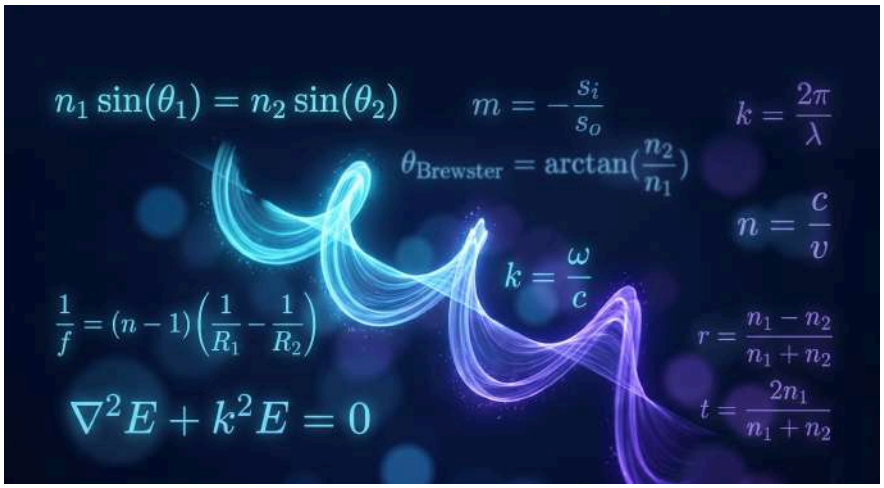
OPTYKA

Szymon Pustelny



Podręcznik akademicki (Wersja 1.1)

Kraków, czerwca 2026

Dla osób dla mnie ansolutnie najważniejszych
Dla moich NAJBLIŻSZYCH

Spis treści

Lista oznaczeń vii

Od autora xii

Wstęp xiii

1 Równanie falowe i fale elektromagnetyczne 1
- 1.1 Równania Maxwella . . . 1
- 1.2 Równanie fali elektromagnetycznej . . . 3
- 1.3 Fale monochromatyczne . . . 5
- 1.4 Zapis zespolony fali elektromagnetycznej . . . 6
- 1.5 Parametry opisujące fale elektromagnetyczne . . . 7
  - 1.5.1 Podstawowe wielkości opisujące fale elektromagnetyczne . 7
  - 1.5.2 Gęstość energii i wektor Poyntinga . . . 10
  - 1.5.3 Ciśnienie światła i pęd fali elektromagnetycznej . . . 12
  - 1.5.4 Moment pędu . . . 14
  - 1.5.5 Oddziaływanie światła z materią a zasady zachowania . . 15
- 1.6 Podsumowanie rozdziału . . . 16

2 Optyka geometryczna 19
- 2.1 Równanie Helmholtza . . . 19
- 2.2 Równanie eikonału i przybliżenie optyki geometrycznej . . . 20
- 2.3 Zasada Fermata . . . 22
- 2.4 Prawa optyki geometrycznej . . . 23
  - 2.4.1 Prawo odbicia . . . 23
  - 2.4.2 Prawo załamania . . . 24
  - 2.4.3 Zjawisko całkowitego wewnętrznego odbicia . . . 24
  - 2.4.4 Prawo odbicia i załamania a zasada Fermata . . . 24
- 2.5 Przybliżenie przyosiowe i macierze ABCD . . . 25
  - 2.5.1 Konwencja znaków (Konwencja kartezjańska) . . . 25
  - 2.5.2 Macierz ABCD dla pustej przestrzeni . . . 25
  - 2.5.3 Macierz ABCD dla odbicia . . . 27
  - 2.5.4 Macierz ABCD dla załamania . . . 29
  - 2.5.5 Propagacja promienia w przestrzeni a macierze ABCD . . 30

2.5.6 Właściwości macierzy ABCD . . . 30
2.6 Odbicie od zwierciadła . . . 31
2.6.1 Równanie soczewkowe dla zwierciadła . . . 31
2.6.2 Powiększenie układu optycznego . . . 33
2.7 Soczewki . . . 33
2.7.1 Propagacja światła w soczewce . . . 33
2.7.2 Równanie soczewkowe dla soczewki . . . 34
2.7.3 Soczewka cienka i wzór szlifierzy . . . 34
2.7.4 Płaszczyzny główne . . . 34
2.7.5 Równanie soczewkowe Newtona . . . 37
2.7.6 Soczewki gradientowe . . . 38
2.8 Diagramy promieni . . . 38
2.9 Rodzaje obrazów optycznych . . . 39
2.10 Inne elementy optyczne . . . 41
2.10.1 Pryzmaty . . . 41
2.10.2 Pryzmaty rozpraszające . . . 41
2.11 Pryzmaty odbijające . . . 42
2.12 Światłowody . . . 42
2.13 Układy optyczne . . . 46
2.13.1 Ogniskowa układu optycznego . . . 46
2.13.2 Ograniczenie wiązek: Przysłony i źrenice . . . 47
2.13.3 Oko . . . 47
2.13.4 Mikroskop . . . 48
2.13.5 Teleskop . . . 49
2.14 Wady soczewek . . . 51
2.14.1 Aberracje . . . 51
2.14.2 Koma . . . 52
2.14.3 Astygmatyzm . . . 52
2.15 Podsumowanie rozdziału . . . 53

