypadku fal cieplnych, ale także w przypadku fal sprężystych wytwarzanych przez źródła zewnętrzne, o ile tylko częstotliwość jest odpowiednio duża. Podobnie jak w przypadku fotonów, średnia energia fononu wyraża się wzorem

$$\langle E \rangle = \frac{h\nu}{\exp(h\nu/kT) - 1}. \quad (14.15)$$

Dzieląc średnią energię $\langle E \rangle$ przez energię fononu $h\nu$ otrzymujemy średnią liczbę fononów o danej energii E

$$f(E) = \frac{1}{\exp(h\nu/kT) - 1}. \quad (14.16)$$

Funkcja $f(E)$ jest funkcją rozkładu fononów względem energii.

14.6. Energia wewnętrzna i ciepło właściwe kryształu

Znajomość średniej energii oscylatora oraz funkcji rozkładu fononów pozwala wyznaczyć energię sieci. Otrzymujemy ją mnożąc średnią energię oscylatora przez liczbę drgań zawartych w przedziale częstotliwości od ν do $\nu+d\nu$ i całkując dane wyrażenie po wszystkich możliwych częstotliwościach

$$U = \int_{\nu_{min}}^{\nu_{max}} \langle E \rangle g(\nu) d\nu. \quad (14.17)$$

Ponieważ ν_{min} jest zwykle bardzo małe, dolną granicę całkowania możemy przyjąć równą zeru, zaś górną granicą jest częstotliwość Debye'a. Podstawiając do wzoru (14.17) wyrażenia na $\langle E \rangle$ i $g(\nu)$ odpowiednio według wzorów (14.15) i (14.3), otrzymujemy

$$U = \int_0^{\nu_0} \frac{h\nu}{\exp(h\nu/kT) - 1} \frac{12\pi\nu^2 d\nu}{\nu^3} V.$$

Ponieważ ze wzoru (14.5) $\nu_0^3 = 3N\nu^3/4\pi V$, zatem

$$U = \frac{9N}{\nu_0^3} \int_0^{\nu_0} \frac{h\nu^3 d\nu}{\exp(h\nu/kT) - 1}. \quad (14.18)$$

Energię drgań cieplnych odnosimy zazwyczaj do jednego mola i wówczas N jest równe stałej Avogadra. Podstawiając $x = h\nu/kT$ oraz wprowadzając charakterystyczną temperaturę Debye'a θ zdefiniowaną jako

$$\theta = \frac{h\nu_0}{kT}, \quad (14.19)$$

otrzymujemy energię drgań cieplnych jednego mola

$$U = 9N_A k\theta \left(\frac{T}{\theta}\right)^4 \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1}. \quad (14.20)$$

Zanim przejdziemy do dyskusji wyrażenia (14.20), omówimy sens fizyczny wprowadzonej temperatury Debye'a. Z klasycznej teorii ciepła wiadomo, że średnia energia ruchu kinetycznego przypadająca na jeden stopień swobody wynosi $kT/2$. Ponieważ dla stanu w sieci musimy jeszcze uwzględnić energię potencjalną wynoszącą również $kT/2$, więc pełna energia na jeden stopień swobody wynosi kT . Jeżeli wartość ta znacznie przekracza $h\nu_0$, to mogą być wzbudzone drgania o wszystkich możliwych częstotliwościach i znika kwantowy charakter fal cieplnych, a wyrażenie na średnią energię oscylatora harmonicznego przechodzi w wyrażenie klasyczne

$$\langle E \rangle = \frac{h\nu}{\exp(h\nu/kT) - 1} = \frac{h\nu}{1 + h\nu/kT + \dots - 1} \cong kT.$$

Natomiast, gdy $h\nu_0 \gg kT$, to drgania o wysokiej częstotliwości nie mogą być wzbudzone. Temperatura Debye'a jest więc granicą temperaturą poniżej której zaczyna się przejawiać kwantowy charakter drgań cieplnych. W tabeli 14.1 podano wartości temperatury Debye'a dla kilku kryształów.

Tabela 14.1. Temperatura Debye'a

Kryształ	Cs	Fe	Si	B	Diament	NaCl
θ [K]	45	460	630	1220	2050	280

Powróćmy teraz do wzoru (14.20). W temperaturach wysokich $kT \gg h\nu_0$ ($\theta/T \gg 1$) słuszne będzie przybliżenie $e^x = 1+x$. Wówczas

$$U = 9N_A k \theta \left(\frac{T}{\theta} \right)^4 \int_0^{\theta/T} \frac{x^3 dx}{1+x-1} = 3RT.$$

A ciepło molowe

$$c_v = \frac{dU}{dT} = 3R, \quad (14.21)$$

gdzie R jest stałą gazową. Uzyskaliśmy znany, klasyczny wynik Dulonga i Petita dla ciepła molowego kryształu. Należało tego oczekiwać, bowiem w wysokich temperaturach zanika kwantowy charakter drgań.

W obszarze niskich temperatur ($T \ll \theta$) górna granica całkowania we wzorze (14.19) może być przyjęta za nieskończoność. Ponieważ

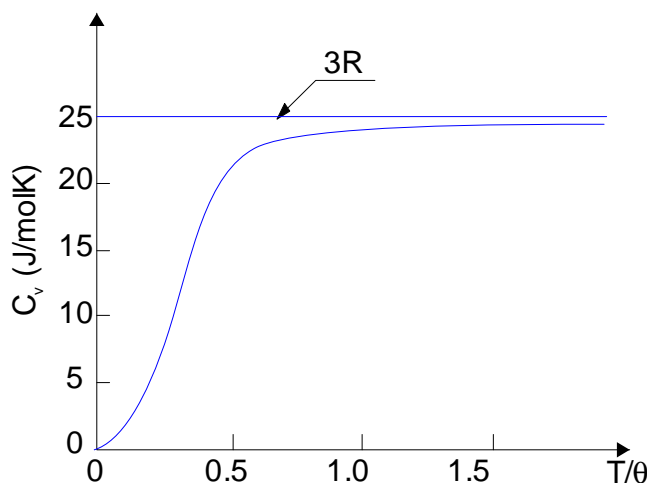
$$\int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15},$$

więc

$$U = 9N_A k \theta \left(\frac{T}{\theta} \right)^4 \int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{3\pi^4}{5} N_A k \theta \left(\frac{T}{\theta} \right)^4,$$

natomiast c_v wynosi

$$c_v = \frac{12\pi^4}{5} N_A k \left(\frac{T}{\theta} \right)^3 \quad (14.22)$$



Rys. 14.16. Zależność ciepła właściwego ciał stałych od stosunku T/θ

Otrzymaliśmy znaną zależność Debye'a wskazującą, że w obszarze niskich temperatur $C \sim T^3$. Zależność ta jest w bardzo dobrej zgodności z doświadczeniem. Na rys. 14.16 przedstawiono zależność c_v od T/θ . Zwróćmy uwagę, że jest to funkcja jednakowa dla wszystkich ciał. Dla $T \rightarrow 0$, także $c_v \rightarrow 0$. W temperaturach bardzo bliskich zera bezwzględnej pozostają tylko drgania zerowe.

14.7. Koncentracja fononów

Koncentrację fononów można stosunkowo prosto obliczyć z wyrażenia (14.18) podstawiając w miejsce N koncentrację atomów n i dzieląc przez energię fononu $h\nu$. Otrzymujemy w ten sposób

$$n_f = \frac{9n}{v_o^3} \int_0^{v_o} \frac{v^2 dv}{\exp(h\nu/kT) - 1} = 9n \left(\frac{T}{\theta} \right)^3 \int_0^{\theta/T} \frac{x^2 dx}{e^x - 1}. \quad (14.23)$$

Podobnie jak w przypadku energii fononów, obliczenia prowadzimy dla dwóch przypadków: temperatur niskich ($T \ll \theta$) i temperatur wysokich ($T \gg \theta$).

Dla $T \ll \theta$ górną granicę całkowania zastępujemy przez nieskończoność i uwzględniając, że

$$\int_0^{\infty} \frac{x^2 dx}{e^x - 1} = \frac{\pi^2}{3},$$

otrzymujemy

$$n_f = 3\pi^2 n \left(\frac{T}{\theta} \right)^3. \quad (14.24)$$

W niskich temperaturach koncentracja fononów jest proporcjonalna do T^3 .

W obszarze wysokich temperatur x jest niewielkie i dlatego $e^x \approx 1 + x$. Podstawiając do wzoru (14.23) i całkując otrzymujemy

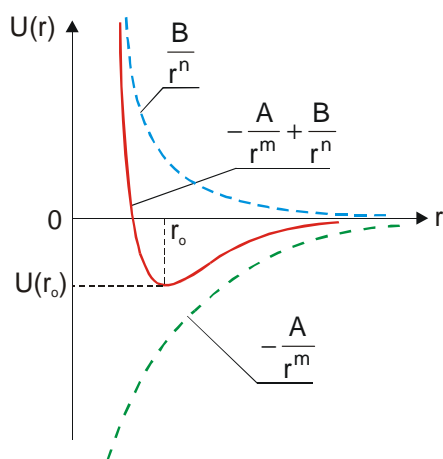
$$n_f = \frac{9}{2} n \frac{T}{\theta}. \quad (14.25)$$

W wysokich temperaturach koncentracja fononów rośnie liniowo z temperaturą.

14.8. Wiązania atomów w kryształach

Zrozumienie wielu procesów zachodzących w ciałach stałych wymaga znajomości charakteru wiązań między atomami występujących w tych ciałach.

Istnienie ciał stałych świadczy o występowaniu sił międzyatomowych, które dla dużych odległości są siłami przyciągania, a dla małych odległości – siłami odpychania. Przy odległościach międzyatomowych rzędu 10^{-10} m siły te równoważą się. Prawie we wszystkich przypadkach znacznie wygodniej operować jest nie pojęciem siły działającej między atomami, ale związanej z nią energią potencjalną. Niezależnie od natury sił międzyatomowych, przebieg energii potencjalnej w funkcji odległości między atomami ma postać przedstawioną na rys. 14.17.



Rys. 14.17. Zależność energii potencjalnej od odległości między atomami.

W stanie równowagi trwałej, który jest osiągnięty w pewnej odległości r_0 , energia osiąga wartość minimalną, a siła jest równa zero. Dla $r > r_0$ pojawia się siła przyciągania, a dla $r < r_0$ – siła odpychania. W tym przypadku, siły te są skierowane w kierunku r_0 . Przy małych wychyleniach atomu z położenia równowagi, energię potencjalną $U(r)$ można rozwinąć w szereg wokół punktu r_0 , zachowując tylko pierwsze wyrazy rozwiązania.

$$U(r) = U(r_0) + \left(\frac{dU}{dr} \right)_{r_0} (r - r_0) + \frac{1}{2} \left(\frac{d^2U}{dr^2} \right)_{r_0} (r - r_0)^2 + \frac{1}{6} \left(\frac{d^3U}{dr^3} \right)_{r_0} (r - r_0)^3 + \dots$$

Energia $U(r)$ może być wyrażona za pomocą następującego wzoru

$$U(r) = -\frac{A}{r^m} + \frac{B}{r^n}, \quad (14.26)$$

gdzie A , B , m i n są stałymi, r jest odległością między atomami. Człon A/r^m określa siły przyciągania, a człon B/r^n – siły odpychania. Ponieważ siły przyciągania są bardziej długi zasięgowe od sił odpychania, n musi być większe od m .

Ponieważ dla $r = r_o$ energia osiąga wartość minimalną, a więc $(dU/dr)_{r_o} = 0$, a człon $(d^2U/dr^2)_{r_o} > 0$. Natomiast $(d^3U/dr^3)_{r_o} < 0$, gdyż w pobliżu punktu r_o siły odpychania rosną szybciej niż maleją siły przyciągania. Oznaczając wychylenie atomu z położenia równowagi przez $r - r_o = x$, a $U(r_o) = -U_o$, otrzymujemy

$$U(x) + U_o = \frac{1}{2} \beta x^2 - \frac{1}{3} \gamma x^3,$$

gdzie $\beta = (d^2U/dr^2)_{r_o}$ oraz $\gamma = 1/2(d^3U/dr^3)_{r_o}$. Siła działająca na atom wynosić więc będzie

$$F = \frac{dU}{dr} = -\beta x + \gamma x^2. \quad (14.27)$$

W pierwszym przybliżeniu (dla bardzo małych wychyleń) siła $F = -\beta x$, tzn. jest to siła harmoniczna. Przybliżenie harmoniczne nie zawsze prowadzi do dobrych wyników (np. wyjaśnienie rozszerzalności cieplnej ciał stałych wymaga uwzględnienia członu γx^2).

Siły odpychające między atomami są głównie spowodowane nakładaniem się powłok elektronowych atomów, co powoduje wzrost energii kinetycznej w atomach. Przy rozważaniu sił odpychania należy również uwzględnić czysto kulombowskie oddziaływanie między jądrami atomów, występujące w przypadku kiedy jądro atomu znajduje się wewnątrz powłoki elektronowej drugiego atomu.

Siły przyciągania mają w zasadzie charakter sił elektromagnetycznych. W zależności od wzajemnego oddziaływania atomów (jonów) wyodrębniamy:

- wiązania jonowe – występujące między atomami różniącymi się znacznie tzw. elektroujemnością (czyli zdolnością atomu do przyciągania elektronów),
- wiązania kowalencyjne – występujące między atomami tego samego pierwiastka,
- wiązania metaliczne – występujące w metalach i ich stopach,
- wiązania molekularne – występujące w kryształach zbudowanych z cząsteczek chemicznych; cząsteczki te są ze sobą związane słabymi siłami zwanymi siłami van der Walsa (ten typ wiązań występuje również w kryształach gazów szlachetnych).

Natura sił działających między atomami i jonami jest często podstawą klasyfikacji kryształów. Mamy więc kryształy jonowe, kowalencyjne, metaliczne i molekularne. Wymienione wyżej rodzaje wiązań rzadko występują w czystej postaci.

Miarą wielkości sił wiążących atomy w kryształach jest energia wiązania, czyli minimalna energia potrzebna do rozdzielenia kryształu na swobodne atomy. Wyraża się ją w elektronowoltach na atom lub w dżulach na mol. Energia ta zawiera się w bardzo szerokich granicach od 0,02 eV/atom dla neonu (temperatura topnienia -249°C) do 8,9 eV/atom dla wolframu (temperatura topnienia -3370°C). Dla kryształów jonowych może być większa (zobacz tabela 14.2)

Tabela 14.2. Energia wiązania dla różnych typów wiązań

Typ wiązania	Przykłady	Energia wiązania	
		kJ/mol	eV/atom
Jonowe	AgCl	987	10,2
	LiF	1014	10,5
	KCl	694	7,2
	NaCl	765	7,9
Kowalencyjne	C	712	7,4
	Ge	374	3,9
	Si	448	4,6
Metaliczne	Cu	338	3,5
	Ag	286	2,9
	Fe	393	4,1
	Pb	197	2,0
Molekularne	Ne	1.9	0,02
	Ar	7.9	0,08
	Kr	3.2	0,12

14.8.1. Wiązania jonowe

Wiązanie jonowe w czystej postaci występuje w kryształach typu NaCl, CsCl, KJ itp, a więc w kryształach utworzonych z silnie elektrododatnich atomów metali i silnie elektroujemnych atomów chlorowców. Wiązania te omówimy na przykładzie kryształu NaCl, którego strukturę przedstawiono na rys. 14.6. W węzłach sieci znajdują się przemiennie jony Na^+ i Cl^- . Powstanie jonów możemy wytłumaczyć w sposób następujący. Energia jonizacji atomu Na wynosi około 5,2 eV, przyłączenie elektronu od atomu chloru jest związane z wydzieleniem energii 3,8 eV. Do przemieszczenia elektronu z atomu sodu do atomu chloru potrzebna jest zatem energia 1,4 eV. Jeżeli więc atomy (jony) znajdują się od siebie w tak dużej odległości, że można zaniedbać ich wzajemne oddziaływanie, to układ $\text{Na}^+ + \text{Cl}^-$ posiada energię o 1,4 eV wyższą od energii układu $\text{Na} + \text{Cl}$.

Przy zmniejszeniu odległości r między jonami coraz większą rolę zaczyna odgrywać oddziaływanie kulombowskie. Wartość energii 1,4 eV = $e^2 / 4\pi\epsilon_0 r$ jest osiągnięta przy odległości $r = 10^{-9}$ m.

Przy mniejszych odległościach układ jonów jest bardziej stabilny niż układ atomów. Przy jeszcze większym zbliżeniu się jonów pojawiają się siły odpychające, o których mówiliśmy poprzednio. W wyniku otrzymujemy przebieg energii przedstawiony na rys. 14.17.

Energia oddziaływania dwóch jonów sieci krystalicznej może być przedstawiona w postaci wzoru

$$U(r) = -\frac{e^2}{4\pi\epsilon_0 r} + \frac{B}{r^n}$$

Parametr n wynosi około 10, tzn. że siły odpychania maleją bardzo szybko ze wzrostem odległości i można ograniczyć się do oddziaływania najbliższych sąsiadów. Natomiast w przypadku bardziej długozasięgowych oddziaływań kulombowskich należy uwzględnić oddziaływanie z dalszymi sąsiadami. W rezultacie energia potencjalna jonu w kryształach wynosi

$$U(r) = -\alpha_M \frac{e^2}{4\pi\epsilon_0 r} + \frac{B}{r^n}, \quad (14.28)$$

gdzie α_M nosi nazwę stałej Madelunga i uwzględnia oddziaływanie kulombowskie jonu z dalszymi sąsiadami.

Aby zorientować się w rzędzie stałej Madelunga, obliczamy ją dla sieci jednowymiarowej złożonej z dodatnich i ujemnych jonów. Energia kulombowska w tym przypadku wynosi

$$U(r) = -2 \frac{e^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots \right) = -\frac{e^2}{4\pi\epsilon_0 r_0} 2 \ln 2.$$

Zatem w tym przypadku stała Madelunga $\alpha_M = 2 \ln 2 = 1,386$. Dla sieci trójwymiarowych stała Madelunga wynosi: dla struktury NaCl – 1,748, dla struktury CsCl – 1,763. Stałą B wyznaczamy z warunku, że w położeniu równowagi energia potencjalna osiąga minimum

$$\left(\frac{dU}{dr} \right)_{r_0} = \frac{\alpha_M e^2}{4\pi\epsilon_0 r_0^2} - \frac{nB}{r_0^{n+1}} = 0,$$

stąd

$$B = \frac{\alpha_M e^2}{4\pi\epsilon_0 n} r_0^{n-1}.$$

Energia potencjalna jednego jonu w położeniu równowagi wynosi więc

$$U(r_0) = -\frac{\alpha_M e^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n} \right).$$

Całkowita energia potencjalna kryształu złożonego z $2N$ jonów jest równa połowie sumy energii poszczególnych jonów (przyjmujemy $1/2$ dlatego, aby energii każdej pary nie liczyć dwukrotnie). W rezultacie

$$U_N = \frac{1}{2} 2NU(r_0) = -\frac{\alpha_M e^2 N}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n} \right).$$

Kładąc w miejsce N liczbę N_A , otrzymujemy molową energię wiązania kryształu jonowego

$$U_m = -\frac{\alpha_M e^2 N_A}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n} \right), \quad (4.29)$$

Podstawiając dane dla kryształu NaCl: $r_o = 2,81 \times 10^{-10}$ m, $n = 9.4$, $\alpha_M = 1,784$ otrzymujemy $U_m = -7,7 \times 10^5$ J/mol.

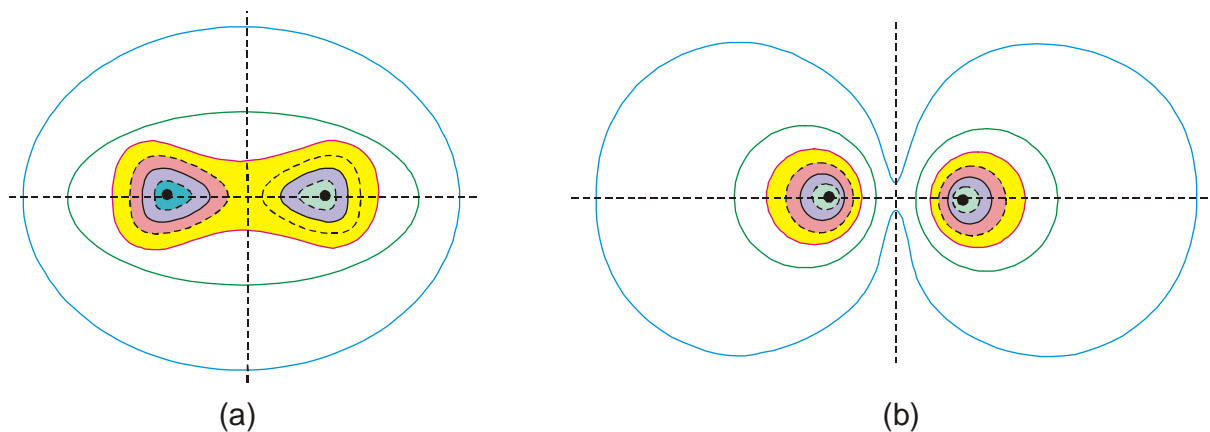
Znaczne wartości energii wiązania powodują, że kryształy jonowe są twarde i mają wysoką temperaturę topnienia. Istotną cechą wiązania jonowego jest to, że nie ma ono wyróżnionego kierunku, gdyż zamknięte powłoki elektronowe mają symetrię sferyczną. Brak elektronów swobodnych powoduje, że kryształy te są złymi przewodnikami ciepła i prądu.

Jak już wspomniano, wiązanie jonowe w czystej postaci występuje w kryształach utworzonych przez atomy I i VII grupy układu okresowego, tzn. w kryształach typu I-VII. Również kryształy typu II-VI wykazują cechy wiązania jonowego. Często jednak trudno jest określić rodzaj wiązania. Na przykład, kryształy typu III-V nie są ani wyraźnie jonowe ani wyraźnie kowalencyjne.

14.8.2. Wiązania kowalencyjne

Wiązania kowalencyjne występują między atomami o identycznych albo o bardzo zbliżonych właściwościach. Naturę tego wiązania omówimy na przykładzie cząstki wodoru.

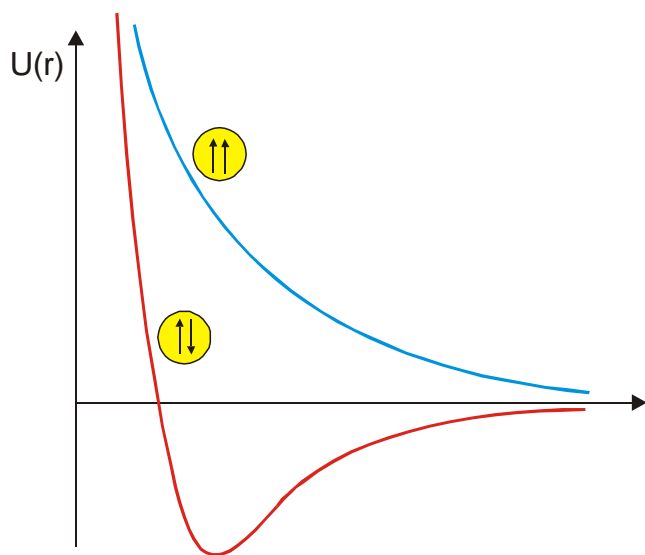
Jeżeli atomy wodoru znajdują się daleko od siebie, to możemy je traktować jako odosobnione, gdyż gęstość chmury elektronowej bardzo szybko maleje ze wzrostem odległości. Gdy elektrony obu atomów znajdują się w stanie $1s$, to funkcja falowa elektronu w tym stanie ma postać typu $\Psi(r) \sim \exp(-r/r_o)$; natomiast prawdopodobieństwo p znalezienia elektronu w odległości r od jądra wynosi $p \sim r^2 \exp(-2r/r_o)$, gdzie $r_o = 0,53 \times 10^{-10}$ m jest promieniem Bohra.



Rys. 14.18. Rozkład gęstości chmury elektronowej w układzie złożonym z dwu atomów wodoru: (a) orbital wiążący, (b) orbital antywiązący. W celu zwiększenia czytelności rysunku zaciemniono obszary, w których gęstość ładunku elektrycznego przekracza pewną wartość jednakową na obu wykresach.

Przy zbliżaniu się dwóch atomów nakrywają się ich chmury elektronowe i rośnie prawdopodobieństwo przejścia elektronu z jednego atomu do drugiego. Przy odpowiednio małej odległości stopień pokrywania się chmur jest tak duży, że można powiedzieć, iż elektrony należą do obu jąder. Mówimy wówczas, że są skolektywizowane lub uwspólnione. Okazuje się

ponadto, że nakrywaniu chmur elektronowych towarzyszy przegrupowanie gęstości elektronowej i zmiana energii układu. Jeżeli spiny elektronów są ustawione przeciwnie, to gęstość chmury elektronowej w przestrzeni między jądrami będzie większa niż ta, którą otrzymalibyśmy w wyniku dodania chmur elektronowych dwóch odosobnionych atomów. Jednocześnie poza tym obszarem gęstość chmury jest mniejsza niż w odosobnionych atomach (rys.14.18). Funkcję falową opisującą gęstość chmury elektronowej utworzonej w ten sposób nazywamy orbitalem wiążącym. Pojawienie się stanu o większej gęstości chmury elektronowej (przy przeciwnie ustawionych spinach) powoduje zmniejszenie energii układu i prowadzi do pojawienia się sił przyciągania. Jeżeli natomiast spiny elektronów są skierowane zgodnie, między atomami pojawiają się siły odpychania. Układ taki może więc znaleźć się w stanie równowagi, a przypisany temu stanowi orbital nazywamy antywiązącym. Oddziaływanie uwarunkowane wymianą elektronów między atomami ma kwantowy charakter i nosi nazwę oddziaływania wymiennego. Siły takiego oddziaływania nazywamy siłami wymiany, a ich energię - energią wymiany.



Rys. 14.19. Zależność energii oddziaływania między atomami od odległości przy równoległym i antyrównoległym ustawieniu spinów.

Na rys. 14.19 przedstawiono energię oddziaływania dwóch atomów w zależności od odległości między nimi przy równoległym i antyrównoległym ustawieniu spinów. W przypadku gdy spiny elektronów są ustawione zgodnie, energia układu wzrasta w sposób ciągły w miarę zbliżania się atomów. Jeżeli natomiast spiny elektronów są ustawione przeciwnie, energia osiąga minimum dla $r = r_0$ a następnie wzrasta. Taka zależność energii od odległości wskazuje, że występują siły przyciągania i odpychania. Siły odpychania są właśnie siłami wymiany. Występowanie minimum energii wskazuje na możliwość istnienia stabilnego układu złożonego z dwóch atomów.

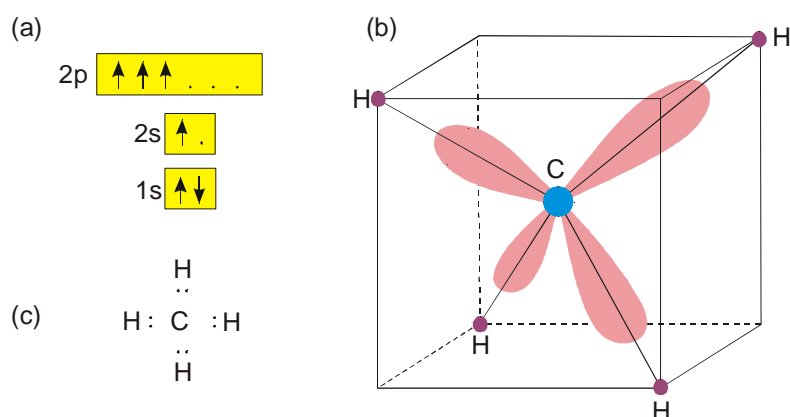
Wyżej omówione wiązanie, które powstaje za pośrednictwem oddziaływania elektronów walencyjnych nazywamy wiązaniem kowalencyjnym.

W kryształach typowe wiązanie kowalencyjne występuje w pierwiastkach czwartej grupy (węgiel o strukturze diamentu, krzem, german). Wszystkie te kryształy posiadają strukturę diamentu.

Atomy IV grupy mają na orbicie walencyjnej cztery elektrony (np. o konfiguracji $2s^2 2p^6$ dla węgla): dwa w stanie s (o spinach przeciwnych) i dwa w stanie p (o spinach zgodnych). Na

skutek oddziaływania między atomami w kryształach, konfiguracja elektronów w atomie zmienia się w ten sposób, że w stanie *s* jest jeden elektron, a w stanie *p* trzy elektrony, przy czym wszystkie cztery elektrony mają spiny równoległe (rys. 14.20). Każdy elektron może uczestniczyć w czterech wiązaniach i będzie tworzył pary elektronowe z czterema sąsiednimi atomami.

W cząsteczce CH_4 orbitale $2s$ i $2p$ atomu węgla nie zachowują swego normalnego kształtu, lecz ulegają zmieszaniu (czyli tzw. hybrydyzacji), w wyniku czego powstaje orbital będący kombinacją liniową jednej funkcji falowej $2s$ oraz trzech funkcji falowych $2p$. Nosi on nazwę orbitalu zhybrydowanego sp^3 . Jego kształt przedstawiono w sposób poglądowy na rys. 14.20b. Chmura elektronowa otaczająca rdzeń atomowy węgla ma cztery równoważne "ramiona" skierowane wzdłuż osi łączących środek czworościanu foremnego (tetraedru) z narożnikami.



Rys. 14.20. Stan walencyjny atomu węgla: (a) konfiguracja elektronowa; (b) orbital atomowy cząsteczki CH_4 , (c) wzór strukturalny cząsteczki CH_4 .

Wiązanie kowalencyjne jest wiązaniem silnym. Ma charakter wysoce kierunkowy, ponieważ powstaje w tym kierunku, w którym skupiona jest największa część chmury elektronowej skolektywizowanych elektronów. Sztywność struktury elektronowej powoduje, że kryształy są twarde i niepodatne na odkształcanie.

W temperaturach bardzo niskich kryształy kowalencyjne są izolatorami, gdyż każdy elektron walencyjny jest związany z atomem. Ponieważ jednak powłoka walencyjna nie jest zamknięta, siły wiążące na tej powłoce nie są zbyt duże. W wysokiej temperaturze energia drgań cieplnych jest wystarczająco duża do zerwania niektórych wiązań i uwolnienia pewnej liczby elektronów. Kryształ staje się przewodnikiem prądu. Wiązanie kowalencyjne jest typowym wiązaniem kryształów półprzewodnikowych.

14.8.3. Wiązania metaliczne

Wiązanie metaliczne występuje w przypadku atomów o małej liczbie elektronów walencyjnych. Wówczas energia wiązania tych elektronów jest niewielka. Przy zbliżaniu się atomów zachodzi nakrywanie się funkcji falowych elektronów walencyjnych. Jeżeli nakrywanie to występuje w różnych kierunkach, to w rezultacie gęstość prawdopodobieństwa znalezienia

elektronów walencyjnych jest praktycznie stała. Oznacza to, że elektrony nie należą do konkretnych atomów i zachowują się jak elektrony swobodne. Można więc powiedzieć, że metal składa się z dodatnio naładowanych jonów znajdujących się w morzu elektronów swobodnych. Siły przyciągania między jonami dodatnimi a ujemnie naładowaną plazmą elektronową przewyższają kulombowskie siły odpychania między jonami, dzięki czemu układ jest stabilny.

Wiązanie metaliczne nie ma charakteru wiązania skierowanego. Powoduje to, że atomy mogą zmieniać swoje położenie, co oznacza, że metale są plastyczne. Jedne atomy można zastąpić drugimi. Tłumaczy to możliwość występowania stopów i łatwość łączenia się metali. Metale są nieprzezroczyste, gdyż promieniowanie zakresu widzialnego jest silnie absorbowane przez swobodne elektrony.

Koncentracja swobodnych elektronów w metalu jest bardzo duża i zawiera się w granicach 1–2 elektronów na atom, a więc jest rzędu 10^{23} cm^{-3} .

14.8.4. Wiązanie molekularne

W przypadku kryształów molekularnych w węzłach sieci krystalicznej znajdują się molekuly połączone ze sobą siłami molekularnymi, zwanymi inaczej siłami van der Walsa. Oddziaływanie tego typu występuje w każdej substancji, jest jednak zdominowane przez inne (omówione wcześniej) oddziaływania. Jeżeli jednak powłoki elektronowe w atomach są całkowicie wypełnione, to siły van der Walsa są głównymi siłami jakie występują między atomami i cząstkami. Sytuacja taka występuje np. w gazach szlachetnych. Wiązania molekularne mają charakter wiązań elektrostatycznych. Wyróżniamy trzy rodzaje sił: orientacyjne, indukcyjne i dyspersyjne.

Siły orientacyjne występują w przypadku, gdy cząstki mają trwałe elektryczne momenty dipolowe. Siły oddziaływania elektrostatycznego starają się ułożyć je w określonym porządku, któremu odpowiadać będzie minimum energii. Ruchy termiczne cząstek burzą to uporządkowanie, dlatego energia tych oddziaływań zmniejsza się ze wzrostem temperatury.

Siły indukcyjne występują w przypadku, kiedy tylko część molekuł ośrodka ma trwałe momenty dipolowe. Pole elektryczne wytworzone przez te dipole indukuje momenty dipolowe w pozostałych niepolarnych cząstkach i powoduje ich przyciąganie.

Oddziaływanie dyspersyjne występuje między molekułami, które nie mają trwałych elektrycznych momentów dipolowych. Ponieważ molekula jest układem dynamicznym, to w poszczególnych chwilach występuje pewne rozsuniecie środków ciężkości ładunków dodatnich i ujemnych. Kierunki tych momentów są w każdej chwili różne. Jednak w wyniku ich oddziaływania zmiany kierunków oddziaływania mogą odbywać się w sposób skoordynowany, co prowadzi do pojawienia się trwałej siły przyciągania.

Na ogół w kryształach występują wszystkie rodzaje oddziaływań, przy czym największa część energii przypada na oddziaływanie orientacyjne i dyspersyjne. Warto również zapamiętać, że

wymienione oddziaływania są oddziaływaniami krótkiego zasięgu. Niewielka wartość tych sił powoduje, że wiązania międzycząsteczkowe łatwo ulegają zniszczeniu na skutek ruchów termicznych. Na skutek tego kryształy cząsteczkowe (hel, azot, wodór) topią się w niskich temperaturach i łatwo parują. Słabe wiązania sprawiają także, że kryształy molekularne są podatne na odkształcenia, a brak swobodnych elektronów powoduje, że są one bardzo złymi przewodnikami ciepła i prądu elektrycznego.

ROZDZIAŁ 15

ELEMENTY FIZYKI STATYSTYCZNEJ

Kryształ możemy traktować jako układ składający się z olbrzymiej ilości cząstek takich jak jądra i elektrony. W takich układach ujawniają się specyficzne prawa statystyczne będące przedmiotem fizyki statystycznej.

W tym rozdziale krótko omówimy niektóre elementy fizyki statystycznej, konieczne dla zrozumienia rozważań dotyczących koncentracji nośników w półprzewodnikach.

Aby jednak uświadomić sobie zasadniczą różnicę w podejściu fizyki statystycznej do układów składających się z dużej liczby cząstek, przypomnijmy sobie podstawowe wiadomości dotyczące termodynamicznego opisu układu.

15.1. Termodynamiczny opis układu

W termodynamicznym opisie właściwości układu składającego się z ogromnej liczby cząstek, układ rozpatrujemy jako obiekt makroskopowy nie interesując się naturą cząstek z których on się składa. Układ taki nosi nazwę układu termodynamicznego. Stan w którym układ może znajdować się dowolnie długo nosi nazwę stanu równowagi termodynamicznej. Stan taki jest w sposób jednoznaczny określony przez zadany zbiór niezależnych parametrów fizycznych, tzw. parametrów stanu. Podstawowymi parametrami stanu są: objętość układu V , ciśnienie p i temperatura T . Często jednak parametry te niewystarczają do opisu stanu układu, np. dla układu składającego się z kilku substancji należy także podać ich koncentrację.

Każda zmiana w układzie termodynamicznym, związana ze zmianą choćby jednego z parametrów stanu, nazywa się procesem termodynamicznym.

U podstaw termodynamiki leżą dwie zasady termodynamiki.

15.1.1. Pierwsza zasada termodynamiki

Zbiór wszystkich postaci energii zawartej w układzie izolowanym nosi nazwę energii wewnętrznej układu U . Energia wewnętrzna jest funkcją stanu układu. Oznacza to, że każdemu stanowi układu odpowiada jedna i tylko jedna jednoznacznie określona wartość energii wewnętrznej. W wyniku oddziaływania z otaczającym ośrodkiem, układ termodynamiczny może otrzymać lub oddać pewną ilość ciepła ΔQ , może wykonać pracę lub może być nad nim wykonana praca ΔW . We wszystkich przypadkach *zmiana energii wewnętrznej układu jest*

równa sumie ilości ciepła i pracy wymienionej przez układ z otoczeniem. Jest to pierwsza zasada termodynamiki. Ilościowo zapiszemy ją w postaci

$$\Delta Q = dU + \Delta W, \quad (15.1)$$

gdzie ΔQ – ilość ciepła pobranego przez układ, dU – zmiana energii wewnętrznej, ΔW – praca wykonana przez układ przeciwko siłom zewnętrznym.

Należy zaznaczyć, że w odróżnieniu od energii wewnętrznej, praca ΔW i ciepło ΔQ zależą nie tylko od stanu początkowego i końcowego układu, lecz także od drogi po której nastąpiła zmiana stanu układu.

15.1.2. Druga zasada termodynamiki

Jeżeli w układzie termodynamicznym proces przebiega nieskończenie powoli ze stanu 1 do stanu 2 i z powrotem do stanu 1 przez te same pośrednie stany (bez jakichkolwiek zmian w otaczającym ośrodku), to tego rodzaju proces nazywamy odwracalnym. Proces odwracalny jest procesem wyidealizowanym. Wszystkie rzeczywiste procesy są nieodwracalne. Nieodwracalne są procesy chłodzenia, mieszania, rozprężania gazów, itp. Procesy nieodwracalne mogą przebiegać tylko w jednym kierunku. Kryterium przebiegu procesów określa druga zasada termodynamiki.

Założmy, że przy nieskończenie małej zmianie stanu układu w temperaturze T układ uzyskuje ciepło ΔQ . Stosunek $\Delta Q/T$, jak się okazuje, jest różniczką zupełną pewnej funkcji S , która podobnie jak energia wewnętrzna, zależy tylko od stanu układu

$$dS = \frac{\Delta Q}{T}. \quad (15.2)$$

Funkcja ta nosi nazwę entropii układu.

Rozpatrzmy zmianę entropii w procesach nieodwracalnych.

W przypadku wymiany ciepła między dwoma ciałami, ogólna zmiana entropii wynosi

$$\Delta S = \frac{Q_1}{T_1} - \frac{Q_2}{T_2} = Q \frac{T_2 - T_1}{T_1 T_2},$$

gdzie Q_1 – ilość ciepła pobranego przez pierwsze ciało, Q_2 – ilość ciepła oddanego przez drugie ciało. Ponieważ $Q_1 = Q_2 = Q$ oraz $T_1 < T_2$, więc

$$\Delta S > 0.$$

Z nierówności tej wynika, że w procesie nieodwracalnym zwiększa się entropia układu.

W procesie odwracalnym $T_2 = T_1$ i $\Delta S = 0$. Zatem wzrost entropii przy nieodwracalnej wymianie ciepła jest większy niż przy odwracalnej.

Przypuśćmy, że ΔQ jest ilością ciepła jaką układ pobrał w temperaturze T . Gdyby ciepło to było pobrane w procesie odwracalnym, to przyrost entropii wynosiłby

$$\Delta S_{od} = \frac{\Delta Q}{T}. \quad (15.3)$$

Natomiast w procesie nieodwracalnym przyrost entropii będzie większy od tej wielkości

$$\Delta S_{nieod} > \frac{\Delta Q}{T}. \quad (15.4)$$

W przypadku układu izolowanego $\Delta Q = 0$, zatem

$$\Delta S_{nieod} > 0. \quad (15.5)$$

Związek (15.5) stanowi kryterium określające ukierunkowanie procesu nieodwracalnego w układzie izolowanym: *proces nieodwracalny przebiega w kierunku wzrostu entropii*.

Łącząc związki (15.3) i (15.4) otrzymamy

$$dS \geq \frac{\Delta Q}{T}. \quad (15.6)$$

Wzór (15.6) wyraża drugą zasadę termodynamiki: *w procesach odwracalnych zachodzących w układzie izolowanym entropia pozostaje stała, a w nieodwracalnych wzrasta*.

Podstawiając ΔQ wyznaczone z (15.6) do (15.1) otrzymujemy podstawowy związek termodynamiczny, który łączy pierwszą i drugą zasadę termodynamiki

$$dU \leq TdS - \Delta W. \quad (15.7)$$

15.1.3. Warunki równowagi termodynamicznej

Zgodnie z drugą zasadą termodynamiki, w układzie izolowanym samorzutnie mogą przebiegać tylko takie procesy w których entropia wzrasta. Proces zanika i układ osiąga stan równowagi gdy entropia osiągnie największą wartość. Dlatego *warunkiem równowagi układu izolowanego jest maksimum entropii*

$$S = S_{max} \quad (\delta S) \leq 0. \quad (15.8)$$

Symbol δ wskazuje, że entropia pozostaje stała lub maleje przy dowolnie małych odchyleniach układu od stanu równowagi.

Warunek równowagi może być podany także w innej postaci. Ze związku (15.7) mamy

$$\Delta W \leq TdS - dU. \quad (15.9)$$

Funkcja

$$F = U - TS \quad (15.10)$$

nosi nazwę energii swobodnej układu.

Podstawiając (15.10) do (15.9) znajdujemy

$$dW \leq -dF . \quad (15.11)$$

Ze wzoru (15.11) wynika, że układ wykonuje maksymalną pracę w warunkach izotermicznych w procesie odwracalnym. Praca ta jest równa ubytkowi energii swobodnej układu.

Jeżeli układ nie wykonuje pracy, to $\Delta W = 0$ i związek (15.11) przyjmuje postać

$$dF \leq 0 . \quad (15.12)$$

Z nierówności (15.12) wynika, że *we wszystkich samorzutnych procesach energia swobodna układu maleje*. Podczas równowagi osiąga ona wartość minimalną. Dlatego warunek równowagi ma postać

$$F = F_{min} \quad (\delta F)_{T,W} \geq 0 . \quad (15.13)$$

Przy dowolnie małych odchyleniach układu z położenia równowagi, energia swobodna układu pozostaje stała (procesy odwracalne) lub wzrasta (procesy nieodwracalne).

15.1.4. Potencjał chemiczny

Ze związku (15.7) wynika, że energia wewnętrzna układu może zmieniać się kosztem wymiany ciepła i pracy. Jednakże energia układu może zmieniać się także w wyniku zmiany liczby cząstek N w układzie, bowiem każda ubywająca cząstka unosi z sobą określoną ilość energii. Dlatego dla procesów odwracalnych związek (15.7) w ogólnej postaci zapisujemy następująco:

$$dU = TdS - pdV + \mu dN . \quad (15.14)$$

Przyjeliśmy tutaj, że układ wykonuje pracę przeciwko ciśnieniu zewnętrznemu $\Delta W = p\Delta V$. Parametr μ nosi nazwę potencjału chemicznego, analogicznie do potencjału elektrostatycznego, który pomnożony przez zmianę ładunku dq (tzn. przez zmianę liczby naładowanych cząstek) wyraża zmianę energii elektrostatycznej ciała.

Z równania (15.14) znajdujemy

$$\mu = \left(\frac{\partial U}{\partial N} \right)_{S,V} . \quad (15.15)$$

Potencjał chemiczny wyraża zmianę energii wewnętrznej układu przy zmianie liczby jego cząstek o jednostkę w warunkach stałej entropii i objętości układu.

Ze wzoru (15.10) znajdujemy

$$dF = dU - TdS - SdT .$$

Podstawiając tu dU z (15.14) otrzymujemy

$$dF = -pdV - SdT + \mu dN, \quad (15.16)$$

stąd zaś

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V}. \quad (15.17)$$

Potencjał chemiczny wyraża również zmianę energii swobodnej układu przy zmianie liczby jego cząstek o jednostkę w warunkach stałej temperatury i objętości.

Rozważmy warunek równowagi układu, w którym całkowita liczba cząstek pozostaje stała, lecz cząstki te mogą przechodzić z jednego ciała do drugiego. Przykładem takiego układu są dwa metale kontaktujące ze sobą i znajdujące się w stałej temperaturze. Oznaczmy potencjał chemiczny gazu elektronowego w pierwszym metalu przez μ_1 , w drugim zaś przez μ_2 . Ponieważ objętość układu pozostaje stała, więc warunkiem jego równowagi, zgodnie z (15.13), jest minimum energii swobodnej, przy którym $(dF)_{T,V} = 0$. Korzystając z wyrażenia (15.16) na dF , otrzymujemy

$$\mu_1 dN = \mu_2 dN,$$

gdzie $\mu_1 dN$ oznacza przyrost energii swobodnej pierwszego przewodnika wskutek przejścia do niego dN elektronów z przewodnika drugiego, $\mu_2 dN$ – ubytek energii w drugim przewodniku. Skracając przez dN , znajdujemy

$$\mu_1 = \mu_2. \quad (15.18)$$

Warunkiem równowagi takiego układu jest równość potencjałów chemicznych obu przewodników.

15.2. Statystyczny opis układu

Drugą metodą opisu właściwości układu składającego się z wielkiej liczby cząstek jest metoda statystyczna.

Stan każdej cząstki opisuje się przez zadanie trzech jej współrzędnych oraz jej składowych pędu. Wydawałoby się, że układając równanie ruchu dla cząstek i rozwiązując je, można uzyskać całą wiedzę o zachowaniu się układu i przewidzieć jego stan w dowolnej chwili. Jednakże takiego rodzaju obliczenia są nie tylko niemożliwe do wykonania (dla gazu o objętości 1m^3 w warunkach normalnych należałoby rozważyć około 10^{26} zależnych od siebie równań ruchu z uwzględnieniem warunków początkowych), lecz także bezużyteczne. W dużych zespołach cząstek nie jest istotny los każdej cząstki oddzielnie, lecz zachodzi konieczność znalezienia innych zależności obowiązujących cały układ. Układ jako całość podlega innym prawom niż prawa rządzące pojedynczą cząstką. Tymi prawami są prawa statystyczne.

Podstawową osobliwością statystycznych prawidłowości jest ich probabilistyczny charakter. Pozwalają one bowiem przewidzieć jedynie prawdopodobieństwo zajścia tego lub innego zdarzenia lub otrzymania tego lub innego wyniku.

Celem fizyki statystycznej jest, między innymi, szukanie najbardziej prawdopodobnego rozkładu cząstek w zależności od ich energii. Rozkład ten nazywamy statystyczną funkcją rozkładu. Jest to funkcja parametryczna. Za parametry termodynamiczne wygodnie jest przyjąć temperaturę i potencjał chemiczny.

Jeżeli znana jest funkcja rozkładu, to możemy znaleźć wartości średnie poszczególnych wielkości fizycznych charakteryzujących cząstki; a więc ich średnią energię kinetyczną, średnią wartość pędu, średnią wartość prędkości, itp. Znajomość różnych średnich wartości parametrów cząstki umożliwia obliczenie niektórych makroskopowych parametrów układu; jak np. ciśnienia, przewodnictwa cieplnego, temperatury, współczynnika dyfuzji, itp.

W zależności od indywidualnych właściwości cząstek tworzących układ, podlegają one odpowiednim rozkładom. Cząstki dzielimy na dwie grupy: fermiony i bozony.

Fermiony charakteryzują się wyraźną dążnością do "samotności". Jeżeli dany stan już jest zajęty przez fermion, to żaden inny fermion danego rodzaju nie może zająć tego stanu. Jest to konsekwencją zasady Pauliego.

Bozony przeciwnie, charakteryzują się dążnością do "łączenia się". Mogą one nieograniczenie zapełniać ten sam stan; przy czym czynią to tym "chętniej", im więcej bozonów znajduje się w tym stanie.

Jak wykazuje doświadczenie, fermiony mają spin połówkowy ($\hbar/2, 3\hbar/2, \dots$), bozony zaś - spin całkowity ($0, \hbar, 2\hbar, \dots$). Spośród cząstek elementarnych fermionami są: elektrony, protony, neutrony, neutrino i inne; a bozonami są: fotony, piony itp. Oprócz tego fermionami są jądra atomów pierwiastków chemicznych składające się z nieparzystej liczby nukleonów; bozonami zaś - jądra składające się z parzystej liczby nukleonów.

15.3. Układy niezwyrodniałe i zwyrodniałe

Okazuje się, że specyficzne właściwości bozonów i fermionów mają zdecydowany wpływ na zachowanie się układu jako całości. Dla ujawnienia się tych specyficznych właściwości konieczne jest aby cząstki "spotykały się" ze sobą dostatecznie często.

Założmy, że na N jednakowych cząstek przypada G różnych stanów, w których może znajdować się pojedyncza cząstka. Za miarę częstości "spotkań" można przyjąć stosunek N/G . Cząstki będą spotykać się rzadko, jeśli spełniony jest warunek

$$\frac{N}{G} \ll 1. \quad (15.19)$$

W takich warunkach liczba nieobsadzonych stanów jest dużo większa od liczby cząstek, a wówczas specyficzne właściwości fermionów i bozonów nie mogą się ujawnić. Każda cząstka

ma do swojej dyspozycji wiele różnych stanów swobodnych i w związku z tym obsadzenie tego samego stanu przez kilka cząstek praktycznie nie występuje.

Układy tego rodzaju noszą nazwę niezwyrodniałych, a warunek (15.19) nazywa się warunkiem niezwyrodnienia.

Jeżeli liczba stanów G jest tego samego rzędu co liczba cząstek N , tzn. gdy spełniony jest warunek

$$\frac{N}{G} \geq 1, \quad (15.20)$$

wtedy zagadnienie jak są obsadzone stany – pojedynczo czy zespołowo, staje się bardzo istotne. W tym przypadku specyficzne właściwości fermionów i bozonów przejawiają się w pełni, a to wywiera znaczny wpływ na właściwości układu jako całości. Układy takie noszą nazwę zwyrodniałych.

Układy zwyrodniałe mogą tworzyć tylko obiekty kwantowo-mechaniczne. Dla spełnienia warunku (15.20) konieczne jest, aby liczba możliwych stanów cząstek (liczba G) była w każdym przypadku skończona. Może to zachodzić tylko w takim przypadku, gdy parametry stanu cząstki zmieniają się w sposób dyskretny.

W przeciwieństwie do tego obiekty klasyczne, dla których parametry stanów zmieniają się w sposób ciągły, mogą tworzyć tylko układy niezwyrodniałe. Układy niezwyrodniałe mogą tworzyć także obiekty kwantowo-mechaniczne, jeżeli tylko spełniony jest warunek (15.19).

15.4. Statystyki: klasyczna i kwantowa

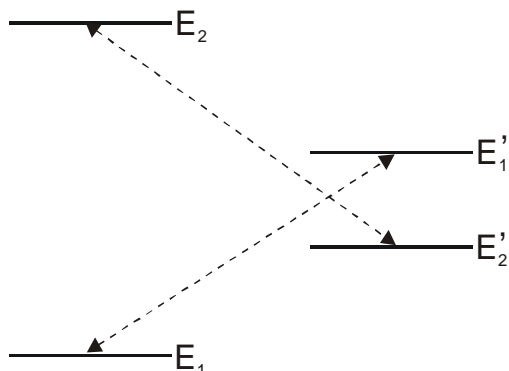
Fizyka statystyczna, która bada właściwości układów niezwyrodniałych, nazywa się statystyką klasyczną. Łączy się ją zazwyczaj z nazwiskami Maxwella i Boltzmanna (statystyka Maxwella-Boltzmanna).

Fizyka statystyczna, która bada właściwości układów zwyrodniałych, nazywa się statystyką kwantową. Wpływ specyficznych właściwości cząstek na właściwości układów zwyrodniałych warunkuje różnicę między zwyrodniałymi układami fermionów i bozonów. W związku z tym rozróżnia się dwie statystyki kwantowe.

Statystykę kwantową bozonów, związaną z nazwiskami Bosego i Einsteina, nazywamy statystyką Bosego-Einsteina.

W statystykach kwantowych występują tylko obiekty kwantowe, podczas gdy w statystyce klasycznej mogą występować zarówno obiekty klasyczne jak i kwantowe. Jeżeli będziemy zmniejszać liczbę cząstek w układzie lub zwiększać liczbę możliwych stanów cząstki, to układ zwyrodniały przekształci się w końcu w układ niezwyrodniały, który niezależnie od swojego charakteru będzie opisany przez statystykę klasyczną Maxwella-Boltzmanna.

W dalszych rozważaniach konieczna będzie znajomość funkcji rozkładu elektronów (fermionów) określająca prawdopodobieństwo obsadzenia stanów energetycznych. Będziemy więc szukać funkcji rozkładu $f(E)$ stosując uproszczone postępowanie, które jednak pozwoli nam bardziej obrazowo uzmysłwić sobie sens fizyczny $f(E)$.



Rys. 15.1. Stany energetyczne początkowe i końcowe elektronów.

Rozważmy sytuację gdy dwa elektrony o energiach E_1 i E_2 w wyniku jakiegoś oddziaływania zmieniły zajmowane stany na stany o energiach E'_1 i E'_2 . Aby tego rodzaju oddziaływanie zachodzące w układzie odosobnionym było możliwe, musi być spełniona zasada zachowania energii

$$E_1 + E_2 = E'_1 + E'_2, \quad (15.21)$$

tzn., że całkowita energia układu nie ulegnie zmianie.

Ilość przejść w jednostce czasu od stanu A do A' jest proporcjonalna do prawdopodobieństwa przejścia między tymi stanami $p(A, A')$ oraz do obsadzenia stanów wyjściowych:

$$N(A, A') = p(A, A') f(E_1) f(E_2).$$

Analogicznie ilość przejść w drugą stronę wynosi:

$$N(A', A) = p(A', A) f(E'_1) f(E'_2).$$

W stanie równowagi termodynamicznej ilość przejść $N(A, A')$ jest równa ilości przejść w drugą stronę $N(A', A)$:

$$p(A, A') f(E_1) f(E_2) = p(A', A) f(E'_1) f(E'_2).$$

Ale na podstawie zasady równowagi szczegółowej, prawdopodobieństwo przejść między dwoma stanami jest niezależne od ich kolejności, czyli $p(A, A') = p(A', A)$. Funkcja rozkładu musi więc spełniać równanie funkcyjne:

$$f(E_1) f(E_2) = f(E'_1) f(E'_2) \quad (15.22)$$

z warunkiem (15.21). Jediną funkcją, która spełnia powyższe założenia, jest funkcja wykładnicza

$$f(E) = Ne^{-E/kT}. \quad (15.23)$$

Jest to funkcja rozkładu Maxwella-Boltzmannna.

Statystyki Maxwella-Boltzmana nie można jednak zastosować do gazu elektronowego w metalach czy półprzewodnikach. Nieprzydatność tej statystyki wynika po prostu z tego, że dla elektronów nie jest obojętne czy stan końcowy jest zajęty, czy też nie. W fizyce klasycznej nie obowiązuje zakaz Pauliego.

Istotna różnica między statystyką kwantową a klasyczną polega na uwzględnieniu zależności liczby przejść od obsadzenia stanu końcowego.

Aby otrzymać poprawny rozkład $f(E)$ dla elektronów w ciele stałym, musimy uwzględnić fakt, że są one fermionami, czyli obowiązuje zakaz Pauliego. W naszych rozważaniach zasada ta znajduje swój wyraz w stwierdzeniu, że liczba przejść ze stanu A do A' jest proporcjonalna nie tylko do obsadzenia stanów wyjściowych, ale również do ilości wolnych miejsc w stanach końcowych, tzn. do $[1 - f(E)]$. Tak więc równanie funkcyjne (15.22) przejdzie w równanie postaci

$$f(E_1)[1 - f(E_1')]f(E_2)[1 - f(E_2')] = f(E_1')[1 - f(E_1)]f(E_2')[1 - f(E_2)]. \quad (15.24)$$

Równanie to można łatwo sprowadzić do równania (15.22), jeżeli wprowadzimy funkcję $\phi(E)$ zdefiniowaną następująco

$$\phi(E) = \frac{f(E)}{1 - f(E)}. \quad (15.25)$$

Przez podzielenie obu stron równania (15.24) przez wszystkie wyrazy w nawiasach kwadratowych, przekształcamy je do postaci:

$$\phi(E_1)\phi(E_2) = \phi(E_1')\phi(E_2'). \quad (15.26)$$

Przypomnijmy, że argumenty funkcji $\phi(E)$ spełniają równanie (15.21).

Sprowadziliśmy więc zagadnienie do poprzedniego. Rozwiązaniem równania (15.26) będzie znów funkcja eksponentialna

$$\phi(E) = Ne^{-E/kT} = e^{-(E-\mu)/kT}. \quad (15.27)$$

Zastąpiliśmy tu stałą N odpowiadającą jej wielkością w wykładniku eksponensa [$N = \exp(\mu/kT)$]. Z rozważań termodynamicznych wynika, że wielkość μ jest potencjałem chemicznym zwanym inaczej energią Fermiego, którą dalej będziemy oznaczać przez E_F . Podstawiając znalezione rozwiązanie do równania (15.25) widzimy, że funkcja rozkładu, której podlegają elektrony, ma postać

$$f(E) = \frac{1}{\exp\left[\frac{(E - E_F)}{kT}\right] + 1}, \quad 0 \leq f(E) \leq 1 \quad (15.28)$$

Jest to dobrze znana funkcja rozkładu Fermiego-Diraca. Jej wartość jest zawarta w przedziale między 0 a 1. Jest to zgodne z intuicją fizyczną, gdyż wartości 0 i 1 oznaczają stany puste i całkowicie obsadzone.

Zanim przejdziemy do dyskusji właściwości tego rozkładu, zastanówmy się jakiemu rozkładowi będą podlegały cząstki należące do grupy bozonów.

Dla bozonów ilość przejść do stanu końcowego jest proporcjonalna do obsadzenia stanu wyjściowego oraz do obsadzenia stanu końcowego +1, tzn. do $[1 + f(E)]$. W tym przypadku obsadzenie stanu końcowego nie zmniejsza, ale zwiększa ilość przejść. Fakt ten nie ma niestety prostej intuicyjnej interpretacji fizycznej, tak jak to było dla fermionów. Żartobliwie można przyrównać bozony do pieniędzy. Im większym kapitałem dysponujemy, tym szybciej go pomnażamy; szybciej rosną duże kapitały niż małe. Wracając do fizyki widzimy, że równanie (15.24) dla bozonów będzie miało postać

$$f(E_1)[1 + f(E_1)]f(E_2)[1 + f(E_2)] = f(E_1)[1 + f(E_1)]f(E_2)[1 + f(E_2)]. \quad (15.29)$$

Powtarzając analogiczne kroki do wykonanych przy wyprowadzaniu statystyki Fermiego-Diraca otrzymujemy, że rozkład Bosego-Einsteina ma postać

$$f(E) = \frac{1}{\exp\left[\frac{(E - E_F)}{kT}\right] - 1}. \quad (15.30)$$

Rozkład ten nie ma ograniczenia od góry, tzn. że $f(E)$ może przybierać wartości dowolne. W związku z tym sens tego rozkładu jest nieco inny aniżeli poprzednio. Daje on bowiem średnią ilość bozonów o energii E w temperaturze T .

Dla fotonów potencjał chemiczny jest równy zeru.

15.5. Gaz doskonały

Gaz doskonały spełnia następujące warunki: objętość cząsteczek gazu jest dużo mniejsza niż objętość zajmowana przez gaz, i zasięg sił działających między dwiema cząsteczkami jest o wiele mniejszy niż średnia odległość międzycząsteczkowa.

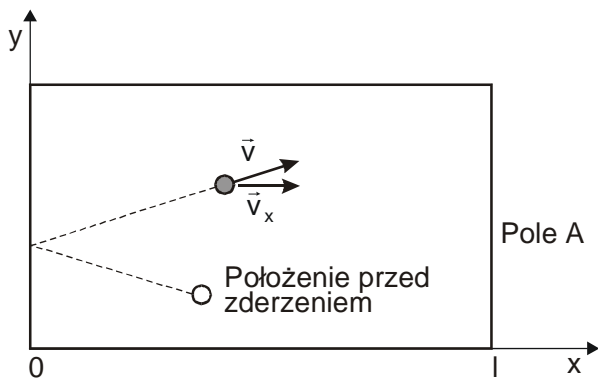
Rozważmy N sprężystych cząsteczek zamkniętych w pudełku o objętości V (rys. 15.2). Średnia siła, którą cząsteczka wywiera na ściankę w czasie Δt , wynosi

$$\bar{F} = \frac{\Delta p_x}{\Delta t}.$$

Zmiana pędu spowodowana zderzeniem jest równa

$$\Delta p_x = mv_x - (-mv_x) = 2mv_x.$$

Ponieważ czas pomiędzy kolejnymi zderzeniami z tą samą ścianką wynosi



Rys. 15.2. Cząsteczka w pudełku o objętości Al tuż po odbiciu od lewej ścianki

$$\Delta t = \frac{2l}{v_x},$$

więc średnia siła działająca na ściankę przez jedną cząsteczkę, jest równa

$$\bar{F} = \frac{(2mv_x)}{(2l/v_x)} = \frac{mv_x^2}{l}.$$

Dla wszystkich N cząsteczek w pudełku, całkowita siła działająca na ściankę wynosi

$$F = \frac{Nm\overline{v_x^2}}{l},$$

gdzie $\overline{v_x^2}$ jest to v_x^2 uśrednione po wszystkich cząsteczkach. Inaczej, $\overline{v_x^2}$ nazywamy średnią kwadratową prędkości w kierunku x . Dzieliąc obie strony ostatniego równania przez pole A ścianki otrzymujemy ciśnienie

$$p = \frac{Nm\overline{v_x^2}}{Al}.$$

A ponieważ Al jest objętością V , więc

$$p = \frac{Nm\overline{v_x^2}}{V},$$

czyli

$$pV = Nm\overline{v_x^2}. \quad (15.31)$$

Widzimy, że iloczyn pV dla gazu jest stały, jak długo stała jest energia kinetyczna cząsteczek. Ponieważ

$$\overline{v^2} = \overline{v_x^2} + \overline{v_y^2} + \overline{v_z^2},$$

a cząsteczki zderzają się w taki sam sposób ze wszystkimi sześcioma ściankami pudełka, więc mamy

$$\overline{v_x^2} = \overline{v_y^2} = \overline{v_z^2},$$

$$\overline{v^2} = 3\overline{v_x^2},$$

czyli

$$\overline{v_x^2} = \frac{\overline{v^2}}{3}.$$

Uwzględniając ostatni związek w równaniu (15.31), otrzymujemy

$$pV = Nm \left(\frac{\overline{v^2}}{3} \right). \quad (15.32)$$

Ostatni związek można przekształcić w inną postać uwzględniając, że temperatura bezwzględna jest definiowana jako wielkość wprost proporcjonalna do średniej energii kinetycznej cząsteczek

$$T = \left(\frac{2}{3k}\right) \frac{\overline{mv^2}}{2} = \left(\frac{2}{3k}\right) \overline{K}, \quad (15.33)$$

gdzie \overline{K} jest średnią energią kinetyczną przypadającą na jedną cząsteczkę. Czynnikiem $(2/3k)$ jest współczynnikiem proporcjonalności, a wielkość $k = 1,38 \times 10^{-23}$ J/K, zwana stałą Boltzmanna, zależy od właściwości wody skoro jednostkę temperatury ustalono poprzez wyznaczenie skali 100-stopniowej między punktami zamarzania i wrzenia wody przy ciśnieniu atmosferycznym.

Eliminując $\overline{v^2}$ z (15.32) i (15.33) otrzymujemy *prawo gazów doskonałych*

$$pV = NkT. \quad (15.34)$$

Prawo gazów doskonałych zastosowane do 1 mola gazu brzmi

$$pV = N_0 k T.$$

Iloczyn $N_0 k$ jest zwykle nazywany stałą gazową $R = N_0 k = (6,02 \times 10^{23})(1,38 \times 10^{-23})$ J/(mol K) = 8,31 J/(mol K).

Ze związku (15.33) otrzymujemy dodatkowo, że średnia energia kinetyczna ruchu postępowego wynosi

$$\frac{1}{2} \overline{mv^2} = \frac{3}{2} kT. \quad (15.35)$$

Odpowiada to trzem stopniom swobody; żeby określić położenie środka cząsteczki potrzeba trzech współrzędnych (x,y,z) . Stąd *średnia energia przypadająca na stopień swobody wynosi $(1/2)kT$ na cząsteczkę*. Jest to tzw. zasada ekwipartycji energii.

15.6. Gaz elektronów swobodnych

Wiele zjawisk fizycznych, zwłaszcza w ciałach stałych, jest uwarunkowanych elektronami swobodnymi. Przykładem może być przewodnictwo elektryczne i cieplne metali. Omówimy teraz podstawowe właściwości gazu elektronów swobodnych. Będziemy gaz ten traktować jako gaz doskonały stanowiący zbiór cząstek, których energia oddziaływania wzajemnego jest mała w porównaniu z ich energią kinetyczną. Zakładamy, że cząstki poruszają się ruchem postępowym.

15.6.1. Przestrzeń fazowa. Funkcja gęstości stanów

W fizyce klasycznej stan ruchu cząstki o trzech stopniach swobody określamy przez podanie trzech współrzędnych i trzech składowych pędu. Dla tej cząstki możemy wprowadzić sześciowymiarową przestrzeń zwaną przestrzenią fazową. Przestrzeń fazową nazywamy też

przestrzeni Γ . Stan ruchu cząstki jest w każdej chwili dokładnie odtworzony przez punkt w przestrzeni fazowej. Jeśli mamy do czynienia z układem złożonym z n cząstek, z których każda ma sześć stopni swobody, to musimy posługiwać się $6n$ -wymiarową przestrzenią fazową. Dla pojedynczej cząstki element objętości $d\Gamma$ w przestrzeni Γ będzie wynosił

$$d\Gamma = dx dy dz dp_x dp_y dp_z = d\Gamma_V d\Gamma_p, \quad (15.36)$$

gdzie $d\Gamma_V = dx dy dz$ jest elementem objętości w przestrzeni współrzędnych, $d\Gamma_p = dp_x dp_y dp_z$ – elementem objętości w przestrzeni pędów. Jeżeli punkt, który jest obrazem cząstki, znajdzie się w elemencie objętości $d\Gamma$, to oznacza, że rozważana cząstka ma współrzędne w przedziale od x do $x + dx$, od y do $y + dy$, od z do $z + dz$ oraz składowe pędu od p_x do $p_x + dp_x$, od p_y do $p_y + dp_y$, od p_z do $p_z + dp_z$.

Załóżmy, że rozpatrywanymi cząstkami są elektrony. Zgodnie z zasadą nieoznaczoności Heisenberga, stan elektronu nie możemy przedstawić za pomocą punktu w przestrzeni fazowej. Możemy natomiast elektronowi przypisać pewien element objętości. Ponieważ $dx dp_x = h$, więc stanowi elektronu w przestrzeni Γ przypiszemy element objętości

$$d\Gamma = dx dy dz dp_x dp_y dp_z = h^3. \quad (15.37)$$

Element objętości $d\Gamma = h^3$ jest komórką fazową w przestrzeni Γ . Jeżeli uwzględnić fakt, że elektrony mogą mieć dwa różne ustawienia spinów, to w elemencie objętości h^3 mogą być dwa elektrony. Objętość komórki elementarnej dla elektronu wynosi zatem $h^3/2$. Fakt, że elektronowi przypisuje się komórkę o objętości $h^3/2$ sprawia, że w określonym przedziale pędów (czy energii) znajduje się ściśle określona liczba komórek elementarnych. Aby wyznaczyć tę liczbę, rozważymy w przestrzeni pędów dwie powierzchnie kuliste o promieniach p i $p + dp$ (rys. 15.3). Element objętości tej warstwy kulistej wynosi

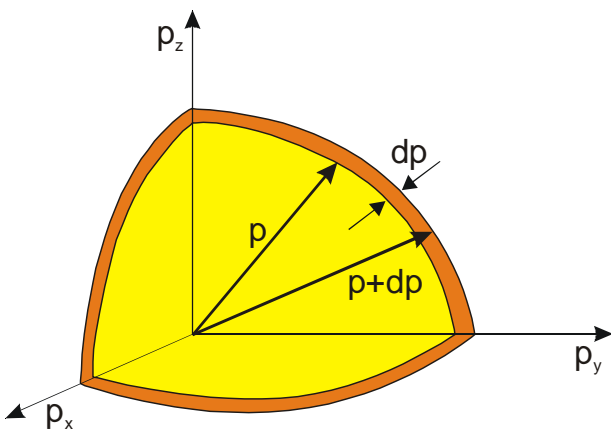
$$d\Gamma_p = 4\pi p^2 dp.$$

Ponieważ w przestrzeni pędów objętość komórki elementarnej wynosi

$$d\Gamma_p = \frac{h^3/2}{dx dy dz} = \frac{h^3/2}{d\Gamma_V},$$

więc w elemencie o objętości $4\pi p^2 dp$ liczba komórek elementarnych wynosi

$$\frac{4\pi p^2 dp}{d\Gamma_p} = \frac{8\pi p^2 dp}{h^3} d\Gamma_V.$$



Rys. 15.3. Element objętości w przestrzeni pędów.

Dalsze rozważania przeprowadzimy dla jednostkowego elementu objętości w przestrzeni współrzędnych. Wobec tego liczba komórek elementarnych znajdująca się w jednostce objętości w przedziale pędów od p do $p + dp$ wynosi

$$g(p)dp = \frac{8\pi p^2}{h^3} dp. \quad (15.38)$$

Funkcję $g(p)$ nazywamy funkcją gęstości stanów w przestrzeni pędów. Jest to liczba komórek elementarnych (liczba stanów) w jednostce objętości w jednostkowym przedziale pędów.

W wielu zagadnieniach wygodniej jest posługiwać się gęstością stanów w przestrzeni energii. Jeżeli przyjmujemy, że $v/c \ll 1$, to dla elektronów swobodnych

$$E = \frac{mv^2}{2} = \frac{p^2}{2m}.$$

Zatem

$$p^2 = 2mE, \quad dp = \frac{1}{2} \sqrt{\frac{2m}{E}} dE.$$

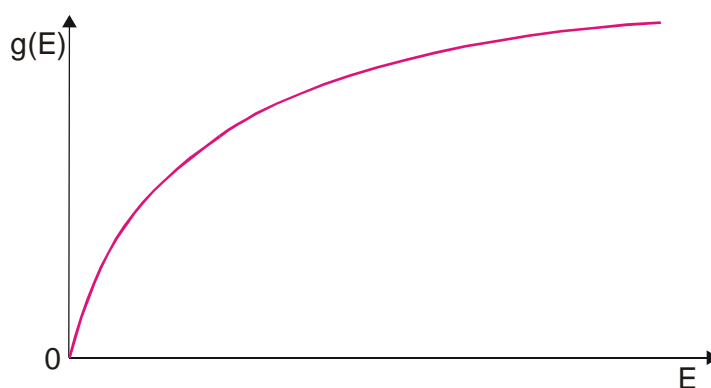
Podstawiając to do wzoru (15.38) otrzymamy wyrażenie na liczbę komórek elementarnych znajdujących się w jednostce objętości w przedziale energii od E do $E + dE$

$$g(E)dE = \frac{4\pi(2m)^{3/2}}{h^3} E^{1/2} dE.$$

Funkcję

$$g(E) = \frac{4\pi(2m)^{3/2}}{h^3} E^{1/2} \quad (15.39)$$

nazywamy funkcją gęstości stanów w przestrzeni energii. Jest to liczba komórek elementarnych (liczba stanów) w jednostce objętości w jednostkowym przedziale energii. W naszych rozważaniach będziemy zazwyczaj stosowali funkcję $g(E)$, nazywając ją krótko funkcją gęstości stanów. Na rys. 15.4 przedstawiono zależność funkcji gęstości stanów od energii.



Rys. 15.4. Zależność funkcji gęstości stanów od energii.

15.6.2. Gaz elektronowy w temperaturze 0 K

Rozważmy teraz gaz elektronowy który oziębamy do 0 K, tzn. zaniedbujemy ruchy termiczne. W takiej sytuacji cząstki klasyczne zajęłyby poziom energetyczny $E = 0$, gdyż nie ma żadnych ograniczeń na liczbę cząstek o danej energii. W przypadku elektronów będzie inaczej. W każdej komórce elementarnej o objętości $h^3/2$ może być tylko jeden elektron, gdyż umieszczenie tam drugiego elektronu oznaczałoby naruszenie zakazu Pauliego. Wobec tego elektrony będą zajmowały komórki o możliwie najmniejszej energii aż do pewnej energii maksymalnej, która będzie funkcją koncentracji elektronów. *Maksymalną energię elektronów w 0 K nazywamy energią Fermiego*. Ponieważ w temperaturze 0 K są obsadzone wszystkie stany energetyczne poniżej energii Fermiego, dlatego liczba elektronów dn w jednostce objętości w przedziale energii od E do $E + dE$ jest równa liczbie komórek. Możemy więc zapisać

$$dn = g(E)dE = \frac{4\pi(2m)^{3/2}}{h^3} E^{1/2} dE .$$

Całkując otrzymujemy

$$n = \frac{4\pi(2m)^{3/2}}{h^3} \int_0^{E_F} E^{1/2} dE = \frac{8\pi(2m)^{3/2}}{3h^3} E_F^{3/2} , \quad (15.40)$$

gdzie E_F jest energią Fermiego, a n – koncentracją elektronów (liczbą elektronów w jednostce objętości). Poniżej energii Fermiego wszystkie stany są obsadzone, powyżej – wszystkie stany są puste. Warto sobie uświadomić, że w przestrzeni pędów elektrony będą znajdowały się wewnątrz kuli o promieniu $p_F = (2mE_F)^{1/2}$. Powierzchnię tej kuli nazywa się powierzchnią Fermiego.

Aby oszacować jakie są wartości energii Fermiego, rozpatrzmy gaz elektronowy w miedzi. Koncentracja elektronów w miedzi jest równa koncentracji atomów

$$n = \rho \frac{N_A}{m_A} ,$$

gdzie ρ jest gęstością, N_A – liczbą Avogadra a m_A masą atomową. Podstawiając dane liczbowe otrzymujemy $n = 8,46 \times 10^{28} \text{ m}^{-3}$. Z wyrażenia (15.40) obliczymy

$$E_F = \frac{h^2}{2m} \left(\frac{3n}{8\pi} \right)^{2/3} , \quad (15.41)$$

otrzymując $E_F = 7,07 \text{ eV}$. Mając E_F możemy policzyć prędkość elektronów na poziomie Fermiego

$$v_F = \sqrt{\frac{2E_F}{m}} . \quad (15.42)$$

Dla rozpatrywanego przypadku $v_F = 1,57 \times 10^6$ m/s.

Z powyższego wynika, że nawet w temperaturze zera bezwzględnego, elektrony mają bardzo duże energie i poruszają się z olbrzymimi prędkościami.

Energia ruchu termicznego cząstki o trzech stopniach swobody jest równa $3kT/2$; w temperaturze $T = 300$ K wynosi 0,039 eV. W temperaturze 0 K elektrony mają energie setki razy większe od energii ruchów termicznych w temperaturze pokojowej. Na przykład, średnia energia kinetyczna określona wzorem

$$\langle E \rangle = \frac{\int E dn}{\int dn} = \frac{\int g(E)E dE}{\int g(E) dE} = \frac{3}{5} E_F, \quad (15.43)$$

wynosi 4,24 eV dla warunków podanych wyżej. Widzimy więc jak istotny jest wpływ zakazu Pauliego na właściwości gazu elektronowego. Aby energia ruchu termicznego była równa E_F , temperatura gazu powinna być rzędu 50000 K. Często gaz elektronowy charakteryzujemy temperaturą Fermiego zdefiniowaną tak, że $E_F = kT_F$.

W tabeli 15.1 podano wartości koncentracji elektronów n , energii Fermiego E_F , prędkości na poziomie Fermiego v_F i temperatury Fermiego T_F dla gazu elektronowego w różnych metalach.

Tabela 15.1. Parametry gazu elektronowego metali

Metal	n [10^{28} m^{-3}]	E_F [eV]	v_F [10^6 m/s]	T_F [10^4 K]
Ag	5,86	5,52	1,39	6,4
Al	18,06	11,69	2,03	13,5
Au	5,90	5,54	1,39	6,4
Cu	8,46	7,07	1,58	8,2
Fe	16,99	11,22	1,99	13,0
K	4,07	4,33	1,24	5,0
Na	2,54	3,16	1,05	3,7
Mg	8,93	7,31	1,60	8,6

15.6.3. Wpływ temperatury na gaz elektronowy

Wiemy, że w temperaturze zera bezwzględnego elektrony będą obsadzały wszystkie stany energetyczne od zera do E_F . Inaczej mówiąc, prawdopodobieństwo obsadzenia przez elektrony stanów o energii $E < E_F$ jest równe jeden, natomiast prawdopodobieństwo obsadzenia stanów o energii $E > E_F$ jest równe zero. Jeśli teraz podniesiemy temperaturę gazu, to elektrony uzyskają dodatkową energię w wyniku ruchów termicznych i zajmą wyższe stany energetyczne. Ponieważ energia ruchów termicznych wynosząca $3kT/2$ jest nawet w temperaturach stosunkowo wysokich znacznie mniejsza od E_F , dlatego tylko elektrony, których energia jest bardzo bliska E_F , mogą

zająć wyższe stany energetyczne. Oznacza to, że prawdopodobieństwo obsadzenia stanów o $E < E_F$ jest mniejsze od jedności, natomiast prawdopodobieństwo obsadzenia stanów o $E > E_F$ jest większe od zera.

Jak już poprzednio wykazaliśmy, funkcja Fermiego-Diraca określająca prawdopodobieństwo obsadzenia poziomu energetycznego o energii E , ma postać (15.28)

$$f(E) = \frac{1}{\exp\left[\frac{(E - E_F)}{kT}\right] + 1}.$$

Zwróćmy uwagę, że dla $T = 0$ K

$$f(E) = \begin{cases} 1, & \text{gd}y E < E_F \\ 0, & \text{gd}y E > E_F \end{cases}.$$

Tak w istocie być powinno gdyż stwierdziliśmy, że w temperaturze 0 K prawdopodobieństwo obsadzenia stanów o $E < E_F$ jest równe 1, a stanów o energii $E > E_F$ jest równe 0.

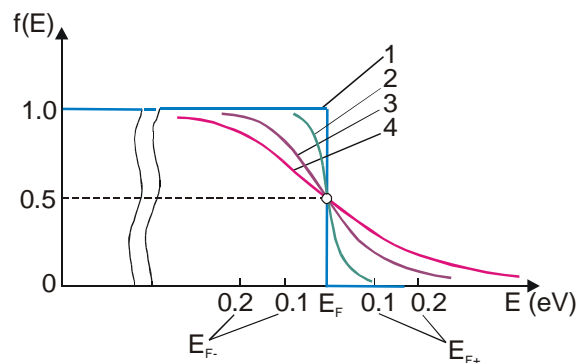
Natomiast dla $T > 0$ K mamy

$$\frac{1}{2} < f(E) < 1, \quad \text{dla } E < E_F$$

$$f(E) = \frac{1}{2}, \quad \text{dla } E = E_F$$

$$0 < f(E) < \frac{1}{2}. \quad \text{dla } E > E_F$$

Oznacza to, że w określonej temperaturze $T > 0$ K, *poziom Fermiego określamy jako poziom energetyczny, którego prawdopodobieństwo obsadzenia wynosi 1/2*. Stany położone poniżej poziomu Fermiego są obsadzone z większym prawdopodobieństwem niż stany położone powyżej tego poziomu. Im poziom energetyczny jest położony wyżej w stosunku do poziomu Fermiego, tym prawdopodobieństwo jego obsadzenia maleje, dążąc do zera. Natomiast im poziom energetyczny jest położony poniżej, tym prawdopodobieństwo jego obsadzenia wzrasta, dążąc do jedności. Przedział energii w którym funkcja rozkładu zmienia się od jedności do zera wynosi kilka kT . Na rys. 15.6 przedstawiono przebieg funkcji $f(E)$ dla kilku temperatur.



Rys. 15.6. Funkcja rozkładu Fermiego-Diraca dla temperatur: 1 – 0 K, 2 – 300 K, 3 – 1000 K i 4 – 1500 K

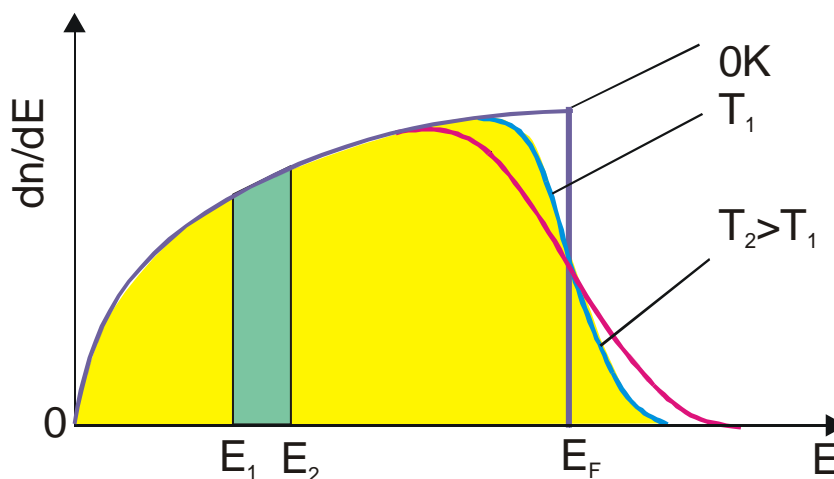
Koncentracja elektronów o energiach w przedziale od E do $E + dE$ wyrazi się wzorem

$$dn = g(E) f(E) dE .$$

Ponieważ energia zmienia się od 0 do ∞ , więc koncentracja elektronów

$$n = \int_0^{\infty} g(E) f(E) dE . \quad (15.43)$$

Na rys.15.7 przedstawiono zależność koncentracji elektronów na jednostkowy przedział dn/dE dla $T = 0$ K, T_1 i $T_2 > T_1$.



Rys. 15.7. Zależność koncentracji elementów na jednostkowy przedział energii od energii w różnych temperaturach. Zaznaczone pola są miarą całkowitej liczby elektronów i liczby elektronów o energii w przedziale od E_1 do E_2 .

Obsadzenie przez elektrony wyższych poziomów energetycznych powoduje, że ich średnia energia rośnie. Z obliczeń wynika, że w temperaturze T średnia energia wyraża się wzorem

$$\langle E \rangle = \frac{3}{5} E_F \left[1 + \frac{5\pi^2}{12} \left(\frac{kT}{E_F} \right)^2 \right] .$$

Ponieważ praktycznie zawsze $E_F \gg kT$, dlatego $\langle E \rangle$ bardzo nieznacznie wzrasta ze wzrostem temperatury.

Powróćmy jeszcze raz do wzoru (15.28) określającego rozkład Fermiego-Diraca. Okazuje się, że przy spełnianiu pewnych warunków rozkład ten może zostać zastąpiony przez rozkład Boltzmann (15.23). Załóżmy, że

$$\exp\left(\frac{E - E_F}{kT}\right) \gg 1. \quad (15.44)$$

Wówczas

$$f(E) = \frac{1}{\exp\left(\frac{E - E_F}{kT}\right) + 1} = \exp\left(-\frac{E - E_F}{kT}\right) = Ne^{-E/kT}, \quad (15.45)$$

gdzie $N = \exp(E_F/kT)$. Oznacza to, że przy spełnieniu warunku (15.44), rozkład Fermiego-Diraca przechodzi w rozkład Boltzmann. Jeśli żądamy, aby warunek (15.44) był spełniony dla każdej wartości energii, to w szczególności winien być spełniony dla $E = 0$. Oznacza to, że

$$e^{-E_F/kT} \gg 1,$$

lub

$$N = e^{E_F/kT} \ll 1. \quad (15.46)$$

Stałą N wyznaczamy w następujący sposób. Korzystając z zależności (15.38) i (15.43), i podstawiając w miejsce $f(E)$ wzór (15.45), otrzymujemy

$$n = \frac{4\pi(2m)^{3/2}}{h^3} N \int_0^\infty E^{1/2} e^{-E/kT} dE.$$

Obliczenia będzie łatwiej przeprowadzić, jeżeli podstawić $E = mv^2/2$, $dE = mv dv$. Mamy wówczas

$$n = \frac{8\pi m^3}{h^3} N \int_0^\infty v^2 e^{-mv^2/2kT} dv.$$

Ponieważ

$$\int_0^\infty v^2 e^{-mv^2/2kT} dv = \frac{\sqrt{\pi}}{4} \left(\frac{2kT}{m}\right)^{3/2},$$

więc

$$n = \frac{2(2\pi mkT)^{3/2}}{h^3} N. \quad (15.47)$$

Ostatecznie otrzymujemy

$$N = \frac{nh^3}{2(2\pi mkT)^{3/2}}. \quad (15.48)$$

Stała N będzie więc duża dla dużych koncentracji n i małych mas cząstek. W tych przypadkach można oczekiwać zwyrodnienia.

Obliczymy wartość stałej N w dwóch przypadkach dla temperatury $T = 300$ K:

- a) wodór w warunkach normalnych: $n = 2,69 \times 10^{25} \text{ m}^{-3}$, $m = 1,66 \times 10^{-27} \text{ kg}$,
- b) gaz elektronowy w miedzi: $n = 8,46 \times 10^{28} \text{ m}^{-3}$, $m = 9,1 \times 10^{-31} \text{ kg}$.

W przypadku (a) dostajemy $N = 1,4 \times 10^{-5}$, natomiast w przypadku (b) $N = 3,4 \times 10^3$. Oznacza to, że dla wodoru spełniony jest warunek (15.46) i można z powodzeniem stosować statystykę

Boltzmann. Podobnie dla innych cząstek gazu doskonałego. Zwyródnienia można się spodziewać w bardzo niskich temperaturach, znacznie niższych od temperatury skraplania.

Zupełnie inaczej przedstawia się sytuacja dla gazu elektronowego w metalach. Bardzo duże n i małe m powodują, że w zwykłych temperaturach $N \gg 1$. Gaz elektronowy w metalach jest zawsze gazem zwyródniałym i podlega rozkładowi Fermiego-Diraca.

Tylko w przypadku gdy koncentracje elektronów są niewielkie, gaz elektronowy jest niezwyrodniały. Sytuacja taka występuje w półprzewodnikach.

ROZDZIAŁ 16

STRUKTURA PASMOWA CIAŁ STAŁYCH

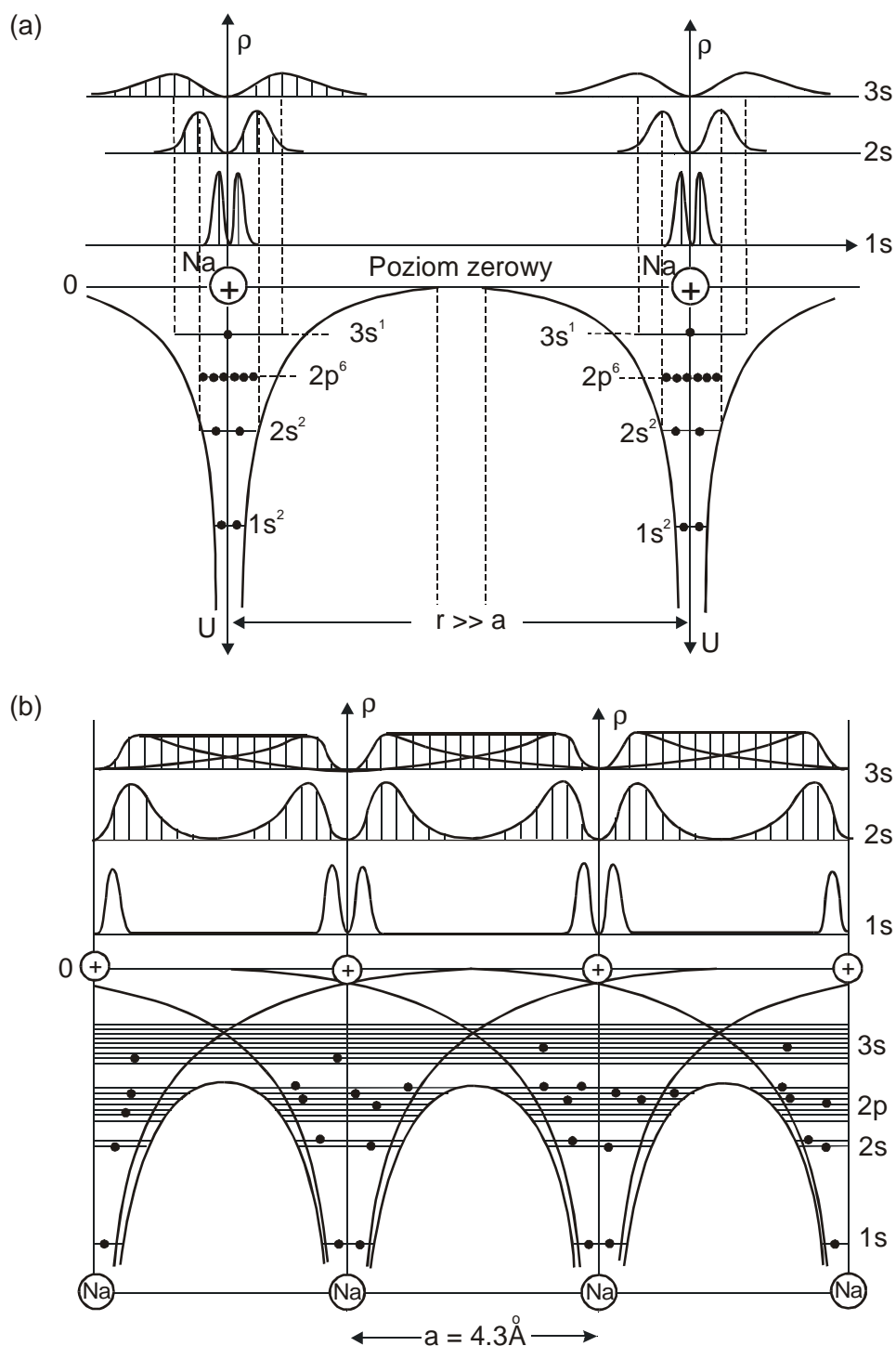
W punkcie 15.6 poznaliśmy podstawowe właściwości gazu elektronowego. Omówimy teraz, w sposób możliwie najprostszy, zachowanie się elektronów w ciele stałym. Jest to niezbędne do zrozumienia wielu właściwości tych ciał, a zwłaszcza półprzewodników.

Właściwości elektronów w ciałach stałych wynikają z ich oddziaływania między sobą i oddziaływania z atomami (jonami) sieci. Jednakże każde z tych oddziaływań jest skomplikowane. Wystarczy sobie bowiem uświadomić, że w 1 cm^3 ciała stałego znajduje się około 10^{23} elektronów i jonów będących w ciągłym ruchu i oddziaływujących na siebie.

W celu przesłedzenia wpływu pola elektrycznego na poziomy energetyczne elektronów w atomie, rozpatrzmy następujący wyidealizowany przykład, który pozwala w sposób jakościowy zrozumieć zachodzące zjawiska. Rozmieścimy N atomów sodu w węzłach sieci przestrzennej odpowiadającej kryształowi sodu, ale w tak dużych odległościach aby można było pominąć ich wzajemne oddziaływanie. W takim przypadku stany energetyczne elektronów będą odpowiadały stanom atomu izolowanego. Na rys. 16.1a pokazano schemat poziomów energetycznych dwóch izolowanych atomów sodu. W atomie tym jak wiadomo, poziomy energetyczne $1s$, $2s$ i $2p$ są zapełnione całkowicie, a poziomy rozmieszczone powyżej $3s$ są wolne. Zauważmy, że bariera potencjału nie pozwala elektronom poruszać się swobodnie między poszczególnymi atomami.

Poddajemy teraz sieć jednorodnemu ściskaniu. W miarę zbliżania się atomów do siebie, wzrasta ich oddziaływanie i w odległościach równych parametrowi sieci ($a = 0,43 \text{ nm}$) energia oddziaływania osiąga wartość minimalną. Na rys. 16.1b pokazano schemat poziomów energetycznych elektronów odpowiadający zbliżeniu atomów. Widzimy, że krzywe potencjalne oddzielające sąsiednie atomy częściowo zachodzą na siebie dając wypadkową krzywą potencjalną, która przebiega poniżej poziomu zerowego. Godnym uwagi jest fakt, że wysokość ta jest nawet nieco mniejsza od początkowego położenia poziomu elektronów walencyjnych $3s$. Dlatego też elektrony walencyjne uzyskują możliwość przechodzenia z jednego atomu do drugiego. Elektrony takie nazywamy swobodnymi, a ich właściwości poprzednio omówiliśmy. Zauważmy, że charakter funkcji falowych tych elektronów jest taki, że zachodzą one na siebie w takim stopniu, że tworzą chmurę o prawie równomiernej gęstości, a to oznacza stan pełnego uwspólnienia elektronów walencyjnych. Natomiast chmury elektronowe wewnętrznych powłok elektronowych atomów nie pokrywają się i stany elektronów wewnętrznych atomów kryształu pozostają w zasadzie takie same jak w atomach izolowanych. Należy ponadto zauważyć, że pojedyncze poziomy atomowe uległy rozszczepieniu na zespoły poziomów zwanych

dozwolonymi pasmami energetycznymi. Ponieważ prędkość ruchu elektronu jest rzędu 10^5 m/s, a rozmiary atomu wynoszą 10^{-10} m, to elektron walencyjny znajduje się w danym węzle w czasie $\Delta t = 10^{-15}$ s co powoduje, że szerokość poziomu energetycznego elektronu walencyjnego ΔE jest w przybliżeniu równa $\Delta E = h / \Delta t \approx 1$ eV.



Rys. 16.1. Zmiana stanu elektronów przy zbliżeniu się atomów: (a) schemat energetyczny dla atomów sodu znajdujących się w odległościach znacznie większych od parametru sieci; (b) ten sam schemat dla atomów sodu znajdujących się w odległościach rzędu parametru sieci.

Jeżeli poziomy energetyczne w atomie swobodnym były $(2l+1)$ -krotnie zdegenerowane (bez uwzględnienia spinu), to odpowiadające im pasma energetyczne będą miały $N(2l+1)$ podpoziomów (N – ilość atomów); np. z poziomu p wytworzy się pasmo składające się z $3N$ podpoziomów mogących zmieścić $6N$ elektronów. Jeżeli dalej uwzględnimy, że 1 g ciała stałego zawiera $N = 10^{22}$ atomów i że każdy pojedynczy poziom energetyczny rozpada się na N poziomów, to przy szerokości pasma rzędu 1 eV odległości między poziomami wynoszą około 10^{-22} eV, co wskazuje, że nie ma możliwości doświadczalnego ich rozróżnienia. O takiej właśnie grupie poziomów mówimy jako o paśmie dozwolonym uważając, że elektrony w tym paśmie mają ciągły rozkład energii. Pasma te są rozdzielone pasmami wzbronionymi.

Wyniki poglądowo przedstawione powyżej uzyskuje się również rozwiązując kwantowomechaniczne zagadnienia ruchu wielu elektronów w polu ruchu jonów. Złożoność tego zagadnienia zmusza nas do wprowadzenia szeregu uproszczeń. Dlatego analizując zachowanie elektronów w sieci stosujemy wiele przybliżeń. Dwoma skrajnymi przeciwstawnymi przybliżeniami są: przybliżenie elektronów prawie swobodnych i przybliżenie elektronów silnie związanych. W obu tych przypadkach otrzymujemy zakresy energii dozwolonych i zabronionych dla elektronów. Kształt pasm energetycznych i ich wypełnienie przez elektrony jest jednym z najpoważniejszych zagadnień w fizyce ciała stałego.

16.1. Funkcja Blocha

Cechą charakterystyczną kryształu doskonałego jest okresowość energii potencjalnej. Okresem tym jest stała sieci a . Dla przypadku jednowymiarowego możemy napisać

$$U(x) = U(x + a). \quad (16.1)$$

Równanie (16.1) jest spełnione jeśli atomy są nieruchome, tzn. zaniedbujemy ruchy termiczne i w kryształach nie ma żadnych defektów. Z okresowości potencjału wynika, że prawdopodobieństwo znalezienia elektronu w punkcie x jest takie samo jak w punkcie $(x+a)$, tzn.

$$|\Psi(x)|^2 = |\Psi(x + a)|^2.$$

Z powyższego wynika, że $\Psi(x+a)$ może się różnić od $\Psi(x)$ tylko czynnikiem o module równym 1. Czynnikiem ten zapiszemy w postaci e^{ika} , gdzie k jest wektorem falowym. Tak więc

$$\Psi(x + a) = \Psi(x) e^{ika}.$$

Funkcję $\Psi(x)$ możemy zapisać też w postaci

$$\Psi(x) = u_k(x) e^{ikx}, \quad (16.2)$$

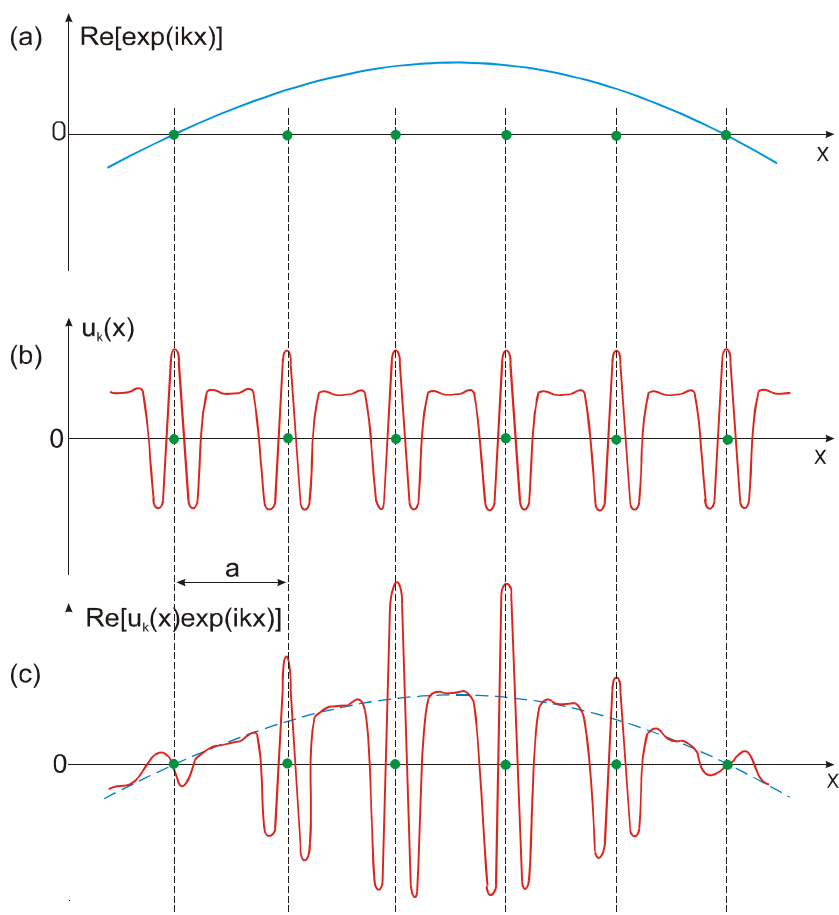
gdzie $u_k(x)$ jest funkcją okresową o okresie a , tzn. $u_k(x) = u_k(x+a)$. Istotnie, zgodnie ze wzorem (16.2) możemy zapisać

$$\Psi(x+a) = u_k(x+a)e^{ik(x+a)},$$

z drugiej strony

$$\Psi(x+a) = \Psi(x)e^{ika} = u_k(x)e^{ik(x+a)}.$$

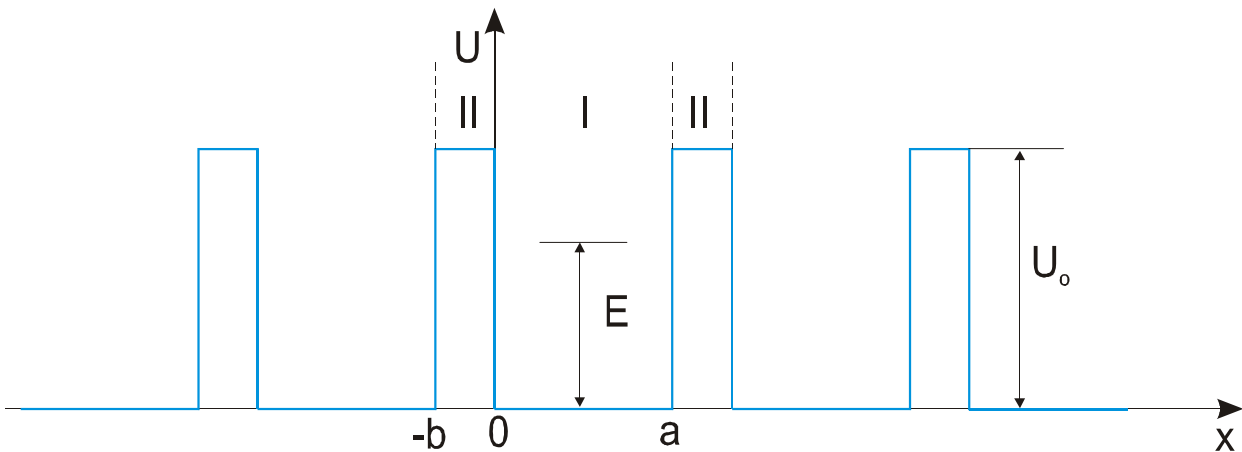
Prawe strony ostatnich dwóch równań są identyczne, jeżeli tylko $u_k(x+a) = u_k(x)$. Konkretna postać funkcji $u_k(x)$ zależy od funkcji $U(x)$. Dla każdej wartości wektora \vec{k} ma ona przy tym inny kształt, dlatego też oznaczamy ją symbolem $u_k(x)$. Funkcje (16.2) nazywamy funkcjami Blocha. Wynika z nich, że funkcja falowa elektronu jest iloczynem funkcji e^{ikx} reprezentującej falę płaską oraz funkcji charakteryzującej się okresowością sieci krystalicznej. Fala płaska odzwierciedla pewną swobodę poruszania się elektronu w kryształach, a czynnik okresowy spowodowany jest oddziaływaniem elektronu z atomami sieci. Na przykład, dla elektronów swobodnych $u(x)$ ma wartość stałą, a dla elektronów silnie związanych z atomem $u_k(x)$ jest zbliżone do funkcji falowej opisującej stan elektronu w atomie odosobnionym. Na rys. 16.2 pokazano przykładowy przebieg funkcji Blocha. Funkcja falowa elektronu w polu periodycznym ma więc postać fali płaskiej zmodulowanej czynnikiem $u_k(x)$ o okresowości sieci.



Rys. 16.2. Przykładowy przebieg jednowymiarowej funkcji Blocha: (a) funkcja modulowana, (b) funkcja modulująca, (c) wypadkowa funkcja falowa.

16.2. Model Kröniga-Penneya

W przybliżeniu elektronów słabo związanych zakładamy, że energia oddziaływania elektronu z siecią jest znacznie mniejsza od energii kinetycznej. Sytuacja taka ma miejsce w przypadku elektronów najslabiej związanych z atomami, czyli elektronów przewodnictwa w metalach. Pole elektryczne wytworzone przez jony jest kompensowane przez pole wszystkich pozostałych elektronów oprócz rozpatrywanego. Niewielkie, nieskompensowane pole ma charakter okresowy o okresie równym stałej sieci. Z okresowości tej wynika, że funkcja falowa opisująca stan elektronu będzie funkcją Blocha.



Rys. 16.3. Rozkład energii potencjalnej w modelu Kröniga-Penneya.

Aby opisać ruch elektronu w polu periodycznym, posłużymy się modelem Kröniga-Penneya. W modelu tym potencjał rzeczywisty w kryształach zastępujemy przez nieskończoną liczbę prostokątnych barier potencjalnych o szerokości b i wysokości U_0 . Studnia potencjału oddzielająca bariery ma szerokość a i energię U_0 (rys. 16.3). Potencjał ten odtwarza wprawdzie w niewielkim stopniu warunki występujące w kryształach, jednak uzyskane rozwiązanie umożliwia sformułowanie pewnych ogólnych wniosków dotyczących zależności dyspersyjnej dla elektronu w ośrodku periodycznym. Potencjał jest więc okresowy o okresie $d = a + b$. Zakładamy, że energia całkowita elektronu $E < U_0$. Równanie Schrödingera ma postać

$$\frac{d^2\Psi}{dx^2} + \frac{2m}{\hbar^2} E\Psi = 0, \quad \text{dla } 0 < x < a \quad (16.3)$$

$$\frac{d^2\Psi}{dx^2} - \frac{2m}{\hbar^2} (U_0 - E)\Psi = 0 \quad \text{dla } -b < x < 0. \quad (16.4)$$

Rozwiązaniem powyższych równań są funkcje

$$\Psi_1 = A_1 e^{i\alpha x} + B_1 e^{-i\alpha x}, \quad \text{dla } 0 < x < a \quad (16.5)$$

$$\Psi_2 = A_2 e^{\beta x} + B_2 e^{-\beta x}, \quad \text{dla } -b < x < 0, \quad (16.6)$$

gdzie

$$\alpha = \sqrt{\frac{2m}{\hbar^2} E}, \quad \beta = \sqrt{\frac{2m}{\hbar^2} (U_0 - E)}.$$

Wiemy już, że jeśli potencjał jest periodyczny, to funkcja $\Psi(x)$ spełnia warunek

$$\Psi(x) = u(x)e^{ikx},$$

skąd

$$u(x) = \Psi(x)e^{-ikx}.$$

Podstawiając do powyższego równania $\Psi(x)$ wyrażone wzorami (16.5) i (16.6), otrzymujemy

$$u_1(x) = A_1 e^{i(\alpha-kx)} + B_1 e^{-i(\alpha+kx)}, \quad (16.7)$$

$$u_2(x) = A_2 e^{(\beta-ik)x} + B_2 e^{-(\beta+ik)x}. \quad (16.8)$$

Warunek ciągłości funkcji $u(x)$ i ich pierwszych pochodnych prowadzi do związków

$$u_1(0) = u_2(0), \quad u_1(a) = u_2(-b)$$

$$\left(\frac{du_1}{dx}\right)_{x=0} = \left(\frac{du_2}{dx}\right)_{x=0}, \quad \left(\frac{du_1}{dx}\right)_{x=a} = \left(\frac{du_2}{dx}\right)_{x=-b}.$$

W ostatnich równaniach skorzystaliśmy z warunku, że funkcja $u(x)$ jest okresowa o okresie $a+b$.

Podstawiając wzory (16.7) i (16.8) do ostatnich równań, otrzymujemy układ czterech równań, w których niewiadome stanowią A_1, A_2, B_1, B_2

$$A_1 + B_1 = A_2 + B_2,$$

$$i(\alpha - k)A_1 - i(\alpha + k)B_1 = (\beta - ik)A_2 - (\beta + ik)B_2,$$

$$A_1 e^{i(\alpha-k)a} + B_1 e^{-i(\alpha+k)a} = A_2 e^{-(\beta-ik)b} + B_2 e^{(\beta+ik)b},$$

$$i(\alpha - k)A_1 e^{i(\alpha-k)a} - i(\alpha + k)B_1 e^{-i(\alpha+k)a} = (\beta - ik)A_2 e^{-(\beta-ik)b} - (\beta + ik)B_2 e^{(\beta+ik)b}.$$

Układ ten posiada niezerowe rozwiązanie tylko wtedy, gdy wyznacznik macierzy współczynników przy A_1, A_2, B_1, B_2 równa się zeru. Czytelnik zechce sprawdzić, że warunek ten prowadzi do równania

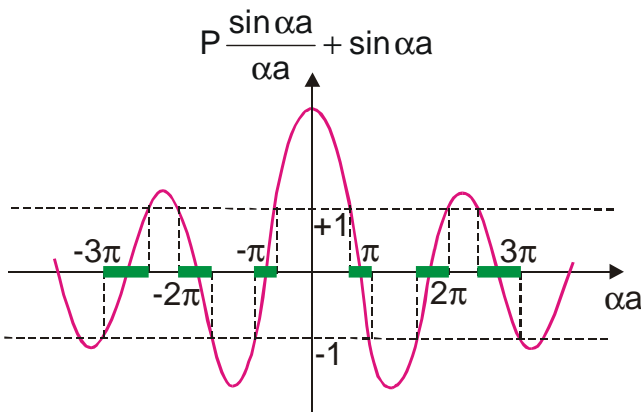
$$\frac{\beta^2 - \alpha^2}{2\alpha\beta} \sinh \beta b \sin \alpha a + \cosh \beta b \cos \alpha a = \cos k(a+b). \quad (16.9)$$

Równanie to nakłada warunki na α i β , a co z tym się wiąże i na dopuszczalne wartości energii. Analiza tego równania jest jednak dość uciążliwa. Dlatego zastosujemy uproszczenie Kröniga i Penneya polegające na założeniu, że U_0 dąży do nieskończoności, natomiast b dąży do zera w taki sposób, że iloczyn $U_0 b$ pozostaje skończony. Oznacza to, że $\beta^2 b$ jest skończone, ale $\beta b \rightarrow 0$. Po uproszczeniu równanie (16.9) ma postać

$$P \frac{\sin \alpha a}{\alpha a} + \cos \alpha a = \cos ka, \quad (16.10)$$

gdzie $P = U_0 m a b / \hbar^2$; przyjęto też, że $\sinh \beta b = \beta b$ i $\cosh \beta b = 1$. Wielkość P można traktować jako miarę trudności pokonywania bariery przez elektron.

Równanie (16.10) wyraża zależność energii elektronu (poprzez α) od liczby falowej k . Na rys. 16.4 przedstawiono rozwiązanie powyższego równania metodą graficzną. Ponieważ $\cos ka$ znajdujący się po prawej stronie równania jest zawarty w przedziale $[-1, 1]$, a więc dopuszczalne mogą być tylko te wartości αa , dla których wartości lewej strony równania (16.10) mieszczą się w przedziale $[-1, 1]$. Wartości αa spełniające ten warunek zaznaczono grubą linią. Widać, że dopuszczalne wartości energii tworzą pewne przedziały, tzw. pasma energetyczne. Poszczególne pasma energetyczne oddzielone są przez obszary o wzbronionych wartościach energii. W miarę wzrostu αa (a więc wartości energii, gdyż a jest ustalone) dozwolone pasma energetyczne stają się coraz szersze. Szerokość pasm zależy także od parametru P .



Rys. 16.4 Wielkość $P \sin \alpha a / \alpha a + \cos \alpha a$ w funkcji αa [patrz równanie (16.10)].

Ze wzrostem P pasma dozwolone zężają się; w szczególności gdy $P \rightarrow \infty$, równanie (16.10) ma rozwiązanie tylko wtedy gdy

$$\sin \alpha a = 0,$$

czyli $\alpha a = n\pi$, a $n = \pm 1, \pm 2, \pm 3, \dots$

Stąd

$$E = \frac{\pi^2 \hbar^2}{2ma^2} n^2.$$

Otrzymujemy więc poziomy dyskretny jak w odosobnionej studni potencjału.

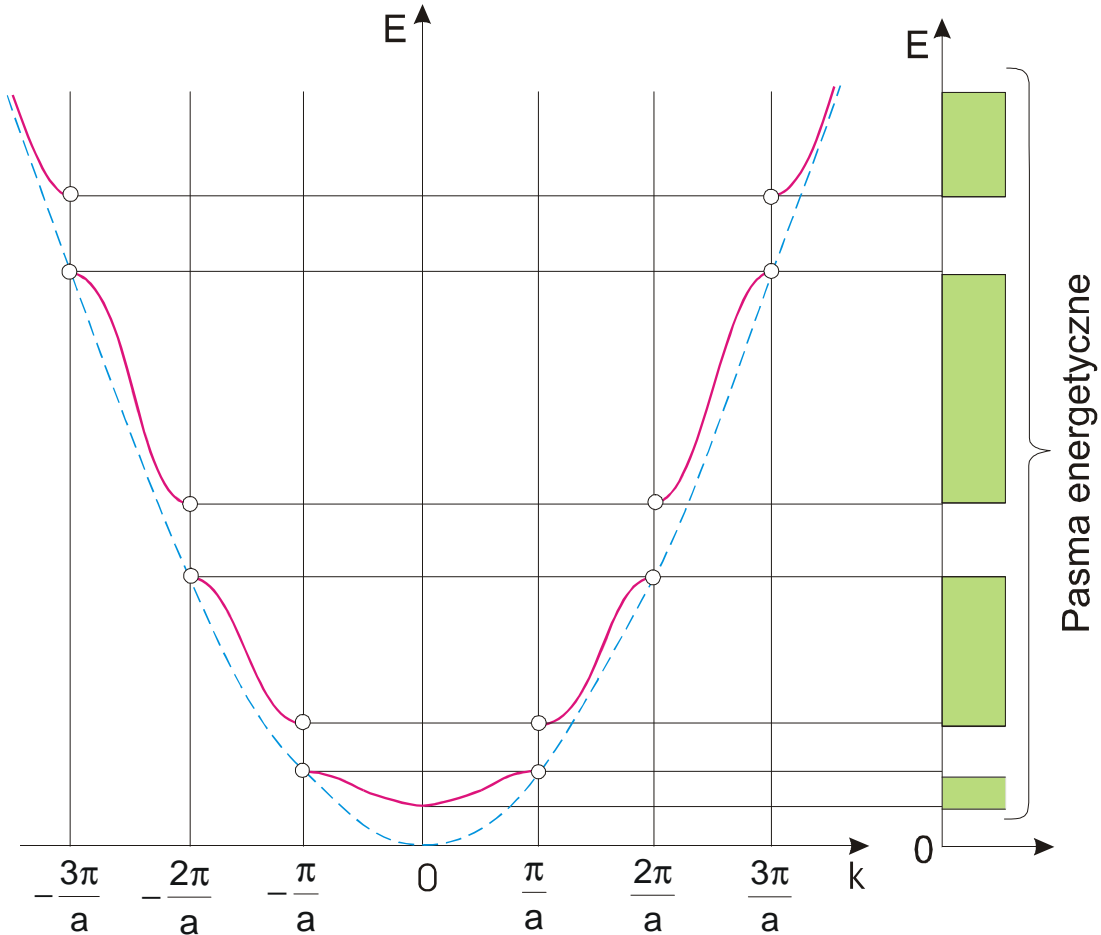
Powróćmy do równania (16.10). Na krańcach pasm energetycznych spełniony jest warunek:

$$\cos ka = \pm 1, \quad \text{czyli} \quad k = n \frac{\pi}{a}, \quad n = \pm 1, \pm 2, \pm 3, \dots$$

Wyrażając k przez długość fali $k = 2\pi/\lambda$ otrzymujemy

$$n\lambda = 2a. \quad (16.11)$$

Jest to wzór Wulfa-Bragga, wyrażający warunek odbicia od płaszczyzn sieciowych w przypadku kiedy kąt poślizgu wynosi 90° . Wynika stąd, że przerwy energetyczne dla elektronu w kryształach występują wówczas, gdy jest spełniony warunek braggowskiego odbicia fal elektronowych od płaszczyzn sieciowych kryształu. Fale o długości wyrażonej wzorem (16.11) ulegają całkowitemu wewnętrznemu odbiciu i nie mogą rozchodzić się w kryształach.

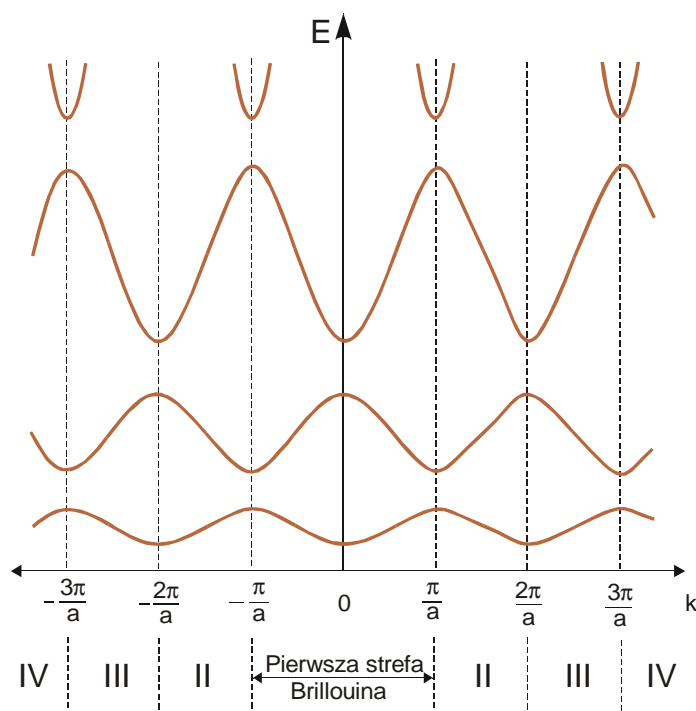


Rys.16.5. Zależność dyspersyjna energii E od liczby falowej k dla elektronu w modelu Kröniga-Penneya. Linia przerywana przedstawia zależność $E(k)$ dla elektronu swobodnego.

Na podstawie wzoru (16.10) można sporządzić wykres przedstawiający zależność $E(k)$ dla ogólnego przypadku (rys. 16.5). Linią przerywaną pokazano tu zależność dyspersyjną dla elektronu swobodnego. Jak widać, wewnątrz każdego pasma energetycznego funkcja $E(k)$ wzrasta monotonicznie, a na brzegach występują ekstrema i nieciągłości tej funkcji. Zauważmy, że na brzegach pasm $dE/dk = 0$, a zatem zgodnie ze wzorem

$$v_g = \frac{1}{\hbar} \frac{dE}{dk}, \quad (16.12)$$

prędkość grupowa jest równa zero. Fala elektronowa jest w tym przypadku falą stojącą, której nie towarzyszy transport energii. Fala ta powstaje w wyniku interferencji fali padającej i odbitej od płaszczyzn sieciowych kryształu.



Rys. 16.6. Pełny wykres zależności $E(k)$ dla modelu Kröniga-Penneya.

Każdej wartości energii elektronu odpowiada nieskończenie wiele wartości wektora falowego. Spośród nich, na rys. 16.5 uwzględniono tylko po jednej parze różniących się znakami; tzn. odpowiadającej przeciwnym kierunkom ruchu elektronu. Pełny wykres uwzględniający wszystkie wartości k (w obranym przedziale) przedstawiono na rys. 16.6. Jeżeli weźmiemy pod uwagę tylko jeden kierunek ruchu elektronu, to jak wynika z rys. 16.6, dowolnie obranej wartości energii elektronu należy przyporządkować następujący zbiór wartości wektora falowego

$$k_n = k + \frac{2\pi}{a}n, \quad (16.13)$$

gdzie $n = 0, \pm 1, \pm 2, \dots$. Jednocześnie możemy zauważyć, że danej wartości wektora falowego odpowiada wiele wartości energii – po jednej w każdym paśmie. Energia jest więc wieloznaczną funkcją wektora falowego k .

Periodyczność zależności dyspersyjnej $E(k)$ jest bezpośrednią konsekwencją symetrii translacyjnej kryształu. Wynika to z faktu, że stany elektronowe o wektorach falowych związanych zależnością (16.13) są sobie równoważne pod każdym względem fizycznym. Możemy się o tym przekonać przekształcając zależność (16.2) do następującej postaci

$$\Psi(x) = \left[u(x) \exp\left(-i \frac{2\pi n}{a} x\right) \right] \exp\left[i \left(k + \frac{2\pi n}{a}\right) x\right].$$

Ponieważ czynnik $\exp(-i2\pi nx/a)$ jest periodyczny z okresem a , zatem całe wyrażenie $u(x)\exp(-i2\pi nx/a)$ może być traktowane jako funkcja modulująca, odpowiadająca stanowi o wartości wektora falowego wyrażonej wzorem (16.13). Możemy więc zapisać

$$\Psi(x) = u_k(x) \exp\left[i\left(k + \frac{2\pi n}{a}x\right)\right]. \quad (16.14)$$

Jeżeli funkcja falowa o postaci $\Psi(x) = u_k(x) \exp(ikx)$ jest rozwiązaniem równania Schrödingera, to także funkcja (16.14) spełnia to równanie. Oba tym stanom odpowiada taki sam rozkład gęstości prawdopodobieństwa znalezienia elektronu, ponieważ

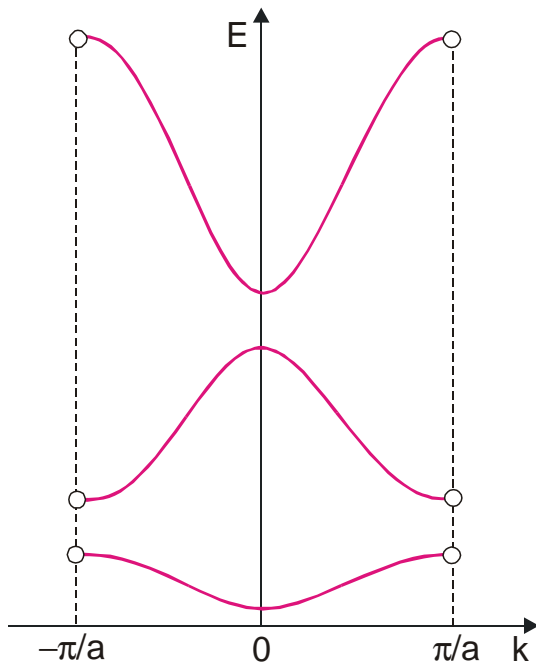
$$\left|u_k(x) \exp\left(-i\frac{2\pi n}{a}x\right)\right| = |u_k(x)|$$

i w obu stanach energia elektronu ma więc tę samą wartość. Funkcja $E(k)$ spełnia warunek

$$E\left(k + \frac{2\pi n}{a}\right) = E(k). \quad (16.15)$$

Periodyczność tej zależności umożliwia ograniczenie przedziału zmienności wektora k do pewnego obszaru, który dla sieci jednowymiarowej jest odcinkiem o długości $2\pi/a$. Środek tego odcinka dogodnie jest obrać w punkcie $k = 0$. Tak zdefiniowany obszar nosi nazwę pierwszej strefy Brillouina. Numeracje dalszych stref pokazano na rys. 16.6. Wykres $E(k)$ ograniczony do pierwszej strefy Brillouina pokazano na rys. 16.7. Wektory falowe należące do tej strefy, tzn. spełniające warunki

$$-\frac{\pi}{a} < k \leq \frac{\pi}{a}, \quad (16.16)$$

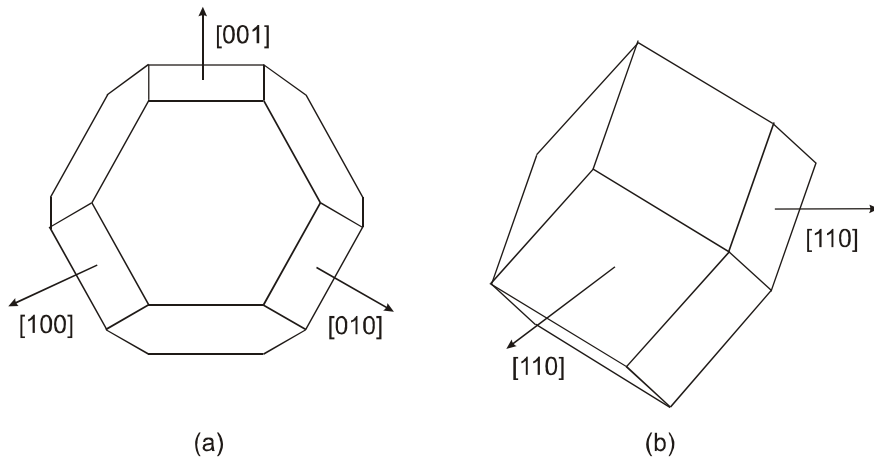


Rys. 16.7 Fragment wykresu zależności $E(k)$ ograniczony do pierwszej strefy Brillouina.

nazywamy zredukowanymi wektorami falowymi. W dalszych rozważaniach będziemy się posługiwać tylko takimi wektorami. Na rys. 16.7 punkty $k = -\pi/a$ i $k = +\pi/a$ są równoważne. Jeżeli więc wektor falowy zmienia się w sposób ciągły, np. wzrasta, co jak zobaczymy dalej zachodzi pod działaniem siły zewnętrznej, to w pewnej chwili dochodzi on do granicy sfery ($k = \pi/a$). Na wykresie przedstawionym na rys. 16.6 można dalej przesuwac się wzdłuż osi k , ale na mocy periodyczności funkcji $E(k)$ jest to równoważne przemieszczeniu się elektronu do punktu $k = -\pi/a$. Tym sposobem elektron kontynuuje swój ruch w obrębie pierwszej strefy Brillouina.

Wszystkie powyższe rozważania przeprowadziliśmy na przykładzie modelu jednowymiarowego. Ma on jednak bezpośredni związek z rzeczywistymi trójwymiarowymi

strukturami krystalicznymi, ponieważ można go uważać za reprezentację zbioru płaszczyzn sieciowych prostopadłych do danego kierunku. Wnioski wynikające z modelu jednowymiarowego mają więc zastosowanie również do stanów elektronowych związanych z wybranym kierunkiem w sieci trójwymiarowej.



Rys. 16.8. Kształt pierwszej strefy Brillouina dla sieci regularnej powierzchniowo (a) i przestrzennie (b) centrowanej.

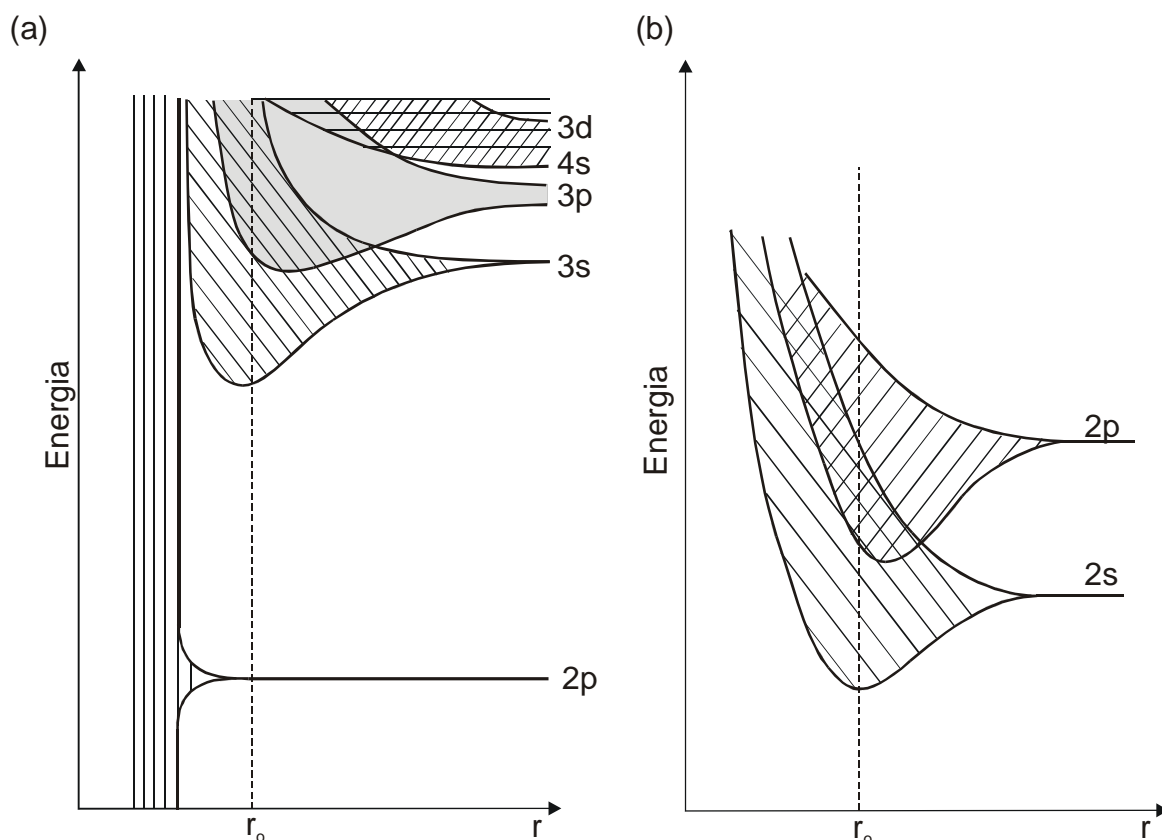
W przypadku sieci trójwymiarowej, strefy Brillouina są zamkniętymi powierzchniami których kształt zależy od struktury kryształu. Na rys. 16.8 pokazano kształt pierwszej strefy Brillouina dla sieci regularnej powierzchniowo i przestrzennie centrowanej.

16.3. Przykłady pasm energetycznych w kryształach

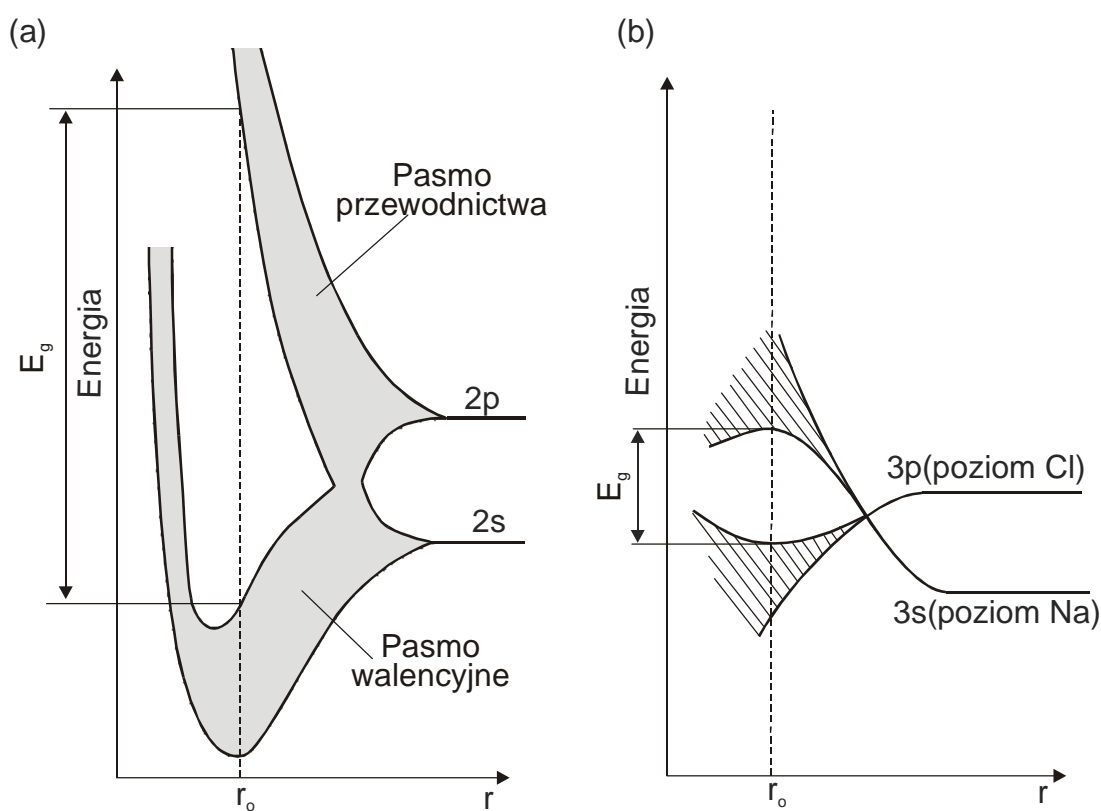
Poniżej omówimy tworzenie się pasm energetycznych w kryształach kilku pierwiastków i ich wypełnienie przez elektrony.

Na rys. 16.9a przedstawiono poziomy energetyczne elektronu w kryształach sodu w funkcji odległości r między jądrami. Odległość między jądrami sodu w sieci krystalicznej wynosi $r_0 = 3.67 \times 10^{-10}$ m. Przy tej odległości poziomy $1s$, $2s$ i $2p$ nie ulegają rozszczepieniu. Poziomy $3s$ jest pierwszym rozszczepionym poziomem energetycznym. Pasma $3s$ i $3p$ zachodzą na siebie. Zauważmy, że rozszczepieniu się poziomów energetycznych towarzyszy również obniżenie wartości energii tych poziomów. W atomie sodu na powłoce walencyjnej znajduje się jeden elektron w stanie $3s$. Zatem w kryształach sodu składających się z N atomów istnieje w stanie $3s$ $2N$ stanów energetycznych do obsadzenia, a liczba elektronów wynosi N . Poza tym, pasmo $3s$ zachodzi na całkowicie puste pasmo $3p$, w którym jest $6N$ stanów. Dlatego pasmo walencyjne jest tylko częściowo wypełnione.

Nieco inna jest sytuacja w kryształach zbudowanych z pierwiastków II grupy układu okresowego (beryl, magnez). W tym przypadku w stanie s są dwa elektrony i stan ten jest całkowicie wypełniony. Jednak w kryształach pasma s i p pokrywają się tak, że w efekcie powstaje pasmo częściowo wypełnione (rys. 16.9b).



Rys. 16.9. Pasma energetyczne w sodzie (a) i w berylu (b) w zależności od odległości między atomami.



Rys. 16.10. Pasma energetyczne w zależności od odległości między jądrami w: (a) diamencie, $r_o = 1,54 \times 10^{-10}$ m, $E_g = 5,6$ eV; (b) NaCl, $r_o = 2,82 \times 10^{-10}$ m, $E_g = 6$ eV;.

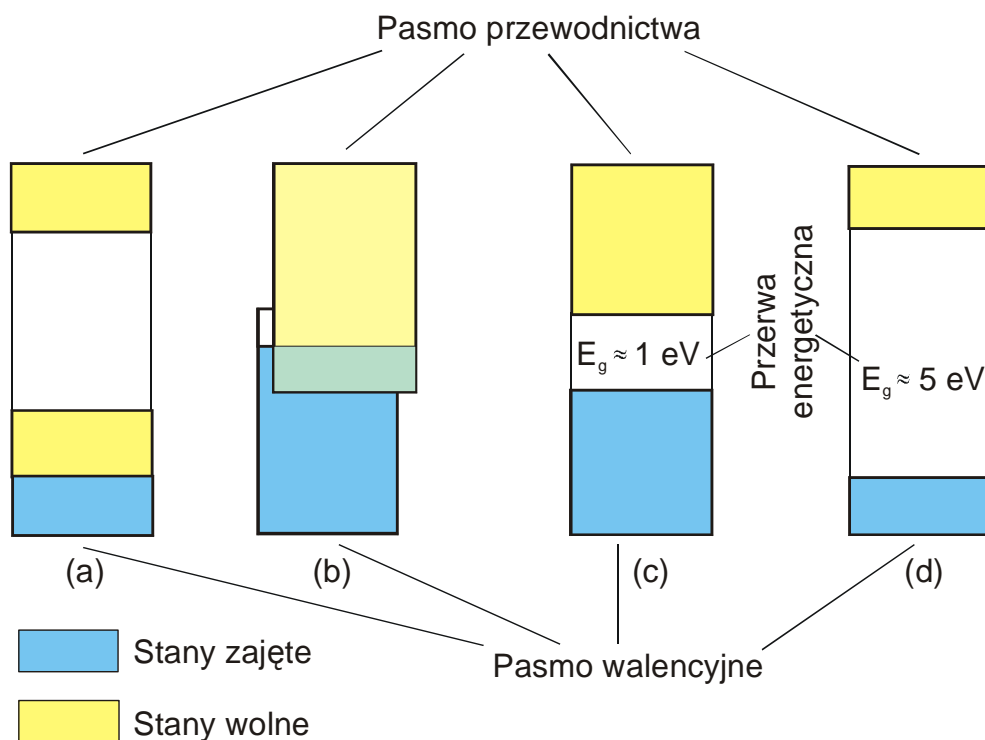
Na rys. 16.10a przedstawiono pasma energetyczne w kryształach diamentu, składającego się z atomów węgla o konfiguracji elektronów $1s^2 2s^2 2p^6$. Na poziomie $2p$ może znajdować się sześć elektronów. Podczas zbliżania się atomów węgla poziomy $2s$ i $2p$ ulegają rozszczepieniu na dwa oddzielne pasma, przy czym w paśmie $2s$ może być $2N$ elektronów, a w paśmie $2p$ – $6N$ elektronów. Dalsze zbliżenie się atomów powoduje nakrywanie się tych pasm. Jednak przy odległości między jądrami w kryształach diamentu równej $r_0 = 1,54 \times 10^{-10}$ m następuje ponowne rozszczepienie na dwa pasma. W każdym z nich znajduje się $2N$ podpoziomów, a liczba elektronów może wynosić $4N$. Cztery elektrony walencyjne zajmują dolne pasmo wypełniając je całkowicie, natomiast pasmo górne pozostaje całkowicie puste. Pasma te rozdzielone są przerwą energetyczną $E_g = 5,6$ eV. Podobnie przebiega tworzenie się pasm energetycznych w kryształach krzemu i germanu.

Jako przykład tworzenia się pasm energetycznych w kryształach jonowych rozpatrzmy NaCl. Przypomnijmy, że atomy Na i Cl mają następującą konfigurację elektronów: $1s^2 2s^2 2p^6 3s^1$ (Na) i $1s^2 2s^2 2p^6 3s^2 3p^5$ (Cl). W izolowanych atomach poziom $3s$ atomu sodu jest położony niżej od poziomu $3p$ atomu chloru. W wyniku zbliżenia atomów, elektrony walencyjne Na przechodzą do pasma $3p$ Cl i zapełniają wszystkie wolne stany. W ten sposób pasmo $3p$ chloru jest całkowicie zapełnione, natomiast pasmo $3s$ sodu – całkowicie puste. Jednocześnie pasmo $3p$ znajduje się poniżej pasma $3s$. Jest to spowodowane oddziaływaniem elektronów należących do jonów Na^+ z jonami Cl^- (wzrost energii) i elektronów należących do jonów Cl^- z jonami Na^+ (zmniejszenie energii). W rezultacie otrzymujemy strukturę pasmową przedstawioną na rys. 16.10b. Przerwa energetyczna kryształu NaCl wynosi 6 eV.

16.4. Przewodniki, półprzewodniki i izolatory

Wzajemne ułożenie pasm energetycznych i ich zapełnienie przez elektrony warunkuje właściwości elektryczne ciał stałych i jest podstawą ich podziału na metale, półprzewodniki i dielektryki.

W przypadku metali istnieje pasmo energetyczne zapełnione tylko częściowo (rys. 16.11a). Poniżej wszystkie pasma są zajęte, powyżej - wszystkie puste. Przykładem takiego ciała jest sód, w którym pasmo $3s$ jest zapełnione tylko do połowy. Podobna sytuacja występuje w innych metalach alkalicznych. Pasma zapełnione tylko częściowo może także powstać w wyniku nałożenia się pasm całkowicie zapełnionych z pasmami pustymi lub częściowo obsadzonymi. Sytuację taką mamy w kryształach zbudowanych z atomów należących do drugiej grupy układu okresowego. Jako przykład weźmy magnez w którym pasma $3s$ i $3p$ nakładają się tak, że część elektronów przechodzi z pasma $3s$ do $3p$ i oba pasma są zapełnione tylko częściowo (rys. 16.11b).



Rys. 16.11. Schemat pasm energetycznych w przypadku metali (a,b), półprzewodników (c) i dielektryków (d).

Wewnątrz częściowo zapełnionego pasma elektrony mogą łatwo przechodzić do nowych, niezapełnionych stanów. Elektrony mogą być zatem pobudzone zewnętrznym polem elektrycznym. Pojawia się wtedy w kryształach uporządkowany ruch elektronów, warunkujący powstanie prądu elektrycznego. Z tych powodów omawiane kryształy są dobrymi przewodnikami prądu.

Drugą grupę ciał stanowią półprzewodniki. Półprzewodnikami nazywamy takie ciała stałe w których w temperaturze 0 K pasmo walencyjne (i pasma niższe) są całkowicie zapełnione, a pasmo przewodnictwa całkowicie puste; przy czym przerwa energetyczna między pasmem walencyjnym a pasmem przewodnictwa jest niewielka – około 1 eV [rys. 16.11c]. Umownie półprzewodnikami przyjęto nazywać wszystkie ciała, których przerwa energetyczna E_g jest mniejsza od 3 eV. Wskutek wzbudzeń cieplnych, niektóre elektrony mogą uzyskać dostateczną energię do pokonania przerwy energetycznej i przejść do pasma przewodnictwa. Elektrony te znajdujące się w paśmie częściowo zapełnionym są elektronami swobodnymi. Im wyższa temperatura, tym większa jest koncentracja elektronów w paśmie przewodnictwa. W większości półprzewodników już w temperaturze pokojowej koncentracja elektronów jest dostatecznie duża i kryształ wykazuje znaczną przewodność elektryczną. Zwróćmy uwagę na to, że pole elektryczne nie jest w stanie nadać elektronom energii wystarczającej do pokonania przerwy energetycznej. Jeśli bowiem średnia droga swobodna elektronu wynosi około 10^{-8} m , to w polu elektrycznym o natężeniu 10^4 V/m uzyskuje on energię 10^{-4} eV . Jest to zdecydowanie zbyt mała energia aby elektron mógł przejść z jednego pasma do drugiego.

Jeżeli przerwa energetyczna jest duża (powyżej 3 eV) wzbudzenia elektronów praktycznie nie występują i ciało jest izolatorem (rys. 16.11d) Jedynie w bardzo silnym polu elektrycznym ($\approx 10^8$ V/m) może nastąpić przebicie izolatora i przerzucenie elektronów z pasma walencyjnego do pasma przewodnictwa.

16.5. Masa efektywna elektronu

Zajmiemy się teraz ruchem elektronów w kryształach. Ruch ten jest znacznie bardziej złożony niż ruch elektronu swobodnego, gdyż oprócz ewentualnej siły zewnętrznej F działającej na elektron, należy jeszcze uwzględnić oddziaływanie elektronu z siecią. Wiemy już, że dla elektronu w kryształach energia nie jest, jak to widać na rys. 16.5, proporcjonalna do k^2 , dlatego też wektora $\hbar\vec{k}$ nie można utożsamiać z pędem elektronu. Wielkość ta ma jednak pewne cechy pędu i w związku z tym nosi ona nazwę pseudopędu lub pędu krystalicznego. Nie będziemy jednak czynili tego rozróżnienia i będziemy posługiwali się nazwą "pęd".

Poniżej wykazemy, że dla opisu ruchu elektronu w kryształach wygodnie jest wprowadzić pojęcie masy efektywnej.

Pamiętamy, że prędkość elektronu v jest równa prędkości grupowej fali de Broglie'a (patrz pkt. 11.3)

$$v = v_g = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{dE}{dk}. \quad (16.17)$$

Różniczkując to wyrażenie względem czasu otrzymujemy przyśpieszenie

$$a = \frac{dv_g}{dt} = \frac{1}{\hbar} \frac{d^2E}{dkdt} = \frac{1}{\hbar} \frac{d^2E}{dk^2} \frac{dk}{dt}. \quad (16.18)$$

Jednocześnie siłę zewnętrzną F działającą na elektron możemy przedstawić w postaci

$$F = \frac{dp}{dt} = \hbar \frac{dk}{dt}. \quad (16.19)$$

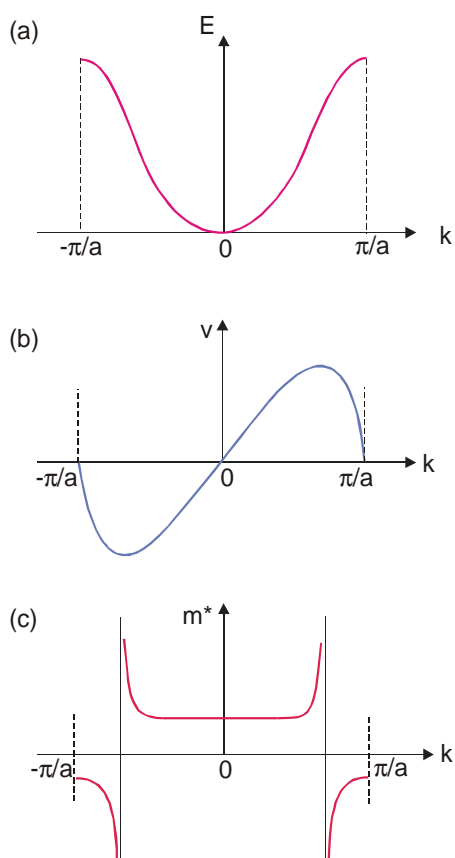
Wyliczając z tego wzoru dk/dt i podstawiając do wzoru (16.18) mamy

$$a = F \frac{1}{\hbar^2} \frac{d^2E}{dk^2}. \quad (16.20)$$

Wzór (16.20) wiąże przyśpieszenie a z siłą F działającą na elektron i wyraża drugą zasadę dynamiki Newtona dla elektronu w sieci krystalicznej. Ze wzoru (16.20) wynika jednocześnie, że elektronowi należy przypisać masę

$$m^* = \hbar^2 \left(\frac{d^2E}{dk^2} \right)^{-1}. \quad (16.21)$$

Masa m^* nosi nazwę masy efektywnej. Przypisując elektronowi znajdującemu się w okresowym polu kryształu masę m^* , możemy go traktować jako swobodny i opisywać jego ruch w polu zewnętrznym tak samo jak ruch zwyczajnego elektronu swobodnego. Oczywiście dla elektronu, dla którego $E = \hbar^2 k^2 / 2m$, wzór (16.21) daje $m^* = m$, czego należało się spodziewać. Należy podkreślić, że masa efektywna nie jest masą w zwykłym tego słowa znaczeniu. Jest tylko sposobem opisu zachowania się elektronu w kryształach.



Rys. 16.12. Zależność E , v i m^* od wektora falowego w pierwszej strefie Brillouina.

Omówionej tu zależności masy efektywnej od wektora falowego można nadać interpretację fizyczną podobną jak przy omawianiu zależności energii od wektora falowego. W środku strefy, gdzie $k = 0$, nie ma praktycznie odbicia fali de Broglie'a, ponieważ daleko jest do spełnienia warunku Bragga. Zewnętrzne pole elektryczne powoduje wzrost energii elektronu, jak to ma miejsce dla elektronu swobodnego. Dla większej wartości k odbicie zaczyna być zauważalne, co oznacza zmianę znaku pędu. W punkcie gdzie d^2E/dk^2 wynosi zero, wzrost pędu elektronu spowodowany działaniem siły zewnętrznej jest dokładnie kompensowany wzrostem ujemnego pędu na skutek zwiększonego udziału odbicia elektronu przez jony sieci. Wypadkowa zmiana pędu jest zatem równa zero, elektron zachowuje się tak, jak gdyby jego masa była nieskończenie wielka. Na granicy strefy masa efektywna jest ujemna, ponieważ zwiększenie udziału odbicia, związane ze zbliżaniem się do warunku całkowitego odbicia, znacznie bardziej zmienia pęd elektronu niż siła zewnętrzna.

Na rys. 16.12(a) przedstawiono zależność $E(k)$ w pierwszej strefie Brillouina. Pierwsza pochodna dE/dk , zgodnie ze wzorem (16.12), określa prędkość elektronu [rys. 16.12(b)], zaś druga pochodna d^2E/dk^2 określa wielkość masy efektywnej [rys. 16.12(c)]. Dla elektronu w środku strefy Brillouina, zależność $E(k)$ jest prawie taka sama jak dla elektronu swobodnego (zob. rys. 16.5) i dlatego $m^* \approx m$. Ze wzrostem k zależność $E(k)$ znacznie odbiega od zależności kwadratowej i masa efektywna staje się coraz większa. W punktach przegięcia krzywej $E(k)$ masa efektywna staje się nieskończenie wielka [rys. 16.12(c)]. Przy dalszym wzroście k prędkość elektronu maleje, a masa przyjmuje wartości ujemne. Fizycznie oznacza to, że dla dużych wartości k elektron zachowuje się tak jak dodatnio naładowana cząstka.

Zwróćmy też uwagę na to, że masa efektywna charakteryzuje pasmo energetyczne a nie elektron. Zależy bowiem od gęstości poziomów energetycznych w paśmie. Zauważymy, że gęstość poziomów energetycznych jest mała, jeżeli energia szybko rośnie ze wzrostem k . Szybkie zmiany $E(k)$ powodują, że druga pochodna d^2E/dk^2 jest duża, a masa efektywna mała. Innymi słowy, jeżeli rozszczepienie jakiegoś poziomu energetycznego jest duże i pasmo jest szerokie (mała gęstość poziomów), to masa efektywna jest niewielka. Odwrotnie, masa efektywna jest duża, gdy pasmo energetyczne jest wąskie.

Masa efektywna jest na ogół wielkością anizotropową, tzn. w różnych kierunkach ma różne wartości. Jest to spowodowane anizotropią struktury pasmowej kryształów.

16.6. Pojęcie dziury

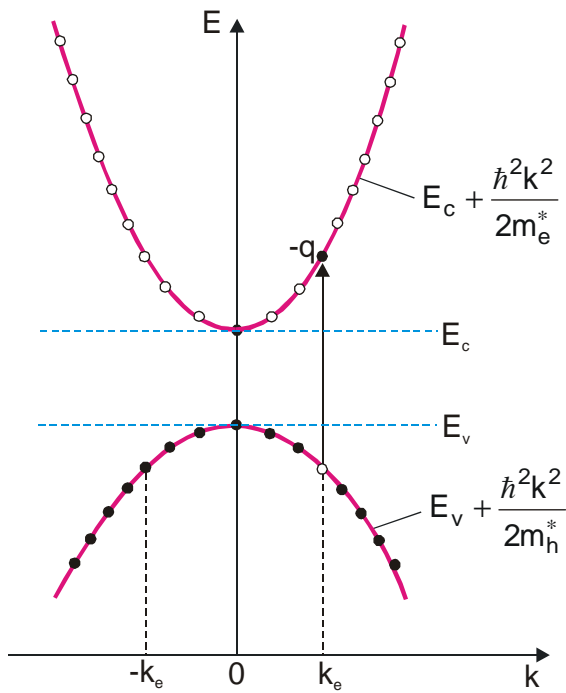
Rozpatrzmy teraz pewien fragment rys. 16.7. Najwyższe pokazane tam pasmo jest pasmem przewodnictwa półprzewodnika, pasmo znajdujące się bezpośrednio pod nim jest zatem pasmem walencyjnym. Niższe pasma nie będą nas interesowały, ponieważ w warunkach normalnych są one całkowicie zajęte i nie mają wpływu na właściwości elektryczne i optyczne półprzewodnika. W temperaturze 0 K, jak już wyżej wyjaśnialiśmy, pasmo walencyjne jest całkowicie obsadzone przez elektrony, a pasmo przewodnictwa – puste. W wyższych temperaturach część elektronów z pasma walencyjnego zostaje wzbudzona do pasma przewodnictwa. Elektrony obsadzają poziomy znajdujące się w pobliżu dna tego pasma. Z tego powodu w paśmie przewodnictwa interesuje nas wyłącznie jego dolny odcinek. Rozwijając funkcję $E(k)$ w szereg potęgowy i uwzględniając jej parzystość, można dla interesującego nas odcinka pasma napisać

$$E(k) = E_c + \left(\frac{d^2E}{dk^2} \right)_{k=0} k^2 + \dots, \quad (16.22)$$

gdzie E_c oznacza energię dna pasma przewodnictwa. Jeżeli ograniczymy się do rozwinięcia dwóch pierwszych wyrazów, to zależność energii od wektora falowego przyjmuje postać formy kwadratowej. Pasma, dla których przybliżenie takie jest dopuszczalne nazywamy pasmami parabolicznymi. Dla takich pasm można zależność (16.22) przedstawić w postaci analogicznej do odpowiedniej zależności dla elektronu swobodnego

$$E(k) = E_c + \frac{\hbar^2 k^2}{2m_e^*}, \quad (16.23)$$

gdzie m_e^* oznacza masę efektywną elektronu w paśmie przewodnictwa. Jest ona zdefiniowana wzorem (16.21) i jak widać, jest dodatnia. Dla pasm parabolicznych masa efektywna jest wielkością stałą, niezależną od energii elektronu.



Rys. 16.13. Wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa.

Przeanalizujemy teraz sytuację występującą w paśmie walencyjnym. Załóżmy, że pod wpływem jakiegoś czynnika zewnętrznego (np. fotonu) jeden elektron z całkowicie wypełnionego pasma walencyjnego został wzbudzony do pasma przewodnictwa, przy czym wartość jego wektora falowego nie uległa zmianie. Przejście takie przedstawione na rys. 16.13, nosi nazwę przejścia prostego (przejścia, przy których zachodzi zmiana wektora falowego, nazywamy przejściami skośnymi). W wyniku zderzeń z niedoskonałościami sieci krystalicznej, rozpatrywany elektron będzie wytracał swą energię kinetyczną, przechodząc na coraz niżej położone poziomy energetyczne osiągając energię minimalną w punkcie $k = 0$.

Rozważając zachowanie się elektronów w prawie całkowicie wypełnionym paśmie walencyjnym, mamy do czynienia z zagadnieniem bardziej skomplikowanym, ponieważ liczba elektronów w tym paśmie jest ogromna. Na domiar złego, część z nich ma masy efektywne ujemne, część dodatnie, a dla niektórych jest ona nieskończenie wielka. Okazuje się jednak, że pasmo, w którym wszystkie stany kwantowe z wyjątkiem jednego są obsadzone przez elektrony, można potraktować jako obsadzone przez pewną quasi-cząstkę, którą nazywamy dziurą. Tej niby-cząstce przypiszemy określony ładunek elektryczny, masę efektywną i energię.

W celu wyznaczenia tych parametrów założmy, że liczba wszystkich stanów kwantowych w paśmie walencyjnym rozpatrywanego kryształu wynosi N . Spośród nich $N-1$ stanów jest obsadzonych przez elektrony. Załóżmy dalej, że elektron wzbudzony do pasma przewodnictwa miał przed wzbudzeniem następujące wartości parametrów dynamicznych: masę efektywną $m^* < 0$, wektor falowy k_e , prędkość v_e i energię E_e .

Jeżeli dziura ma reprezentować cały zespół elektronów w prawie całkowicie wypełnionym paśmie, to musimy przypisać jej pęd równy sumie pędów tych elektronów. Oznaczając wektor falowy dziury symbolem k_h możemy napisać

$$\hbar k_h = \sum_{i=1}^{N-1} \hbar k_i = \sum_{i=1}^N \hbar k_i - \hbar k_e. \quad (16.24)$$

Do sumy włączyliśmy tu pęd brakującego elektronu, a następnie od całości odjęliśmy ten pęd. Pierwszy składnik w ostatnim wzorze jest równy zero, ponieważ z parzystości $E(k)$ wynika, że dla każdego zbioru o wektorze falowym k , istnieje stan symetryczny o wektorze $-k$. Wpływa

stąd wniosek, że wektor falowy dziury jest co do wartości równy wektorowi falowemu wzbudzonego elektronu, ma natomiast w stosunku do niego przeciwny znak.

$$k_h = -k_e. \quad (16.25)$$

Widać to bezpośrednio na rys. 16.13. Ponieważ przyczynki wnoszone do sumy (16.24) przez elektrony znajdujące się w symetrycznie położonych stanach znoszą się wzajemnie, wypadkowy pęd wszystkich elektronów w paśmie walencyjnym jest równy pędowi jedyne występującego w tym paśmie "elektronu bez pary".

Podobne rozważania można przeprowadzić odnośnie energii. Analogicznie do wzoru (16.24) możemy więc napisać

$$E_h(k) = \sum_{i=1}^N E_i - E_e(k). \quad (16.26)$$

Rozróżnienie między k_h i k_e jest zbędne w tym przypadku, ponieważ energia nie zależy od zwrotu wektora k , tylko od k^2 . Pierwszy składnik we wzorze (16.26) jest wielkością stałą, jest to bowiem energia elektronów w całkowicie zapełnionym paśmie. Składnik ten można pominąć, ponieważ poziom względem którego jest mierzona energia, może być ustalony w sposób dowolny. Ze wzoru (16.26) wynika, że im większa jest energia nie obsadzonego stanu przez elektron w paśmie walencyjnym, tym jest mniejsza energia dziury. Wobec tego wykres przedstawiający zależność dyspersyjną $E_h(k_h)$ dla dziur otrzymamy zatem z odpowiedniego wykresu dla elektronów w paśmie walencyjnym zmieniając zwrot wektora falowego i kierunek osi energii (rys. 16.13). Energia dziury jest tym większa, im niżej na rysunku znajduje się obsadzony przez nią stan kwantowy.

Całkowicie zapełnione pasmo walencyjne można traktować, operując pojęciem dziur, jako pasma puste, tzn. nie zawierające dziur, a wyrwanie elektronu z tego pasma – jako kreację dziury. Wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa nazywamy zatem generacją pary elektron-dziura.

Przy analizie zjawisk zachodzących w półprzewodnikach, dziury w paśmie walencyjnym traktuje się analogicznie jak elektrony w paśmie przewodnictwa. Uważa się je więc za nośniki ładunku, których właściwości dynamiczne są reprezentowane przez masę efektywną.

Z kształtu funkcji $E_h(k_h)$ wynika, że masa efektywna dziury jest dodatnia i równa, co do wartości bezwzględnej, masie efektywnej elektronu w paśmie walencyjnym, tzn.

$$m_h^*(k) = -m_e^*(k). \quad (16.27)$$

Ładunek elektryczny dziur jest także dodatni, co wynika z następującego rozumowania. Dla brakującego elektronu w paśmie walencyjnym możemy napisać

$$m_e^* v_e = \hbar k_e, \quad (16.28)$$

zaś dla dziury

$$m_h^* v_h = \hbar k_h. \quad (16.29)$$

Opierając się na wzorach (16.26) i (16.27) dochodzimy do wniosku, że dziura porusza się z prędkością v_h równą co do wartości i kierunku prędkości v_e elektronu, który znajdowałby się w nie obsadzonym stanie w paśmie walencyjnym. Ponieważ prąd elektryczny wytwarzany przez elektrony całkowicie zapełnionego pasma jest równy zeru, zatem usunięcie z tego pasma jednego nośnika o ładunku $-e$ poruszającego się z prędkością v_e jest równoważne wprowadzeniu do tego pasma nośnika o ładunku $+e$ poruszającego się z tą samą prędkością.

Analogicznie do wzoru (16.23) energia elektronu dla pasma walencyjnego ma postać

$$E(k) = E_v - \frac{\hbar^2 k^2}{2m_h^*},$$

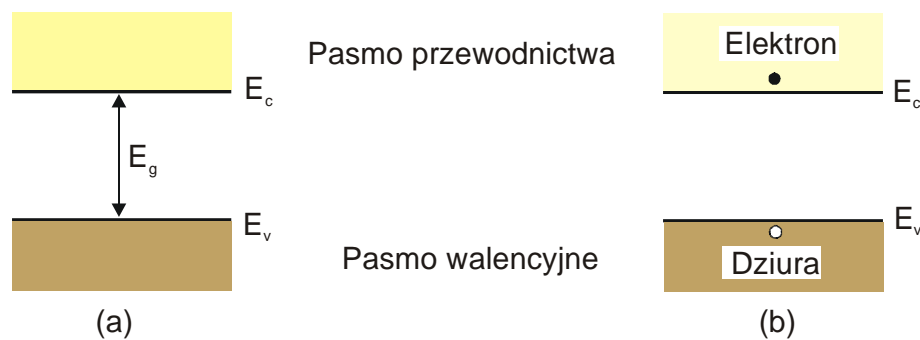
gdzie m_h^* jest masą efektywną dziur w paśmie walencyjnym (indeks h wiąże się z angielską nazwą dziury – *hole*). Energię dziury wyrażamy natomiast wzorem

$$E_h(k) = E_v + \frac{\hbar^2 k^2}{2m_h^*}. \quad (16.30)$$

Zarówno we wzorze (16.23) i (16.30) pierwsze składniki reprezentują energię potencjalną, drugie zaś energię kinetyczną.

Tak więc dziura zachowuje się tak jak "normalna" cząstka, ma bowiem dodatnią masę efektywną. Podobnie jak elektron w paśmie przewodnictwa, dąży do tego aby zająć stan o najniższej energii. Należy zwrócić uwagę, że pojęcie dziury wprowadziliśmy tylko w przypadku pasm całkowicie zapełnionych.

Przy rozpatrywaniu działania większości elementów półprzewodnikowych nie zachodzi konieczność posługiwania się zależnościami dyspersyjnymi przedstawionymi na rys. 16.13. Do tego celu wystarcza najczęściej uproszczony model pasmowy półprzewodnika pokazany na rys. 16.14. W takim modelu można również pokazać wzbudzenie elektronu z pasma walencyjnego do pasma przewodnictwa.



Rys. 16.14. Generacja pary elektron-dziura: (a) stan obsadzenia pasma przed generacją, (b) stan obsadzenia po generacji.

W literaturze operuje się często modelem dziury, w którym jest ona traktowana jako zerowe wiązanie przy określonym atomie sieci krystalicznej. Wskutek wytworzonego "wolnego miejsca" staje się możliwy ruch elektronów tworzących wiązanie międzyatomowe w kryształ. Ruch dziury jest w tym modelu traktowany jako "przeskakiwanie" zerowego wiązania od atomu do atomu w wyniku przemieszczania się elektronów walencyjnych w przeciwnym kierunku. Model ten jest bardzo pogładowy, zawiera jednak pewne nieścisłości – przypisuje bowiem dziurze określoną lokalizację, co jest niezgodne z zasadą nieoznaczoności Heisenberga. "Przeskakiwanie" na wolne miejsca zachodzi w przestrzeni wektora falowego, a nie w przestrzeni rzeczywistej.

ROZDZIAŁ 17

PODSTAWY FIZYKI PÓŁPRZEWODNIKÓW

Jak już poprzednio zaznaczyliśmy, półprzewodnikami nazywamy materiały, które w temperaturze zera bezwzględnego mają całkowicie obsadzone pasmo walencyjne i całkowicie puste pasmo przewodnictwa, a szerokość pasma zabronionego nie przekracza 3 eV. Przy takiej wartości przerwy energetycznej, pod wpływem wzbudzeń termicznych część elektronów uzyskuje wystarczającą energię do pokonania przerwy energetycznej i przechodzi do pasma przewodnictwa, gdzie stają się swobodnymi nośnikami prądu. Jednocześnie w paśmie walencyjnym powstają dziury. W półprzewodnikach występują więc dwa rodzaje nośników: elektrony i dziury, stąd przy przepływie prądu mówimy o przewodnictwie elektronowym i dziurowym. Konduktywność materiałów półprzewodnikowych zmienia się w przedziale od 10^{-8} do $10^6 (\Omega\text{cm})^{-1}$.

Grupa \Rightarrow Okres \Downarrow	II	III	IV	V	VI	VII
II		B 1.1	C 5.2			
III			Si 1.1	P 1.5	S 2.5	
IV			Ge 0.7	As 1.2	Se 1.7	
V			Sn 0.08	Sb 0.12	Te 0.36	J 1.25
VI						

Rys. 17.1. Rozmieszczenie półprzewodników w układzie okresowym pierwiastków. Obok symbolu pierwiastka podano wartość przerwy energetycznej w eV.

Do półprzewodników zaliczamy 12 pierwiastków (B, C, Si, P, S, Ge, As, Se, Sn, Sb, Te, J), których rozmieszczenie w układzie okresowym Mendelejewa przedstawiono na rys. 17.1 podając jednocześnie szerokości przerw energetycznych E_g . Zauważmy, że szerokość przerwy energetycznej wykazuje określoną prawidłowość. Wartość E_g wzrasta w każdym okresie przy przejściu do pierwiastka o większej liczbie atomowej Z . Maleje natomiast w każdej grupie przy przejściu do pierwiastków o większym Z . Jest to uwarunkowane strukturą zewnętrznych powłok elektronowych atomów. Na przykład, przechodząc w danej grupie z góry na dół spotykamy atomy, które mają wypełnione powłoki o większej wartości głównej liczby kwantowej n .

Jednocześnie ze wzrostem n występuje większe oddziaływanie sąsiednich atomów, co w efekcie daje szerokie pasmo powstałe z rozszczepienia takiego poziomu. Odpowiada to oczywiście mniejszej przerwie energetycznej.

Spośród pierwiastków największe znaczenie jako materiał półprzewodnikowy ma krzem. Półprzewodnikami są także liczne związki chemiczne. Do takich należą związki pierwiastków grupy III i V, oznaczone symbolem $A^{III}B^V$, np. GaAs, GaP, GaN; związki typu $A^{II}B^{VI}$ – CdS, ZnO i wiele innych. Kryształy mieszane zazwyczaj zbudowane z atomów dwóch lub trzech pierwiastków (lub cząstek związków chemicznych) są również półprzewodnikami, np. Ge_xSi_{1-x} $Hg_{1-x}Cd_xTe$ (wskaźnik x oznacza skład molowy). W Tabeli 17.1 podano przykładowo wartości przerw energetycznych niektórych półprzewodników. Do półprzewodników należą także liczne związki organiczne.

Tabela 17.1. Wartości szerokości przerw energetycznych niektórych półprzewodników

Półprzewodnik	E_g w 300 K (eV)
Pierwiastki	
Si	1,12
Ge	0,67
Diament	5,30
α -Sn	0,08
Związki $A^{III}B^V$	
GaAs	1,35
GaP	2,24
InAs	0,36
InSb	0,18
GaN	3,39
Związki $A^{II}B^{VI}$	
CdS	2,42
CdSe	1,73
CdTe	1,50
ZnTe	2,25
$Hg_{1-x}Cd_xTe$	0–1,5

Ze względu na właściwości fizyczne półprzewodniki dzieli się na samoistne i domieszkowane.

17.1. Półprzewodniki samoistne

W półprzewodnikach samoistnych w warunkach równowagi termodynamicznej, elektrony w paśmie przewodnictwa pojawiają się wyłącznie wskutek wzbudzenia z pasma walencyjnego. Skoro zatem wszystkie elektrony w paśmie przewodnictwa pochodzą z pasma walencyjnego, w półprzewodniku samoistnym koncentracje elektronów i dziur są jednakowe i równe koncentracji samoistnej n_i

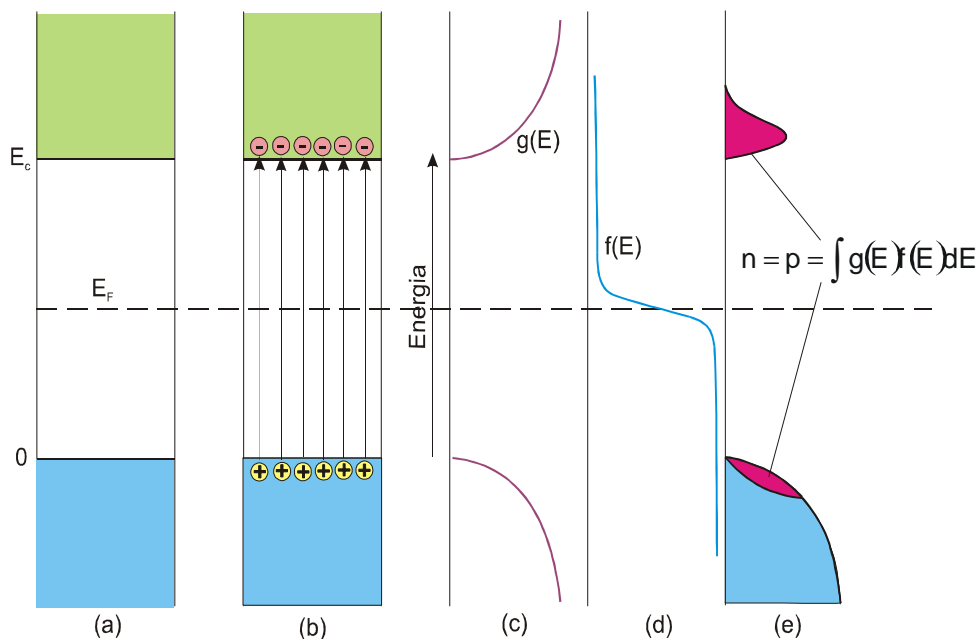
$$n = p = n_i \quad (17.1)$$

Idealny półprzewodnik samoistny charakteryzuje się doskonałą strukturą sieci krystalicznej, nie zawiera więc obcych atomów (zanieczyszczeń i domieszek) ani też innych defektów strukturalnych.

W większości półprzewodników koncentracja elektronów przewodnictwa jest mała w porównaniu z liczbą stanów w paśmie przewodnictwa i dlatego funkcja rozkładu Fermiego-Diraca $f(E)$ [zob. wzór (15.28)] jest znacznie mniejsza od jedności. Możemy wówczas zapisać dla elektronów

$$f_e(E) = \frac{1}{\exp\left[\frac{E - E_F}{kT}\right] + 1} \approx e^{-(E - E_F)/kT}. \quad (17.2)$$

Funkcja rozkładu Fermiego-Diraca została zastąpiona przez funkcję rozkładu Boltzmann. Oznacza to, że gaz elektronowy w półprzewodnikach jest gazem niezwyrodniałym. Mówiąc inaczej, warunek (17.2) będzie spełniony tylko wówczas, jeżeli $(E - E_F)$ jest znacznie większe od kT .



Rys. 17.2. Schematyczne przedstawienie modelu półprzewodnika samoistnego. Pasma walencyjne i pasmo przewodnictwa w 0 K (a) i $T > 0$ (b). Funkcje gęstości stanów (c), funkcje rozkładu (d) oraz koncentracja nośników (e).

Koncentracja elektronów w przedziale energii od E do $E+dE$ dana jest zależnością (zob. pkt 15.6.3)

$$dn = g_e(E)f_e(E)dE, \quad (17.3)$$

gdzie $g_e(E)$ jest gęstością stanów dla elektronów. Całkowitą liczbę elektronów w jednostce objętości otrzymujemy całkując wyrażenie (17.3) po całej szerokości pasma przewodnictwa.

Górna granica może być przesunięta do nieskończoności, gdyż funkcja wykładnicza (17.2) szybko maleje ze wzrostem E . Fizycznie oznacza to, że prawdopodobieństwo obsadzenia przez elektrony najwyższych poziomów jest praktycznie równe zero. Należy jeszcze zauważyć, że funkcja gęstości stanów w tym przypadku jest proporcjonalna do $(E - E_c)^{1/2}$, gdyż najmniejsza wartość energii elektronu w paśmie przewodnictwa wynosi E_c (rys. 17.2). Po uwzględnieniu powyższych uwag możemy napisać:

$$n = \frac{4\pi(2m_e^*)^{3/2}}{h^3} \int_{E_c}^{\infty} (E - E_c)^{1/2} e^{-(E-E_F)/kT} dE, \quad (17.4)$$

gdzie m_e^* jest masą efektywną elektronu. Podstawiając $x = [(E - E_c)/kT]^{1/2}$, całkę w relacji (17.4) można sprowadzić do całki

$$\int_0^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{4}.$$

Wykonując proste przekształcenia otrzymujemy

$$n = \frac{2}{h^3} (2\pi m_e^* kT)^{3/2} e^{-(E_c - E_F)/kT} = N_c e^{-(E_c - E_F)/kT}. \quad (17.5)$$

Wielkość

$$N_c = \frac{2}{h^3} (2\pi m_e^* kT)^{3/2}, \quad (17.6)$$

nazywamy efektywną gęstością stanów w paśmie przewodnictwa. Podstawiając stałe występujące we wzorze (17.6) mamy:

$$N_c = \frac{2}{h^3} (2\pi m_e^* kT)^{3/2} \left(\frac{m_e^*}{m}\right)^{3/2} T^{3/2} = 4.81 \times 10^{21} \left(\frac{m_e^*}{m}\right)^{3/2} T^{3/2}, \quad [\text{m}^{-3}] \quad (17.7)$$

gdzie m oznacza masę elektronu swobodnego.

Obliczymy teraz koncentrację dziur w paśmie walencyjnym. Zauważmy, że funkcja rozkładu dla dziur $f_h(E)$ wyraża się wzorem

$$f_h(E) = 1 - f_e(E) = \frac{1}{\exp\left[\frac{(E_F - E)}{kT}\right] + 1}, \quad (17.8)$$

gdź prawdopodobieństwo zajęcia określonego stanu przez dziurę jest równoważne prawdopodobieństwu nie zajmowania tego stanu przez elektron. Podobnie jak w przypadku elektronów

$$f_h(E) = \frac{1}{\exp\left[\frac{(E_F - E)}{kT}\right] + 1} = e^{-(E_F - E)/kT}. \quad (17.9)$$

Przyjmując górną granicę całkowania równą zero, dolną – minus nieskończoność, otrzymujemy wyrażenie na koncentrację dziur w paśmie walencyjnym

$$p = \int_{-\infty}^0 g_h(E) f_h(E) dE = \frac{4\pi(2m_h^*)^{3/2}}{h^3} \int_{-\infty}^0 (-E)^{1/2} e^{-(E_F - E)/kT} dE, \quad (17.10)$$

gdzie m_h^* jest masą efektywną dziury. Przypomnijmy, że począwszy od wierzchołka pasma walencyjnego w dół, energia przyjmuje wartości ujemne (rys. 17.2). Postępując podobnie jak przy wyprowadzeniu wyrażenia (17.5) otrzymujemy:

$$p = \frac{2}{h^3} (2\pi m_h^* kT)^{3/2} e^{-E_F/kT} = N_v e^{-E_F/kT}. \quad (17.11)$$

Wielkość N_v jest efektywną gęstością stanów w paśmie walencyjnym. Liczbowo N_v wynosi

$$N_v = \frac{2}{h^3} (2\pi m k)^{3/2} \left(\frac{m_h^*}{m}\right)^{3/2} T^{3/2} = 4.81 \times 10^{21} \left(\frac{m_h^*}{m}\right)^{3/2} T^{3/2} \quad [\text{m}^{-3}] \quad (17.12)$$

Mnożąc przez siebie wyrażenia na koncentrację elektronów (17.5) i dziur (17.11) mamy

$$np = N_v N_c e^{-E_c/kT}.$$

Przyjmując, że wierzchołek pasma walencyjnego ma energię równą zero, E_c jest równe szerokości przerwy energetycznej E_g . Wobec tego

$$np = N_v N_c e^{-E_g/kT}. \quad (17.13)$$

Warto zwrócić uwagę, że wyrażenie to nie zawiera poziomu Fermiego i dla określonego półprzewodnika jest funkcją tylko temperatury. Iloczyn np jest taki sam dla półprzewodnika samoistnego jak i domieszkowanego. Ponieważ dla półprzewodnika samoistnego $n = p = n_i$, pierwiastkując wzór (17.13) otrzymujemy ostatecznie wyrażenie na koncentrację nośników w półprzewodniku samoistnym

$$n_i = \frac{2(2\pi kT)^{3/2}}{h^3} (m_e^* m_h^*)^{3/4} e^{-E_g/2kT}. \quad (17.14)$$

Tak więc w półprzewodnikach samoistnych koncentracja dziur bardzo szybko wzrasta ze wzrostem temperatury i maleje ze wzrostem szerokości przerwy energetycznej. Jest to wynik oczywisty. Im temperatura jest wyższa, tym większa jest energia ruchu termicznego i wzrasta prawdopodobieństwo przejścia elektronów z pasma walencyjnego do pasma przewodnictwa. Z kolei wzrost przerwy energetycznej powoduje, że w ustalonej temperaturze mniej elektronów ma energię wystarczającą do pokonania tej przerwy. W tabeli 17.1 podano podstawowe parametry dla kilku najważniejszych półprzewodników.

Tabela 17.2. Podstawowe parametry niektórych półprzewodników w T = 300 K

Parametr	Si	Ge	GaAs	InSb
Koncentracja atomów lub molekuł [m^{-3}]	5×10^{28}	$4,4 \times 10^{28}$	$2,2 \times 10^{28}$	
Stała sieci krystalicznej [nm]	0,543	0,566	0,565	0,648
Gęstość [kg/m^3]	2330	5330	5320	5775
Temperatura topnienia [$^{\circ}\text{C}$]	1420	957	1238	530
Szerokość przerwy energetycznej [eV]	1,12	0,67	1,43	0,18
Efektywna gęstość stanów:				
N_c [m^{-3}]	$2,8 \times 10^{25}$	$1,0 \times 10^{25}$	$4,7 \times 10^{23}$	
N_v [m^{-3}]	$1,0 \times 10^{25}$	$6,0 \times 10^{24}$	$7,0 \times 10^{24}$	
Koncentracja samoistna [m^{-3}]	$1,45 \times 10^{16}$	$2,4 \times 10^{19}$	10^{13}	$1,9 \times 10^{22}$
Ruchliwość nośników półprzewodnika samoistnego [m^2/Vs]:				
elektronów	0,13	0,39	0,86	8,0
dziur	0,05	0,19	0,025	0,08
Względna przenikalność dielektryczna	11,9	16,0	13,1	12

Warto pamiętać, że w temperaturze pokojowej koncentracja nośników jest bardzo mała w porównaniu z koncentracją atomów. Na przykład, w krzemie w $T = 300 \text{ K}$ na jeden swobodny nośnik przypada 3×10^{12} atomów. Dla porównania, w metalach koncentracje swobodnych elektronów i koncentracje atomów są porównywalne.

W tabeli 17.3 podano zależności n_i od T dla Si, Ge i GaAs. Wynika z niej, że koncentracja samoistna silnie wzrasta ze wzrostem temperatury, przy czym zmiany te są większe dla półprzewodników o większej przerwie energetycznej.

Tabela 17.3. Zależność koncentracji samoistnej od temperatury

T [K]	n_i [m^{-3}]		
	Si	Ge	GaAs
200	$6,1 \times 10^{10}$	$5,5 \times 10^{16}$	10^6
300	$1,5 \times 10^{16}$	$2,4 \times 10^{19}$	10^{13}
500	$9,2 \times 10^{19}$	$7,7 \times 10^{21}$	$6,1 \times 10^{17}$
700	$1,0 \times 10^{22}$	$2,7 \times 10^{23}$	$1,7 \times 10^{20}$

Z porównania wzorów (17.5) i (17.11) możemy obliczyć położenie poziomu Fermiego w półprzewodniku samoistnym

$$E_F = \frac{E_g}{2} + \frac{3}{4} kT \ln \frac{m_h^*}{m_e^*}. \quad (17.15)$$

W temperaturze 0 K poziom Fermiego przypada dokładnie w środku przerwy energetycznej i nie zmienia się ze zmianą temperatury, o ile $m_e^* = m_h^*$. Jeżeli natomiast masy efektywne są różne, to poziom Fermiego przesuwają się przy wzroście temperatury w kierunku pasma, któremu odpowiada mniejsza masa efektywna. Jednak w większości półprzewodników samoistnych zmiany położenia poziomu Fermiego pod wpływem temperatury można pominąć.

17.2. Półprzewodniki domieszkowe

W praktycznych zastosowaniach bardzo ważną rolę odgrywają półprzewodniki domieszkowe, tzn. półprzewodniki zawierające celowo wprowadzone atomy obce. Domieszkowanie w istotny sposób wpływa na właściwości elektryczne półprzewodników. Mechanizm działania domieszkowania omówimy na przykładzie pięciowartościowych atomów arsenu i trójwartościowych atomów indu wprowadzonych do germanu.

German ma strukturę diamentu, w której każdy z czterech elektronów walencyjnych bierze udział w jednym z czterech wiązań z najbliższymi sąsiadami. Jeżeli teraz pięciowartościowy atom arsenu zastąpi w sieci atom germanu, to cztery z pięciu elektronów walencyjnych biorą udział w wiązaniu (rys. 17.3). Pozostały piąty elektron nie uczestniczy w wiązaniu i jest związany z dodatnim polem domieszki siłami kulombowskimi. Siły te są jednak bardzo osłabione, gdyż oddziaływanie występuje w ośrodku o przenikalności dielektrycznej ϵ . Wówczas energia wiązania elektronu z jądrem maleje ϵ^2 razy, a promień orbity zwiększa się ϵ -krotnie. Dla germanu $\epsilon = 16$ i energia jonizacji domieszki jest 16^2 razy mniejsza niż energia jonizacji atomu swobodnego i wynosi setne części elektronowolta. W tabeli 17.4 podano Poniżej podano energię jonizacji domieszek pięciowartościowych w krzemie i germanie. Większe wartości energii jonizacji domieszek w krzemie są uwarunkowane mniejszą wartością przenikliwości dielektrycznej ($\epsilon = 11.7$). Pewien wpływ na energię jonizacji mają również wartości mas efektywnych.

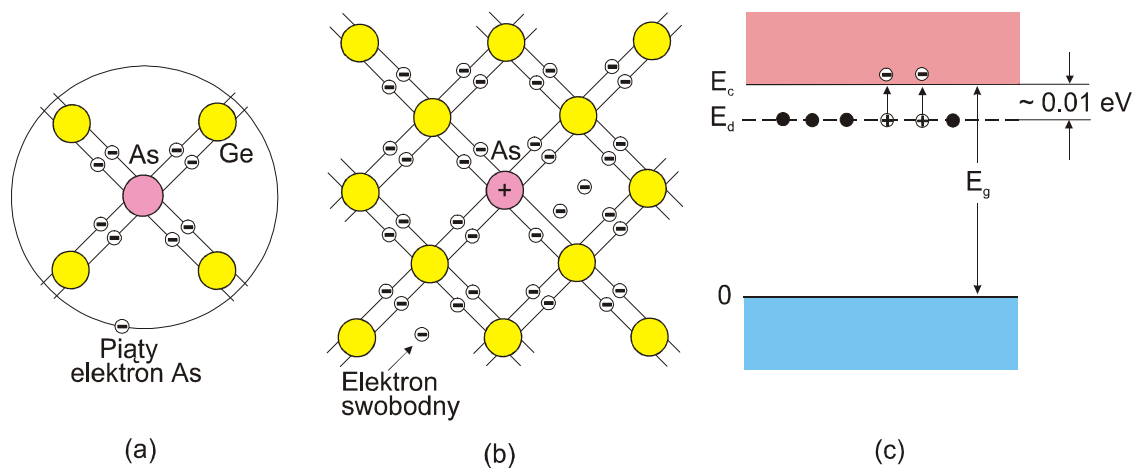
Tabela 17.4. Energia jonizacji domieszek pięciowartościowych w Si i Ge

Domieszka	E_j [eV]	
	Si	Ge
P	0,044	0,012
As	0,049	0,0127
Sb	0,039	0,0096

Tabela 17.5. Energia jonizacji domieszek trójwartościowych w Si i Ge

Domieszka	E_a [eV]	
	Si	Ge
B	0,045	0,0104
Al	0,057	0,0102
Ga	0,065	0,0108
In	0,16	0,0112

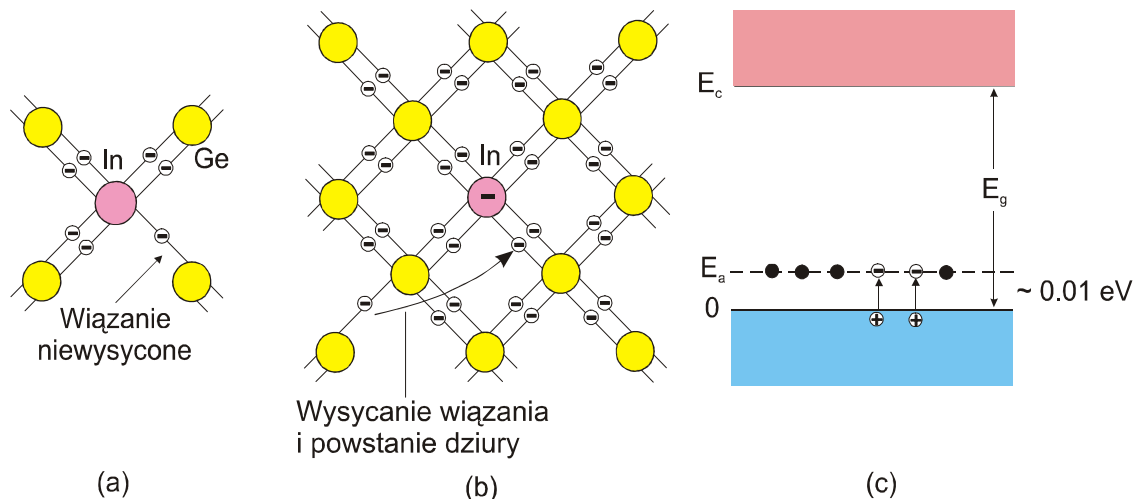
Ze względu na strukturę pasmową, istnienie domieszki o bardzo małej energii jonizacji oznacza pojawienie się dodatkowego poziomu w obszarze przerwy energetycznej, leżącego blisko dna pasma przewodnictwa. Energię tego poziomu będziemy oznaczać E_d (rys. 17.3c). $E_j = E_c - E_d$ jest oczywiście energią jonizacji domieszki. Dostarczenie elektronom energii $E_c - E_d$ powoduje ich przejście do pasma przewodnictwa. Powstałe jony domieszki są nieruchome i nie uczestniczą w przewodnictwie elektrycznym. Tego rodzaju domieszki nazywamy donorami, a wnoszone przez nie poziomy – poziomami donorowymi.



Rys. 17.3. Półprzewodnik typu n. Część atomów germanu zostaje zastąpiona przez pięciwartościowe atomy domieszki As: (a) w temperaturze 0 K atom domieszki jest niezjonizowany, piąty elektron nie biorący udziału w wiązaniu ma energię wiązania wynoszącą setne części elektronowolta, a promień odbity - kilkadziesiąt Å; (b) w temperaturze $T > 0$ K następuje jonizacja atomów domieszki i powstanie swobodnych elektronów, (c) położenie poziomu donorowego E_d .

Ze względu na małą wartość energii jonizacji, znacznie łatwiej jest otrzymać w paśmie przewodnictwa elektrony pochodzące od domieszki niż nośniki samoistne. W takim półprzewodniku dominować będzie przewodnictwo elektronowe. Mówimy też, że elektrony są nośnikami większościowymi, a dziury – mniejszościowymi. Półprzewodniki takie nazywamy półprzewodnikami typu p.

Rozpatrzmy teraz przypadek, kiedy w sieci germanu znajduje się atom trójwartościowy, np. ind (rys. 17.4). Wówczas jedno z wiązań pozostaje niewysyczone, gdyż atom taki ma o jeden elektron mniej niż atom germanu. Wiązanie to może być uzupełnione dowolnym elektronem z innego germanu. Przejście takie wymaga bardzo małej ilości energii. W tabeli 17.5 podano wartości energii potrzebnej do wysycenia wiązania atomu domieszki w krzemie i germanie.

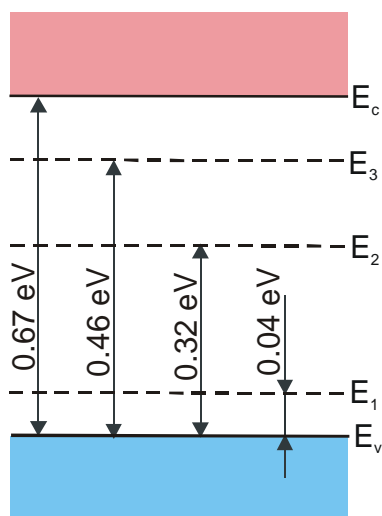


Rys. 17.4. Półprzewodnik typu p. Część atomów germanu zostaje zastąpiona przez trójwartościowe atomy domieszki In: (a) w temperaturze 0 K atom domieszki ma czwarte wiązanie niewysyczone; (b) w temperaturze $T > 0$ K następuje wysycenie wiązania atomu domieszki i jednocześnie pojawienie się dziury; (c) położenie poziomu akceptorowego E_a .

Biorąc pod uwagę strukturę pasmową półprzewodnika oznacza to pojawienie się dodatkowego poziomu energetycznego E_a leżącego blisko wierzchołka pasma walencyjnego. Ze względu na niewielką wartość energii potrzebnej do wysycenia wiązania, proces ten występuje już w bardzo niskich temperaturach.

Elektron, który wysycza wiązanie w atomie domieszki, zostawia jednocześnie dziurę w tym węźle. Miejsce to może zająć nowy elektron. W rezultacie takich procesów, dziura będzie przesuwać się w kierunku przeciwnym względem ruchu elektronu. W ujęciu struktury pasmowej oznacza to pojawienie się dziury w paśmie walencyjnym. Jednocześnie elektrony związane z atomami domieszki tracą możliwość przemieszczania się. Domieszki tego typu nazywamy akceptorowymi, a wnoszone poziomy – poziomami akceptorowymi. Na rys. 17.4c poziom ten oznaczyliśmy przez E_a . Półprzewodniki takie nazywamy półprzewodnikami typu p. Nośnikami większościowymi będą w tym przypadku dziury, a mniejszościowymi – elektrony. Typowe wartości koncentracji domieszek znajdują się w przedziale 10^{22} – 10^{25} m⁻³.

Dotychczas mówiliśmy o poziomach donorowych i akceptorowych leżących na bardzo niewielkich odległościach od granic pasm. Dlatego są one nazywane poziomami płytkimi. Niektóre domieszki w półprzewodnikach powodują powstawanie poziomów domieszkowych położonych daleko od granicy pasm energetycznych. Poziomy takie noszą nazwę głębokich. Dla przykładu na rys. 17.5 pokazano poziomy domieszkowe atomów miedzi wprowadzonych do



Rys. 17.5. Poziomy domieszkowe w germanie.

germanu. W temperaturze 0 K atom miedzi jest obojętny. Ze wzrostem temperatury wychwytuje on z pasma walencyjnego germanu początkowo jeden, potem dwa a w końcu trzy elektrony przekształcając się w ten sposób w trójwartościowy ujemnie naładowany jon. W paśmie walencyjnym germanu tworzą się przy tym dziury.

Głębokie poziomy domieszkowe odgrywają dużą rolę w przebiegu procesów rekombinacji nośników, co będzie szerzej omówione w następnym rozdziale.

17.3. Koncentracja nośników w półprzewodniku domieszkowym

Weźmy pod uwagę półprzewodnik typu n o koncentracji domieszki donorowej N_d i wartości energii jonizacji $E_j = E_c - E_d$ (rys. 17.3). Elektrony w paśmie przewodnictwa będą występować na skutek przejść poziom donorowy-pasmo przewodnictwa i pasmo walencyjne-pasmo przewodnictwa. Można więc napisać, że koncentracja elektronów swobodnych wynosi

$$n = N_d^+ + p, \quad (17.16)$$

gdzie N_d^+ jest koncentracją zjonizowanych domieszek, p – koncentracją dziur w paśmie walencyjnym, która jest równa koncentracji elektronów w paśmie przewodnictwa powstałych na skutek przejść elektronów z pasma walencyjnego. Ponieważ energia jonizacji domieszek jest znacznie mniejsza od szerokości przerwy energetycznej można przyjąć, że w niskich temperaturach wystąpi tylko jonizacja termiczna domieszek. Ze wzrostem temperatury istotnego znaczenia nabiera jonizacja atomów materiału bazowego (przejścia pasmo-pasmo).

Można wyróżnić dwa zakresy temperatury:

(a) temperatury niskie, w których $n \leq N_d$;

(b) temperatury wysokie, w których $n = N_d + p$.

Przekonamy się później, że pojęcie temperatur niskich i wysokich jest związane z koncentracją domieszek, energią jonizacji i szerokością przerwy energetycznej.

Temperatury niskie ($n \leq N_d$)

Koncentrację elektronów w paśmie przewodnictwa możemy zapisać w postaci

$$n = N_d^+ = N_d - N_d^o, \quad (17.17)$$

gdzie N_d^o i N_d^+ oznaczają odpowiednio koncentracje domieszek niezjonizowanych i zjonizowanych. N_d^o , czyli koncentrację elektronów na poziomie E_d można zapisać w postaci:

$$N_d^o = N_d f(E_d), \quad (17.18)$$

gdzie $f(E_d)$ jest prawdopodobieństwem obsadzenia przez elektron poziomu energetycznego E_d . Można wykazać, że funkcja $f(E_d)$ ma postać

$$f(E_d) = \frac{1}{\frac{1}{2} \exp\left[\frac{(E_d - E_F)}{kT}\right] + 1}. \quad (17.19)$$

Różnica między funkcją określoną wzorem (17.19) a funkcją rozkładu Fermiego-Diraca wynika stąd, że na poziomie E_d może znajdować się jeden elektron, podczas gdy funkcja rozkładu Fermiego-Diraca określa prawdopodobieństwo obsadzenia poziomu energetycznego przez dwa elektrony o przeciwnie ustawionych spinach.

Na podstawie wzorów (17.17), (17.18), (17.19) otrzymujemy

$$n = N_d [1 - f(E_d)] = \frac{N_d}{2 \exp\left(\frac{E_F - E_d}{kT}\right) + 1}. \quad (17.20)$$

Z powyższego wzoru nie możemy jeszcze wyznaczyć n , gdy nie znamy E_F . Ale jednocześnie wiemy, że

$$n = N_c \exp\left[\frac{-(E_c - E_F)}{kT}\right]. \quad (17.21)$$

Porównując dwa ostatnie wzory mamy

$$N_c e^{-(E_c - E_F)/kT} = \frac{N_d}{2 \exp[(E_F - E_d)/kT] + 1}. \quad (17.22)$$

Z tego wzoru można obliczyć E_F i podstawić do wyrażenia (17.21), aby obliczyć n . Jednakże otrzymane wyrażenie ma stosunkowo skomplikowaną postać i jest trudne do dyskusji. Dlatego rozpatrzmy dwa graniczne przypadki: $E_F - E_d \gg kT$ i $E_F - E_d \ll kT$. Jeżeli przyjąć typową wartość $E_F - E_d = 0.1$ eV, to pierwszy warunek jest spełniony w temperaturach 30–40 K, zaś drugi warunek jest spełniony w temperaturze pokojowej.

Jeżeli $E_F - E_d \gg kT$, to wówczas wzór (17.22) może być zapisany w postaci

$$N_c e^{-(E_c - E_F)/kT} = \frac{1}{2} N_d e^{-(E_F - E_d)/kT}. \quad (17.23)$$

Wyliczając E_F otrzymujemy

$$E_F = \frac{E_d + E_c}{2} + \frac{kT}{2} \ln \frac{N_d}{2N_c}. \quad (17.24)$$

Podstawiając wyrażenie na E_F do wzoru (17.21) otrzymujemy wyrażenie na koncentrację elektronów w półprzewodniku typu n w obszarze bardzo niskich temperatur

$$n = \sqrt{\frac{N_d N_c}{2}} e^{-E_g/2kT}. \quad (17.25)$$

Ze wzoru (17.24) wynika, że w temperaturze 0 K poziom Fermiego leży pośrodku między dnem pasma przewodnictwa i poziomem donorowym. W miarę wzrostu temperatury poziom Fermiego podnosi się, osiąga maksimum a następnie obniża się (rys. 17.6a).

W rozpatrywanym zakresie temperatur koncentracja elektronów jest proporcjonalna do pierwiastka kwadratowego z koncentracji domieszek. Logarytm z n zmienia się praktycznie liniowo ze wzrostem temperatury, przy czym tangens kąta nachylenia jest równy $E_g/2k$. Okazuje się, że nachylenie to zależy także od koncentracji domieszek, ale sprawę tę wyjaśnimy później.

Przy dalszym wzroście temperatury zostają zjonizowane wszystkie domieszki i wówczas koncentracja elektronów

$$n = N_c e^{-(E_c - E_F)/kT} = N_d, \quad (17.26)$$

nie zależy od temperatury i równa jest koncentracji domieszek. Zakres temperatury w którym koncentracja elektronów jest stała i nie zależy od temperatury, nazywa się zakresem nasycenia. Dolną granicę obszaru nasycenia wyznaczymy z warunku, aby koncentracja elektronów była równa koncentracji domieszek, tzn.

$$n = N_d = \left(\frac{N_d N_c}{2} \right)^{1/2} e^{-E_j/2kT}. \quad (17.27)$$

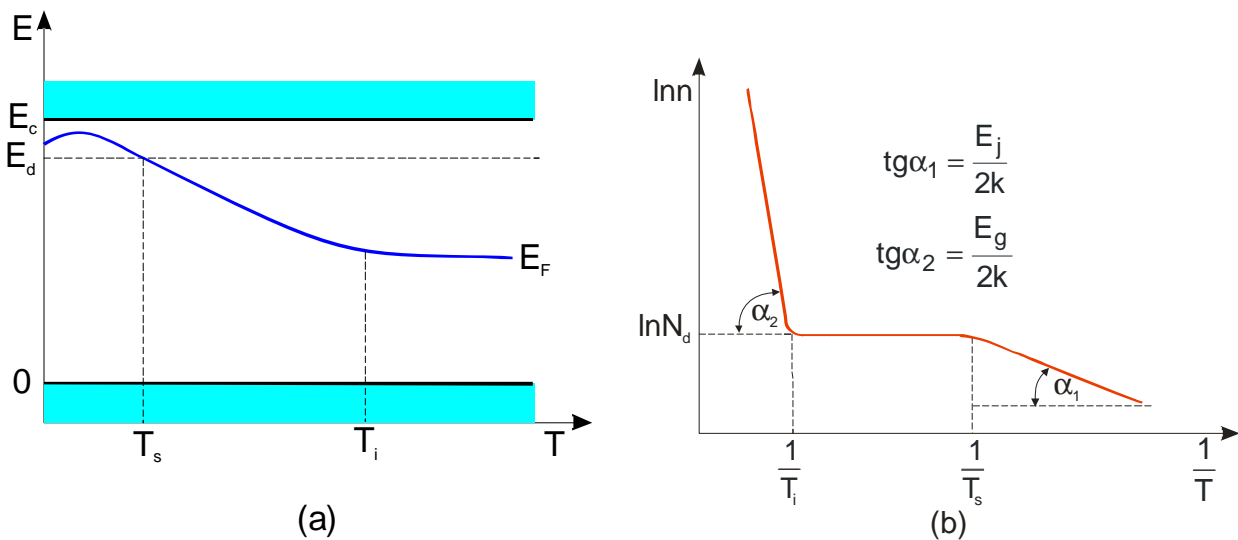
Wyliczając dolną granicę obszaru nasycenia T_s , otrzymujemy

$$T_s = \frac{E_j}{k \ln[N_c(T_s)/2N_d]}. \quad (17.28)$$

Koncentrację nośników mniejszościowych wyliczymy z zależności $np = n_i^2$, skąd

$$p = \frac{n_i^2}{n} = \frac{n_i^2}{N_d} = \frac{N_c N_v}{N_d} e^{-E_g/kT}. \quad (17.29)$$

Tak więc w obszarze nasycenia koncentracja nośników mniejszościowych bardzo szybko wzrasta ze wzrostem temperatury.



Rys. 17.6. Schematyczne przedstawienie temperaturowej zależności poziomu Fermiego (a) i koncentracji nośników (b) w półprzewodniku typu n.

Ze wzoru (17.26) możemy jeszcze wyliczyć poziom Fermiego

$$E_F = E_c + kT \ln \frac{N_d}{N_c}. \quad (17.30)$$

Ponieważ w rozpatrywanym zakresie temperatur $N_d < N_c$, więc E_F obniża się ze wzrostem temperatury.

Temperatury wysokie

Dalszy wzrost temperatury powoduje, że istotną rolę odgrywa jonizacja atomów materiału bazowego (przejścia pasmo-pasmo) i koncentracja elektronów uwarunkowana tą jonizacją staje się porównywalna z koncentracją domieszki. Wówczas możemy napisać:

$$n = N_d + p = N_d + \frac{n_i^2}{n}. \quad (17.31)$$

Wyliczając n otrzymujemy

$$n = \frac{N_d + \sqrt{N_d^2 + 4n_i^2}}{2}, \quad (17.32)$$

natomiast p

$$p = \frac{n_i^2}{n} = \frac{2n_i^2}{N_d + \sqrt{N_d^2 + 4n_i^2}}. \quad (17.33)$$

Jeżeli $N_d \gg n_i$, co bardzo często ma miejsce, to $n = N_d$; natomiast jeżeli $n_i \gg N_d$, to $n = n_i$ i półprzewodnik zachowuje się jak samoistny. Nierówność $n_i \gg N_d$ będzie spełniona w wysokich temperaturach, gdyż wówczas koncentracja elektronów otrzymana w wyniku jonizacji materiału podstawowego (przejść pasmo-pasmo) może przekraczać koncentrację domieszek. Koncentracja nośników będzie wykładniczo rosnącą funkcją temperatury. Logarytm z n zmienia się liniowo z odwrotnością temperatury, przy czym tangens kąta nachylenia wynosi $E_g/2k$ [rys.17.6(b)].

Temperatura T_i w której następuje przejście półprzewodnika domieszkowanego w stan przewodnictwa samoistnego, jest górną granicą obszaru nasycenia. Dla półprzewodnika samoistnego

$$n = n_i = (N_c N_v)^{1/2} e^{-E_g/2kT}.$$

Ponieważ w obszarze nasycenia $n = N_d$, zatem górną granicę zakresu nasycenia otrzymamy z zależności

$$N_d = (N_c N_v)^{1/2} e^{-E_g/2kT_i}, \quad (17.34)$$

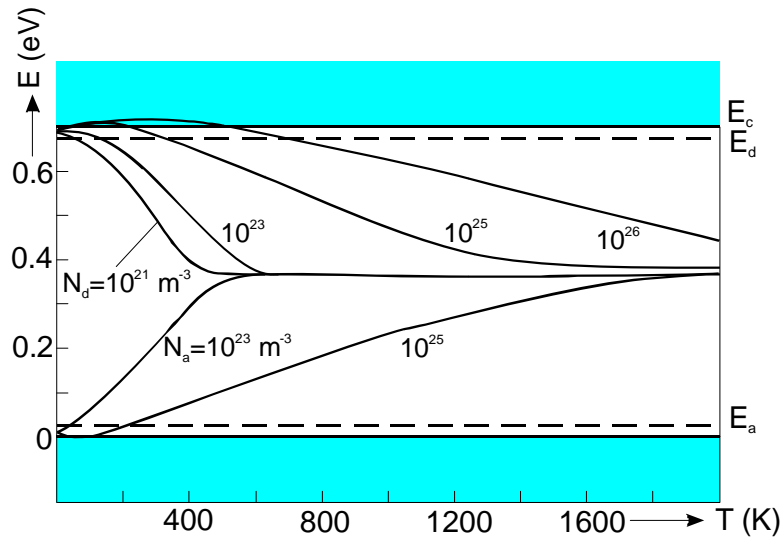
stąd

$$T_i = \frac{E_g}{k \ln[N_c N_v / N_d^2]}. \quad (17.35)$$

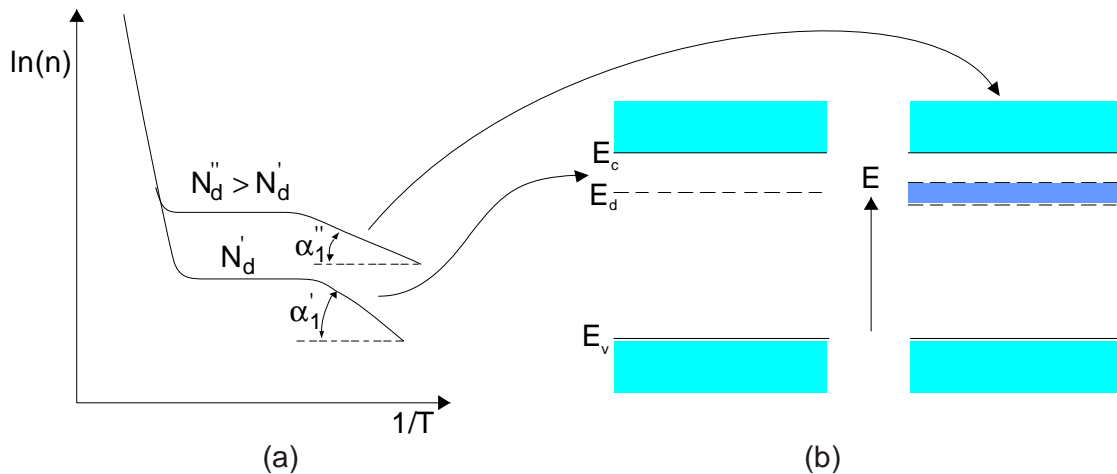
Temperatura przejścia w obszar przewodnictwa domieszkowanego jest tym wyższa im większa jest koncentracja domieszki i większa szerokość przerwy energetycznej. Np. dla Ge z $E_i = 0.01$ eV i $N_d = 10^{22} \text{ m}^{-3}$, $T_i = 35$ K oraz $T_s = 400$ K; zaś dla Si przy tej samej koncentracji donorów z $E_j = 0.04$ eV, $T_s = 86$ K oraz $T_i = 780$ K. Zauważmy, że nasycenie jest osiągalne już w

temperaturach bardzo niskich. Przejście do obszaru samoistnego następuje na ogół w temperaturach znacznie wyższych od pokojowych. W obszarze rozpatrywanych temperatur poziom Fermiego obniża się do połowy przerwy energetycznej. Ilościową ilustracją przeprowadzonych rozważań są rysunki 17.7 i 17.8.

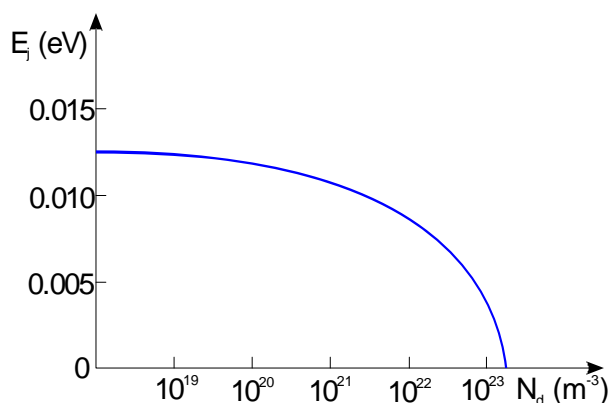
Na rys. 17.7 przedstawiono zależność położenia poziomu Fermiego od temperatury przy różnych koncentracjach domieszek w germanie. Należy zwrócić uwagę, że nawet przy niewielkich koncentracjach domieszek przejście w obszar przewodnictwa samoistnego następuje w wysokich temperaturach, znacznie wyższych od temperatury pokojowej.



Rys. 17.7. Zależność położenia poziomu Fermiego od temperatury w germanie przy różnych koncentracjach domieszek. Krzywe górne odpowiadają półprzewodnikowi typu n, krzywe dolne – półprzewodnikowi typu p.



Rys. 17.8. Zależność $n = n(T)$ przy różnych koncentracjach domieszki donorowej (a) i rozszczepienie poziomu donorowego przy wzroście koncentracji domieszki (b).



Rys. 17.9. Zależność energii jonizacji domieszki As w germanie od koncentracji N_d .

Na rys. 17.9 pokazano zależność energii jonizacji domieszki donorowej (arsenu) w germanie. Zmniejszenie energii jonizacji ze wzrostem koncentracji domieszek występuje dlatego, że przy dużych koncentracjach domieszki pojawia się oddziaływanie między obcymi atomami, co powoduje rozszczepianie pierwotnego poziomu energetycznego w pasmo. Jego szerokość jest tym większa, im mniejsza jest odległość pomiędzy atomami domieszki, czyli im większa jest ich koncentracja (rys. 17.9). Przy dostatecznie dużej koncentracji domieszki pasmo to zachodzi na pasmo przewodnictwa. Podobnie jak w metalach, poziom Fermiego będzie w tym przypadku najwyższym poziomem energetycznym obsadzonym przez elektrony. Gaz elektronowy w takich półprzewodnikach jest gazem zwyrodniałym, a półprzewodniki takie noszą nazwę zwyrodniałych. Do osiągnięcia stanu zwyrodnienia konieczne są koncentracje domieszek 10^{25} – 10^{26} m^{-3} .

Analogiczne rozważania można przeprowadzić dla półprzewodnika typu p. Położenie poziomu Fermiego i koncentrację nośników większościowych w obszarze niskich temperatur opisują wyrażenia:

$$E_F = \frac{E_a}{2} - \frac{kT}{2} \ln \frac{N_a}{N_v}, \quad (17.36)$$

$$p = \left(\frac{N_a N_v}{2} \right)^{1/2} e^{-E_a/2kT}, \quad (17.37)$$

w obszarze nasycenia

$$E_F = kT \ln \frac{N_v}{N_a}, \quad (17.38)$$

$$p = N_a, \quad (17.39)$$

w obszarze samoistnym

$$p = \frac{N_a + \sqrt{N_a^2 + 4n_i^2}}{2}, \quad (17.40)$$

Na rys. 17.8 przedstawiono zależność koncentracji elektronów od temperatury przy różnych koncentracjach domieszki. Zauważmy, że ze wzrostem koncentracji domieszek nachylenie krzywej $n(T)$ w obszarze niskich temperatur zmniejsza się. Wskazuje to na zmniejszenie energii jonizacji ze wzrostem koncentracji domieszek, co istotnie ma miejsce.

$$E_F = \frac{E_g}{2} + \frac{kT}{2} \ln \frac{N_v}{N_c}, \quad (17.41)$$

gdzie N_a jest koncentracją domieszki akceptorowej, N_v – efektywną gęstością stanów w paśmie walencyjnym. Zależność położenia poziomu Fermiego od temperatury w półprzewodniku typu p przedstawiono na rys. 17.7. Koncentracja dziur zmienia się z temperaturą w sposób analogiczny do koncentracji elektronów w półprzewodniku typu n.

17.4. Ruchliwość nośników w półprzewodnikach

Ruchliwość nośników μ definiujemy jako stosunek prędkości unoszenia (dryfu) v_d do natężenia zewnętrznego pola elektrycznego E

$$\mu = \frac{v_d}{E}. \quad (17.42)$$

Gdyby na nośniki działała tylko siła zewnętrzna pola elektrycznego, to prędkość unoszenia wzrastałaby z upływem czasu. W rzeczywistości elektrony doznają zderzeń z fononami i domieszkami sieci. Efektem tych zderzeń jest ustalenie się pewnego stanu równowagi, który charakteryzuje się ustaloną prędkością unoszenia. Można przyjąć, że im większa będzie prędkość dryfu, tym częściej nośniki będą ulegały zderzeniu. Przy założeniu, że siła hamująca jest proporcjonalna do prędkości dryfu, równanie ruchu dla elektronu ma postać

$$m_e^* \frac{d\vec{v}_d}{dt} = -e\vec{E} - \frac{1}{\tau} m_e^* \vec{v}_d, \quad (17.43)$$

gdzie τ jest pewną stałą o wymiarze czasu, której fizyczną interpretację podamy dalej. Przy założeniu, że w chwili włączenia pola elektrycznego $v_d = 0$, rozwiązaniem powyższego równania jest funkcja

$$\vec{v}_d(t) = \vec{v}_d \left(1 - e^{-t/\tau}\right), \quad (17.44)$$

gdzie

$$\vec{v}_d = -\frac{e\vec{E}}{m_e^*} \tau \quad (17.45)$$

jest prędkością unoszenia w stanie ustalonym. Jeżeli natomiast po ustaleniu się stanu równowagi wyłączymy pole elektryczne to równanie (17.43) ma postać

$$\frac{d\vec{v}_d}{dt} = -\frac{\vec{v}_d}{\tau},$$

którego rozwiązaniem jest funkcja

$$\vec{v}_d(t) = \vec{v}_d e^{-t/\tau}. \quad (17.46)$$

Z wyrażenia tego wynika, że po wyłączeniu pola zewnętrznego prędkość unoszenia, a więc i gęstość prądu, maleje wykładniczo w czasie. Przyczyną tego jest oczywiście rozproszenie elektronów na fononach i domieszkach. Liczba elektronów dających swój wkład do prądu będzie zmniejszała się w czasie, przy czym

$$n(t) = n(0)e^{-t/\tau}, \quad (17.47)$$

gdzie $n(0)$, $n(t)$ – koncentracja elektronów o prędkości dryfu v_d odpowiednio w chwili $t = 0$ i po czasie t od momentu wyłączenia pola elektrycznego. Jest to jednocześnie liczba elektronów, które przeżyły bez zderzenia czas t .

Ażeby nadać interpretację fizyczną stałej τ , obliczymy średni czas między zderzeniami. Z równania (17.47) wynika, że liczba elektronów, które bez zderzeń przeżyły czas t i uległy zderzeniom w przedziale czasu od t do $t + dt$ wynosi

$$dn = \frac{1}{\tau} n(0) e^{-t/\tau} dt.$$

Mnożąc powyższe wyrażenie przez czas t otrzymujemy łączny czas życia tych elektronów bez zderzeń

$$\Delta t = \frac{1}{\tau} n(0) e^{-t/\tau} t dt.$$

Całkując to wyrażenie i dzieląc przez liczbę wszystkich elektronów $n(0)$ otrzymujemy średni czas upływający między zderzeniami

$$\langle t \rangle = \frac{\int_0^{\infty} (1/\tau) n(0) e^{-t/\tau} t dt}{n(0)} = \frac{1}{\tau} \int_0^{\infty} e^{-t/\tau} t dt = \tau. \quad (17.48)$$

Widać, że stała τ jest średnim czasem między zderzeniami.

Mając prędkość dryfu v_d , możemy wyznaczyć gęstość prądu

$$\vec{j} = -en\vec{v}_d = \frac{e^2}{m_e^*} n\vec{E}\tau,$$

gdzie n jest koncentracją elektronów, e – ładunkiem elementarnym.

Ponieważ jednocześnie z prawa Ohma mamy

$$\vec{j} = \sigma\vec{E},$$

więc konduktywność wynosi

$$\sigma = \frac{e^2 n \tau}{m_e^*}, \quad (17.49)$$

a w związku z tym, zgodnie z (17.42)

$$\mu_e = \frac{e\tau}{m_e^*}. \quad (17.50)$$

Wtedy

$$\sigma_e = en\mu_e. \quad (17.51)$$

Podobnie dla dziur

$$\sigma_h = ep\mu_h. \quad (17.52)$$

Na podstawie rozważań przeprowadzonych uprzednio możemy napisać

$$\mu = \frac{e}{m} \tau = \frac{e}{m} \frac{\langle l \rangle}{\langle v \rangle}, \quad (17.53)$$

gdzie τ nazywamy czasem relaksacji zderzeniowej, natomiast m , $\langle l \rangle$ i $\langle v \rangle$ odpowiednio oznaczono masę, średnią drogę swobodną i średnią prędkość nośników.

Wartość średniej drogi swobodnej jest uwarunkowana mechanizmami rozpraszania. W półprzewodnikach o szerszej przerwie energetycznej dominują dwa mechanizmy rozpraszania nośników: rozpraszanie na fononach i rozpraszanie na zjonizowanych domieszkach. Pierwszy mechanizm odgrywa decydującą rolę w temperaturach powyżej 100–150 K, kiedy koncentracja fononów jest duża; drugi natomiast – w temperaturach niskich. Przy rozpraszaniu na fononach

$$\langle l \rangle \sim T^{-1}, \quad (17.54)$$

natomiast przy rozpraszaniu na domieszkach

$$\langle l \rangle \sim \frac{v^4}{N}, \quad (17.55)$$

gdzie N jest koncentracją domieszek.

Z wyjątkiem półprzewodników o dużej koncentracji domieszek, gaz elektronowy w półprzewodnikach jest gazem niezwyrodniałym i średnia prędkość nośników jest wyrażona wzorem

$$\langle v \rangle = \sqrt{\frac{8kT}{\pi m}}. \quad (17.56)$$

Podstawiając wartości liczbowe otrzymujemy $\langle v \rangle = 10^5$ m/s, przy $T = 300$ K. Jest to wartość prędkości kilkadziesiąt razy mniejsza niż prędkość elektronów na poziomie Fermiego w metalach. Można więc oczekiwać, że ruchliwość nośników w półprzewodnikach będzie znacznie większa niż w metalach.

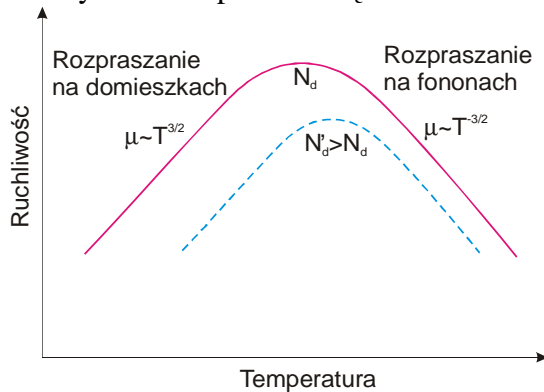
Z zależności (17.53), (17.54) i (17.56) wynika, że temperaturowa zależność ruchliwości nośników w półprzewodnikach niezwyrodniałych przy rozpraszaniu na fononach ma postać

$$\mu \sim T^{-3/2}. \quad (17.57)$$

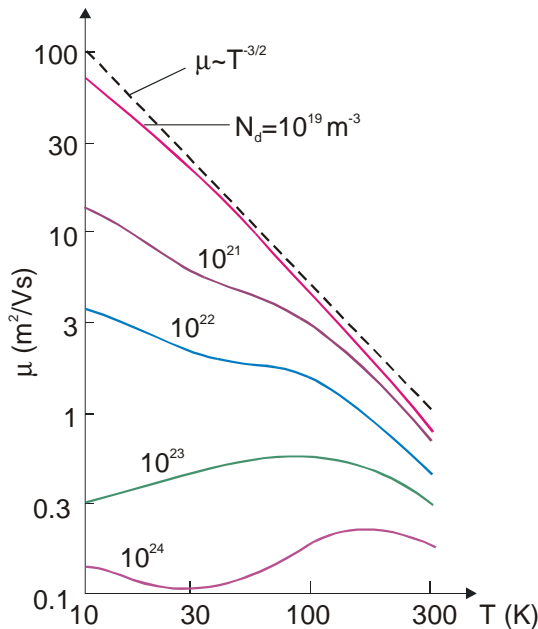
Natomiast z zależności (17.53) i (17.55) otrzymujemy, że przy rozpraszaniu na domieszkach $\mu \sim v^3$. Ponieważ jednocześnie $v \sim T^{1/2}$, więc

$$\mu \sim T^{3/2}. \quad (17.58)$$

Reasumując można powiedzieć, że w półprzewodnikach niezwyrodniałych ze wzrostem temperatury ruchliwość wzrasta proporcjonalnie do $T^{3/2}$, następnie przechodzi przez maksimum i zmniejsza się proporcjonalnie do $T^{-3/2}$. Ze wzrostem koncentracji domieszek ruchliwość nośników maleje, a maksimum ruchliwości przesuwają się w kierunku wyższych temperatur. Jest to zrozumiałe, ponieważ im większa koncentracja domieszek, tym bardziej wpływają one na ruchliwość nośników, nawet w wyższych temperaturach. Na rys. 17.10 przedstawiono schematycznie temperaturową zależność ruchliwości nośników od temperatury.



Rys. 17.10. Teoretyczna zależność ruchliwości nośników od temperatury w półprzewodniku domieszkowym.

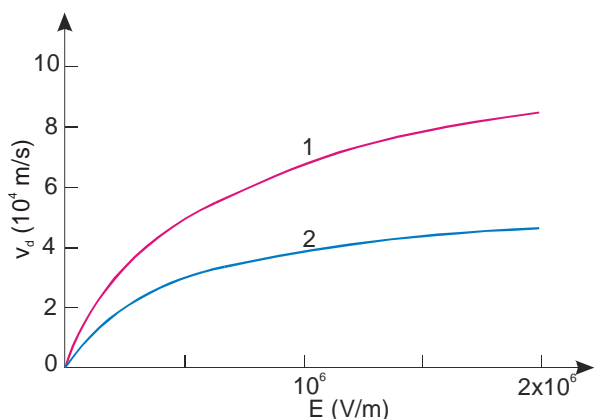


Rys. 17.11. Zależność ruchliwości elektronów od temperatury w germanie przy różnych koncentracjach domieszki donorowej.

Dane eksperymentalne wskazują, że zależności temperaturowe ruchliwości nie są ściśle opisane zależnościami (17.57) i (17.58) ze względu na występowanie innych mechanizmów rozpraszania. Problemu tego nie będziemy jednak bliżej omawiać.

Na rys. 17.11 przedstawiono zależność ruchliwości od temperatury i koncentracji domieszek w germanie. Natomiast w tabeli 17.6 podano wartości ruchliwości nośników w kilku półprzewodnikach w temperaturze 77 i 300 K. Warto zwrócić uwagę, że ruchliwość elektronów jest na ogół większa niż ruchliwość dziur, co jest spowodowane głównie większą masą efektywną dziur. W ostatnich kolumnach tabeli 17.6 podano doświadczalne wartości wykładnika potęgi n temperaturowej zależności ruchliwości przy rozpraszaniu na fononach ($\mu \sim T^n$).

Ruchliwość nośników zależy również od natężenia zewnętrznego pola elektrycznego, przy czym zależność ta występuje, jeżeli natężenie pola jest duże. Wówczas nośniki uzyskują prędkości unoszenia porównywalne z prędkościami w ruchu chaotycznym. Nośniki mają energię większą od równowagowej energii cieplnej. Wzrost



Rys. 17.12. Zależność prędkości dryfu elektronów 1 i dziur 2 w krzemie w zależności od natężenia pola elektrycznego.

całkowitej prędkości powoduje, że częściej występują zderzenia z niedoskonałościami sieci, co w efekcie prowadzi do zmniejszenia prędkości dryfu. Na rys.17.12 przedstawiono zależność prędkości dryfu elektronów i dziur od natężenia pola elektrycznego w krzemie. Przy polach o natężeniu 10^6 – 10^7 V/m prędkość dryfu osiąga ustaloną wartość.

Tabela 17.6. Eksperymentalne wartości ruchliwości elektronów i dziur

Półprzewodnik	μ_e (m ² /Vs)		μ_h (m ² /Vs)		n	
	77K	300K	77K	300K	Elektrony	Dziury
GaAs	2,10	0,85	0,42	0,04	1,0	2,1
Ge	3,71	0,38	4,37	0,18	1,6	2,3
InSb	120,00	7,80	1,00	0,07	1,6	2,1
Si	4,55	0,13	1,16	0,05	2,6	2,3

17.5. Przewodnictwo elektryczne półprzewodników

Znajomość koncentracji i ruchliwości nośników pozwoli nam omówić przewodnictwo elektryczne półprzewodników. Biorąc pod uwagę fakt, że w półprzewodnikach transport ładunku jest spowodowany zarówno elektronami jak i dziurami, konduktywność wyraża się wzorem:

$$\sigma = |e|(n\mu_e + p\mu_h), \quad (17.59)$$

gdzie n , μ_e , p , μ_h są odpowiednio koncentracjami i ruchliwościami elektronów i dziur.

17.5.1. Półprzewodniki samoistne

W tym przypadku $n = p = n_i$ i konduktywność wynosi

$$\sigma = |e|n_i(\mu_e + \mu_h), \quad (17.60)$$

gdzie

$$n_i = \frac{2}{h^3} (2\pi kT)^{3/2} (m_e^* m_n^*)^{3/4} e^{-E_g/2kT}.$$

Ruchliwość nośników z wyjątkiem temperatur bardzo niskich ($< 100\text{K}$) jest w przybliżeniu proporcjonalna do $T^{-3/2}$. Dzięki temu zależność potęgowa od temperatury upraszcza się tak, że wzór (17.60) może być zapisany w postaci

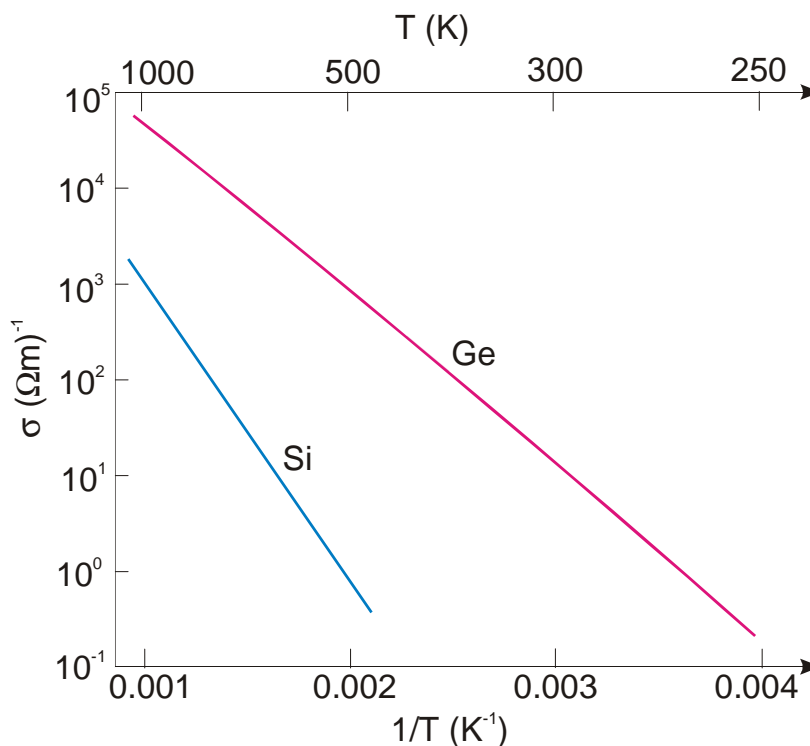
$$\sigma = \sigma_0 e^{-E_g/2kT}, \quad (17.61)$$

gdzie σ_0 jest dla danego półprzewodnika wartością stałą.

Zależność (17.61) wygodnie jest przedstawić w skali półlogarytmicznej

$$\ln \sigma = \ln \sigma_0 - \frac{E_g}{2kT}. \quad (17.62)$$

Wykres logarytmiczny konduktywności w funkcji odwrotności temperatury jest linią prostą. Tangens nachylenia tej prostej wynosi $E_g/2k$. Pozwala to wyznaczyć szerokość przerwy energetycznej z pomiarów temperaturowej zależności σ . Na rys. 17.13 przedstawiono wykres zależności σ do $1/T$ dla krzemu i germanu otrzymany na podstawie danych eksperymentalnych. Widać, że jakościowo wykresy te są w dobrej zgodności z zależnością (17.62).



Rys. 17.13. Zależność konduktywności od temperatury dla czystego germanu i krzemu.

Z powyższego wynika, że konduktywność elektryczna półprzewodników bardzo szybko zmienia się z temperaturą, przy czym zmiany te są uwarunkowane zmianami koncentracji nośników. Jest to dość istotna różnica między półprzewodnikami a metalami. W metalach koncentracja nośników jest stała, a temperaturowa zależność σ jest spowodowana tylko temperaturową zależnością ruchliwości nośników.

W danej temperaturze konduktywność zależy od szerokości pasma zabronionego; im szersze pasmo wzbronione, tym mniejsza koncentracja nośników, mniejsza konduktywność. Poniżej podano wartości E_g i σ dla kilku półprzewodników w 300 K.

Tabela 17.7. Konduktywność wybranych półprzewodników w 300 K

	InSb	PbS	Ge	Si
E_g [eV]	0,18	0,37	0,67	1,12
σ [Ωm] ⁻¹	2×10^4	30	2.3	1.6×10^{-3}

17.5.2. Półprzewodniki domieszkowe

Podobnie jak w przypadku półprzewodników samoistnych, również w półprzewodnikach domieszkowych zależność konduktywności od temperatury jest uwarunkowana głównie zależnością koncentracji nośników od temperatury. W związku z tym wyróżniamy trzy zakresy temperatury.

Temperatury niskie

Koncentracja nośników wzrasta z temperaturą zgodnie z zależnością (17.25)

$$n = \sqrt{\frac{N_d N_c}{2}} e^{-E_j/2kT}, \quad N_c \sim T^{3/2}.$$

Jednocześnie w tym obszarze temperatury dominuje rozpraszanie na domieszkach i ruchliwość $\mu \sim T^{3/2}$. Zaniedbując słabą zależność potęgową od temperatury

$$\sigma = \sigma_o e^{-E_j/2kT},$$

lub

$$\ln \sigma = \ln \sigma_o - \frac{E_j}{2k} \frac{1}{T}.$$

Na rys. 17.14a obszarowi temu odpowiada odcinek "ab". Nachylenie odcinka wynosi $E_j/2k$. W tym obszarze, konduktywność wzrasta z temperaturą na skutek termicznej jonizacji domieszek i jest uwarunkowana tylko nośnikami większościowymi.

Temperatury umiarkowane (obszar nasycenia)

Koncentracja nośników większościowych pozostaje praktycznie stała i równa koncentracji domieszki. Jednocześnie w tym obszarze temperatur o ruchliwości decyduje już właściwie rozproszenie na fononach i dlatego $\mu \sim T^{-3/2}$. Oznacza to, że konduktywność maleje ze wzrostem temperatury, choć zmiany te są stosunkowo niewielkie. Obszarowi temu na rys. 17.14a odpowiada odcinek "bc".

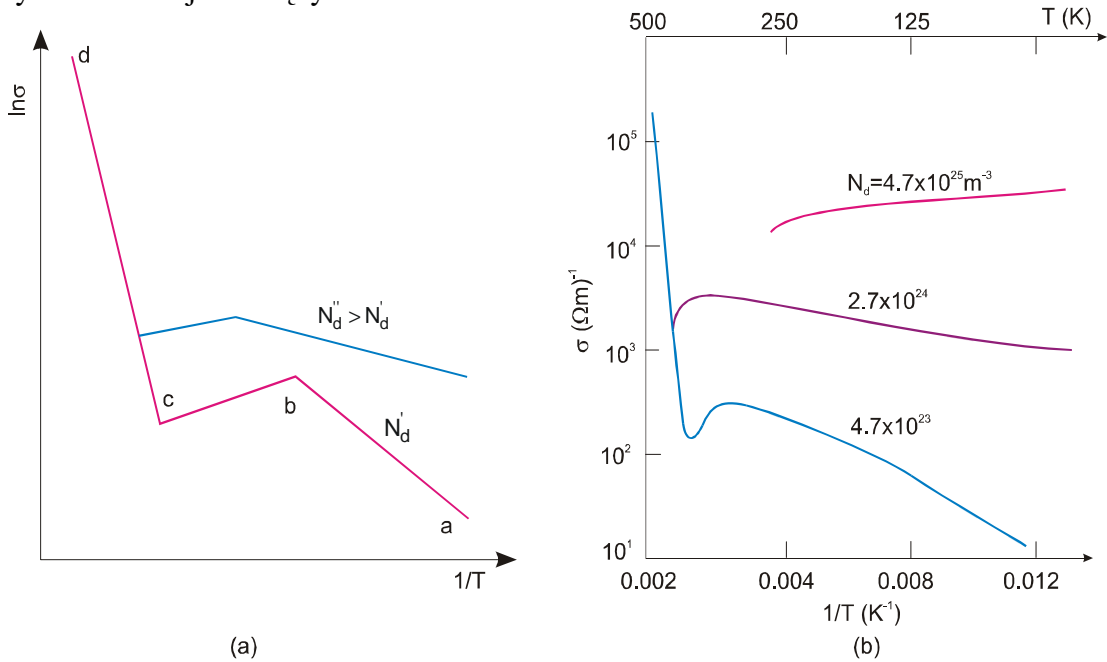
Temperatury wysokie

W temperaturach wysokich następuje jonizacja termiczna atomów materiału podstawowego, koncentracja nośników jest praktycznie taka jak w materiale samoistnym. Ruchliwość nośników maleje z temperaturą proporcjonalnie do $T^{-3/2}$. Zatem konduktancja

$$\sigma = \sigma_0 e^{-E_g/2kT}$$

Na rys. 17.14a odpowiada to odcinkowi "cd", tworzącemu z osią odciętych kąt α taki, że $\text{tg}\alpha = E_g/2k$. Im większa jest koncentracja domieszek, tym ich całkowita jonizacja następuje w wyższej temperaturze, jednocześnie w wyższej temperaturze następuje przejście w obszar samoistny.

Dla półprzewodnika zwyrodniałego koncentracja nośników pozostaje stała, a temperaturowe zmiany konduktancji zależą tylko od zmian ruchliwości.



Rys. 17.14. (a) Schematyczny przebieg zależności $\sigma = \sigma(T)$; (b) zależność $\sigma = \sigma(T)$ dla krzemu domieszkowanego fosforem.

Na rys. 17. 14b przedstawiono zależność σ od T dla krzemu domieszkowanego fosforem, przy różnych koncentracjach domieszki. Przebieg krzywych doświadczalnych potwierdza teoretyczną analizę.

ROZDZIAŁ 18

ZJAWISKA NIERÓWNOWAGOWE W PÓŁPRZEWODNIKACH

W tym rozdziale rozpatrzmy niektóre elementarne zagadnienia fizyki półprzewodników, których znajomość jest niezbędna do zrozumienia działania przyrządów półprzewodnikowych. Działanie wszelkich elementów półprzewodnikowych (takich jak np. złącze p-n, tranzystor) opiera się na wykorzystaniu zjawisk zachodzących w stanie zakłóconej równowagi termodynamicznej. W półprzewodniku, w odróżnieniu od metalu, można za pomocą czynników zewnętrznych zmieniać koncentrację nośników ładunku, czyli wytwarzać w nich lub wprowadzić do nich tzw. nośniki nadmiarowe elektronów i dziur Δn i Δp . Po ustaleniu działania tych czynników, półprzewodnik powraca samorzutnie do stanu równowagi termodynamicznej. Proces ten nazywamy ogólnie relaksacją.

18.1. Relaksacja dielektryczna

Rozpatrzmy jednorodny półprzewodnik, w którym nastąpiło lokalne zakłócenie neutralności dielektrycznej, np. na skutek wystąpienia fluktuacji gęstości nośników prądu pojawił się obszar nieskompensowanego ładunku elektrycznego o gęstości ρ . Oddziaływanie elektrostatyczne przywróci oczywiście po pewnym czasie neutralność kryształu. Określimy kinetykę tego procesu przyjmując przypadek jednowymiarowy.

Korzystamy z równania Poissona [zobacz wzór (7.22)]

$$\operatorname{div} E = \frac{\rho}{\epsilon_0 \epsilon_r}. \quad (18.1)$$

Dla modelu jednowymiarowego

$$\frac{dE}{dx} = \frac{\rho}{\epsilon_0 \epsilon_r}, \quad (18.2)$$

a równanie zachowania ładunku (ciągłości prądu) będzie miało postać

$$-\frac{dj}{dx} = \frac{d\rho}{dt}, \quad (18.3)$$

gdzie j jest gęstością prądu. Dzieliąc stronami wyrażenia (18.2) i (18.3) otrzymujemy równanie

$$\frac{1}{\rho} \frac{d\rho}{dt} = \frac{d \ln \rho}{dt} = -\frac{1}{\varepsilon_o \varepsilon_r} \frac{dj}{dE},$$

którego rozwiązaniem jest

$$\rho(t) = \rho(0)e^{-t/\tau_M}. \quad (18.4)$$

Wielkość

$$\tau_d = \varepsilon_o \varepsilon_r \left(\frac{dj}{dE} \right)^{-1} \quad (18.5)$$

nazywana jest czasem relaksacji dielektrycznej lub stałą czasową Maxwella.

Gdy spełnione jest prawo Ohma, to wówczas $dj/dE = \sigma$, gdzie σ jest konduktywnością materiału i wtedy

$$\tau_d = \frac{\varepsilon_o \varepsilon_r}{\sigma}. \quad (18.6)$$

Dla krzemu z $\varepsilon_r \approx 12$ w najbardziej typowym przypadku $\sigma = 100 (\Omega\text{m})^{-1}$ a zatem $\tau_d \approx 10^{-12}$ s. Dla metali czas ten jest jeszcze krótszy i wynosi 10^{-15} s. Zatem wszelkie zakłócenie lokalnej neutralności elektrycznej materiału zanika wykładniczo z czasem relaksacji dielektrycznej, który zależy głównie od wartości konduktywności.

18.2. Procesy generacji i rekombinacji

Wiadomo, że w stanie równowagi termodynamicznej, w dowolnym makroskopowym obszarze półprzewodnika obowiązuje warunek neutralności dielektrycznej, tj.

$$N_d - N_a + p_o - n_o = 0, \quad (18.7)$$

przy czym n_o i p_o oznaczają koncentrację elektronów i dziur w stanie równowagi. Należy pamiętać, że dla półprzewodników w stanie równowagi $n_o p_o = n_i^2$.

Gdy teraz półprzewodnik znajduje się w stanie nierównowagi, to koncentracje nośników wynoszą

$$n = n_o + \Delta n, \quad (18.8)$$

$$p = p_o + \Delta p, \quad (18.9)$$

gdzie Δn i Δp oznaczają koncentrację dodatkowych elektronów i dziur w stanie nierównowagi. Te dodatkowe nośniki nazywa się nadmiarowymi lub nierównowagowymi.

Zajmiemy się teraz zjawiskami występującymi w półprzewodniku przy zaburzeniu stanu równowagi termodynamicznej. Stan nierównowagi może być spowodowany działaniem różnych czynników zewnętrznych. Najważniejsze z nich to:

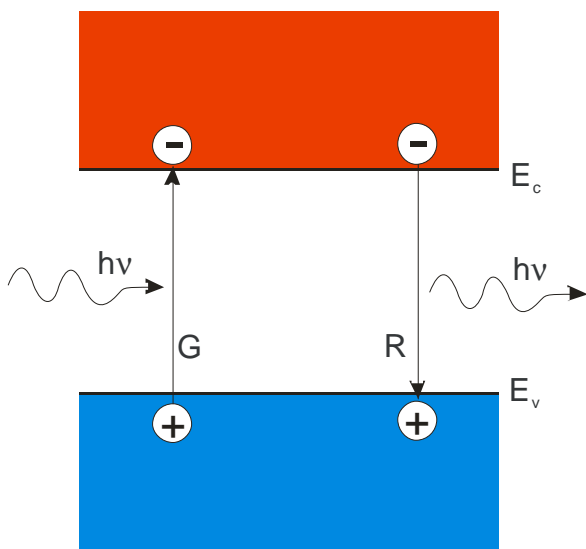
- generacja par elektron-dziura wywołana absorpcją promieniowania elektromagnetycznego o energii fotonów $\geq E_g$; zjawisko to nazywa się inaczej efektem fotoelektrycznym,
- wstrzykiwanie (iniekcja) lub wyciąganie (ekstrakcja) nośników prądu elektrycznego przez kontakt metal-półprzewodnik lub złącze p-n (np. przy odpowiedniej polaryzacji zewnętrznej złącza p-n do obszaru typu n są dostarczane dodatkowe dziury z przylegającego obszaru typu p – mamy do czynienia ze zjawiskiem wstrzykiwania dziur),
- jonizacja atomów półprzewodnika zachodząca w wyniku działania silnych pól elektrycznych.

W sposób najbardziej ogólny stan nierównowagi można określić jako taki stan, w którym

$$np \neq n_i^2. \quad (18.10)$$

Jeżeli do danego obszaru półprzewodnika nośniki są dostarczone, to $np > n_i^2$ i mówi się wówczas o wstrzykiwaniu nośników. Natomiast w przypadku gdy $np < n_i^2$, mówi się o wyciąganiu nośników.

Wróćmy jeszcze raz do generacji par elektron-dziura wywołanej przejściami międzypasmowymi pod wpływem padającego promieniowania. W tym przypadku mamy $\Delta n = \Delta p$ i warunek neutralności elektrycznej jest spełniony, a powrót do stanu równowagi zachodzi przez inny mechanizm relaksacyjny niż ten, który jest opisany w poprzednim punkcie. Proces ten, zaniku nadmiarowych par elektron-dziura, nazywa się rekombinacją nośników prądu.



Rys. 18.1. Generacja pary elektron-dziura wywołana absorpcją fotonu o energii $h\nu$, oraz bezpośrednia rekombinacja tej pary z emisją fotonu.

Najprostszym możliwym sposobem rekombinacji jest bezpośrednia "anihilacja" elektronu i dziury, w której uwolniona energia jest uniesiona przez emitowany foton (rys. 18.1). Rekombinujące nośniki muszą spełnić prawo zachowania energii i pędu. Ten typ rekombinacji nazywamy rekombinacją promienistą. W półprzewodnikach z szeroką przerwą energetyczną główny kanał rekombinacyjny związany jest z niedoskonałościami kryształu, przez które rozumie się defekty sieci, domieszki chemiczne i powierzchnię. Dla określenia tych wszystkich "słabych punktów" w kryształach poprzez które zachodzi rekombinacja, używa się ogólnego terminu: centra rekombinacji.

18.2.1. Rekombinacja bezpośrednia

Weźmy pod uwagę jednorodny półprzewodnik w którego objętości są generowane równomiernie pary elektron-dziura, np. w wyniku absorpcji światła. G jest szybkością generacji wyrażającą liczbę par elektron-dziura generowanych w jednostce czasu w jednostce objętości (ma wymiar $\text{m}^{-3}\text{s}^{-1}$). Jeżeli w kryształ nie ma pola elektrycznego ani nie występują gradienty koncentracji nośników, to wówczas zmiany koncentracji elektronów i dziur w odpowiednich pasmach opisane są równaniem

$$\frac{dn}{dt} = G - R_e \quad ; \quad \frac{dp}{dt} = G - R_h, \quad (18.11)$$

gdzie R_e i R_h są szybkościami rekombinacji (zdefiniowanymi w analogiczny sposób jak G – wymiar także $\text{m}^{-3}\text{s}^{-1}$) elektronów i dziur.

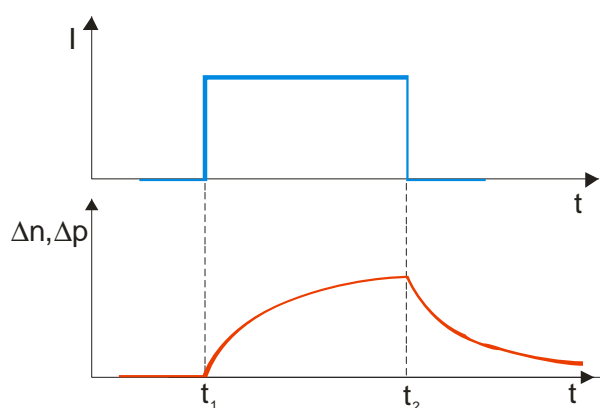
Przy bezpośredniej "anihilacji" elektronu i dziury zachodzi oczywiście $R_e = R_h = R$ gdzie R jest wówczas szybkością rekombinacji pary elektron-dziura i wówczas równania (18.11) przyjmują postać

$$\frac{dn}{dt} = \frac{dp}{dt} = G - R. \quad (18.12)$$

W najprostszym przypadku małego zaburzenia można przyjąć, że szybkość rekombinacji jest proporcjonalna do nadmiarowej koncentracji par

$$R = \frac{\Delta n}{\tau} = \frac{\Delta p}{\tau}, \quad (18.13)$$

gdzie współczynnik τ nazywa się czasem życia pary, a τ^{-1} ma wówczas znaczenie prawdopodobieństwa rekombinacji.



Rys. 18.2. Prostokątny impuls światła generujący pary elektron-dziura i fotoodpowiedź półprzewodnika (zmiana koncentracji nadmiarowych nośników prądu w funkcji czasu).

Rozpatrzmy przypadek, gdy generacja par rozpoczyna się w momencie $t = 0$ oraz spełnione są warunki początkowe: $\Delta n = \Delta p = 0$ dla $t = 0$. Rozwiązaniem równań (18.12) i (18.13) opisującym narastanie koncentracji w funkcji czasu jest wyrażenie

$$\Delta n(t) = \Delta p(t) = G\tau(1 - e^{-t/\tau}). \quad (18.14)$$

Zatem koncentracja nadmiarowych par w stanie ustalonym ($t \rightarrow \infty$) wynosi

$$\Delta n = \Delta p = G\tau.$$

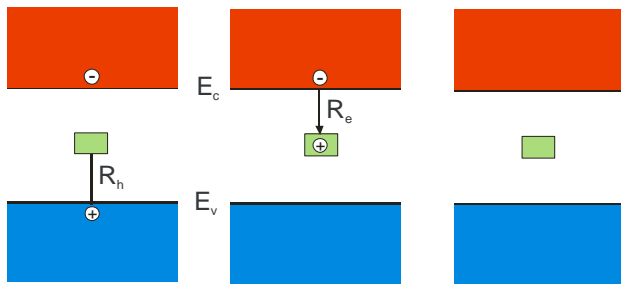
Zakładamy teraz, że kryształ znajduje się w stanie ustalonym opisanym powyższym wyrażeniem i w chwili $t = 0$ generacja ustaje. W tym przypadku rozwiązaniem równań (18.12) i (18.13) z warunkiem początkowym $\Delta n = \Delta p = G\tau$ dla $t = 0$, jest wyrażenie

$$\Delta n(t) = \Delta p(t) = G\tau \exp(-t/\tau), \quad (18.15)$$

które opisuje zanik par elektron-dziura w funkcji czasu. Sytuacja opisana powyżej występuje na przykład wtedy, gdy kryształ jest oświetlony impulsem światła generującym pary elektron-dziura. "Fotoodpowiedź" półprzewodnika w tym przypadku zobrazowana jest na rys. 18.2.

Z wyrażenia (18.15) wynika, że czas życia nośników nadmiarowych można zdefiniować jako przedział czasu, w którym koncentracja nośników nadmiarowych maleje e-krotnie.

18.2.2. Rekombinacja przez centra. Pułapkowanie



Rys. 18.3. Dwustopniowa rekombinacja pary elektron-dziura poprzez centrum, z którym związany jest poziom energetyczny leżący wewnątrz przerwy energetycznej.

Najczęściej rekombinacja nośników prądu zachodzi poprzez pewne centra, z którymi związane są zlokalizowane poziomy energetyczne wewnątrz przerwy energetycznej półprzewodnika. W takim przypadku rekombinacja nadmiarowej pary elektron-dziura jest procesem dwustopniowym. Jeden z nośników prądu (zazwyczaj nośnik mniejszościowy) jest wychwycony przez centrum jako pierwszy, a następnie zostaje wychwycony nośnik przeciwnego znaku dopełniając akt rekombinacji (rys. 18.3). Wówczas na ogół $R_e \neq R_h$ i wtedy zamiast

jednego parametru czasowego charakteryzującego proces rekombinacji należy wprowadzić dwa: czas życia elektronu w paśmie przewodnictwa τ_e i czas życia dziury w paśmie walencyjnym τ_h .

Jeżeli można przyjąć, że

$$R_e = \frac{\Delta n}{\tau_e} \quad ; \quad R_h = \frac{\Delta p}{\tau_h} \quad (18.16)$$

to wówczas narastanie koncentracji elektronów i dziur opisane jest wyrażeniami analogicznymi do (18.14)

$$\begin{aligned} \Delta n(t) &= G\tau_e \left(1 - e^{-t/\tau_e}\right), \\ \Delta p(t) &= G\tau_h \left(1 - e^{-t/\tau_h}\right). \end{aligned} \quad (18.17)$$

Zatem nadmiarowe koncentracje w stanie ustalonym będą

$$\Delta n = G\tau_e \quad ; \quad \Delta p = G\tau_h, \quad (18.18)$$

a zanik tych koncentracji po wyłączeniu generacji w chwili $t = 0$ jest opisany funkcjami

$$\begin{aligned} \Delta n(t) &= G\tau_e e^{-t/\tau_e}, \\ \Delta p(t) &= G\tau_h e^{-t/\tau_h}. \end{aligned} \quad (18.19)$$

Ponieważ w ogólności $\tau_e \neq \tau_h$, to w stanie ustalonym $\Delta n \neq \Delta p$. W takim przypadku pewna koncentracja nadmiarowych ładunków Δm , zostaje zlokalizowana (wychwycona) na centrach. Stosując konwencję w której Δm oznacza koncentrację dodatnich ładunków, otrzymuje się z warunku neutralności elektrycznej półprzewodnika

$$\Delta n = \Delta m + \Delta p.$$

Zjawisko lokalizowania nadmiarowych elektronów lub dziur na centrach nazywa się pułapkowaniem nośników prądu. Zawsze więc, gdy czasy życia elektronów i dziur różnią się między sobą, występuje zjawisko pułapkowania.

Analizę teoretyczną rekombinacji przez centra wykonali Shockley, Read oraz Hall (centra generacyjno-rekombinacyjne są nazywane często centrami SRH - od pierwszych liter nazwisk tych uczonych). W tym miejscu przytoczymy tylko jeden z najważniejszych wyników tej analizy. Okazuje się, że prawdopodobieństwo wychwytu nośnika przez centrum można określić przez podanie przekrojów czynnych na wychwyty elektronu i dziury, σ_e i σ_h , które wyrażają się poprzez czasy życia

$$\tau_e^{-1} = v_{th}\sigma_e N \quad ; \quad \tau_h^{-1} = v_{th}\sigma_h N, \quad (18.20)$$

gdzie v_{th} jest uśrednioną prędkością termiczną nośników, a N oznacza koncentrację centrów. Notowane w literaturze wartości przekrojów czynnych dla izolowanych centrów zawierają się w bardzo szerokich granicach od 10^{-12} do 10^{-20} cm².

18.2.3. Rekombinacja powierzchniowa

Na powierzchni półprzewodnika może zachodzić rekombinacja nośników prądu przez tak zwane stany powierzchniowe. Tego typu proces można scharakteryzować parametrem zwanym szybkością rekombinacji powierzchniowej S . Parametr ten można zdefiniować w sposób następujący

$$j/e = \Delta n S, \quad (18.21)$$

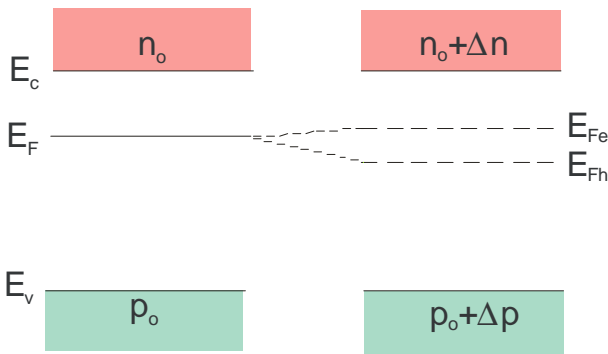
gdzie j jest gęstością prądu elektronowego lub dziurowego, jaki dopływa do powierzchni dla podtrzymania stanu ustalonego, w którym koncentracja nadmiarowych nośników prądu w objętości wynosi Δn . Prędkość rekombinacji powierzchniowej zależy w istotny sposób od obróbki powierzchni krysztalu.

18.3. Quasi-poziomy Fermiego

W stanie zakłóconej równowagi termodynamicznej rozkład koncentracji nośników w półprzewodniku nie podlega rozkładowi Fermiego-Diraca. W tych warunkach nie można więc posługiwać się pojęciem poziomu Fermiego. Okazuje się jednak, że nierównowagowe koncentracje elektronów i dziur można opisać formalnie analogicznymi wyrażeniami, jakie były wprowadzone dla koncentracji w stanie równowagi termodynamicznej. Należy wówczas zamiast jednego parametru występującego w wyrażeniach (17.5) i (17.11) - poziomu Fermiego E_F wprowadzić dwa niezależne parametry: quasi-poziom Fermiego dla elektronów E_{Fe} i quasi-poziom Fermiego dla dziur E_{Fh} . Można wówczas napisać

$$n = n_o + \Delta n = N_c \exp[-(E_c - E_{Fe})/kT], \quad (18.22)$$

$$p = p_o + \Delta p = N_v \exp(-E_{Fh}/kT). \quad (18.23)$$



Rys. 18.4. Quasi poziomy Fermiego w przypadku występowania nadmiarowych elektronów i dziur.

Wyrażenia te można traktować jako definicję wielkości E_{Fe} i E_{Fh} . Ich sens jest jednak znacznie głębszy. Przy ustalonej równowadze między nośnikami prądu w obrębie każdego z pasm E_{Fe} i E_{Fh} mają termodynamiczne znaczenie potencjałów chemicznych, odpowiednio dla elektronów i dziur. Po ustaleniu się równowagi między nośnikami prądu w obu pasmach zachodzi $E_{Fe} = E_{Fh} = E_F$, to znaczy równowaga termodynamiczna całego układu została przywrócona.

Zastanowimy się teraz jak będą ułożone quasi-poziomy Fermiego elektronów i dziur w stosunku do równowagowego poziomu Fermiego, po wprowadzeniu do półprzewodnika nadmiarowej koncentracji nośników prądu. Zakładając, że $\Delta n = \Delta p$ otrzymujemy z wyrażen (18.22) i (17.5) dla elektronów

$$E_{Fe} - E_F = kT \ln\left(1 + \frac{\Delta n}{n_o}\right), \quad (18.24)$$

$$E_{Fh} - E_F = -kT \ln\left(1 + \frac{\Delta p}{p_o}\right). \quad (18.25)$$

Biorąc pod uwagę materiał typu n w którym $n_o \gg p_o$ widać, że w wyniku generacji nadmiarowych par elektron-dziura, quasi-poziom Fermiego dla dziur zostaje zepchnięty silnie

w dół w stosunku do E_F , podczas gdy quasi-poziom Fermiego dla elektronów przesuwają się jedynie nieznacznie w górę (rys. 18.4). Dla materiału typu p sytuacja będzie odwrotna.

18.4. Transport nośników w półprzewodnikach

Transport nierównowagowych nośników prądu jest opisany trzema układami równań:

- równaniami gęstości prądów,
- równaniami ciągłości strumieni nośników prądu,
- równaniem Poissona (18.1).

Na powyższych równaniach opiera się tzw. fenomenologiczna teoria transportu sformułowana w 1950 r przez Roosbroecka. Teoria ta jest słuszna dla znacznej większości materiałów stosowanych w przyrządach półprzewodnikowych.

18.4.1. Prądy unoszenia i dyfuzji

Z rozważań przeprowadzonych w rozdziale 7 wiemy, że gęstość prądu elektrycznego można zapisać w postaci sumy gęstości prądu przewodzenia (ruch nośników) j_c i prądu przesunięcia

$$j = j_c + \epsilon_o \epsilon_r \frac{\partial E}{\partial t}. \quad (18.26)$$

Okazuje się, że prąd przesunięcia w metalach i półprzewodnikach nie odgrywa istotnej roli w bardzo szerokim zakresie częstotliwości sygnałów elektrycznych, bowiem szybkość relaksacji następującej po zakłóceniu lokalnej objętości ładunku elektrycznego określona jest czasem relaksacji dielektrycznej. Stąd wynika, że dla częstotliwości poniżej 10^{12} Hz składową prądu przesunięcia można pominąć. Pozostaje więc do rozpatrzenia składowa prądu przewodzenia.

W poprzednim rozdziale wykazaliśmy, że gęstość prądu unoszenia tworzonego przez strumień nośników poruszających się wskutek działania siły elektrycznej jest równy

$$j_u = env_d = en\mu E.$$

W przypadku udziału obu rodzajów nośników w prądzie unoszenia

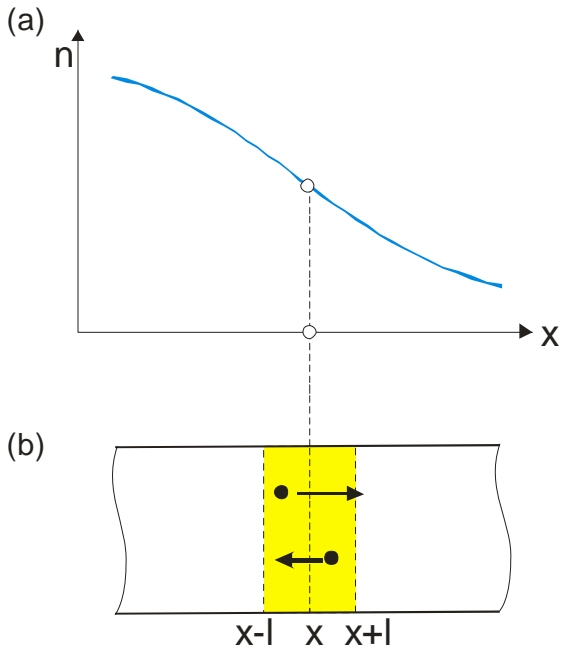
$$j_u = |e|(n\mu_e + p\mu_h)E = \sigma E, \quad (18.27)$$

a stąd

$$\sigma = |e|(n\mu_e + p\mu_h). \quad (18.28)$$

W tym przypadku założyliśmy równomierny rozkład koncentracji nośników w półprzewodniku.

Okazuje się jednak, że w półprzewodniku często na skutek działania czynników zewnętrznych (w przyrządach półprzewodnikowych) rozkład koncentracji nośników prądu jest nierównomierny. Wskutek chaotycznego ruchu nośników następuje proces wyrównywania się koncentracji co w konsekwencji prowadzi do pojawienia się prądu dyfuzyjnego.



Rys. 18.5. Dyfuzyjny przepływ nośników: (a) koncentracja elektronów, (b) strumienie elektronów przechodzące przez płaszczyznę x .

W celu wyznaczenia gęstości prądu dyfuzyjnego posłużymy się modelem jednowymiarowym. Założymy, że rozkład koncentracji elektronów w półprzewodniku jest opisany funkcją $n(x)$ (rys. 18.5a) oraz, że wszystkie ruchy elektronów odbywają się w kierunku równoległym do osi x . Rozpatrzmy prostopadłościenną próbkę półprzewodnika o przekroju jednostkowym (rys. 18.5b). Liczba elektronów przechodzących przez płaszczyznę x w kierunku zgodnym z kierunkiem osi x w czasie τ równym średniemu czasowi między kolejnymi zderzeniami elektronu z siecią krystaliczną wynosi $n(x-l)l/2$, gdzie l – średnia droga swobodna elektronu, a $n(x-l)$ – koncentracja elektronów w płaszczyźnie $x-l$. Czynnikiem $1/2$ jest wynikiem uwzględnienia faktu, że połowa elektronów znajdujących się po lewej stronie płaszczyzny x porusza się w prawo, druga zaś w lewo.

Analogiczna liczba elektronów przechodzących przez rozpatrywaną płaszczyznę w przeciwnym kierunku wynosi $n(x+l)l/2$. Gęstość elektronowego prądu dyfuzyjnego jest proporcjonalna do różnicy tych liczb. Uwzględniając to, że elektron ma ładunek $-e$, znajdujemy, że szukana gęstość wynosi

$$j_{ed}(x) = \frac{el}{2\tau} [n(x+l) - n(x-l)]. \quad (18.29)$$

Przyjmując, że w rozważanym obszarze półprzewodnika rozkład koncentracji elektronów może być aproksymowany funkcją liniową współrzędnej x , możemy napisać

$$n(x+l) = n(x) + \frac{dn}{dx}\bigg|_x l,$$

$$n(x-l) = n(x) - \frac{dn}{dx}\bigg|_x l,$$

co po podstawieniu do równania (18.29) daje

$$j_{ed}(x) = \frac{el^2}{\tau} \frac{dn}{dx}\bigg|_x.$$

Średnia droga swobodna elektronu jest związana z jego prędkością termiczną v_{th} zależnością

$$l = v_{th} \tau,$$

co prowadzi do zależności

$$j_{ed}(x) = ev_{th}^2 \tau \frac{dn}{dx} \Big|_x. \quad (18.30)$$

Gęstość elektronowego prądu dyfuzyjnego jest więc proporcjonalna do gradientu koncentracji nośników, przy czym współczynnik proporcjonalności jest zależny od intensywności ruchów termicznych nośników w półprzewodniku.

Wzór (18.30) zapisuje się zwykle w postaci,

$$j_{ed} = eD_e \frac{dn}{dx}, \quad (18.31)$$

przy czym D_e nosi nazwę współczynnika dyfuzji elektronów. W rozpatrywanym tu modelu jednowymiarowym współczynnik ten równa się $v_{th}^2 \tau$.

Jeżeli uwzględnimy, że ruchy termiczne elektronów odbywają się we wszystkich trzech wymiarach, to otrzymamy nieco inną postać wzoru na D_e , to jest

$$D_e = \frac{1}{3} v_{th}^2 \tau.$$

Ponieważ na jeden stopień swobody cząstki przypada średnio energia $kT/2$, więc na trzy stopnie elektronu przypada $3kT/2$, co można wyrazić wzorem

$$\frac{1}{2} m_e^* v_{th}^2 = \frac{3}{2} kT,$$

czyli

$$v_{th} = \sqrt{\frac{3kT}{m_e^*}}. \quad (18.32)$$

Opierając się na tej zależności, można otrzymać

$$D_e = \frac{e\tau}{m_e^*} \frac{kT}{e}.$$

Pierwszy czynnik w tym wyrażeniu jest równy ruchliwości elektronów [zob. wzór (17.53)]. Stąd

$$D_e = \frac{kT}{e} \mu_e. \quad (18.33)$$

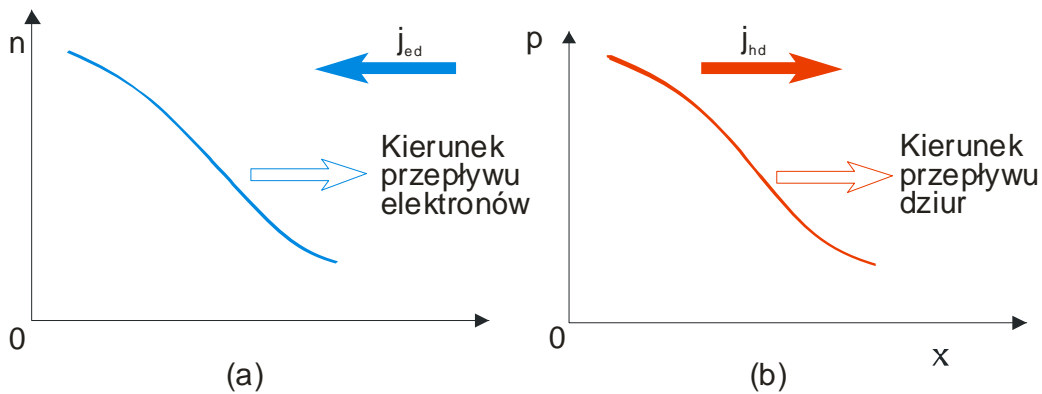
Związek ten i analogiczna zależność dla dziur, noszą nazwę relacji Einsteina. Ze wzoru (18.33) wynika w sposób oczywisty zależność współczynnika dyfuzji od rodzaju przewodnictwa, koncentracji domieszek i temperatury. Wzór ten jest słuszny dla dowolnego

półprzewodnika niezdegenerowanego, a w odniesieniu do nośników mniejszościowych – także dla półprzewodnika zdegenerowanego.

Analogiczne rozumowanie prowadzi do następującego wzoru dla gęstości prądu dyfuzyjnego dziur

$$j_{hd} = -eD_h \frac{dp}{dx}. \quad (18.34)$$

Różnica znaku wynika z różnicy znaków ładunku elektronów i dziur. Jak pokazano na rys. 18.6, zarówno elektrony jak i dziury poruszają się w kierunku ich malejących koncentracji. Prąd dziurowy płynie jednakże w przeciwnym kierunku niż prąd elektronowy



Rys.18.6. Kierunek prądu wywołanego dyfuzją elektronów (a) i dziur (b).

W ogólnym przypadku jeżeli prędkości ruchów uporządkowanych (pod wpływem pola elektrycznego) są dostatecznie małe w stosunku do prędkości termicznych, to procesy unoszenia i dyfuzji nośników zachodzą niezależnie od siebie. W tych warunkach gęstości prądu elektronowego i prądu dziurowego wynoszą odpowiednio

$$j_e = en\mu_e E + eD_e \frac{dn}{dx}, \quad (18.35)$$

$$j_h = ep\mu_h E - eD_h \frac{dp}{dx}. \quad (18.36)$$

Prąd całkowity jest więc sumą składowych prądu unoszenia i prądu dyfuzyjnego. Dla przypadku trójwymiarowego

$$j_e = en\mu_e E + eD_e \text{grad}n, \quad (18.37)$$

$$j_h = ep\mu_h E - eD_h \text{grad}p. \quad (18.38)$$

Są to więc równania gęstości prądów elektronowego i dziurowego.

18.4.2. Równanie ciągłości

Jednym z ważnych wniosków wypływających z równań Maxwella jest równanie ciągłości. Równanie to jest wyrazem podstawowego prawa przyrody – zasady zachowania ładunku. Orzeka ono, że ładunek elektryczny dowolnego układu pozostaje niezmienny w czasie, jeżeli układ jest elektrycznie odosobniony. Ładunek ten może więc zmienić się tylko wtedy, gdy przez powierzchnię otaczającą układ przepływa prąd elektryczny.

Z równania (7.21) mamy

$$\operatorname{rot} \vec{H} = \frac{\partial \vec{D}}{\partial t} + \vec{j} \quad (18.39)$$

Ponieważ dla każdego pola wektorowego dywergencja rotacji jest równa zero, zatem na mocy równania (18.39) mamy

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

A ponieważ zgodnie z równaniem (7.22)

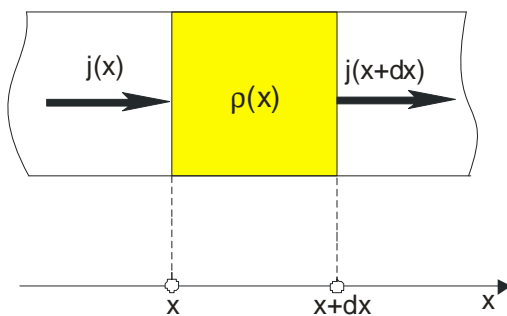
$$\operatorname{div} \vec{D} = \rho,$$

więc

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{j} = 0. \quad (18.40)$$

Jest to równanie ciągłości w postaci różniczkowej.

Dla przepływu jednowymiarowego równanie ciągłości można wyprowadzić na podstawie rys. 18.7. Weźmy pod uwagę zaznaczony na rysunku obszar ograniczony płaszczyznami x oraz $x+dx$ i oznaczmy symbolem A pole przekroju poprzecznego. Prąd wpływający do tego obszaru jest równy $Aj(x)$, a prąd wypływający z niego $-Aj(x+dx)$. Z kolei ładunek zawarty w tym obszarze wynosi $A\rho(x)dx$. Ze względu na to, że ładunki elektryczne nie mogą ginąć ani też powstać z niczego, szybkość zmian ładunku zawartego w danym obszarze musi być równa różnicy prądów wpływającego i wypływającego



$$A \frac{\partial \rho(x)}{\partial t} dx = A[j(x) - j(x+dx)] \quad (18.41)$$

z którego wynika

$$\frac{\partial \rho}{\partial t} = -\frac{\partial j}{\partial x}. \quad (18.42)$$

Rys. 18.7. Jednowymiarowy przepływ prądu w półprzewodniku.

Otrzymany wzór jest równoważny wyrażeniu (18.40) dla przypadku jednowymiarowego.

Równanie ciągłości odnosi się do całkowitego ładunku elektrycznego znajdującego się w rozważanym obszarze. Nie można go natomiast stosować w tej formie dla jednego rodzaju

nośników w półprzewodniku. Chcąc stosować równanie ciągłości dla jednego rodzaju nośników, np. dla dziur, należy w nim uwzględnić wszystkie mechanizmy mające wpływ na ilość tych nośników. Zamiast równania (18.41) musimy więc napisać (zob. rys. 18.8)

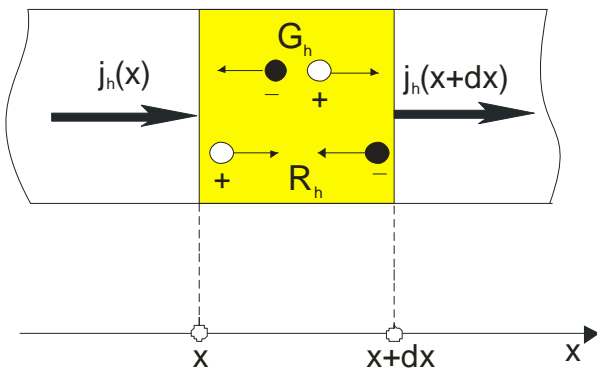
$$Ae \frac{\partial p}{\partial t} dx = A[j_h(x) - j_h(x + dx)] + Ae(G_h - R_h)dx. \quad (18.43)$$

Stąd otrzymuje się równanie ciągłości dla dziur

$$\frac{\partial p}{\partial t} = -\frac{1}{e} \frac{\partial j_h}{\partial x} + G_h - R_h, \quad (18.44)$$

gdzie G_h i R_h oznaczają odpowiednio szybkości generacji i rekombinacji dziur. Szybkości te różnią się na ogół od analogicznych wielkości G_e i R_e dla elektronów. Równanie ciągłości dla elektronów ma zatem postać

$$\frac{\partial n}{\partial t} = \frac{1}{e} \frac{\partial j_e}{\partial x} + G_e - R_e. \quad (18.45)$$



Rys. 18.8. Czynniki wpływające na koncentrację dziur przy przepływie jednowymiarowym.

Dla uświadomienia sobie ogólnych prawidłowości zjawiska, rozpatrzmy jednowymiarowy przypadek transportu nośników prądu w długiej belce półprzewodnikowej, do której przyłożona jest różnica potencjałów. Zakładamy dla uproszczenia, że $G_e = G_h = G$ i $R_e = R_h = R$ oraz pomijamy ewentualną rekombinację na powierzchni. Wówczas równania ciągłości (18.44) i (18.45) z uwzględnieniem wyrażeń na prądy (18.37) i (18.38) przyjmują postać:

$$\frac{\partial n}{\partial t} = G - R + \mu_e \frac{\partial}{\partial x}(nE) + D_e \frac{\partial^2 n}{\partial x^2}, \quad (18.46)$$

$$\frac{\partial p}{\partial t} = G - R - \mu_h \frac{\partial}{\partial x}(pE) + D_h \frac{\partial^2 p}{\partial x^2}, \quad (18.47)$$

gdzie $E = E(x)$ jest natężeniem pola wzdłuż próbki. Ponieważ w nieobecności nierównowagowych nośników lokalna neutralność elektryczna półprzewodnika jest zachowana, to równanie Poissona można napisać następująco

$$\frac{dE}{dx} = \frac{e}{\epsilon_0 \epsilon_r} (\Delta p - \Delta n). \quad (18.48)$$

Analityczne rozwiązanie układu równań (18.46)-(18.48) staje się możliwe przy upraszczającym założeniu quasi-neutralności półprzewodnika. Jeśli nierównowagowe koncentracje nośników są małe, na przykład $\Delta n \approx \Delta p \ll n$ w materiale typu n, to można przyjąć, że ładunek nadmiarowych nośników mniejszościowych jest dokładnie lokalnie skompensowany ładunkiem nadmiarowych nośników większościowych. Wówczas $\Delta n = \Delta p$ i z równania (18.48) otrzymujemy $E = \text{const}$. Tym samym została zaniedbana lokalna modyfikacja pola elektrycznego w kryształach wywołana wprowadzonymi nośnikami prądu.

Dla tego przypadku równania (18.46) i (18.47) można zapisać w postaci

$$\frac{\partial(\Delta n)}{\partial t} = G - R + \mu_e E \frac{\partial(\Delta n)}{\partial x} + D_e \frac{\partial^2(\Delta n)}{\partial x^2}, \quad (18.49)$$

$$\frac{\partial(\Delta p)}{\partial t} = G - R - \mu_h E \frac{\partial(\Delta p)}{\partial x} + D_h \frac{\partial^2(\Delta p)}{\partial x^2}. \quad (18.50)$$

Wykorzystaliśmy tu fakt, że w wyrażeniach różniczkowych można opuścić stałe koncentracje równowagowe. Jeżeli pomnożymy stronami równanie (18.49) przez $\mu_h p$, równanie (18.50) przez $\mu_e n$, a następnie dodamy je do siebie, otrzymujemy wówczas tak zwane równanie ambipolarne

$$\frac{\partial(\Delta n)}{\partial t} = G - R + \frac{p-n}{(n/\mu_h) + (p/\mu_e)} E \frac{\partial(\Delta n)}{\partial x} + \frac{p+n}{(p/D_e) + (n/D_h)} \frac{\partial^2(\Delta n)}{\partial x^2}. \quad (18.51)$$

Uwzględniliśmy tu równość $\Delta n = \Delta p$ oraz zależność Einsteina. Otrzymane w ten sposób równanie ambipolarne (18.51) ma analogiczną postać jak równanie ciągłości (18.49) opisujące transport nierównowagowych elektronów, a różni się od niego jedynie wartościami współczynników występujących przy odpowiednich pochodnych przestrzennych. W miejscu występującej w równaniu (18.49) ruchliwości μ_e , pojawiła się w równaniu (18.51) wielkość

$$\mu^* = \frac{p-n}{\frac{n}{\mu_e} + \frac{p}{\mu_h}}, \quad (18.52)$$

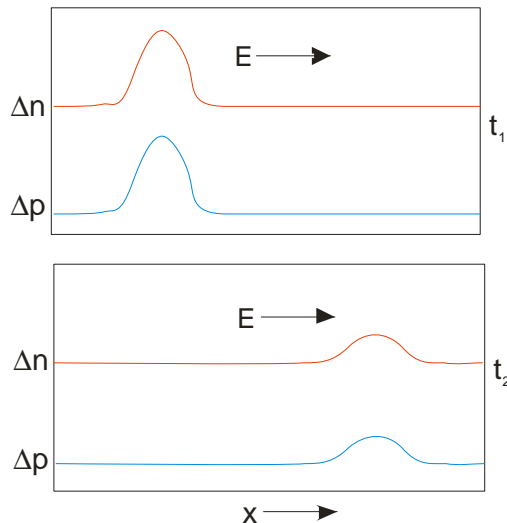
natomiast w miejscu D_e występuje wyrażenie

$$D^* = \frac{p+n}{\frac{p}{D_e} + \frac{n}{D_h}}. \quad (18.53)$$

Tak więc, transport nierównowagowej koncentracji nośników prądu w półprzewodniku jest opisany takimi samymi równaniami jak ruch elektronów, z tym że należy przypisać mu ambipolarną ruchliwość μ^* [równanie (18.52)] oraz ambipolarny współczynnik dyfuzji D^* [równanie (18.53)], które są w ogólności funkcjami koncentracji nośników prądu.

18.5. Dyskusja ambipolarnego równania transportu

Rozpatrzmy przypadek, gdy w długiej próbce półprzewodnikowej, w której istnieje pole E , w niewielkim obszarze między x_l a x_l+dx zostały wygenerowane nadmiarowe pary elektron-dziura. "Chmura" nadmiarowych nośników prądu jest unoszona polem elektrycznym wzdłuż próbki z prędkością μ^*E , a równocześnie rozptywa się dzięki dyfuzji oraz zanika w czasie na skutek rekombinacji (rys. 18.9).



Rys. 18.9. Unoszenie chmury nośników prądu polem elektrycznym w półprzewodniku

Przy małych nadmiarowych koncentracjach dla materiału typu n, gdy $n \gg p$, otrzymujemy z równań (18.52) i (18.53)

$$\mu^* \approx -\mu_h \quad ; \quad D^* \approx D_h$$

natomiast dla materiału typu p, gdy $p \gg n$

$$\mu^* \approx \mu_e \quad ; \quad D^* \approx D_e.$$

W przypadku półprzewodnika samoistnego gdzie $n = p$

$$\mu^* = 0 \quad ; \quad D^* = \frac{2D_e D_h}{D_e + D_h}.$$

Z powyższych relacji wynika, że w półprzewodniku niesamoistnym chmura nadmiarowych nośników prądu porusza się tak jak nośniki mniejszościowe, natomiast w materiale samoistnym chmura nadmiarowych nośników prądu zawiera ładunki obydwu znaków; pole elektryczne działa na nie w przeciwnych kierunkach. Wytwarza się pewna polaryzacja ładunków w wyniku czego zmniejsza się nieco natężenie pola w obszarze samej chmury. W przypadku małych nadmiarowych koncentracji można zaniedbać niedokładnie skompensowany ładunek i związane z tym zakłócenie pola elektrycznego stosując przybliżenie quasi-neutralności półprzewodnika. Im większa jest koncentracja nośników większościowych, tym łatwiej dopasowują się one do lokalnego odchylenia koncentracji nośników mniejszościowych i tym dokładniej kompensują ich ładunek. Ruch chmury nośników nierównowagowych w domieszkowanym półprzewodniku zachodzi zatem tak jak ruch nośnika mniejszościowego.

18.6. Dyfuzja nadmiarowych nośników prądu

Przypadek "czystej" dyfuzji nadmiarowych nośników prądu następuje przy jednorodnej generacji i nieobecności zewnętrznego pola elektrycznego. Weźmy pod uwagę półprzewodnik, którego jedna z płaskich powierzchni oświetlona jest silnie absorbowanym promieniowaniem, generującym tuż przy powierzchni nadmiarowe pary elektron-dziura (rys. 18.10). Jeżeli generacja jest jednorodna na powierzchni, to zagadnienie można uważać za jednowymiarowe.

Ambipolarne równanie transportu przyjmuje postać

$$\frac{\partial(\Delta n)}{\partial t} = D^* \frac{\partial^2(\Delta n)}{\partial x^2} - R. \quad (18.54)$$

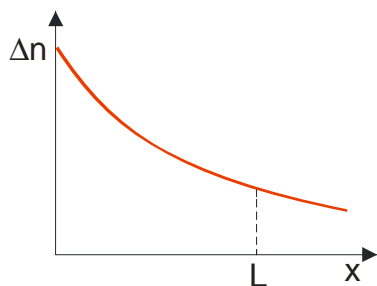
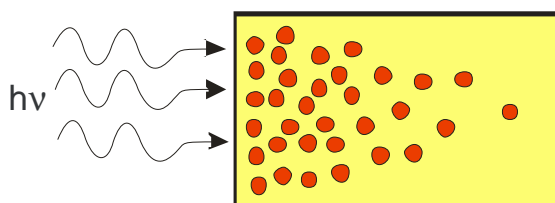
Dla stanu ustalonego, przyjmując $R = \Delta n/\tau$ równanie to sprowadza się do następującego

$$\frac{d^2(\Delta n)}{dx^2} - \frac{\Delta n}{D^* \tau} = 0, \quad (18.55)$$

którego rozwiązaniem mającym sens fizyczny jest funkcja

$$\Delta n = \Delta p = \Delta p(0) \exp\left[-\frac{x}{(D^* \tau)^{1/2}}\right], \quad (18.56)$$

gdzie $\Delta p(0)$ oznacza nadmiarową koncentrację par nośników prądu przy oświetlonej powierzchni kryształu (rys. 18.10a). Występujący w wyrażeniu parametr $L = (D^* \tau)^{1/2}$ nazywa się średnią drogą dyfuzji nośników prądu.



Rys. 18.10. Dyfuzja nadmiarowych nośników prądu w głąb półprzewodnika w przypadku jednowymiarowym.

Pojęcie drogi dyfuzji ma następujące znaczenie fizyczne – jest to odległość na której koncentracja nośników nadmiarowych maleje e-krotnie. W przypadku silnie domieszkowanych półprzewodników otrzymuje się, że dla $n \gg p$

$$L = L_h = (D_h \tau)^{1/2},$$

oraz dla $p \gg n$

$$L = L_e = (D_e \tau)^{1/2}.$$

Dyfuzja nadmiarowych koncentracji zachodzi wówczas tak jak dyfuzja nośników mniejszościowych.

18.7. Półprzewodnik niejednorodny

Półprzewodnikiem niejednorodnym nazywamy półprzewodnik o nierównomiernym rozkładzie koncentracji domieszek. Taka sytuacja jest typowa w wielu przyrządach półprzewodnikowych. W stanie równowagi termodynamicznej prądy dziurowy i elektronowy są równe zeru.

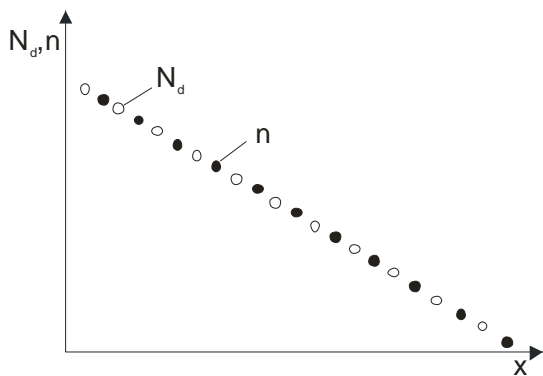
Omówione wcześniej równanie gęstości prądu elektronów i dziur są słuszne dla dowolnych rozkładów koncentracji domieszek. W przypadku półprzewodnika w którym koncentracja domieszki zmienia się w jednym kierunku x , równania te przyjmują postać

$$j_e = e\mu_e n(x)E + eD_e \frac{dn(x)}{dx}, \quad (18.57)$$

$$j_h = e\mu_h p(x)E - eD_h \frac{dp(x)}{dx}. \quad (18.58)$$

Należy tu jednak pamiętać o tym, że w półprzewodniku niejednorodnym zarówno współczynniki dyfuzji D_e i D_h , jak i ruchliwości μ_e i μ_h zmieniają się wraz ze zmianą koncentracji domieszek. Występujące w tych wzorach pole elektryczne E w ogólności składa się z pola wewnętrznego (związanego z niejednorodnością) i pola od napięcia doprowadzonego z zewnątrz.

W celu uproszczenia dalszych równań zakładamy, że w danym obszarze niejednorodnym współczynniki dyfuzji i ruchliwości są stałe.



Rys. 18.11. Rozkład koncentracji elektronów w niejednorodnym półprzewodniku typu n w warunkach równowagi termodynamicznej.

Rozpatrzmy dla przykładu półprzewodnik typu n, w którym $n \gg p$, zaś koncentracja domieszki donorowej N_d zmienia się wzdłuż osi x , tak jak na rys. 18.11. Jeżeli założyć, że wszystkie atomy domieszki są zjonizowane, wówczas rozkład elektronów w półprzewodniku jest taki jak rozkład domieszki [$n(x) = N_d(x)$]. Powstały rozkład koncentracji nośników większościowych wywołuje prąd dyfuzji tych nośników

$$j_e = eD_e \frac{dn(x)}{dx} ..$$

W warunkach równowagi termodynamicznej ten prąd dyfuzyjny musi być równoważony przez równy mu i przeciwnie skierowany prąd unoszenia

$$j = en(x)\mu_e E(x), \quad (18.59)$$

co wymaga powstania odpowiedniego pola wewnętrznego $E_w(x)$. Zatem

$$en(x)\mu_e E_w(x) = eD_e \frac{dn(x)}{dx}. \quad (18.60)$$

Uwzględnivszy przyjęte wcześniej założenia, że $n(x) = N_d(x)$ oraz relacje Einsteina, z równania (18.60) otrzymujemy wartość pola wewnętrznego

$$E_w(x) = -\frac{kT}{e} \frac{1}{N_d(x)} \frac{dN_d(x)}{dx}. \quad (18.61)$$

Pole to ma kierunek przeciwny do gradientu koncentracji domieszki w półprzewodniku typu n i zgodny w półprzewodniku typu p.

ROZDZIAŁ 19

FIZYKA JĄDROWA

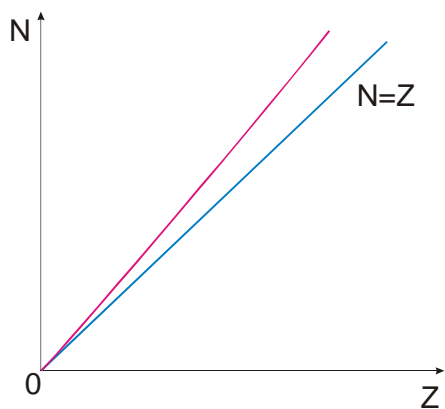
Fizyka jądrowa jest nauką o budowie jądra atomowego, jego przemianach promieniotwórczych i jego oddziaływaniach z innymi jądrami i cząstkami. Rozdział ten poświęcamy głównie wyjaśnieniu struktury i właściwości jąder atomowych. Przedstawimy też praktyczne zastosowania oddziaływań jądrowych, takich jak reakcji rozszczepienia i syntezy jądrowej. Opierając się na prawach podstawowych będziemy próbowali przewidzieć rozmiary, energię wiązania, poziomy energetyczne i inne właściwości jąder atomowych.

19.1. Pojęcia wstępne

Jądra można podzielić na *dwie grupy: jądra trwałe, czyli stabilne i jądra nietrwałe, czyli promieniotwórcze*. Każde jądro składa się z protonów i neutronów związanych siłami jądrowymi. Ponieważ neutron i proton mają prawie taką samą masę i bardzo zbliżone właściwości, więc obydwie określa się nazwą nukleon ($m_n = 1,674 \times 10^{-27}$ kg; $m_p = 1,672 \times 10^{-27}$ kg, czyli $m_p = 1836,1m_e$ i $m_n = 1838,6m_e$). Nazwa nuklid jest stosowana dla każdego jądra większego niż nukleon.

Nuklid pierwiastka o ogólnym symbolu chemicznym X , mający liczbę nukleonów A , w tym liczbę protonów Z , oznaczamy symbolem ${}^A_Z X$. *Nuklidy o tej samej liczbie Z noszą nazwę izotopów*, a o tej samej liczbie A – izobarów. Jądra zawierające jednakowe liczby neutronów, czyli $N = A - Z = \text{const}$, nazywamy izotonami. *Liczbę Z nazywamy liczbą atomową lub liczbą porządkową, a liczbę A – liczbą masową*. W miarę wzrostu liczby atomowej Z , liczba neutronów coraz bardziej przewyższa liczbę protonów w jądrami kolejnych pierwiastków układu okresowego (patrz rys. 19.1). Wartość liczby A dla jądra atomowego jest bardzo bliska masie odpowiadającego mu atomu. Nuklid ${}^{12}_6 C$ został wybrany dla zdefiniowania jednostki masy atomowej (j.m.a. = $1,66 \times 10^{-27}$ kg) – wynosi ona dokładnie 12. Dlatego mówimy, że skala masy atomowej opiera się na jądrze ${}^{12}_6 C$. Jednostka masy atomowej, w przeliczeniu na energię zgodnie ze wzorem Einsteina $E = mc^2$, jest równoważna energii 931 MeV.

Znanych jest obecnie 105 pierwiastków chemicznych, z których przeważająca większość posiada izotopy. Istnieje około 300 izotopów trwałych i powyżej 1000 nietrwałych (promieniotwórczych).



Rys. 19.1. Liczba neutronów N w miarę wzrostu liczby atomowej Z coraz bardziej przewyższa liczbę protonów w jądrach kolejnych pierwiastków.

Nukleony posiadają spin połówkowy, tj. $\frac{1}{2}\hbar$. Spin jądra złożonego z parzystej liczby nukleonów jest liczbą całkowitą (w jednostkach \hbar). Natomiast spin jądra złożonego z nieparzystej liczby cząstek jest liczbą połówkową (w jednostkach \hbar). Jądra atomowe oprócz spinu mają także momenty magnetyczne. Jednostką momentu magnetycznego jądra jest wprowadzony, analogicznie do magnetronu Bohra, tzw. magnetron jądrowy

$$\mu_n = \frac{e\hbar}{2m_p} \approx 10^{-3} \mu_B, \quad (19.1)$$

który jest zdefiniowany takim samym wzorem jak magnetron Bohra, tylko masa elektronu m_e jest zastąpiona masą protonu m_p . Magnetron jądrowy jest $m_p/m_e \approx 1836$ razy mniejszy od wartości magnetronu Bohra.

19.2. Rozmiar jądra

Najprostszym sposobem określenia rozmiaru i rozkładu masy jądra jest rozpraszanie dyfrakcyjne wysokoenergetycznych cząstek. Wynikiem doświadczeń Rutherforda (1911 r.) polegających na rozpraszaniu cząstek α przez cienkie folie złota, było odkrycie jądra atomowego i oszacowanie jego rozmiarów (promień mniejszy od 10^{-14} m). Później jako wiązki o małej długości fali stosowano wysokoenergetyczne protony i neutrony. Wyniki tych pomiarów wykazują, że średni promień dla wszystkich jąder oprócz najmniejszych jest dany wzorem

$$R \approx (1.2 \times 10^{-15} \text{ m}) A^{1/3}. \quad (19.2)$$

W fizyce jądrowej i fizyce cząstek elementarnych wielkość 10^{-15} m pojawia się często i wprowadzono dla niej nazwę fermi.

Jeżeli potraktujemy jądro jako kroplę cieczy jądrowej to liczba cząstek na jednostkę objętości wynosi

$$N = \frac{A}{(4/3)\pi R^3} = 1.38 \times 10^{44} \frac{\text{nukleonów}}{\text{m}^3}. \quad (19.3)$$

Gęstość masy to iloczyn tej liczby i masy nukleonu

$$\rho = Nm_p = (1.38 \times 10^{44}) (1.67 \times 10^{-27}) \text{ kg/m}^3 = 2.3 \times 10^{17} \text{ kg/m}^3. \quad (19.4)$$

Wynika stąd, że 1 cm^3 materii jądrowej ważyłby 230 milionów ton. Zauważmy również, że gęstość materii jądrowej nie zależy od rozmiarów jądra, ponieważ jego objętość jest proporcjonalna do liczby masowej A .

19.3. Energia wiązania jądra. Defekt masy

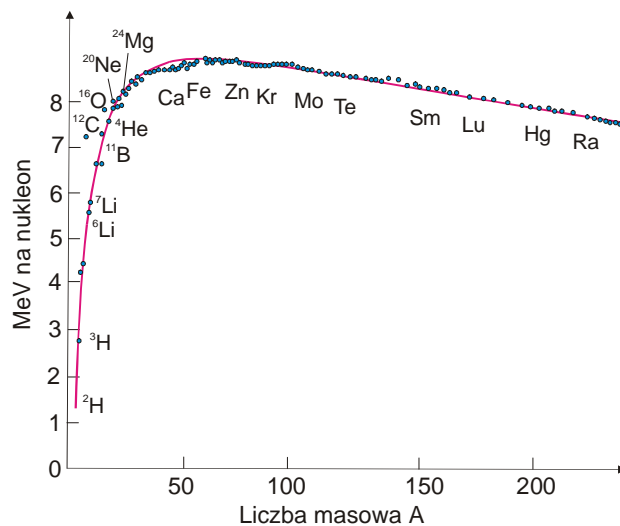
Jądro można sobie wyobrazić jako kulistą, zwartą grupę neutronów i protonów. Odpychającym siłom kulombowskim między protonami przeciwdziałają silne, specyficzne siły jądrowe, które działają między wszystkimi nukleonami. Związana jest z nimi energia wiązania jądra. Energią wiązania nukleonu w jądrze nazywamy pracę potrzebną na usunięcie danego nukleonu z jądra bez nadania mu energii kinetycznej. *Całkowita energia wiązania jądra jest określona jako praca potrzebna na rozłożenie jądra na jego składowe nukleony bez nadania im energii kinetycznej.*

Energię wiązania jądra określimy za pomocą "niedoboru masy" (defektu masy) jądra. Dokładne pomiary mas atomowych poszczególnych nukleonów pokazały, że ich masa jest zawsze większa o kilka dziesiątych procent od masy jądra utworzonego z tych nukleonów.

Zgodnie ze wzorem Einsteina

$$E_w = \Delta M c^2 = c^2 [Zm_p + (A - Z)m_n - M_j], \quad (19.5)$$

gdzie ΔM jest defektem masy, czyli różnicą między sumą mas swobodnych nukleonów i masą jądra.



Rys. 19.2. Energia wiązania przypadająca na jeden nukleon w funkcji liczby masowej.

Średnią energię wiązania jednego nukleonu w jądrze otrzymamy dzieląc całkowitą energię wiązania jądra E_w przez liczbę nukleonów A zawartych w jądrze. Wyniki doświadczalne pokazane są na rys. 19.2. Widzimy, że krzywa osiąga płaskie maksimum o wysokości 8,7 MeV w pobliżu $A = 60$, opada natomiast dla mniejszych i większych wartości A , przy czym spadek

dla malejących A nie jest monotoniczny, lecz wykazuje ostre maksimum dla liczb masowych będących wielokrotnościami 4 (4, 12, 16,.....). Występują wzrosty energii wiązania dla magicznych liczb nukleonów: 2, 8, 20, 28, 50, 82, 126, 152 (Z lub N).

Zauważono, że wśród nuklidów trwałych (niepromieniotwórczych) przeważają jądra parzysto-parzyste, tzn. o parzystych liczbach zarówno protonów jak i neutronów, co jest uwidocznione w Tabeli 19.1.

Tabela 19.1. Rozkład trwałych nuklidów według rodzaju ich parzystości

A	N	Z	Liczba trwałych nuklidów
parzyste	parzyste	parzyste	161
	nieparzyste	nieparzyste	4
nieparzyste	parzyste	nieparzyste	50
	nieparzyste	parzyste	54

19.4. Modele struktury jądra atomowego

Powstało szereg modeli jąder atomowych, stanowiących uproszczone teorie o ograniczonym zakresie stosowalności. Najbardziej rozpowszechnionymi modelami są:

- model kropłowy,
- model powłokowy,
- model kolektywny.

Każdy z tych modeli jest w stanie wyjaśnić pewne właściwości, nie odwołując się do wszystkich szczegółowych właściwości sił jądrowych.

19.4.1. Model kropłowy

Model kropłowy, zaproponowany przez N. Bohra w 1939 r., powstał na gruncie podobieństw między jądrem atomowym a kroplą cieczy, takich jak:

- stała gęstość materii jądrowej i energii wiązania na jeden nukleon,
- stała gęstość cieczy nieściśliwej i stała jej energia wiązania na jednostkę objętości.

Model ten pozwolił na opisanie szeregu właściwości jąder, takich jak masa i energia wiązania oraz związanych z tym zagadnień stabilności, a ponadto pewnych faktów z fizyki rozszczepienia jądra.

W modelu kropłowym jądro przyjmuje się w przybliżeniu jako kulę, której gęstość jest wewnątrz stała i raptownie spada do zera przy przekraczaniu powierzchni. Promień kuli jest proporcjonalny do $A^{1/3}$, pole powierzchni do $A^{2/3}$, a objętość kuli zaś do A .

Ponieważ średnia energia wiązania E_n przypadająca na jeden nukleon w jądrze jest w przybliżeniu stała, więc w pierwszym przybliżeniu energia wiązania jest proporcjonalna do liczby masowej A , czyli

$$E_w \approx \alpha A, \quad (19.6)$$

gdzie α jest współczynnikiem proporcjonalności.

Zapis (19.6) jest równoważny stwierdzeniu, że każdy nukleon w jądrze podlega identycznemu oddziaływaniu. W rzeczywistości tak nie jest, ponieważ nukleony leżące w pobliżu powierzchni jądra podlegają siłom napięcia powierzchniowego, tj. siłom skierowanym w głąb jądra. Należy więc we wzorze (19.6) uwzględnić poprawkę na energię napięcia powierzchniowego. Ponadto wiemy, że energia wiązania maleje dla dużych i małych liczb masowych. Rzeczywistą energię wiązania możemy zapisać w postaci następującego wzoru

$$E_w = E_0 + E_1 + E_2 + E_3 + E_4, \quad (19.7)$$

gdzie E_0 określone wzorem (19.6).

Poprawka na napięcie powierzchniowe powinna mieć wartość ujemną i być proporcjonalna do powierzchni jądra $4\pi R^2$. Ponieważ promień jądra $R = R_0 A^{1/3}$, zatem poprawkę tę możemy napisać w postaci

$$E_1 = -\beta \sqrt[3]{A^2}, \quad (19.8)$$

gdzie β jest stałym współczynnikiem.

Druga poprawka uwzględnia siły odpychania elektrostatycznego między protonami znajdującymi się w jądrze. Jądro można porównać do naładowanej kropli kuli. Ładunek ten obniża energię wiązania jądra. Wobec tego, odpowiednia poprawka powinna być proporcjonalna do kwadratu ładunku, czyli do kwadratu liczby atomowej Z^2 , a odwrotnie proporcjonalna do promienia jądra $R = R_0 A^{1/3}$, czyli

$$E_2 = -\gamma \frac{Z^2}{\sqrt[3]{A}}, \quad (19.9)$$

gdzie γ jest stałym współczynnikiem.

Wzór na energię wiązania powinien również odzwierciedlać obserwowane prawidłowości zależności energii wiązania od liczby nukleonów w jądrze. W przypadku lekkich jąder, stabilne są jądra o równej liczbie protonów i neutronów ($Z = N$). Natomiast jądra ciężkie charakteryzują się przewagą liczby neutronów nad liczbą protonów. Narastanie sił kulombowskich, związane ze wzrostem liczby protonów, powinno być kompensowane przewagą liczby neutronów. Poprawka ta powinna obniżać energię wiązania, gdy wystąpi przewaga liczby neutronów nad liczbą protonów, lub odwrotnie, zatem

$$E_3 = -\varepsilon \frac{(N-Z)^2}{A} = -\varepsilon \frac{(A-2Z)^2}{A}, \quad (19.20)$$

gdzie ε jest stałym współczynnikiem.

Ostatnia poprawka do wzoru na energię wiązania w kropłowym modelu jądra uwzględnia fakt, że jądra o parzystej liczbie nukleonów jednego rodzaju są silniej związane i w związku z tym są bardziej stabilne w porównaniu z jądrami o nieparzystej liczbie nukleonów. Odpowiednia poprawka ma postać

$$E_4 = \delta A^{-3/4}, \quad (19.21)$$

gdzie δ jest stałym współczynnikiem.

Po uwzględnieniu wszystkich wyżej omówionych poprawek całkowita energia wiązania jądra w modelu kropłowym wyrazi się wzorem

$$E_w = \alpha A - \beta \sqrt[3]{A^2} - \gamma \frac{Z^2}{\sqrt[3]{A}} - \varepsilon \frac{(A-2Z)^2}{4A} + \delta A^{-3/4}. \quad (19.22)$$

Wzór (19.22) nosi nazwę półempirycznego wzoru Weizsäckera (z 1935 r.), w którym:

$$\begin{array}{lll} \alpha = 15,75 \text{ MeV} & \beta = 17,8 \text{ MeV} & \gamma = 0,71 \text{ MeV} \\ \varepsilon = 94,8 \text{ MeV} & \delta = \pm 34 \text{ MeV} & \end{array}$$

Znając energię wiązania jądra łatwo możemy określić defekt masy jądra dzieląc wzór (19.22) przez kwadrat prędkości światła. Masę dowolnego jądra zawierającego Z protonów i $A-Z$ neutronów możemy obliczyć odejmując od sumy mas protonów i neutronów defekt masy. Zatem

$$M(A, Z) = Zm_p + (A-Z)m_n - \Delta m. \quad (19.23)$$

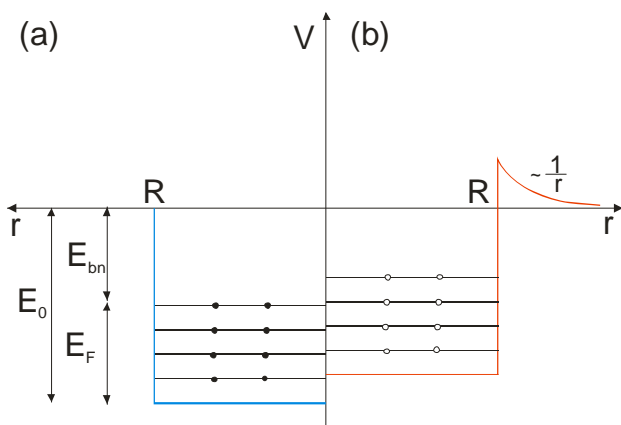
Podstawiając do tego wzoru odpowiednie wartości w jednostkach masy atomowej, otrzymamy wzór na masę dowolnego atomu

$$\begin{aligned} M(A, Z) = & 0.99176A - 0.00084Z + 0.019\sqrt[3]{A^2} \\ & + 0.00076 \frac{Z^2}{\sqrt[3]{A}} + 0.025 \frac{(A-2Z)^2}{A} - \delta A^{-3/4}. \end{aligned} \quad (19.24)$$

Korzystając z zależności (19.24) można obliczyć masę atomu z dokładnością nie mniejszą niż 0.01%. Dokładność obliczonych mas wzrasta w miarę wzrostu A , ponieważ dla jąder ciężkich model kropłowy jądra daje bardzo dobre przybliżenia.

19.4.2. Model powłokowy

Model powłokowy jądra atomowego powstał w analogii do istniejących powłok elektronowych atomu. Jak wiemy, istnieją atomy bardzo trwałe pod względem chemicznym i nieaktywne – są to atomy gazów szlachetnych. Atomy tych gazów mają całkowicie obsadzone powłoki elektronowe, przy czym zapelnienie następnej powłoki prowadzi do utworzenia atomu kolejnego pierwiastka gazu szlachetnego. Jak już poprzednio zaznaczono, wśród jąder atomowych szczególną trwałością wyróżniają się jądra o pewnych, ściśle określonych liczbach protonów i neutronów, nazywanych liczbami magicznymi. W związku z tym wysunięto hipotezę, że pod względem energetycznym budowa jądra podobna jest do budowy atomu, nukleony w jądrze poruszają się tak, że ich stany o danej energii tworzą poszczególne powłoki. Stwierdzono, że powłoki energetyczne jądra obsadzone są przez protony i neutrony w sposób niezależny (patrz rys. 19.3). Dla jądra składającego się z wielu nukleonów głębokość studni dla neutronu wynosi około 40 MeV. Na proton działa jeszcze dodatkowy potencjał sił kulombowskich i dlatego wewnątrz jądra ściany jego studni potencjału są podniesione, a na zewnątrz potencjał maleje z odległością jak $1/r$. Nie są możliwe zderzenia między nukleonami, ponieważ w zderzeniach winna następować zmiana energii, a wszystkie poziomy poniżej energii Fermiego są zapelnione. Możliwe są jedynie takie zderzenia, w których nukleony zamieniają się miejscami, ale za względu na nierozróżnialność cząstek efekt wypadkowy jest zerowy. A więc silne oddziaływanie nukleonów w jądrze nie musi się wiązać z wieloma zderzeniami i z korelacją ich ruchów, jak to przewiduje model kroplowy; nukleony mogą poruszać się niezależnie.



Rys. 19.3. Uproszczony wykres energii potencjalnej neutronu (a) i protonu (b) wewnątrz jądra o promieniu R .

Ze względu na większą trwałość, izotopy o magicznych liczbach nukleonów występują w przyrodzie w szczególnie dużych ilościach. Zaliczamy do nich między innymi takie pierwiastki jak: ${}^4_2\text{He}$, ${}^{16}_8\text{O}$, ${}^{40}_{20}\text{Cu}$, ${}^{208}_{82}\text{Pb}$.

Do opisu ruchu nukleonów wewnątrz jądra stosuje się prawa mechaniki kwantowej, co potwierdzają np. przemiany promieniotwórcze jąder, w których emitowana cząstka lub foton gamma γ unosi zawsze ściśle określoną porcję energii. Oznacza to, że energia jądra może przyjmować tylko pewne ściśle określone wartości, co jest zgodne z prawami mechaniki kwantowej.

19.4.3. Model kolektywny

W modelu powłokowym poszczególne składniki jądra poruszają się niezależnie od siebie. Model kropłowy sugeruje coś wręcz przeciwnego, gdyż w kropli cieczy nieściśliwej ruch każdej cząstki składowej jest ściśle skorelowany z ruchami sąsiadujących z nią cząstek. W celu usunięcia sprzeczności między tymi modelami wprowadzono inny doskonalszy model o szerszym zakresie stosowalności. Jest nim kolektywny model jądra zaproponowany w 1953 r. przez A. Bohra (syna twórcy modelu atomu) i B. Mottelzona. W modelu tym nukleony poruszają się niezależnie (jak w modelu powłokowym) w polu sił jądrowych. Jednak potencjał związany z tymi siłami nie jest potencjałem statycznym (jak w modelu powłokowym), lecz jest potencjałem o zmiennym kształcie. Deformacje tego potencjału odzwierciedlają pewne skorelowane ruchy nukleonów w jądrze.

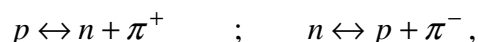
19.5. Siły jądrowe

Trwałe wiązania między nukleonami w jądrze wskazują na istnienie w jądrach atomowych specyficznych sił, tzw. sił jądrowych (oddziaływania silne). Przedstawimy poniżej specyficzne cechy tych sił:

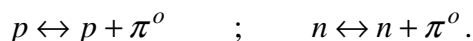
- Oddziaływania te nie możemy wyjaśnić za pomocą klasycznych oddziaływań grawitacyjnych czy elektromagnetycznych. Te oddziaływania są zbyt słabe by wyjaśnić dużą energię wiązania nukleonów w jądrze. Gdyby założyć oddziaływanie grawitacyjne, to energia wiązania między nukleonami byłaby 10^{38} razy mniejsza od wartości otrzymywanej eksperymentalnie. Również założenie o elektrycznym charakterze sił jądrowych jest nierealistyczne; są one nadal słabe a ponadto występują tylko między naładowanymi nukleonami (protonami) i w dodatku są odpychające.
- Siły jądrowe są przyciągające.
- Są krótkozasięgowe. Znikają na odległości około 2×10^{-15} m, co oznacza, że nukleony będące w większej odległości nie oddziałują ze sobą za pośrednictwem sił nukleonowych.
- Siły jądrowe są niezależne od ładunku. Oznacza to, że nie rozróżniają one protonów od neutronów. Biorąc pod uwagę oddziaływanie protonu z protonem, należy oczywiście uwzględnić dodatkowo kulombowskie odpychanie między nimi. O ładunkowej niezależności sił jądrowych świadczą np. wyniki badań rozpraszania nukleonów na deuterze. Świadczą o niej również dane dotyczące energii wiązania par jąder zwierciadlanych, tzn. takich jąder, dla których liczba protonów w jednym jest równa liczbie neutronów w drugim. Weźmy najprostszy przykład tego typu jąder: 3_1H i 3_2H ; w obu jądrach mamy 3 nukleony, z tym, że w 3_1H – zwanym trytonem – mamy $2n$ i $1p$, a w 3_2H odwrotnie – $2p$ i $1n$. Energie wiązania obu tych jąder są następujące: 3_1H – 8,49 MeV, 3_2H – 7,72 MeV. Jeżeli oddziaływanie między nukleonami jest ładunkowo niezależne, to energie wiązania obu tych

jąder powinny być takie same. Ponieważ jednak w jądrze ${}^3_2\text{H}$ mamy dwa protony, to jego energia wiązania powinna być mniejsza o energię odpychania obu protonów. Proste obliczenia energii oddziaływania kulombowskiego dają 0,72 MeV, co jest w dobrej zgodności z różnicą energii wiązania obu jąder wynoszącą 0,77 MeV.

- Siły jądrowe wysycają się, co wskazuje że każdy nukleon oddziałuje z najbliższymi sąsiednimi nukleonami. Gdyby każdy z A nukleonów jądra oddziałował ze wszystkimi $A-1$ pozostałymi, wówczas energia wiązania jądra musiałaby być proporcjonalna do $A(A-1)$, a zatem zależeć od kwadratu nukleonów. Ponieważ energia wiązania jest proporcjonalna do A wnosimy, że nukleon nie oddziałuje ze wszystkimi pozostałymi, lecz że oddziaływanie jądrowe ulega wysyceniu już na kilku wiązaniach. Mamy więc analogię do wysycania wiązań chemicznych; np. wiązanie między atomami wodoru wysyca się już dla dwu cząstek H_2 , a energia wiązania kropli ciekłego wodoru jest w przybliżeniu równa sumie energii wiązania wysyconych struktur H_2 czyli proporcjonalna do liczby atomów wodoru.
- Siły jądrowe zależą od orientacji spinów oddziałujących nukleonów. Wynika stąd, że nie są siłami centralnymi.
- Siły jądrowe są siłami wymiennymi. W przypadku wiązań chemicznych ich wysycanie tłumaczy się tym, że wiązanie atomów w molekułe polega na ciągłej wymianie elektronów między atomami. Przez analogię, można założyć, że również siły jądrowe są wymienne i że istnieje cząstka wymieniana między nukleonami. Przez analogię z oddziaływaniem elektromagnetycznym (gdzie kwantami tego oddziaływania są fotony), również oddziaływanie jądrowe powinno realizować się poprzez jakąś cząstkę. Jest nią cząstka zwana mezonem π (pion). Istnienie takiej cząstki zostało potwierdzone eksperymentalnie. Mezonową teorię sił jądrowych przedstawił w 1935 r. H. Yukawa. Nazwa "mezon" pochodzi stąd, że cząstka ta powinna mieć masę pośrednią pomiędzy elektronem i protonem (grec. *mesos* znaczy pośredni). Siły jądrowe między protonem i neutronem możemy wyjaśnić przez wirtualną wymianę pionu naładowanego



natomiast siły między dwoma neutronami lub między dwoma protonami przez wymianę pionu neutralnego



Masę spoczynkową mezonu π łatwo można oszacować z zasady nieoznaczoności Heisenberga i danych dotyczących sił jądrowych. Prędkość mezonu nie może być większa od prędkości światła. Przyjmując, że $\Delta E = m_\pi c^2$ a czas przelotu pomiędzy dwoma nukleonami oddalonymi o $R \approx 2 \times 10^{-15}$ m będzie wynosił co najmniej $\Delta t = R/c$, otrzymujemy

$$m_\pi \approx \frac{\hbar}{Rc} \approx 200m_e , \tag{19.25}$$

gdzie m_e jest masą spoczynkową elektronu. W 1947 r. odkryto mezony w promieniowaniu kosmicznym i stwierdzono, że ich masa wynosi $274m_e$. Na określenie zależności potencjału sił jądrowych od odległości, Yukawa zaproponował następujące wyrażenie

$$V(r) = V_0 \frac{\exp(-r/r_0)}{r/r_0}, \quad (19.26)$$

gdzie r_0 jest zasięgiem sił jądrowych.

- Nukleony oraz mezony (poza pionami są jeszcze inne mezony) to cząstki oddziałujące silnie – hadrony. Obecnie przyjmuje się, że hadrony nie są niepodzielne – składają się z mniejszych cząstek, kwarków. Przyjmuje się, że istnieje sześć kwarków i sześć antykwarków, chociaż dotychczas nikt nie wykazał doświadczalnie istnienia swobodnego kwarku. Być może, że istnieją tylko w postaci związanej jako składniki, np. protonu czy innego hadronu. Przyjmuje się, że cząstką wymienną w oddziaływaniu kwarków jest gluon.

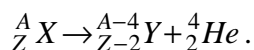
19.6. Przemiany jądrowe

Jądra atomowe niektórych pierwiastków są nietrwałe i samorzutnie przekształcają się w jądra innych pierwiastków, czemu towarzyszy wysyłanie różnego rodzaju promieniowania. Przemiany (rozpady) te noszą historyczne nazwy przemian α , β i γ , a towarzyszące im promieniowania – odpowiednio promieniowań α , β i γ . Promieniowanie α stanowi strumień jąder helu (${}^4_2\text{He}$), promieniowanie β – strumień elektronów, a promieniowanie γ – strumień kwantów promieniowania elektromagnetycznego o dużej energii. Naturalną radioaktywność obserwujemy w przypadku jąder ciężkich, które w układzie okresowym pierwiastków zajmują miejsca za ołowiem. Jednakże istnieją także jądra o mniejszej liczbie masowej wykazujące naturalną radioaktywność. Są to izotopy potasu ${}^{40}_{19}\text{K}$, węgla ${}^{14}_6\text{C}$, rubidu ${}^{87}_{37}\text{Rb}$, ziem rzadkich – lantanu, samaru, lutetu, a także indu i renu.

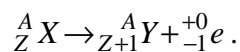
19.6.1. Promieniotwórczość naturalna

Zjawisko naturalnej promieniotwórczości zostało odkryte w związkach uranu w 1896 roku przez A.H. Becquerela. P. Curie i M. Skłodowska-Curie odkryli, że smolista ruda uranowa emituje promieniowanie 4-krotnie większe od promieniowania uranu a następnie odkryli dwa nowe pierwiastki radioaktywne (polon ${}^{210}_{84}\text{Po}$ i rad ${}^{226}_{88}\text{Ra}$) wysyłające cząstki α , β i γ .

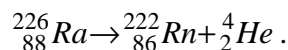
We wszystkich przemianach jądrowych spełnione są prawa zachowania ładunku, liczby nukleonów, energii i pędu. Ze względu na pierwsze i drugie z nich w przemianie α jądro pochodne ma liczbę masową i atomową mniejszą odpowiednio o 4 i 2 od odpowiednich liczb jądra macierzystego



W przemianie β liczba masowa nie ulega zmianie, a liczba atomowa zmienia się o 1 (rozpad β przemieszcza pierwiastek w tablicy Mendelejewa o jedno miejsce w prawo)



Reguły te stanowią tzw. prawo przesunięć F. Soddy'ego i K. Fajansa. Ilustracją tych przemian jest np. rozpad radu, który odbywa się z wydzieleniem radonu i cząstki α



Istnieje również promieniotwórczość sztuczna; okazuje się, że jądra atomowe utworzone w wyniku reakcji jądrowych są w wielu przypadkach promieniotwórcze. Jądra takie nazywa się więc sztucznie promieniotwórczymi.

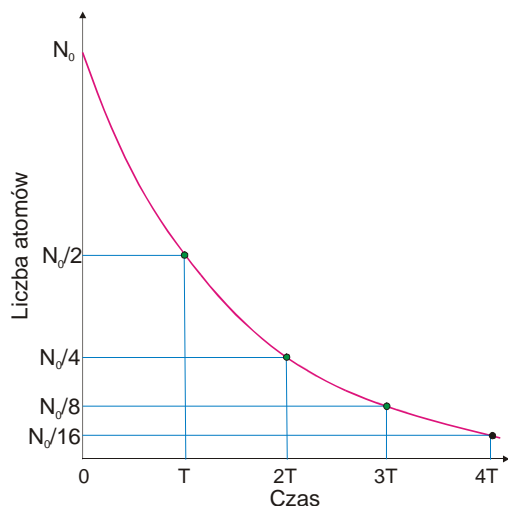
Proces rozpadu promieniotwórczego jest procesem samoistnym. Liczba dN rozpadających się jąder w przedziale czasu t i $t + dt$, powinna być proporcjonalna do przedziału czasu dt i do liczby N jąder, które do chwili t jeszcze się nie rozpadły:

$$dN = -\lambda N(t) dt , \quad (19.27)$$

gdzie λ jest stałą rozpadu. Znak minus wskazuje, że dN jest liczbą ujemną, gdyż N maleje z czasem. Całkując powyższe równanie, otrzymujemy

$$N(t) = N_0 e^{-\lambda t} . \quad (19.28)$$

We wzorze tym N_0 jest liczbą jąder promieniotwórczych w chwili początkowej, a $N(t)$ liczbą jąder, które po czasie t nie uległy jeszcze rozpadowi. Zależność (19.28) jest nazywana prawem rozpadu promieniotwórczego i została wykreślona na rys. 19.4.



Rys. 19.4. Krzywa opisująca prawo rozpadu promieniotwórczego. Wykres obejmuje przedział czasu odpowiadający czterem okresom połowicznego zaniku.

Liczba dN żyje łącznie przez czas równy $t \frac{dN}{dt} = t \lambda N(t) dt$. Stąd średni czas życia jądra pierwiastka promieniotwórczego wynosi

$$\tau = \frac{1}{N_0} \int_0^{\infty} \lambda t N(t) dt = \lambda \int_0^{\infty} t e^{-\lambda t} dt = \frac{1}{\lambda} . \quad (19.29)$$

Widzimy więc, że stała rozpadu jest odwrotnością średniego czasu życia danego pierwiastka promieniotwórczego, a więc czasu, w którym liczba jąder promieniotwórczych, która nie uległa rozpadowi maleje e-krotnie.

Czas połowicznego zaniku $T_{1/2}$ jest definiowany jako wartość t , dla której $N = N_0/2$. Wstawiając to do równania (19.28), otrzymujemy

$$\frac{N_0}{2} = N_0 \exp(-\lambda T_{1/2}),$$

czyli

$$T_{1/2} = \frac{\ln 2}{\lambda} = \tau \ln 2. \quad (19.30)$$

W przypadku gdy jądro pochodne rozpadu promieniotwórczego jest również jądrem promieniotwórczym, wówczas powstaje cały łańcuch przemian promieniotwórczych, zwany również rodziną promieniotwórczą.

Istnieją cztery rodziny promieniotwórcze; ich nazwy pochodzą od nazw pierwszych, macierzystych pierwiastków:

- rodzina toru ${}^{232}_{90}\text{Th} \rightarrow {}^{208}_{82}\text{Pb}$ ($T_{1/2} = 1,39 \times 10^{10}$ lat),
- rodzina uranu ${}^{238}_{92}\text{U} \rightarrow {}^{206}_{82}\text{Pb}$ ($T_{1/2} = 4,5 \times 10^9$ lat),
- rodzina aktynu ${}^{235}_{89}\text{Ac} \rightarrow {}^{207}_{82}\text{Pb}$ ($T_{1/2} = 7,1 \times 10^8$ lat),
- rodzina neptunu ${}^{237}_{93}\text{Np} \rightarrow {}^{209}_{83}\text{Bi}$ ($T_{1/2} = 2,2 \times 10^6$ lat).

Przejścia pomiędzy poszczególnymi członami rodzin promieniotwórczych odbywają się poprzez kolejne rozpady α , β i kończą się na trwałych izotopach ołowiu (trzy rodziny). Czasy połowicznego zaniku podane są powyżej w nawiasach. Ponieważ liczba masowa zmienia się tylko w przejściach α , liczby atomowe wszystkich pośrednich członków szeregów promieniotwórczych można opisać wzorem

$$A = 4n + m, \quad (19.31)$$

gdzie n jest liczną całkowitą, a $m = 0, 2$ i 3 , odpowiednio dla szeregu toru, uranu i aktynu.

Rodzina neptunu ($m = 1$) kończy się na jądrze bizmutu (charakteryzuje się najkrótszym czasem połowicznego zaniku). Ponieważ wiek Ziemi ocenia się na 5×10^9 lat, dlatego rodzina neptunu w stanie naturalnym już nie istnieje. Oczywiście, jest możliwe sztuczne utworzenie jądro macierzystego tego szeregu – ${}^{237}_{93}\text{Np}$, a przez jego rozpad wszystkie jądra pochodne.

19.6.2. Rozpad α

Rozpad α , z wyjątkiem nielicznych przypadków, jest charakterystyczny tylko dla jąder ciężkich o $A > 200$, u których ze wzrostem liczby masowej maleje energia wiązania pojedynczego nukleonu. Wyzwalana energia, przy zmniejszeniu liczby nukleonów A w jądrze o jeden nukleon, jest znacznie mniejsza od energii wiązania pojedynczego nukleonu w jądrze, dlatego wysłanie neutronu lub protonu jest niemożliwe (ich energia poza jądrem jest zerowa). Natomiast emisja jądra helu jest energetycznie możliwa, ponieważ energia wiązania

przypadająca na jeden nukleon w tym jądrze wynosi około 7,1 MeV. Ogólnie można powiedzieć, że *rozpad α jest możliwy, jeżeli suma energii wiązania jądra otrzymanego po rozpadzie i cząstki α jest większa od energii wiązania jądra macierzystego.*

Załóżmy, że jądro ${}^A_Z X$ w wyniku rozpadu α przekształciło się w inne jądro ${}^{A-4}_{Z-2} Y$. Jeżeli energia wiązania E_2 jądra ${}^{A-4}_{Z-2} Y$ przekracza energię wiązania E_1 jądra wyjściowego ${}^A_Z X$ o więcej niż wartość energii wiązania cząstki α , tzn. $\Delta E > 28,4$ MeV, to rozpad α jest możliwy. Zwiększenie energii wiązania nukleonów oznacza równocześnie zmniejszenie energii spoczynkowej jądra o wartość E_α , wydzielającej się przy rozpadzie α . Energia ta wyzwala się w postaci energii kinetycznej produktów rozpadu. Chcemy znaleźć związek między energią cząstki α a energią odrzutu jądra E_M . Korzystając z prawa zachowania energii i pędu możemy napisać

$$E_\alpha = E_m + E_M = \frac{mv_\alpha^2}{2} + \frac{Mv_j^2}{2}, \quad (19.32)$$

$$mv_\alpha = Mv_j, \quad (19.33)$$

gdzie m , M – odpowiednio masa cząstki α i jądra powstającego po rozpadzie, a v_α i v_j – prędkości cząstki α i jądra odrzutu.

Z ostatnich dwóch równań mamy

$$\frac{E_m}{E_M} = \frac{M}{m}, \quad (19.34)$$

zaś energia cząstki α jest równa

$$E_m = E_\alpha \frac{M}{M + m}, \quad (19.35)$$

a energia odrzutu

$$E_M = E_\alpha \frac{m}{M + m}. \quad (19.36)$$

Ze wzoru (19.35) wynika, że gdy masa cząstki α jest dużo mniejsza od masy jądra powstałego po rozpadzie, to praktycznie całą energię rozpadu unosi cząstka α .

Uwzględniając warunek bilansu mas przed i po rozpadzie, mamy

$$M(A, Z) > M(A - 4, Z - 2) + m, \quad (19.37)$$

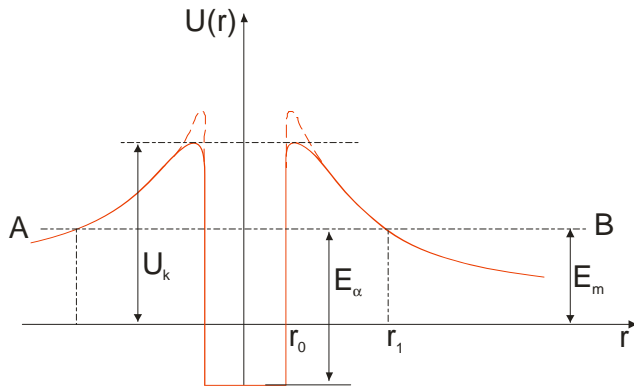
gdzie $M(A, Z)$ jest masą jądra przed rozpadem, $M(A - 4, Z - 2)$ masą jądra po rozpadzie, a m masą jądra helu.

Energia rozpadu jest określona wzorem

$$E_\alpha = [M(A, Z) - M(A - 4, Z - 2) - M_\alpha]c^2, \quad (19.38)$$

co oznacza, że *energia rozpadu α jest równa różnicy mas atomu macierzystego i całkowitej masy produktów rozpadu, pomnożonej przez kwadrat prędkości światła.*

Energie rozpadu α zmieniają się w granicach od 4 do 9 MeV. Praktycznie cała energia rozpadu unoszona jest przez cząstkę α . Na rys. 19.5 przedstawiono energię potencjalną układu cząstka α – jądro. Dla odległości większych od promienia jądra, jest to energia kulombowskiego odpychania. Maksymalna wartość energii kulombowskiej wynosi zwykle około 30 MeV. Nasuwa się pytanie, w jaki sposób cząstka α wychodzi na zewnątrz jądra pokonując obszar o wyższej energii potencjalnej. Z punktu widzenia mechaniki klasycznej taki proces jest niemożliwy.



Rys. 19.5. Potencjalna bariera kulombowska: U_k wysokość bariery kulombowskiej, E_α energia kinetyczna cząstki α w jądrze, E_m energia kinetyczna cząstki α po rozpadzie (na zewnątrz jądra).

Wyjaśnienia tego problemu dostarcza mechanika kwantowa tzw. efektem tunelowym (patrz punkt 11.6). Prawdopodobieństwo przejścia cząstki α przez barierę określa wzór

$$p \propto \exp\left(-\frac{4\pi}{h} \sqrt{2m} \int_{r_0}^{r_1} \sqrt{U(r) - E_\alpha} dr\right). \quad (19.39)$$

Liczbę zderzeń n cząstki z barierą w jednostce czasu obliczymy zakładając, że cząstka α porusza się wewnątrz bariery potencjalnej o promieniu r_0 z prędkością v . Mamy wówczas

$$n = \frac{v}{2r_0}. \quad (19.40)$$

Prawdopodobieństwo λ przeniknięcia cząstki przez barierę potencjalną w jednostce czasu będzie iloczynem dwóch ostatnich wyrażeń

$$\lambda = np, \quad (19.41)$$

co nazywamy stałą rozpadu.

Na szerokości studni potencjału, $2r_0$, powinna odłożyć się całkowita długość fali de Broglie'a, przy czym minimalna prędkość cząstki α powinna odpowiadać długości fali równej $2r_0$, czyli

$$\frac{h}{mv} = 2r_0. \quad (19.42)$$

Z zależności (19.40) i (19.42) wynika, że liczba zderzeń cząstki α z barierą w jednostce czasu wynosi

$$n = \frac{h}{4mr_0^2}. \quad (19.43)$$

Uwzględniając we wzorze (19.41) zależności (19.39) i (19.43) otrzymujemy wyrażenie na stałą rozpadu

$$\lambda = \frac{h}{4mr_0^2} \exp\left(-\frac{4\pi}{h} \sqrt{2m}\right) \int_{r_0}^{r_1} \sqrt{U(r) - E_\alpha} dr. \quad (19.44)$$

Wzór ten wskazuje na zależność między stałą rozpadu a energią cząstki α ; stała rozpadu jest tym większa, im wyższa jest energia cząstki α i im niższa jest wysokość bariery. Energia cząstek α zawiera się w małym przedziale energii, od 4 do 9 MeV, natomiast stałe rozpadu zmieniają się w bardzo szerokich granicach od 10^5 do 10^{-16} s⁻¹. Zależność stałej rozpadu od energii została określona empiryczną regułą w 1912 r. przez Geigera i Nutalla, wg której

$$\ln \lambda = A + \frac{B}{\sqrt{E_\alpha}}, \quad (19.45)$$

gdzie A i B są stałymi współczynnikami, różnymi dla poszczególnych rodzin promieniotwórczych.

Cząstki α oddziałują z materią poprzez takie procesy jak: jonizacja, wzbudzenie, dysocjacja cząsteczek. Cząstki te, spośród cząstek α , β i γ , odznaczają się największym stopniem jonizacji; na drodze 1 mm wytwarzają około 3×10^3 par jonów w powietrzu w warunkach normalnych. Ponieważ ich strata energii jest duża, stąd mały ich zasięg, np. w powietrzu do około 10 cm, w Al zasięg jest znikomy. Zdolność jonizacyjna cząstek α jest największa na końcowych odcinkach drogi (około 2,5 razy), co tłumaczy się wzrostem prawdopodobieństwa trafienia molekuł gazu przy mniejszej prędkości cząstek α .

19.6.3. Rozpad β

W rozpadach β rozróżniamy emisję β^- (elektron ${}_{-1}^0e$) bądź β^+ (pozyton ${}_{+1}^0e$). W pierwszym przypadku jądro pochodne ma liczbę protonów zwiększoną o 1 ($\Delta Z = +1$) w stosunku do jądra macierzystego, a w drugim zmniejszoną o 1 ($\Delta Z = -1$). W obydwu przypadkach liczba masowa nie ulega zmianie. Przemiany te możemy zapisać w postaci:

- ${}^A_Z X \rightarrow {}^A_{Z+1} Y + {}_{-1}^0 e,$
- ${}^A_Z X \rightarrow {}^A_{Z-1} Y + {}_{+1}^0 e.$

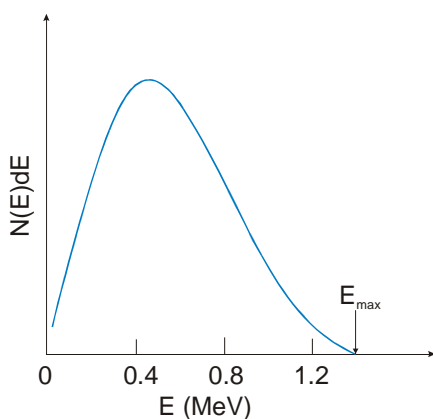
Protonowo-neutronowa budowa jądra wyklucza emisję elektronów z jądra, ponieważ ich tam nie ma. Dlatego teoretyczne wytłumaczenie rozpadu β stało się trudnym problemem fizyków jądrowych.

Emisja cząstek β następuje w procesie przejścia atomu z jednego stanu w drugi (mamy sytuację podobną do emisji fotonu przez atom). Elektrony i pozytony rodzą się w procesie przejścia danego nukleonu z jednego stanu kwantowego w drugi, np. z neutronowego w protonowy z wyemitowaniem elektronu.

Pomiary eksperymentalne wykazały ciągłe widmo energetyczne rozpadu β . Fakt ten trudno jest wytłumaczyć biorąc pod uwagę dyskretność poziomów energetycznych jądra. Na rys. 19.6 przedstawiono widmo energetyczne elektronów emitowanych przez radioaktywny potas ${}_{19}^{40}\text{K}$. Energia ta zawiera się w przedziale od E do E_{\max} ; nie jest możliwe występowanie elektronów o energii większej niż E_{\max} . Badania eksperymentalne wskazują, że dla wszystkich źródeł maksymalna energia cząstek β jest równa różnicy energii poziomów energetycznych jąder, które je emitują. Dlatego też jądra β -aktywne, które przed i po rozpadzie β mają w pełni określone energie, mogą emitować ciągłe widmo energetyczne elektronów.

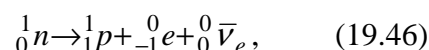
Kolejnym problemem w wyjaśnieniu rozpadu β jest spin jądra. Elektron emitowany przez jądro ma spin połówkowy ($\hbar/2$). Wynikałoby stąd, że przy rozpadzie β powinien ulegać zmianie spin jądra. Jednakże taka zmiana nie następuje.

Obydwie te poważne trudności w wyjaśnieniu rozpadu β (związane z ciągłym charakterem widma energetycznego i ze spinem) zostały usunięte przez Pauliego w 1931 r. Założył on mianowicie, że podczas każdego aktu rozpadu β jądro emituje nie jedną lecz dwie cząstki. Oprócz elektronu, emitowana jest jeszcze jedna cząstka obojętna elektrycznie o zerowej masie spoczynkowej i spinie połówkowym $\hbar/2$. Cząstkę tę nazwano neutrinem (${}^0_0\nu_e$) lub antyneutrinem (${}^0_0\bar{\nu}_e$). Hipoteza neutrina rozwiązała trudności w teorii rozpadu β . Brakującą energię, różnicę między energią maksymalną elektronów a faktyczną ich energią unoszą antyneutrino. Doświadczalnego potwierdzenia istnienia neutrina dokonali w 1956 G.F. Reins i C.L. Cowan.



Rys. 19.6. Widmo energetyczne elektronów emitowanych przez ${}_{19}^{40}\text{K}$.

Tak więc obecnie rozpad β^- wyjaśnia się następującą przemianą jądrową



która wskazuje, że neutron zamienia się na proton, powstające przy tym elektron i antyneutrino są emitowane z jądra. Proton pozostaje w jądrze pochodnym i dlatego liczba masowa A nie ulega zmianie.

Natomiast rozpad β^+ wyjaśnia się przemianą protonu w neutron



Cząstki β oddziałują na materię głównie w procesach jonizacji oraz powstawania promieniowania hamowania. Charakteryzują się mniejszym stopniem jonizacji w porównaniu z cząstkami α , stąd ich zasięg w powietrzu dochodzi do kilkunastu metrów, a w Al do 4 mm. Prędkość najbardziej energetycznych cząstek β wynosi około $0.9c$.

19.6.4. Promieniowanie γ

Promieniowanie γ jest emitowane przez wzbudzone jądra pierwiastka pochodnego, powstającego z rozpadu α , bądź β . Jądra te przechodząc do niższych stanów energetycznych emitują fotony promieniowania γ . Podobnie jak widmo energetyczne cząstek α , widmo promieniowania γ jest widmem liniowym. Emisji promieniowania γ może towarzyszyć emisja promieniowania X .

Jeżeli na jądro pada promieniowanie γ o takiej częstotliwości ν , że energia kwantu $h\nu$ jest równa różnicy energii między jednym ze stanów wzbudzonych i podstawowym stanem energetycznym, to może nastąpić rezonansowe pochłanianie promieniowania γ . Jądro pochłania kwant γ o takiej częstotliwości, jaką ma foton γ , wysyłany przez to jądro podczas przejścia jądra z danego wzbudzonego stanu energetycznego do stanu podstawowego. Jednak praktyczną realizację rezonansowego pochłaniania utrudnia zjawisko odrzutu jądra atomowego. Przy emisji kwantu γ , energia kinetyczna jądra odrzutu $E_k = p^2/2M$, gdzie M jest masą jądra a p – pędem. Z kolei pęd jądra musi być równy pędowi kwantu γ , co zapiszemy jako $p = p_\gamma = h\nu/c$. Dlatego

$$E_k = \frac{(h\nu)^2}{2Mc^2}.$$

Jeżeli jądro atomowe podczas emisji kwantu γ traci energię ΔE_0 , to

$$\Delta E_0 = E_k + h\nu,$$

stąd

$$h\nu = \Delta E_0 - \frac{(hc)^2}{2Mc^2}. \quad (19.48)$$

Taką w rzeczywistości energię posiada emitowany foton γ przez odosobnione jądro. Podczas absorpcji fotonu γ jądro również doznaje odrzutu i pochłania tylko ten foton γ , który posiada energię

$$h\nu = \Delta E_0 + \frac{(h\nu)^2}{2Mc^2}. \quad (19.49)$$

W 1958 r. R. Mössbauer opracował metodę emisji i absorpcji praktycznie bezodrzutowej. Mianowicie, gdy jądro nie będzie swobodne, lecz np. będzie "utwierdzone" w sieci krystalicznej, to energia odrzutu udziela się całej sieci i wówczas M równa się całej masie sieci krystalicznej. Dlatego składnik $(h\nu)^2/2Mc^2$ w wyrażeniach (19.48) i (19.49) można pominąć, co oznacza, że odrzut nie występuje, więc $h\nu = \Delta E_0$. Mamy więc emisję bądź absorpcję bezodrzutową (mesbauerowską).

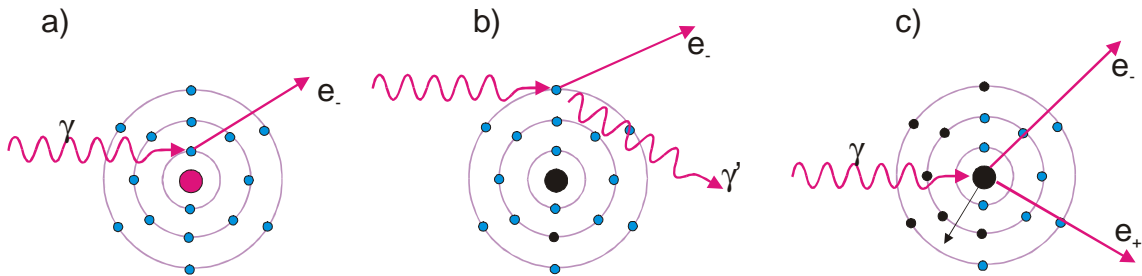
Efekt Mössbauera znalazł zastosowanie w dokładnych pomiarach bardzo małych zmian energii w bardzo subtelnych i ważnych zjawiskach we współczesnej fizyce. Podczas zmiany wzajemnej odległości między krystalicznym źródłem promieni γ i krystalicznym odbiornikiem, występuje dopplerowska zmiana częstotliwości absorbowanych kwantów γ i ich energia $h\nu$

dopasowuje się do określonej różnicy energii poziomów energetycznych jądra, następuje rezonansowe pochłanianie, co ujawnia detektor promieniowania γ

Promieniowanie γ oddziałuje na materię poprzez trzy zjawiska:

- zjawisko fotoelektryczne,
- zjawisko Comptona,
- zjawisko powstawania par.

Rysunek 19.7 przedstawia ich graficzny obraz. Pierwsze dwa zjawiska omówiono w rozdziale 10.



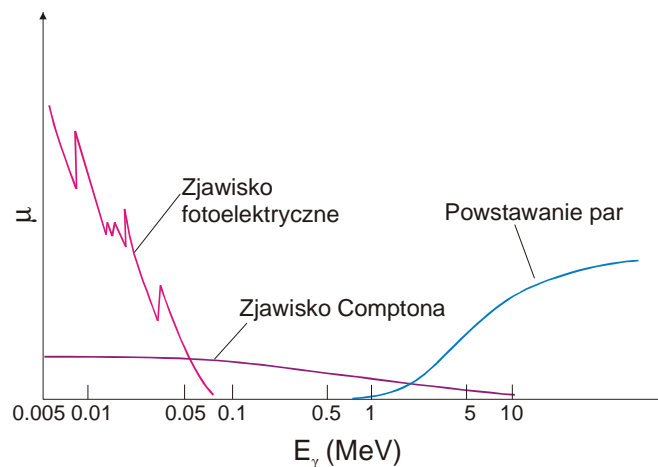
Rys. 19.7. Zjawiska oddziaływania kwantów γ z materią: (a) zjawisko fotoelektryczne, (b) zjawisko Comptona, (c) zjawisko tworzenia pary elektron-pozyton.

W zjawisku tworzenia pary elektronowej (ogólniej pary cząstka-antycząstka) kwant promieniowania γ ulega całkowitej absorpcji, a pojawia się para cząstek; elektron i pozyton. Z prawa zachowania energii wynika, że minimalna energia padającego fotonu musi być równa $2m_0c^2 = 1,02$ MeV.

Wyżej wymienione trzy zjawiska są przyczyną pochłaniania wiązki promieniowania γ w materii. Natężenie wiązki zmienia się z wzorem

$$I = I_0 \exp(-\mu x), \quad (19.50)$$

gdzie I_0 oznacza początkowe natężenie wiązki, I – natężenie po przejściu warstwy absorbenta o grubości x , a μ – całkowity współczynnik pochłaniania. Udział poszczególnych zjawisk zależy od energii, co przedstawia rys. 19.8.



Rys. 19.8. Zależność współczynników absorpcji promieniowania γ od energii dla zjawiska fotoelektrycznego, Comptona i tworzenia par.

Z powyższego rysunku wynika, że udział zjawiska fotoelektrycznego w pochłanianiu szybko maleje ze wzrostem energii fotonów γ . Granicą jest energia najsilniej związanych elektronów w atomach. Udział efektu Comptona jest znacznie mniejszy niż zjawiska fotoelektrycznego, jednakże maleje on wolniej i jest dominujący dla energii w przedziale energii od 0,08 MeV do około 2 MeV. Dla energii fotonów powyżej 10 MeV udział zjawiska Comptona jest praktycznie zerowy. Natomiast udział zjawiska tworzenia par elektronowych wzrasta ze wzrostem energii fotonów γ , dążąc do stałej wartości dla bardzo dużych energii fotonów γ .

19.7. Reakcje jądrowe

19.7.1. Charakterystyka reakcji jądrowych

Pierwszą reakcję jądrową (sztuczną przemianę jądrową) odkrył Rutherford w 1919 r. bombardując jądra azotu cząstkami α (emitowanymi przez jądra ${}^{214}_{84}\text{Po}$). Reakcję tę możemy zapisać



Ogólnie reakcję jądrową zapisujemy w postaci



gdzie a jest cząstką bombardującą, X – jądrem-tarczą, Y i b odpowiednio jądrem i cząstką powstałymi w wyniku reakcji. Często stosuje się uproszczony zapis $X(a,b)Y$ lub gdy interesuje nas tylko typ reakcji – (a,b) .

Dla wyżej zapisanej reakcji, zasada zachowania energii ma postać

$$E_a + m_a c^2 + m_X c^2 = (E_b + m_b c^2) + (E_Y + m_Y c^2), \quad (19.53)$$

gdzie E_i i m_i są odpowiednio energią kinetyczną i masą spoczynkową i -tego elementu reakcji. Przy tym przyjęliśmy, że $E_X = 0$, tzn. rozpatrujemy układ w którym jądro X spoczywa.

Ciepło reakcji określimy jako

$$Q = [(m_a + m_X) - (m_b + m_Y)]c^2. \quad (19.54)$$

Uwzględniając równanie (19.53) możemy ostatnie równanie zapisać jako

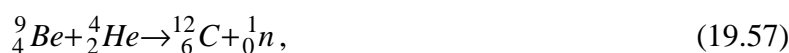
$$Q = E_Y + E_b - E_a. \quad (19.55)$$

Gdy $Q < 0$, to mamy reakcję endotermiczną, czyli całkowita energia spoczynkowa reagentów jest mniejsza od całkowitej energii produktów reakcji. Reakcja tego typu odbywa się kosztem energii kinetycznej cząstki padającej. Przykładem takiej reakcji jest reakcja przeprowadzona przez Rutherforda. Dla tej reakcji z równania (19.54) uzyskujemy $Q = -1,17$ MeV. Reakcja ta zachodzi ponieważ cząstka α emitowana przez ${}^{214}_{84}\text{Po}$ ma energię 7,5 MeV.

Gdy $Q > 0$, to mamy reakcję egzotermiczną. Reakcja taka może nastąpić przy dowolnej energii cząstki padającej.

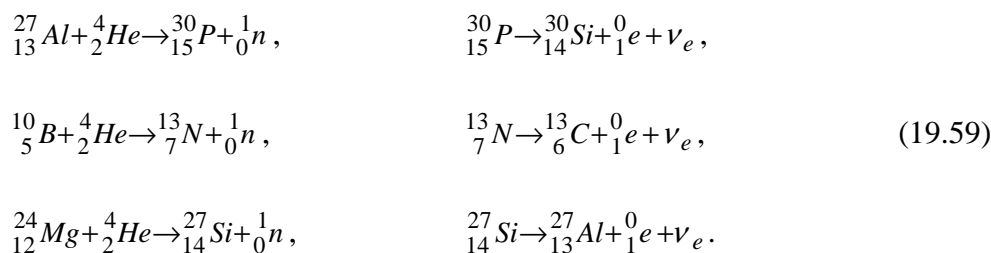
Cząstkami stosowanymi do przeprowadzenia reakcji jądrowych mogą być protony (1_1p lub 1_1H), deuterony (2_1H lub 2_1D), neutrony (1_0n), fotony (kwanty γ), trytony (3_1H lub 3_1T). Rzadziej stosuje się w tym celu ciężkie jądra; muszą posiadać dostatecznie dużą energię początkową, aby mogły zbliżyć się do jądra bombardowanego na odległość sił jądrowych.

W 1930 r. Bohe i Becker przeprowadzili reakcje jądrowe w których odkryto neutron:



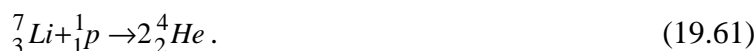
Jednak identyfikacji neutronu dokonał Chadwick w 1932 r. [w reakcji (19.57)] i jemu przypisuje się odkrycie neutronu.

W 1934 r. Fryderyk i Irena Joliot-Curie przeprowadzili reakcje jądrowe, w których odkryli pierwiastki sztucznie promieniotwórcze:



W jądrach sztucznie promieniotwórczych (${}^{30}_{15}P$, ${}^{13}_7N$, ${}^{27}_{14}Si$) zachodzi przemiana protonu w neutron według schematu opisanego wzorem (19.47). Cząstka 0_1e o dodatnim ładunku jednostkowym ma masę równą masie elektronu, spin równy $\hbar/2$ i nazywa się pozytonem.

Napiszemy jeszcze dwie reakcje jądrowe:



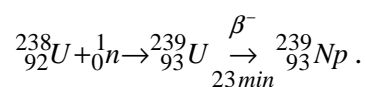
Reakcja (19.60) prowadzi do powstawania radioaktywnego węgla C-14 i zachodzi w atmosferze ziemskiej. Ten izotop węgla jest β -promieniotwórczym (${}^{14}_6C \rightarrow {}^{14}_7N + {}^0_{-1}e + {}^0_0\bar{\nu}_e$) o półokresie zaniku 5100 lat.

Drugą reakcją [wzór (19.61)] przeprowadzili w 1932 r. Cockroft i Walton i była to pierwsza reakcja jądrowa przeprowadzona przy pomocy cząstek przyśpieszanych w akceleratorze.

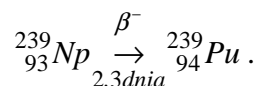
19.7.2. Reakcje rozszczepienia

Spośród wszystkich cząstek zdolnych do wywołania reakcji jądrowych, najbardziej skuteczne są neutrony. Neutrony z łatwością wnikają do jądra, ponieważ nie działa na nie kulombowska bariera potencjału. Wzbudzone jądro złożone, które powstaje w wyniku wychwytu neutronu, ulega różnym przemianom jądrowym z których najważniejsze to reakcje typu (n, α) , (n, p) , $(n, 2n)$. W przypadku jąder najcięższych możliwe jest także rozszczepienie wywołane neutronami.

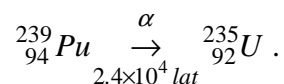
Absorpcja rezonansowa powolnych neutronów przez najbardziej rozpowszechniony izotop uranu ${}^{238}_{92}\text{U}$ prowadzi do powstania radioaktywnego izotopu ${}^{239}_{93}\text{U}$, który ulegając przemianie β^- przekształca się w izotop neptunu ${}^{239}_{93}\text{Np}$:



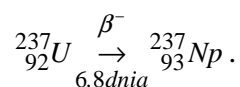
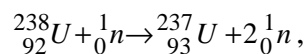
Z kolei jądro izotopu ${}^{239}_{93}\text{Np}$ o okresie połowicznego rozpadu 2,3 dnia, przekształca się w pluton ${}^{239}_{94}\text{Pu}$ według reakcji



Pluton ${}^{239}_{94}\text{Pu}$ jest cennym paliwem jądrowym, dzięki wydajnemu rozszczepianiu się pod wpływem neutronów termicznych odgrywa on podstawową rolę przy otrzymywaniu energii jądrowej. Pluton ${}^{239}_{94}\text{Pu}$ jest α -radioaktywny i ma bardzo duży półokres rozpadu (24 000 lat). Zmienia się na stabilny izotop uranu ${}^{235}_{92}\text{U}$:



Zaznaczmy jeszcze, że reakcja $(n, 2n)$ przekształca izotop uranu ${}^{238}_{92}\text{U}$ w sztuczny izotop radioaktywny ${}^{237}_{92}\text{U}$. Ten izotop z kolei emitując cząstkę β^- przekształca się w izotop neptunu ${}^{237}_{93}\text{Np}$, który jest α -promieniotwórczy i posiada długi półokres rozpadu $2,2 \times 10^6$ lat:

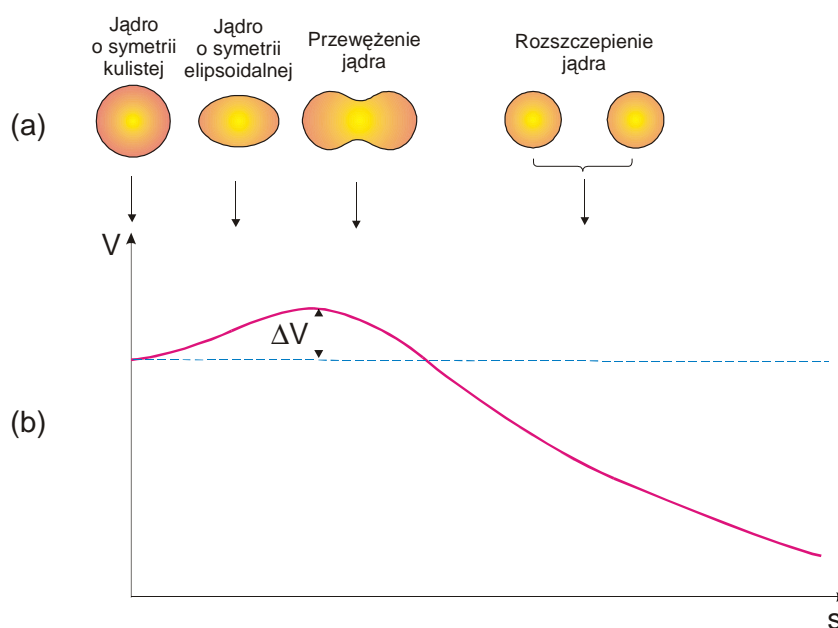


Izotop ${}^{237}_{93}\text{Np}$ daje początek jednej z rodzin promieniotwórczych opisanych poprzednio.

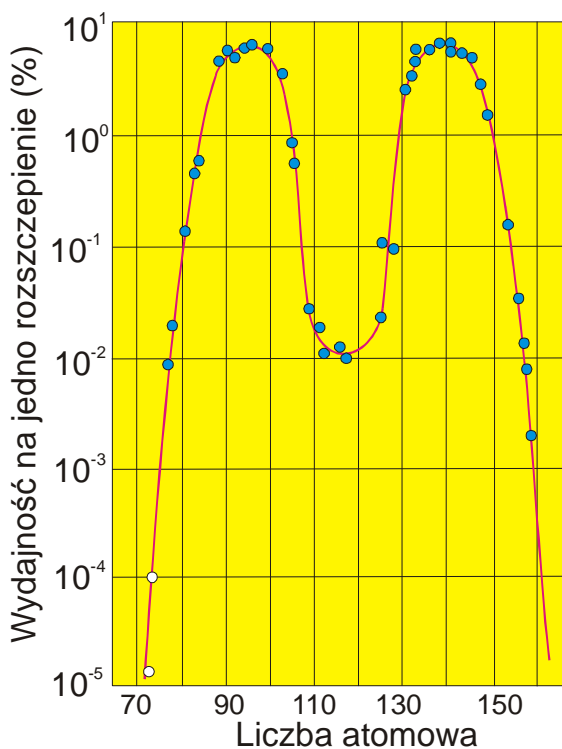
Poprzednio stwierdziliśmy, że rozpad ciężkich jąder atomowych na dwie części jest korzystny energetycznie. Wynika to z kształtu zależności średniej energii wiązania przypadającej na jeden nukleon od liczby atomowej (patrz rys. 19.2). Poczynając od $A \approx 60$, energia ta maleje ze wzrostem wartości A i np. dla ciężkiego jądra uranu wynosi około 7,6 MeV,

podczas gdy dla jąder o masie o połowę mniejszej wynosi około 8,3 MeV. Tak więc, gdyby jądro $^{238}_{92}\text{U}$ rozpadło się na dwie równe części, to wydzieliłoby się około 0,7 MeV na każdy nukleon, czyli rzędu 200 MeV na każde jądro uranu. Zysk energetyczny jest więc duży i wydawać by się mogło, że jądra te samorzutnie powinny ulegać temu procesowi. Tak jednak nie jest. Otóż podobnie jak w przypadku rozpadu α , czynnikiem warunkującym reakcję jest prawdopodobieństwo przeniknięcia cząstki przez barierę kulombowską, które w przypadku rozszczepienia jest znikomo małe. *Energię, jakiej należy dostarczyć jądro, aby mogło ulec rozszczepieniu nazywa się energią progową rozszczepienia lub energią aktywacji.*

Proces rozszczepienia ciężkiego jądra można wyjaśnić na podstawie modelu kropłowego jądra. Według tego modelu jądro, podobnie jak kropla, ma kształt kulisty. Przed rozdzieleniem się na dwie części jądro musi przejść przez pośrednie etapy zniekształceń (rys. 19.9), polegające na tym, że jądro o symetrii kulistej zmienia początkowo swój kształt na elipsoidalny, następnie ulega przewężeniu i po osiągnięciu stanu krytycznego rozpada się na dwie części. Zwiększenie powierzchni wskutek zniekształceń powoduje wzrost energii napięcia powierzchniowego jądra oraz nieznaczne zmniejszenie energii odpychania kulombowskiego ponieważ ładunki przemieszczają się na niewielką odległość i przeciwdziałają tym odkształceniom. Rys. 19.9 przedstawia również przebieg energii potencjalnej jądra w funkcji parametru s , określającego stopień wydłużenia jądra. Obliczenia wykazują, że wysokość maksimum energii liczonego w stosunku do jądra wyjściowego (parametr ΔV) wynosi kilka MeV. Wystarczy więc dostarczyć tej energii (przez pochłonięcie protonu, neutronu lub kwantu γ) aby nastąpił rozpad jądra na dwie części. Wysokość bariery ΔV wynosi około 6,5 MeV dla jądra $^{235}_{92}\text{U}$ i około 7,0 MeV dla jądra $^{238}_{92}\text{U}$. W reakcji rozszczepienia najbardziej prawdopodobny jest podział jądra na dwa fragmenty o stosunku mas około 3:2 (patrz rys. 19.10).

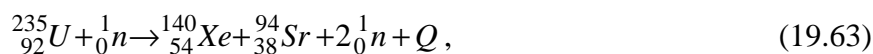
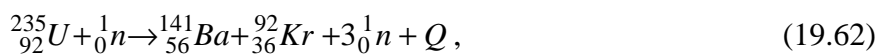


Rys. 19.9. Proces rozszczepienie jądra wg modelu kropłowego: (a) konfiguracja układu, (b) energia potencjalna w funkcji parametru rozszczepienia s .



Rys. 19.10. Zależność procentowego udziału rozkładu mas fragmentów przypadająca na jedno rozszczepienie jądra uranu ${}^{235}_{92}\text{U}$ pod wpływem promieniowania neutronowego.

Reakcję rozszczepienia po raz pierwszy przeprowadził E. Fermi w 1934 r. Zauważył, że uran napromieniowany przez neutrony daje produkty radioaktywne o kilku okresach połowicznego zaniku. Początkowo przypuszczano, że w wyniku napromieniowania powstają pierwiastki transuranowe (pozauranowe). W 1939 r. Hahn i Strassmann wykazali na podstawie dokładnej analizy radiochemicznej, że napromieniowanie uranu prowadzi do pojawienia się pierwiastków chemicznych ze środka układu okresowego. Następnie Frisch i Meitner podali wyjaśnienie tego zjawiska jako rozszczepienie jądra na dwa jądra ze środkowej części układu okresowego pierwiastków. Przykłady reakcji rozszczepienia są przytoczone poniżej:



gdzie Q jest energią jądrową (200 MeV), wydzielającą się w jednym akcie rozszczepienia jednego jądra uranu.

19.7.3. Reakcje łańcuchowe. Reaktor atomowy

Zjawisko rozszczepienia jąder uzyskało ogromne znaczenie praktyczne dzięki możliwości wykorzystania go do produkcji energii. Możliwość ta wiąże się z występowaniem tzw. reakcji łańcuchowej. W procesie rozszczepienia uranu powstaje średnio 2.5 nowego neutronu. W bryle uranu te neutrony zostaną przechwycone przez następne jądra. Jeżeli średnio więcej niż jeden z nich wywoła dalsze rozszczepienia, to liczba powstających neutronów i rozszczepionych jąder będzie się samorzutnie zwiększać.

Do opisu reakcji łańcuchowej wprowadzimy współczynnik powielania k , określony jako stosunek liczby neutronów w dwu następujących po sobie generacjach. Warunkiem podtrzymania reakcji łańcuchowej jest aby $k \geq 1$. Przypuśćmy, że liczba neutronów rozszczepialnych w danym ogniwie reakcji łańcuchowej jest n , zaś w następnym ogniwie już kn

neutronów. Przyrost dn liczby neutronów jest równy $dn = kn - n = n(k - 1)$. Przyrost liczby neutronów w jednostce czasu, czyli szybkość rozwoju reakcji łańcuchowej wynosi

$$\frac{dn}{dt} = \frac{n(k - 1)}{\tau},$$

stąd

$$\frac{dn}{n} = \frac{k - 1}{\tau} dt,$$

a po scałkowaniu tego równania otrzymamy

$$n = n_0 \exp\left[\frac{(k - 1)t}{\tau}\right], \quad (19.64)$$

gdzie n_0 jest liczbą neutronów w chwili $t = 0$, a τ jest czasem dzielącym od siebie dwa ogniwa reakcji łańcuchowej.

Z równania (19.64) wynikają ważne wnioski:

- gdy $(k - 1) > 0$, czyli gdy $k > 1$, $n > n_0$, czyli liczba neutronów wchodzących w reakcję rozszczepienia rośnie; taką reakcję łańcuchową nazywamy reakcją rozwijającą się,
- gdy $(k - 1) = 0$, czyli gdy $k = 1$ i $n = n_0$, czyli liczba neutronów nie zmienia się i reakcję łańcuchową nazywamy stacjonarną,
- gdy $(k - 1) < 0$, $n < n_0$, czyli liczba neutronów maleje z upływem czasu; taka reakcja nazywa się reakcją łańcuchową zanikającą.

Uran naturalny posiada 0,7% izotopu $U-235$; resztę stanowi izotop $U-238$. Uran $U-238$ rozszczepia się tylko pod wpływem neutronów szybkich, natomiast $U-235$ może rozszczepiać się pod wpływem neutronów o dowolnej energii. Neutrony powstające po rozszczepieniu charakteryzują się energią (lub szybko osiągają ją w zderzeniach) niższą od energii progowej rozszczepienia $U-238$ co powoduje, że neutrony te będą silnie pochłaniane przez $U-238$ nie prowadząc do reakcji rozszczepienia. Wynika z tego konieczność wzbogacania masy rozszczepialnej o $U-235$ i zastosowanie moderatorów (spowalniaczy) neutronów.

Zmniejszenie objętości i masy materiału rozszczepialnego powoduje zwiększenie ucieczki neutronów poza obszar masy rozszczepialnej. *Najmniejsza masa, w której jest możliwe rozwinięcie reakcji łańcuchowej nazywa się masą krytyczną* (wówczas $k = 1$). W celu zmniejszenia strat neutronów i zmniejszenia masy krytycznej, materiał rozszczepialny otacza się reflektorem neutronów. Reflektor stanowią substancje stosowane na moderatory takie jak: H_2O , D_2O , grafit, beryl czy związki berylu.

Urządzenia w których przeprowadza się kontrolowane łańcuchowe reakcje rozszczepienia, nazywają się reaktorami jądrowymi lub atomowymi. Pierwszy reaktor jądrowy zbudowany został pod kierunkiem E. Fermiego w Chicago w 1942 r. Jako paliwo jądrowe stosuje się: ${}^{235}_{92}U$, ${}^{233}_{94}Pu$, ${}^{238}_{92}U$ i ${}^{232}_{90}Th$. Energia neutronów rozszczepialnych uranu $U-235$ wynosi około 0,7 MeV. Warunkiem rozwinięcia reakcji łańcuchowej w termicznych reaktorach jądrowych (w

tych reaktorach wykorzystuje się neutrony termiczne) jest zastosowanie moderatora w celu zmniejszenia rezonansowego wychwytu neutronów przez jądra uranu $U-238$. Paliwo jądrowe (najczęściej w postaci prętów) umieszczane jest wewnątrz moderatora, np. w graficie lub ciężkiej wodzie. Pierwsze reaktory zawierały około 45 ton uranu i około 450 ton grafitu (służącego jako moderator i reflektor neutronów).

Podczas rozruchu reaktora jądrowego rośnie ilość wydzielanej energii (reaktor stanowi układ nadkrytyczny; $k > 1$) i kiedy zostanie osiągnięta żądana moc reaktora, musi on przejść w stan krytyczny ($k = 1$) i ten stan musi być utrzymany przez cały czas pracy reaktora. W celu zmniejszenia współczynnika powielania neutronów do rdzenia reaktora wprowadza się pręty z materiałów, które silnie pochłaniają neutrony termiczne, np. z boru lub kadmu.

Rozróżniamy reaktory badawcze i techniczne; m.in. reaktory do produkcji pierwiastków promieniotwórczych, stosowanych w medycynie i technice, reaktory do wytwarzania energii (elektrownie jądrowe). Ważną rolę spełniają reaktory wytwarzające paliwo jądrowe, stosowane w bombach atomowych i reaktorach jądrowych.

19.7.4. Reakcje termonuklearne

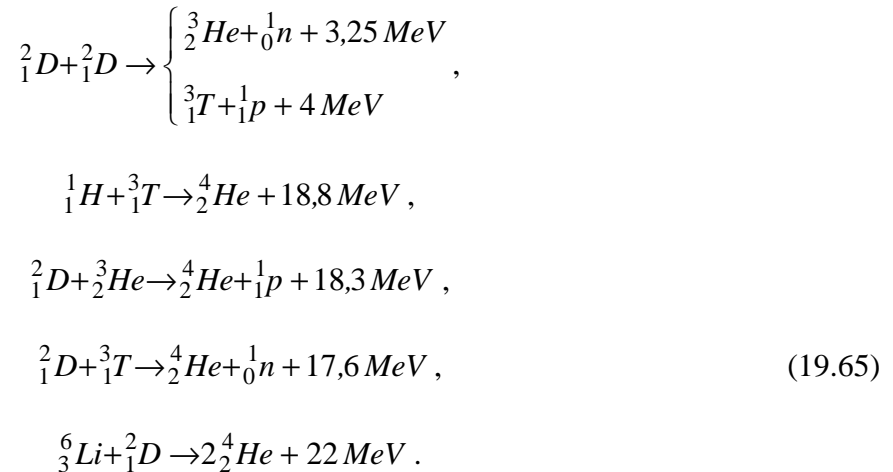
Z rys. 19.2 widzimy, że masa lub energia spoczynkowa dwóch lekkich jąder jest większa niż masa jądra powstającego po ich połączeniu. Np. dwa deuteriumy mogą się połączyć tworząc jądro helu. Masa jądra helu jest o około 24 MeV mniejsza, niż suma mas deuteriumów i energia 24 MeV wyzwoli się przy takim połączeniu (syntezie) jąder. Jeżeli dwa deuteriumy łączą się tworząc jądro helu, to 0,6% masy spoczynkowej jest zamieniane na energię. Podobnie, w procesie syntezy czterech protonów na jądro helu uwalnia się energia około 26 MeV, co stanowi zamianę około 7% początkowej masy spoczynkowej na energię. Widzimy więc, że gdyby proces syntezy mógł być wykorzystywany do produkcji energii mielibyśmy metodę sześć razy wydajniejszą niż rozszczepienie uranu. Poza tym mamy do dyspozycji niemal nieograniczone i niedrogi źródło deuteriumu w wodzie mórz i oceanów, znacznie łatwiej dostępne niż źródła materiałów do reakcji rozszczepienia.

Procesy syntezy pierwiastków są realizowane we Wszechświecie. Stanowią one główne źródło energii gwiazd; ich temperatura jest rzędu 10^7 K. Taka temperatura jest warunkiem zajścia reakcji termonuklearnych, tj. reakcji syntezy jąder. Cząstki dodatnie charakteryzują się wysoką energią kinetyczną wystarczającą do pokonania sił kulombowskiego odpychania umożliwiając przez to ich zbliżenie na odległość działania sił przyciągania jądrowego. Słońce w 80% składa się z wodoru a swoją energię zawdzięcza reakcjom termonuklearnym przebiegającym w zamkniętym cyklu reakcji.

Z chwilą zrealizowania w warunkach ziemskich produkcji energii w wyniku rozszczepienia ciężkich jąder, zaczęto zastanawiać się nad możliwością uzyskania w tych warunkach produkcji energii w reakcjach termojądrowych z lekkimi pierwiastkami. Pierwsze praktyczne

urzeczywistnienie tych możliwości nastąpiło w tzw. bombie wodorowej. Bardzo wysokie temperatury osiągnięte przy wybuchu zwykłej bomby atomowej stworzyły warunki konieczne do uzyskania reakcji termojądrowej. Bomba atomowa jest więc zapalnikiem dla przebiegających również wybuchowo reakcji syntezy. Nadal poszukuje się rozwiązań przeprowadzenia reakcji termojądrowych w sposób kontrolowany, dzięki czemu energia ta mogłaby być wykorzystywana do celów pokojowych.

Jako paliwo termojądrowe można wykorzystać szereg substancji. Poniżej podano kilka przykładów reakcji syntezy:



Podstawowym warunkiem uzyskania kontrolowanej reakcji termojądrowej jest utrzymanie stabilnej, wysokiej temperatury dziesiątek i setek milionów stopni w pewnej objętości deuteru lub jego mieszaniny z trytem. Przy tak wysokiej temperaturze gaz będzie całkowicie zjonizowany; taki stan gazu nazywamy plazmą. Zauważmy, że jeżeli problem kontrolowanej reakcji termojądrowej zostanie rozwiązany, to ludzkość rozwiąże problem braku energii na zawsze.

Obecnie na świecie pracuje około 400 elektrowni atomowych o łącznej mocy około 1000 GW. Obserwuje się jednak spadkową tendencję rozwoju energetyki atomowej. Jest to wynikiem problemów związanych z jej rozwojem, takich jak: bezpieczeństwo, odpady promieniotwórcze (a dokładniej ich pozbywanie się), nieprolifercja (rozprzestrzenienie się broni jądrowej), likwidacja obiektów po zakończeniu produkcji (po 30 latach), opłacalność ekonomiczna. Spadkową tendencję obserwuje się również w produkcji bomb jądrowych (procesy rozbrojeniowe i ekonomiczne).

Bomba atomowa składa się z dwu lub więcej kawałków uranu ${}^{235}\text{U}$ (lub plutonu ${}^{239}\text{Pu}$, który zachowuje się podobnie), przy czym masa każdego z tych kawałków jest mniejsza od masy krytycznej. Wobec tego w kawałkach nie może rozwijać się reakcja łańcuchowa (masa krytyczna ${}^{235}\text{U}$ wynosi 43 kg; 50 kg tej masy stanowi więc stan nadkrytyczny – wybuchowy; czas wybuchu wynosi 2,4 μs). Jeżeli kawałki te złączą się, to powstaje bryła ${}^{235}\text{U}$ o rozmiarach przewyższających krytyczne i następuje wybuchowa reakcja lawinowa. Energia wydzielana przy wybuchu bomby atomowej jest porównywalna do energii wyzwalanej przy wybuchu 20–

100 tysięcy trójnitrotoluenu (TNT). W chwili wybuchu bomby wytwarza się w jej wnętrzu temperatura rzędu 10 milionów K, po czym spada.

W pierwszych chwilach po wybuchu powstają w niewielkiej objętości warunki zbliżone do panujących we wnętrzu gwiazd, a więc warunki umożliwiające przeprowadzenie reakcji termonuklearnej. Wykorzystano to do produkcji jeszcze potężniejszego narzędzia niszczenia bomby wodorowej. W temperaturze rzędu 10^7 K zachodzi reakcja opisana wzorem (19.65), która stanowi podstawę działania bomb wodorowych. Energia wyzwalana przez bomby wodorowe odpowiada energii otrzymanej przy wybuchu od kilku do kilkunastu milionów ton TNT (MT TNT).

Trzecim rodzajem bomby nuklearnej jest bomba wodorowo-uranowa. W tym przypadku bomba wodorowa otoczona jest masą uranu ^{238}U (w postaci płaszcza), który ulega rozszczepieniu pod wpływem szybkich neutronów powstających w dużych ilościach w reakcji termojądrowej [wzór (19.65)]. W ten sposób powstaje dodatkowa energia oraz chmura promieniotwórczych produktów rozszczepienia, wśród których najbardziej długowieczne są ^{90}Sr (okres połowicznego rozpadu 28 lat), ^{137}Cs (27 lat), ^{144}Ce (290 dni) i ^{95}Zr (65 dni). Bomby wodorowo-uranowe (^{238}U) mają moc około 100 MT TNT.

W arsenałach broni nuklearnej USA i Rosji znajduje się około 20–30 tysięcy pocisków z głowicami jądrowymi różnego typu. W całym świecie jest ich około 50 tysięcy, a ich moc jest wystarczająca aby wielokrotnie zniszczyć życie na Ziemi.

Pierwsze bomby atomowe wyprodukowano pod kierownictwem Oppenheimera i z udziałem Comptona, Fermiego, Lawrence'a i Bethego. Dnia 15 lipca 1945 r. wybuchła na pustyni w New Mexico (USA) pierwsza w historii ludzkości bomba atomowa, a już 6 sierpnia tego roku podobna bomba spadła na Hiroszime (Japonia) zabijając około 130 tysięcy ludzi. 28 kwietnia 1986 r. doszło na największej katastrofy jądrowej na skutek awarii reaktora atomowego w Czarnobylu (Ukraina).

ROZDZIAŁ 20

ASTROFIZYKA I KOSMOLOGIA

W tym rozdziale poznamy uderzające przykłady uniwersalności praw przyrody. Postaramy się wyjaśnić jak mogły powstać gwiazdy i planety, skąd pochodzi ogromna energia wysyłana przez gwiazdy i co się dzieje z gwiazdą, gdy jej źródło energii wyczerpuje się. Zajmiemy się opisem takich obiektów jak biały karzeł, gwiazda neutronowa i czarna dziura. Pozwoli nam to zrozumieć właściwości stanów materii w wielkiej skali. Zajmiemy się także budową i prawami rządzącymi Wszechświatem.

20.1. Pojęcia wstępne

Układ Słoneczny, do którego należy Ziemia, jest mikroskopijną cząstką Układu Drogi Mlecznej, naszej Galaktyki, która z kolei jest nieporównywalnie mniejszą częścią Wszechświata. Można się o tym przekonać porównując liczby zebrane w tabeli 20.1.

Tabela 20.1. Podstawowe parametry Galaktyki

Nazwa wielkości	Wartość liczbowa
Średnica	$9,2 \times 10^{20}$ m
Grubość średnia	$6,2 \times 10^{16}$ m
Średnica jądra Galaktyki	$1,55 \times 10^{17}$ m
Masa Galaktyki	$2,2 \times 10^{41}$ kg
Masa Słońca	2×10^{30} kg
Prędkość Słońca w ruchu obrotowym Galaktyki	250 km/s
Promień orbity Słońca	3×10^{20} m
Okres obrotu Słońca	7×10^{15} s = $2,2 \times 10^8$ lat
Liczba gwiazd w Galaktyce	10^{11}

Galaktyka ma kształt dysku o średnicy około 10^{21} m i grubości w części środkowej około 6×10^{19} m. Jej grubość maleje wzdłuż promienia. Centralną część Galaktyki otacza halo galaktyczne, które jest pyłowo-gazowym obłokiem skupiającym około 2% całej masy Galaktyki.

Wszechświat jest wypełniony galaktykami i gromadami galaktyk znajdujących się w ciągłym ruchu obrotowym i postępowym. Jego ogrom uzmysławiają dane liczbowe zawarte w tabeli 20.2. Czas istnienia Wszechświata jest oceniany na około 15 miliardów lat. W 1990 r. wysłano

w przestrzeń kosmiczną teleskop im. E. Hubble'a celem zebrania nowych danych eksperymentalnych o Wszechświecie (zasięg obserwacji wzrósł blisko dziesięciokrotnie).

Tabela 20.2. Podstawowe parametry Wszechświata

Nazwa wielkości	Wartość liczbowa
Średnica	3×10^{26} m
Objętość	$1,5 \times 10^{79}$ m ³
Masa (10^{79} nukleonów)	$1,6 \times 10^{52}$ kg
Średnia gęstość	2×10^{28} kg/m ³
Liczba galaktyk we Wszechświecie	10^{12}
Średnia masa galaktyki	$1,6 \times 10^{41}$ kg
Średnia liczba gwiazd w galaktyce	10^{10}
Czas istnienia Wszechświata	$1,5 \times 10^{10}$ lat

Wszechświat cechuje się niestabilnością; może się rozszerzać lub skupiać. Obecnie znajduje się w stanie ekspansji. Fakty eksperymentalne świadczą, że Wszechświat rozszerza się z prędkościami proporcjonalnymi do wzajemnej odległości galaktyk i ich gromad. Takie zachowanie sugeruje, że we wczesnych stadiach rozwoju Świata materia była skupiona w małej objętości, a proces jej rozszerzania nastąpił w wyniku eksplozji. Eksplozja ta nosi nazwę *Wielkiego Wybuchu* (ang. Big Bang) i wszystkie dalsze etapy rozwoju Wszechświata odnosimy do tego hipotetycznego początku.

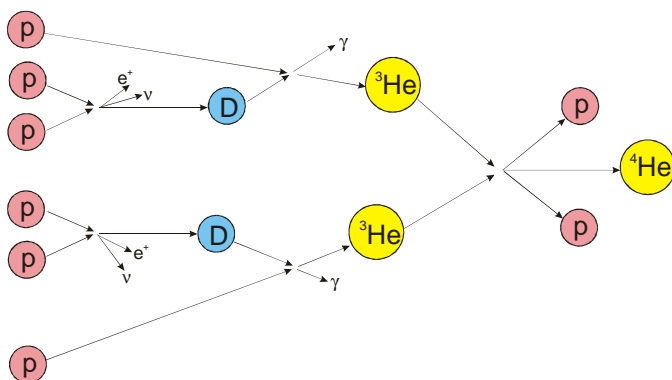
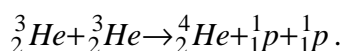
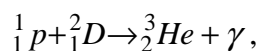
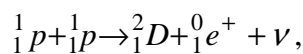
20.2. Gwiazdy

Większość teorii kosmologicznych przyjmuje, że głównym składnikiem gwiazd był wodór. Tworzenie się gwiazd rozpoczyna się od powstania chmury wodoru, która zaczyna kurczyć się pod wpływem przyciągania grawitacyjnego. W miarę jak atomy wodoru zbliżają się do siebie, rośnie ich energia kinetyczna, czyli temperatura gazu. Z kolei nagrzana masa gazu osiąga ciśnienie które hamuje dalsze zapadanie grawitacyjne chmury. Jednak, gdy energia zmniejszy się na skutek promieniowania elektromagnetycznego, to zapadanie grawitacyjne postępuje nadal aż do momentu pojawienia się nowego źródła energii temu przeciwdziałającemu, jakim są reakcje termojądrowe. Jak wiemy reakcje te zaczynają zachodzić gdy temperatura wynosi około 10^7 K.

Jeżeli masa początkowa jest mała, to zapadanie postępuje aż do chwili, kiedy atomy zaczynają stykać się ze sobą. Wtedy powstaje planeta, tak jak w przypadku Ziemi. Jeżeli masa początkowa jest większa, to wtedy ciśnienia i gęstości są dostatecznie duże, aby powodować nakładanie się funkcji falowych w wyniku czego powstaje plazma. Przykładem obiektu należącego do tej kategorii jest planeta Jowisz. W tym przypadku ciśnienie grawitacyjne jest

równoważone przez ciśnienie kwantowo-mechaniczne plazmy (ciśnienie to jest omawiane w pkt. 20.3).

Można wykazać, że jeśli masa początkowa jest większa niż 0,08 masy Słońca, to osiągnięta temperatura jest dostatecznie wysoka aby wywołać reakcje termojądrowe typu



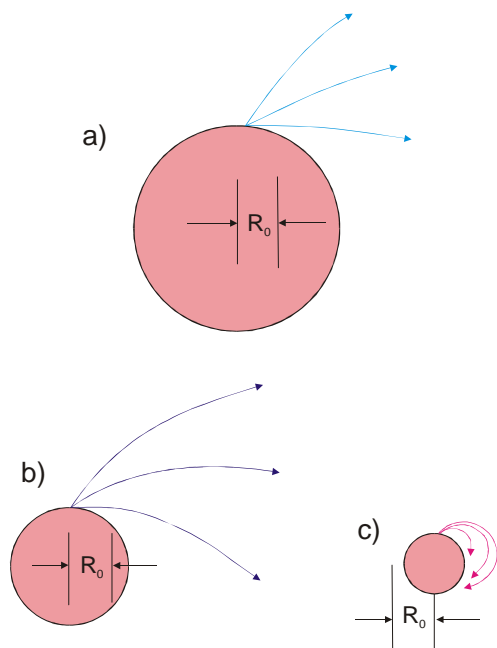
Rys. 20.1. Schemat cyklu wodorowego. Sześć protonów jest użytych do wytworzenia 4_2He , dwóch protonów, dwóch pozytonów, dwóch neutrin i dwóch kwantów γ .

Ten ciąg reakcji termojądrowych pokazany na rys. 20.1 nazywany jest cyklem wodorowym powodującym produkcję energii Słońca i innych gwiazd bogatych w wodór. W wyniku tego cyklu sześć protonów jest zużytych do emisji cząstki α , dwóch pozytonów, 2 neutrin i 2 fotonów o całkowitej energii kinetycznej około 26 MeV.

Gwiazdy o masie większej od Słońca mają wyższą temperaturę (cechują się większą jasnością oraz bardziej niebieską barwą światła) i szybciej spalają zawarty w nich wodór. Gdy zapas wodoru wyczerpie się, to gwiazda promieniuje nadal i zaczyna się zapadać. Sądzi się, że w miarę zmniejszania się promienia gwiazdy, połowa wyzwolanej energii grawitacyjnej zużywa się na promieniowanie a druga część ogrzewa wewnątrz gwiazdy do wyższej temperatury. Zapadanie i ogrzewanie gwiazdy trwa tak długo, aż temperatura osiągnie wartość kreuującą reakcje termojądrowe dla helu (przemiana w węgiel, tlen i neon). W wyniku tych procesów większość materiału we wnętrzu gwiazdy zamienia się w żelazo ${}^{56}Fe$, który jest izotopem najbardziej stabilnym wśród wszystkich jąder (dalsze reakcje jądrowe musiałyby zużyć energię zamiast ich dostarczać).

20.2. Czarne dziury

Jeżeli przyjmiemy, że nie występują inne siły odpychania lub ciśnienia, to gwiazda będzie się wciąż zapadać. Istnieje jednak pewien graniczny promień R_0 , zwany promieniem Schwarzschilda, poniżej którego nie możemy już zobaczyć gwiazdy. Sytuacja taka jest pokazana na rys. 20.2c, gdzie wszystkie fotony spadają z powrotem na powierzchnię gwiazdy.



Rys. 20.2. Gwiazdy o jednakowych masach i trzech różnych promieniach. Dla (a) i (b) promień jest większy od R_0 , gdzie $R_0 = 2GM/c^2$. Dla (c) promień jest mniejszy niż R_0 i pole grawitacyjne jest tak silne, że fotony nie mogą uciec z gwiazdy. W każdym przypadku narysowano tory fotonów wysyłanych pod kątami 0° , 30° i 60° do powierzchni gwiazdy.

Oceniając grawitacyjną energię potencjalną fotonu wysyłanego z powierzchni gwiazdy, $U = -GMm/R$ (M jest masą gwiazdy a $m = \varepsilon/c^2$ jest masą grawitacyjną fotonu o energii ε), możemy wyznaczyć R_0 . Jeżeli więc $\varepsilon < GMm/R$, to foton nie może oddalić się od gwiazdy. Warunek graniczny dla R ma postać

$$\varepsilon = \frac{GMm}{R_0} = \frac{GM}{R_0} \left(\frac{\varepsilon}{c^2} \right),$$

co prowadzi do związku

$$R_0 = \frac{GM}{c^2}.$$

W tych rozważaniach nie uwzględniliśmy ogólnej teorii względności, którą należy uwzględniać jeżeli energia grawitacyjna jest porównywalna z energią całkowitą. Poprawne obliczenia dają nieco inny wynik

$$R_0 = \frac{2GM}{c^2}. \tag{20.1}$$

Szacując tą wielkość dla masy równej masie Słońca (2×10^{30} kg) otrzymujemy $R_0 = 3$ km. Jeżeli gwiazda zapada się tak, że jej promień jest mniejszy od $2GM/c^2$, to żadna cząstka ani światło przez nią emitowane nie mogą dotrzeć do Ziemi (lub innego odległego obserwatora). Jednakże pole grawitacyjne tej gwiazdy nadal jest odczuwalne; cząstki lub światło nadal będą spadać na gwiazdę – stąd nazwa "czarna dziura".

20.3. Ciśnienie kwantowo-mechaniczne

W punkcie 20.2 wspomnieliśmy o efekcie związanym z siłami odpychającymi, który przyczynia się do wzrostu energii kinetycznej i ciśnienia przy dużych gęstościach. Z zasady Pauliego wynika, że cząstki o spinie połówkowym (np. elektrony, protony czy neutrony) nie mogą zajmować dowolnych stanów pędowych i energetycznych, i muszą zajmować kolejne stany aż do tak zwanego pędu (energii) Fermiego, p_F . Jeżeli N takich cząstek znajduje się w objętości V , to średnia odległość w przestrzeni pomiędzy cząstkami jest rzędu $(V/N)^{1/3}$. Z zasady nieoznaczoności Heisenberga wynika nieokreśloność pędu cząstki $\sim h(V/N)^{1/3}$. Wyrażenie dla p_F dane jest wzorem [porównaj z (15.41)]

$$p_F = h \left(\frac{3N}{8\pi V} \right)^{1/3}. \quad (20.2)$$

Ogólne wyrażenie relatywistyczne określające energię Fermiego ma postać

$$K_F = \left(p_F^2 c^2 + m^2 c^2 \right)^{1/2} - mc^2. \quad (20.3)$$

Dla małych $p_F \ll mc$, cząstka jest nierelatywistyczna i wtedy

$$K_F = \frac{p_F^2}{2m} = \frac{h^2}{8m} \left(\frac{3N}{\pi V} \right)^{2/3}. \quad (20.4)$$

Iloczyn ciśnienia i objętości jest dany wyrażeniem (15.32)

$$pV = N \left(\frac{m \overline{v^2}}{3} \right) = \frac{2}{3} N \overline{K}.$$

A ponieważ średnia energia kinetyczna cząstki $K = (3/5)K_F$ (patrz pkt. 15.6.2), więc

$$pV = \frac{2}{3} N \left(\frac{3}{5} K_F \right),$$

czyli

$$p = \frac{2}{5} \frac{NK_F}{V} = \frac{h^2}{20m} \left(\frac{3}{\pi} \right)^{3/5} \left(\frac{N}{V} \right)^{5/3}. \quad (20.5)$$

Jest to ciśnienie kwantowo-mechaniczne nierelatywistycznego gazu Fermiego.

W przypadku cząstek nierelatywistycznych, $p_F \gg mc$ i energia Fermiego wynosi

$$K_F = p_F c = \frac{1}{2} hc \left(\frac{3N}{\pi V} \right)^{1/3}, \quad (20.6)$$

a całkowita energia kinetyczna wszystkich cząstek wynosi $(3/4)NK_F$.

20.4. Ewolucja gwiazd

Zimna wypalona gwiazda będzie nadal promieniowała energię i kurczyła się nawet po wypaleniu paliwa jądrowego, aż do osiągnięcia najmniejszej możliwej wartości energii $E(R)$.

Rozważymy teraz gwiazdę składającą się z atomów o liczbie masowej A . Niech N będzie całkowitą liczbą nukleonów. Mamy wtedy N/A jąder, każde zawiera Z protonów i $(A-Z)$ neutronów. Oznaczmy $x = Z/A$, które dla rozważanych jąder wynosi około 1/2. Całkowita masa gwiazdy wynosi $M \approx Nm_p$, gdzie m_p jest masą protonu. Całkowita liczba elektronów i protonów jest $N_e = xN$.

Jak poprzednio zaznaczono, całkowita energia gwiazdy składa się z całkowitych energii Fermiego cząstek oraz grawitacyjnej energii potencjalnej. Dla kuli o jednorodnej gęstości, energia potencjalna $U = -(3/5)GM^2/R$. Jeżeli temperatura gwiazdy jest na tyle niska, że można zaniedbać energię termiczną w porównaniu z energią kwantowo-mechaniczną, to całkowita energia gwiazdy jest dana wzorem

$$E(R) = N_e \bar{K}_e + U = \frac{3}{5} N_e (K_F)_e - \frac{3}{5} \frac{GM^2}{R}.$$

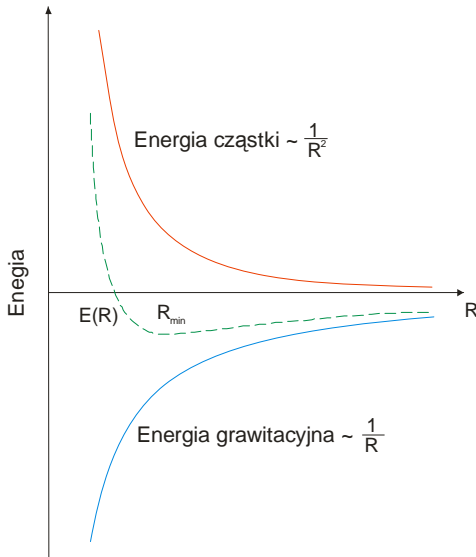
Po uwzględnieniu wzoru nierelatywistycznego (20.4) (stosujemy ten wzór ponieważ mamy do czynienia z zapadaniem się gwiazdy o dostatecznie małej masie po wyczerpaniu się paliwa jądrowego) podstawiając m_e zamiast m , xN zamiast N oraz $V = (4/3)\pi R^3$, otrzymuje

$$E(R) = \frac{3}{5} \frac{xN}{R^2} \frac{h^2}{8m_e} \left(\frac{9}{4\pi^2} xN \right)^{2/3} - \frac{3}{5} G \frac{N^2 m_p^2}{R}. \quad (20.7)$$

Funkcja $G(R)$ jest pokazana na rys. 20.3. Najniższą wartość energii znajdziemy przyrównując pochodną dE/dR do zera. Z tego warunku otrzymujemy

$$R = \frac{xh^2}{4m_e} \left(\frac{9}{4\pi^2} xN \right)^{2/3} \frac{1}{GNm_p^2}. \quad (20.8)$$

Jeżeli następnie tą wartość wstawimy do równania ((20.7), to zauważymy, że grawitacyjna energia wiązania ($-U$) jest dwa razy większa niż całkowita energia kinetyczna [pierwszy człon we wzorze (20.7)].



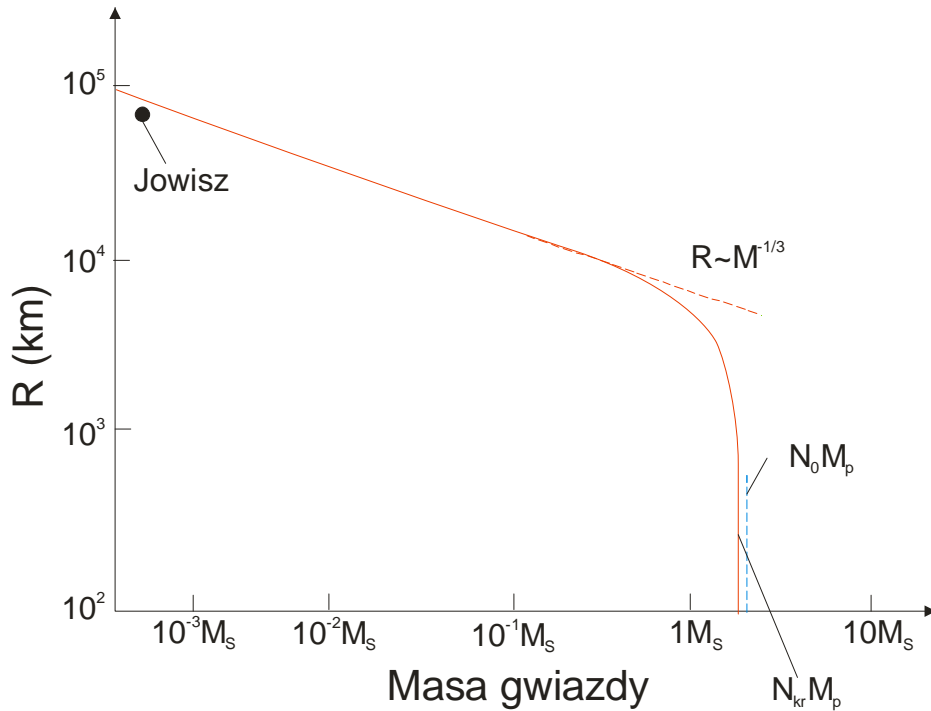
Rys. 20.3. Dwa wkłady do całkowitej energii białego karła. Górna krzywa opisuje całkowitą energię kinetyczną elektronów, a dolna krzywa grawitacyjną energię potencjalną gwiazdy. Suma obydwu krzywych oznaczona linią przerywaną ma minimum dla $R = R_{min}$.

Dla typowego karła o masie 15% mniejszej niż masa Słońca ($N = 10^{57}$, $x = 1/2$) z równania (20.8) otrzymujemy $R \approx 8000$ km, co odpowiada gęstości $\rho \approx 3 \times 10^6$ g/cm³. Czyli białe karły mają rozmiary zbliżone do Ziemi, ale są milion razy cięższe. Okazuje się, że planeta Jowisz może być rozważana jako biały karzeł.

Jeżeli użyjemy w równaniu (20.7) dokładnego, relatywistycznego wzoru na $(K_F)_e$ według wzoru (20.6), to otrzymamy wynik przedstawiony na rys. 20.4 dla R w funkcji masy. Wtedy równanie (20.7) zastępujemy równaniem

$$E(R) = \frac{3}{4} \frac{xN}{R} \frac{hc}{2} \left(\frac{9}{4\pi^2} xN \right)^{1/3} - \frac{3}{5} G \frac{N^2 m_p^2}{R}. \quad (20.9)$$

W tym przypadku zależność od R jest taka sama dla obydwu składników i drugi człon staje się dominujący, gdy Nm_p staje się dostatecznie duże. Gdy N przekroczy wartość krytyczną N_{kr} , energia $E(R)$ maleje ze zmniejszaniem promienia. Nie ma położenia równowagi dla żadnej wartości promienia, jeśli nie występują energia i ciśnienie termiczne. Oznacza to, że dostatecznie ciężkie gwiazdy mogą nadal kurczyć się i promieniować.



Rys. 20.4. Wielkość R , która minimalizuje sumę energii Fermiego dla elektronów i energii grawitacyjnej w funkcji masy gwiazdy M . Przyjęto jednorodną gęstość gwiazdy i relatywistyczne wyrażenie na energię Fermiego. Dla małych mas uzyskujemy przybliżenie nierelatywistyczne opisane równaniem (20.8). Zauważmy ograniczenie na masę dane równaniem (20.10), gdy elektrony stają się relatywistyczne.

Oceniemy teraz krytyczną liczbę nukleonów N_{kr} dla której składniki równania (20.9) są sobie równe. Wtedy otrzymujemy

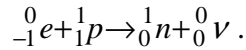
$$N_{kr} \approx \frac{(125\pi)^{1/2}}{4} x^2 \left(\frac{hc}{Gm_p^2} \right)^{3/2}.$$

We wzorze tym pojawia się bezwymiarowa wielkość N_0

$$N_0 = \left(\frac{hc}{Gm_p^2} \right)^{3/2}. \quad (20.10)$$

Dokładne wyznaczenie N_{kr} wymaga przeprowadzenia całkowania numerycznego i daje wynik $N_{kr} = 0.7(x/0.5)^2 N_0$. Odpowiadająca temu masa $N_{kr}m_p$ zwana *masą Chandrasekhara* jest maksymalną jaką może mieć biały karzeł, który będzie stygł do stabilnego stanu o skończonym promieniu i gęstości. Masa ta jest około 40% większa niż masa Słońca $M_s = 0.49N_0m_p$.

Przy dostatecznie dużej gęstości [patrz wzór (20.6)] energie kinetyczne elektronów są wysokie. W takich warunkach może zachodzić reakcja



Jest to odwrotny rozpad β . Dokładny obraz procesu przemiany zwykłej materii w materię bogatą w neutrony jest skomplikowany. Obliczenia wykazują, że neutrony są znacznie liczniejsze niż protony przy gęstościach 10^{11} g/cm³. Mamy więc do czynienia z gwiazdą neutronową. Chcemy znaleźć stan równowagi dla określonego R i x dla ustalonej całkowitej liczby nukleonów N , przy czym $N_p = N_e = xN$. Stosujemy równanie (20.6), ponieważ w każdej gwiazdzie neutronowej elektrony są skrajnie relatywistyczne.

Z równania (20.2) po przekształceniu możemy uzyskać

$$\frac{(xN)m_p}{V} = \frac{8\pi}{3} \left(\frac{m_e c}{h} \right)^3 m_p,$$

zastępując N_e przez xN a p_F przez $m_e c$ i mnożąc obie strony równania przez m_p . Z ostatniego równania uzyskujemy wzór na gęstość

$$\rho_r = \frac{8\pi m_p}{3x} \left(\frac{m_e c}{h} \right)^3. \quad (20.11)$$

Neutrony o znacznie większej masie m_p niż elektrony pozostaną nierelatywistyczne dopóki gęstość będzie mniejsza niż wartość graniczna wyznaczona dla neutronów z równania analogicznego do (20.11)

$$(\rho_r)_n = \frac{8\pi m_p}{3(1-x)} \left(\frac{m_e c}{h} \right)^3 = \frac{6 \times 10^{15}}{1-x} \text{ g/cm}^3. \quad (20.12)$$

Dopóki gęstość jest mniejsza niż wartość graniczna wyznaczona z ostatniego równania, dopóty możemy używać równania (20.4) dla energii neutronów, a energię protonów możemy zaniedbać, ponieważ stanowią one niewielką część jądra. Aby otrzymać zależność $E(R,x)$ dodajemy całkowitą energię kinetyczną neutronów do równania (20.9). Wobec tego

$$\begin{aligned} E(R,x) &= \text{energia neutronów} + \text{energia elektronów} + \text{energia grawitacyjna} = \\ &= \frac{3}{5} \frac{(1-x)N}{R^2} \frac{h^2}{8m_p} \left[\frac{9}{4\pi^2} (1-x)N \right]^{2/3} + \frac{3}{4} \frac{xN}{R} \frac{hc}{2} \left(\frac{9}{4\pi^2} xN \right)^{1/3} - \frac{3}{5} G \frac{N^2 m_p^2}{R}. \end{aligned} \quad (20.13)$$

Minimum, zależności $E(R,x)$ można wyznaczyć graficznie lub rozwiązać układ równań

$$\frac{\partial E}{\partial R} = 0, \quad \text{oraz} \quad \frac{\partial E}{\partial x} = 0.$$

Ponieważ x jest bardzo małe, więc dobre przybliżenie otrzymamy przyjmując najpierw $x = 0$ w równaniu $\partial E/\partial R = 0$ i wyznaczyć R , a potem tę wartość podstawiamy do równania $\partial E/\partial x = 0$ i otrzymać x . Podstawiając $x = 0$ do równania $\partial E/\partial R = 0$ mamy

$$R = \frac{h^2}{4m_p} \left(\frac{9}{4\pi^2} N \right)^{2/3} \frac{1}{GNm_p^2}. \quad (20.14)$$

Wyrażenie to jest podobne do równania (20.8), ale liczbowo ma mniejszą wartość ze względu na masę nukleonu w mianowniku zamiast masy elektronu. Z tego powodu promień gwiazdy neutronowej będzie około 1000 razy mniejszy niż promień białego karła. Dla gwiazdy o masie Słońca $n = 1,2 \times 10^{57}$ i z równania (20.14) otrzymujemy $R = 12,6$ km, a gęstość wynosi $2,4 \times 10^{14}$ g/cm³. Jeżeli gwiazda o masie Słońca zapadłaby się do kuli o promieniu około 25 km, to jej gęstość byłaby podobna jak gęstość materii wewnątrzjądrowej. Zawartość protonów w tym gigantycznym jądrze atomowym wynosiłaby $x = 0,005$ (otrzymujemy to wstawiając $R = 12,6$ km do $\partial E/\partial x = 0$). Wynika z tego, że około 99,5% nukleonów tego gigantycznego jądra atomowego stanowią neutrony. Materia jądrowa w stanie normalnym nie zawiera elektronów, jednak w gwiazdzie neutronowej jest ich 0,5%.

W równaniu (20.13) zaniedbaliśmy wpływ sił jądrowych. Siły te w rzeczywistości występują, ponieważ gęstość wewnątrz gwiazdy jest porównywalna z gęstością wewnątrzjądrową. Okazuje się jednak dla gwiazdy podobnej do Słońca, że składowe przyciągające i odpychające tych sił w przybliżeniu równoważą się i dlatego nasze przybliżenia są zbieżne z obliczeniami dokładnymi.

Przemiana wypalonej gwiazdy w gwiazdę neutronową jest wydarzeniem gwałtownym, zwanym *eksplozją supernowej*. Eksplozja supernowej w 1054 r. mogła być widoczna w ciągu dnia. Z prawa zachowania momentu pędu wynika, że nagłe zapadnięcie się gwiazdy powoduje wzrost częstości wirowania do dziesiątek obrotów na minutę. Dla przykładu, gwiazda z centrum Mgławicy Kraba jest gwiazdą neutronową wirującą 30 razy na sekundę. Z tego powodu wielkie chmury gazu, które niosą dużą część momentu pędu gwiazdy, są odrzucane. Jeżeli na powierzchni gwiazdy występują zakłócenia pokrewne plamom na Słońcu, to emitowane promieniowanie tworzy przestrzeń wokół gwiazdy działającą jak wielki reflektor.

Na skutek szybkiego obrotu, wewnątrz gwiazdy widziane w radioteleskopach jak i teleskopach optycznych "mruga". Tak mrugająca gwiazda neutronowa jest nazywana *pulsarem*. Kluczowym testem dla hipotezy gwiazd neutronowych jest sprawdzenie czy obserwuje się zmniejszenie energii ruchu obrotowego równe wypromieniowanej energii.

Podsumowując, trzema możliwymi wynikami ewolucji gwiazd mogą być: biały karzeł, gwiazda neutronowa lub czarna dziura. Aby powstała czarna dziura, gwiazda musi osiągnąć masę krytyczną. Obliczenia tej masy są skomplikowane i jak dotychczas niejednoznaczne. Masy te szacuje się na od 0,7 do 4 mas Słońca.

W naszej galaktyce umiera rocznie od jednej do kilku gwiazd (przybliżoną ocenę otrzymamy dzieląc całkowitą liczbę gwiazd przez wiek galaktyki, co zgadza się z dokładnością do czynnika

5 z powyższą oceną). Białych karłów jest bardzo wiele, bo około 10% gwiazd w galaktyce. Eksplozja supernowej jest zjawiskiem niezwykłym i rzadkim (około jedno na sto ginących gwiazd). W ostatnich latach pojawiło się wiele dowodów eksperymentalnych istnienia gwiazd neutronowych.

Grawitacyjne zapadanie się gwiazdy neutronowej lub czarnej dziury powinno nastąpić gwałtownie. Promień gwiazdy powinien maleć w końcowej fazie z prędkością porównywalną z prędkością światła. Takie przyspieszenie masy musi wywołać emisję fali grawitacyjnej (wynika to z ogólnej teorii względności) analogicznie do fali elektromagnetycznej wysyłanej przez przyspieszony ładunek. Oczekuje się, że precyzyjne detektory fal grawitacyjnych powinny eksperymentalnie potwierdzić eksplozje supernowej czy tworzenie czarnej dziury.

20.5. Kosmologia

Kosmologia jest dziedziną nauki o budowie i prawach rządzących Wszechświatem.

20.5.1. Prawo Hubble'a

W latach dwudziestych ubiegłego wieku E. Hubble dokonując spektralnych pomiarów promieniowania odległych galaktyk zauważył, że linie w ich widmach promieniowania są przesunięte w kierunku czerwieni. Zgodnie ze zjawiskiem Dopplera (patrz pkt. 8.15), długość fali mierzonej λ jest związana z długością fali λ_0 wysyłanej ze źródła, zależnością

$$\lambda = \lambda_0 \left(1 \pm \frac{v}{c} \right), \quad (20.15)$$

gdzie v jest prędkością radialną źródła promieniowania. Znak plus odnosi się do przypadku, gdy zachodzi zjawisko przesunięcia ku czerwieni (wtedy źródło oddala się od obserwatora). Znak minus gdy źródło zbliża się do obserwatora.

Analiza danych eksperymentalnych wykazała, że stosunek prędkości radialnej $v = dr/dt$ do promienia r jest wielkością niezależną od r . Na tej podstawie w 1929 r. E. Hubble stwierdził, że *prędkości radialne oddalających się galaktyk są proporcjonalne do ich odległości r*

$$v = Hr. \quad (20.16)$$

Zależność tą nazywamy prawem Hubble'a, gdzie H jest stałą Hubble'a.

Wszechświat wg modelu Hubble'a zachowuje się jak równomiernie nadmuchiwany kulisty balonik (patrz rys. 20.5), którego promień rośnie liniowo w czasie z prędkością v

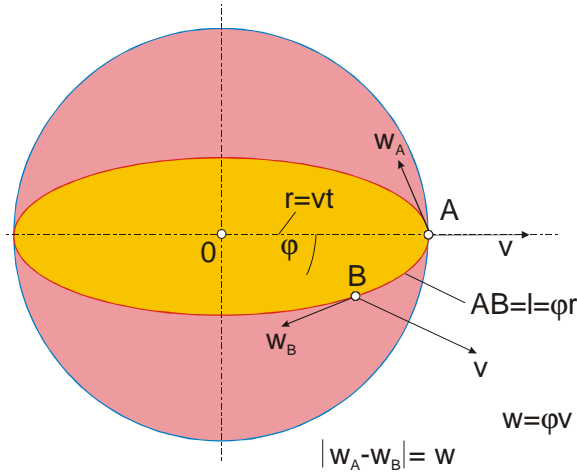
$$r = vt. \quad (20.17)$$

Czas t , po upływie którego Wszechświat osiągnął rozmiary kuli o promieniu r obliczymy z zależności

$$t = \frac{r}{v}. \quad (20.18)$$

w której wstawiając za v prawo Hubble'a, otrzymamy wiek Wszechświata

$$t = H^{-1}. \quad (20.19)$$



Z takiego modelu Wszechświata wynika również, że punkty poruszające się z (20.17) mają również odpowiednie prędkości styczne, co oznacza, że punkty leżące na danej powierzchni kuli oddalają się wzajemnie. Jeżeli dwa punkty na powierzchni kuli o promieniu r są w odległości l od siebie, to oddalają się od siebie z prędkością

$$w = \frac{l}{t} = Hl, \quad (20.20)$$

Rys. 20.5. Model Wszechświata jako nadmuchiwany balon o promieniu wzrastającym liniowo w czasie. Wówczas również prędkość oddalających się punktów A i B na powierzchni balonu rośnie ze wzrostem ich wzajemnej odległości.

a więc z prędkością proporcjonalną do ich wzajemnej odległości l . Uwzględniając, że $l = \varphi r$, gdzie φ jest miarą łukową kąta, pod którym obserwator znajdujący się w początku układu O rejestruje punkty A i B, otrzymamy związek

$$w = \varphi v. \quad (20.21)$$

Powyższy model, pomimo swej prostoty, ułatwia zrozumienie zachowania się materii we Wszechświecie oraz budowę Wszechświata i jego ewolucję. Główna trudność polega jednak na wyznaczeniu stałej Hubble'a. Szacuje się obecnie, że $H \approx 2 \times 10^{-18} \text{ s}^{-1}$. Wiek Wszechświata (20.19) zapisuje się w zmodyfikowanej formie

$$t = KH^{-1},$$

gdzie wartość K przyjmuje się zwykle 1/2, 2/3 lub 1 w zależności od przyjętego modelu zachowania się materii w czasie ewolucji Wszechświata.

Z modelu Hubble'a wynikają istotne wnioski:

- Wszechświat powstał w wyniku eksplozji, którą nazywamy Wielkim Wybuchem,
- Wielki Wybuch nastąpił około $T = KH^{-1} \approx 5 \times 10^{17} \text{ s} \approx 10^{10}$ lat temu. Czas dzielący nas od tego momentu nazywamy wiekiem Wszechświata, który szacujemy w przedziałach od 10^{10} do 2×10^{10} lat.
- Możemy obliczyć odległość r badanej galaktyki jaką przebyła od początku jej powstania

$$r = \frac{v}{H} = \frac{z}{H} c.$$

Obecnie obserwowane obiekty pozagalaktyczne mają wartość $z = 4$ co oznacza, że te obiekty poruszają się z prędkością $v \approx 0.9c$.

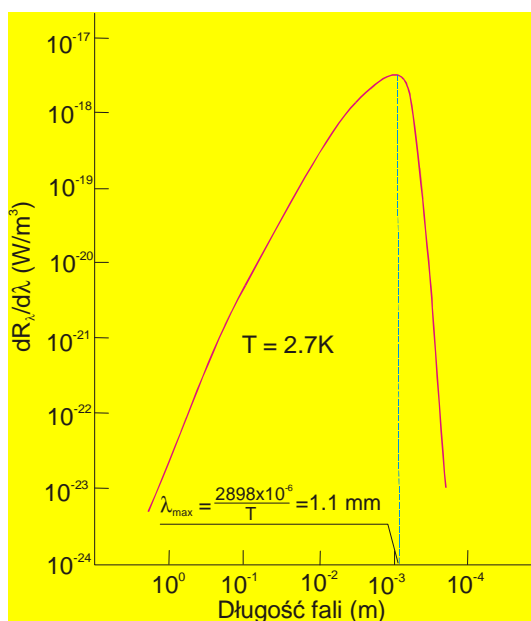
- Możemy oszacować promień Wszechświata

$$R_0 = Tc = 1,5 \times 10^{26} \text{ m.}$$

Prawo Hubble'a stało się fundamentalnym odkryciem astronomicznym i obecnie stanowi podstawę współczesnej kosmologii.

20.5.2. Promieniowanie ciepłe Wszechświata

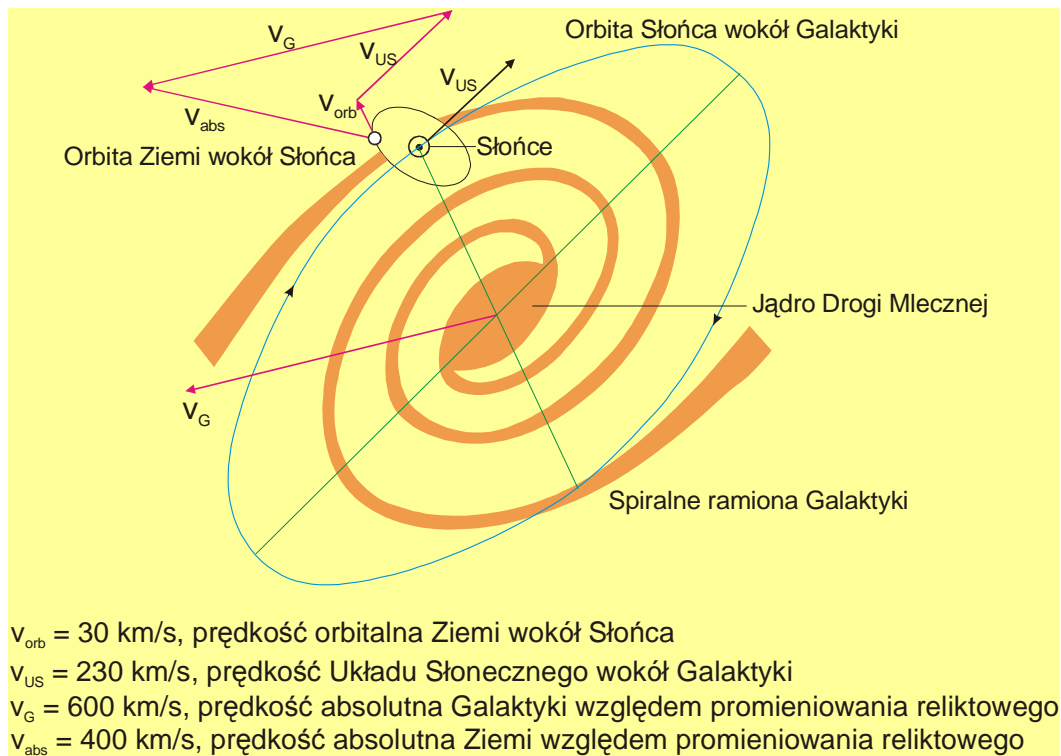
Promieniowanie ciepłe Wszechświata odkryli przypadkowo w 1965 r. A. Penzias i R. Wilson za co otrzymali Nagrodę Nobla w 1978 r. Istnienie tego promieniowania potwierdza teorię Wielkiego Wybuchu i jest pozostałością po wczesnych etapach ewolucji Wszechświata. Z tego powodu nosi ono również nazwę promieniowania relikтового (szczątkowego); jest też nazywane tłem promieniowania mikrofalowego Wszechświata.



Rys. 20.6. Reliktowe widmo promieniowania ciepłego Wszechświata odpowiadające promieniowaniu ciała doskonale czarnego o temperaturze $T = 2.7 \text{ K}$.

Rozkład widmowy promieniowania relikтового pokrywa się z widmem promieniowania ciała doskonale czarnego o temperaturze 2.7 K , co pokazano na rys. 20.6. Gęstość fotonów wynosi około $5 \times 10^8 \text{ m}^{-3}$, gęstość energii – $4 \times 10^{-14} \text{ J/m}^3$, a maksimum natężenia promieniowania, zgodnie z prawem Wiena, przypada na długość fali $1,1 \text{ mm}$. Cechą charakterystyczną promieniowania relikтового jest jego idealna izotropowość względem dowolnego kierunku w przestrzeni. Tę jego cechę wykorzystano do pomiaru bezwzględnych prędkości Ziemi, Układu Słonecznego i Galaktyki względem układu w którym promieniowanie reliktowe ma charakter izotropowy.

Na rys. 20.7 pokazano schematycznie orbitę Układu Słonecznego w Galaktyce i wektory prędkości ruchów orbitalnych oraz prędkości wypadkowej względem promieniowania relikтового. Prędkość bezwzględna Układu Słonecznego wynosi obecnie około 400 km/s , a prędkość Galaktyki około 600 km/s . Układ Słoneczny bierze udział w ruchu obrotowym Galaktyki poruszając się po orbicie o promieniu około $2,6 \times 10^{20} \text{ m}$ z prędkością około 230 km/s . Okres obiegu Układu Słonecznego wokół jądra Galaktyki, czyli rok galaktyczny Słońca, wynosi około 240 milionów lat.



Rys. 20.7. Układ Słoneczny w Galaktyce.

20.5.3. Model Wszechświata

Rozważymy budowę najprostszego modelu Wszechświata przyjmując, że materia we Wszechświecie jest rozmieszczona równomiernie o gęstości ρ_0 . Założenie to jest uzasadnione tym, że obserwacje astronomiczne wskazują dostępny Wszechświat jest jednakowo wypełniony materią. Masę zawartą w kuli o promieniu r określimy zależnością

$$M = \rho_0 V = \frac{4}{3} \pi \rho_0 r^3. \quad (20.22)$$

Jeżeli przyjmiemy, że z powierzchni kuli oddala się ciało o masie m z prędkością v , to zgodnie z zasadą zachowania energii mechanicznej możemy przyjąć, że energia całkowita jest sumą energii kinetycznej $mv^2/2$ i energii potencjalnej $mgh = GmM/r$. Odnosząc rozważania do jednostki masy m , otrzymujemy

$$\frac{E}{m} = \frac{v^2}{2} - \frac{GM}{r}. \quad (20.23)$$

Jeżeli teraz do równania w miejsce v wstawimy prawo Hubble'a, a w miejsce M zależność (20.22), to po prostych przekształceniach otrzymujemy

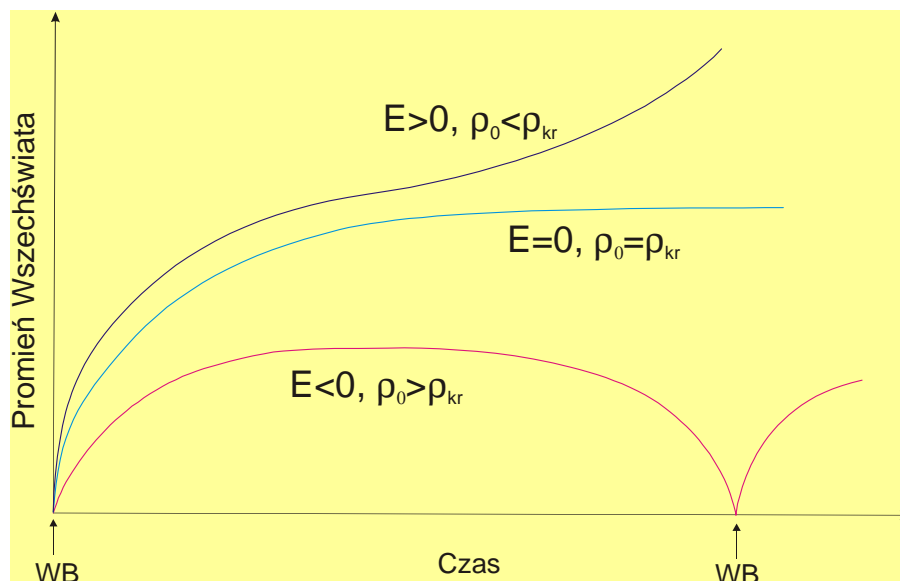
$$\frac{E}{m} = \frac{4\pi G}{3} \left(\frac{3}{8\pi G} H^2 - \rho_0 \right) r^3. \quad (20.24)$$

Pierwszy człon w nawiasie

$$\rho_{kr} = \frac{3}{8\pi G} H^2 \quad (20.25)$$

nazywać będziemy gęstością krytyczną. Jeżeli średnia gęstość masy we Wszechświecie ρ_0 jest mniejsza od masy krytycznej, to $E > 0$ i Wszechświat będzie rozszerzał się nieustannie. Jeżeli zaś znak nierówności będzie przeciwny; to znaczy gdy $\rho_{kr} < \rho_0$, to proces rozszerzania zatrzyma się i Wszechświat zacznie kurczyć się. Podstawiając w zależności wartości stałych ($H = 2 \times 10^{-18} \text{ s}^{-1}$ i $G = 6.672 \times 10^{-11} \text{ Nm}^2/\text{kg}^2$, otrzymamy $\rho_{kr} \approx 10^{-32} \text{ kg/m}^3$.

Godnym podkreślenia jest fakt, że proces rozszerzenia lub kurczenia Wszechświata nie zależy od jego rozmiaru r . Obecnie wiemy, że Wszechświat jest w stanie ekspansji, natomiast jeszcze nie wiemy, czy $\rho_0 < \rho_{kr}$, czy też $\rho_0 > \rho_{kr}$. Wyznaczenie ρ_0 pozwoli przewidzieć, jakie będą losy Wszechświata, czy będzie on rozszerzał się wiecznie czy też zacznie po pewnym czasie kurczyć się. Nie można też wykluczyć, że Wszechświat podlega okresowym procesom Wielkiego Wybuchu powtarzających się np. co kilkadziesiąt miliardów lat świetlnych. Na rys. 20.8 schematycznie przedstawiono procesy możliwych ewolucji Wszechświata według wyżej rozpatrywanego modelu.



Rys. 20.8. Proces ewolucji Wszechświata według modelu zakładającego, że gęstość materii we Wszechświecie jest rozłożona równomiernie. E oznacza energię całkowitą jednostki masy, ρ_0 gęstość materii we Wszechświecie w chwili obecnej, ρ_{kr} – gęstość krytyczna, a WB – moment Wielkiego Wybuchu.

Ewolucja Wszechświata stanowi wdzięczny temat wielu hipotez przedstawianych również w sposób popularnonaukowy w polskiej wersji czasopisma *Scientific American* zatytułowanej *Świat Nauki* (wydawnictwo Prószyński i S-ka). Przytoczmy przykładowo opis początków Wszechświata wg pracy P. Jamsa, E. Peeblesa, D.N. Schramma, E.L. Turnera i R.G. Krona [*Świat Nauki*, Nr 12 (1994), s. 27–32]:

W pewnej chwili, około 15 miliardów lat temu, cała materia i energia, którą możemy dziś obserwować, skupiona w obszarze wielkości dziesięciocentówki zaczęła się błyskawicznie rozszerzać i stygnąć. Gdy temperatura spadła do temperatury 100 milionów razy większej niż panuje we wnętrzu Słońca, siły przyrody nabrały obecnych cech. W tym okresie elementarne cząstki zwane kwarkami poruszały się swobodnie w morzu energii. Gdy Wszechświat rozszerzył się i ostygł jeszcze tysiąckrotnie, cała materia zajmowała obszar wielkości Układu Słonecznego.

W tym momencie kwarki zostały uwięzione we wnętrzu protonów i neutronów. Gdy Wszechświat powiększył się znowu tysiąc razy, protony i neutrony połączyły się ze sobą, tworząc jądra atomowe, m.in. jądra helu i deuteru. Wszystko to wydarzyło się w ciągu pierwszej minuty po Wielkim Wybuchu. Wciąż jeszcze było zbyt gorąco, aby jądra mogły połączyć się z elektronami. Neutralne atomy pojawiły się dopiero wtedy, gdy Wszechświat miał 300 tysięcy lat i był tylko 1000 razy mniejszy niż obecnie. Neutralne atomy zaczęły skupiać się, tworząc chmury gazu, z których później powstały gwiazdy. Gdy Wszechświat osiągnął jedną piątą obecnej wielkości, gwiazdy uformowały już grupy, w których można było rozpoznać młode galaktyki.

Gdy Wszechświat był tylko dwa razy mniejszy niż jest dzisiaj, w reakcjach syntezy jądrowej w gwiazdach powstała już większość ciężkich pierwiastków, z których zbudowane są planety. Nasz Układ Słoneczny jest względnie młody; ukształtował się 5 miliardów lat temu, gdy Wszechświat osiągnął już dwie trzecie obecnej wielkości. Ciągły proces formowania gwiazd doprowadził do wyczerpania zapasów gazu w galaktykach i populacja gwiazd zaczęła zanikać. Za 15 miliardów lat gwiazdy w rodzaju Słońca będą dość rzadkie i dla obserwatorów takich jak my Wszechświat stanie się znacznie mniej gościnnym miejscem.

Model Wielkiego Wybuchu ma pewne ograniczenia i nie wyjaśnia wielu fundamentalnych tajemnic. Nie wiemy jak wyglądał Wszechświat zanim rozpoczęła się ekspansja. Co stanie się w przyszłości, gdy ostatnia gwiazda wyczerpie zapas paliwa jądrowego. Nikt nie zna odpowiedzi na te pytania.