3 Fale płaskie, sferyczne i gaussowskie **57**
3.1 Wprowadzenie . . . 57
3.2 Fale płaskie . . . 57
3.3 Fale sferyczne . . . 59
3.4 Fale cylindryczne . . . 61
3.5 Wiązka gaussowska . . . 64
3.5.1 Fale przyosiowe i przybliżenie wolnej obwiedni . . . 64
3.6 Macierze ABCD dla wiązki gaussowskiej . . . 70
3.7 Wiązki Hermita-Gaussa . . . 72
3.8 Wiązki Laguerre'a-Gaussa . . . 73
3.9 Podsumowanie rozdziału . . . 76

4 Wzory Fresnela **79**
4.1 Zachowanie pola elektrycznego i magnetycznego na granicy ośrodków . . . 79
4.2 Transformacja pola elektrycznego światła na granicy dwóch ośrodków . . . 81
4.3 Stałość częstotliwości fali . . . 82
4.4 Przestrzenna orientacja promienia padającego, odbitego i załamanego . . . 82
4.5 Prawo odbicia i załamania . . . 83
4.6 Wzory Fresnela . . . 84
4.6.1 Wektor **E** prostopadły do płaszczyzny padania (polaryzacja *s*) . . . 84
4.6.2 Wektor **E** równoległy do płaszczyzny padania (polaryzacja *p*) . . . 86
4.6.3 Zależność współczynników Fresnela od kąta padania . . . 87
4.6.4 Kąt Brewstera . . . 89
4.6.5 Całkowite wewnętrzne odbicie i fala zanikająca . . . 90
4.6.6 Relacje Stokesa . . . 93
4.6.7 Reflektancja i transmitancja . . . 94
4.7 Odbicie na granicy dielektryk-przewodnik . . . 97
4.7.1 Równanie falowe w przewodnikach . . . 97
4.7.2 Efekt naskórkowy i głębokość wnikania . . . 100
4.7.3 Wzory Fresnela w przewodnikach . . . 102
4.8 Podsumowanie rozdziału . . . 103

5 Interferencja **106**
5.1 Superpozycja fal . . . 106
5.2 Przykłady nakładania się fal . . . 107
5.2.1 Dwie fale o jednakowych częstotliwościach i wektorach falowych . . . 107
5.2.2 Dwie fale o identycznych częstotliwościach rozchodzące się w przeciwnych kierunkach . . . 107
5.2.3 Dwie fale o jednakowych częstotliwościach rozchodzące się pod kątem . . . 109
5.2.4 Dwie fale o nieco różnych częstotliwościach i liczbach falowych, ale tym samym kierunku propagacji . . . 110
5.2.5 Dwie fale o jednakowych częstotliwościach, tym samym kierunku propagacji i ortogonalnych polaryzacjach liniowych 111
5.2.6 Dwie fale o jednakowych częstotliwościach, identycznych polaryzacjach kołowych i przeciwnych wektorach falowych 112
5.2.7 Dwie fale o jednakowych częstotliwościach, ortogonalnych polaryzacjach kołowych i przeciwnych wektorach falowych 113
5.3 Prędkość fali . . . 113
5.3.1 Prędkość fazowa . . . 113
5.3.2 Prędkość grupowa . . . 113
5.3.3 Prędkość fazowa w przewodnikach . . . 115

5.4 Interferencja konstruktywna i destruktywna . . . 115
5.5 Interferencja przez podział amplitudy . . . 118
5.5.1 Interferometr Michelsona . . . 118
5.5.2 Interferometr Michelsona a źródła rozciągłe . . . 120
5.5.3 Interferometr Macha-Zehndera . . . 121
5.5.4 Interferometr Jamina . . . 123
5.6 Interferencja przez podział frontu falowego . . . 123
5.6.1 Eksperyment Younga . . . 123
5.6.2 Lustro Lloyda . . . 124
5.6.3 Bipryzmat Fresnela . . . 125
5.7 Długość i czas spójności . . . 125
5.8 Interferencja światła białego . . . 127
5.9 Podsumowanie rozdziału . . . 128

6 Interferencja wielowiązkowa **133**
6.1 Wprowadzenie . . . 133
6.2 Interferometr Fabry'ego-Perota . . . 133
6.2.1 Natężenie światła transmitowanego przez interferometr Fabry'ego-Perota . . . 134
6.2.2 Natężenie światła odbitego od interferometru Fabry'ego-Perota . . . 138
6.2.3 Współczynnik finezji . . . 140
6.2.4 Finezja rezonatora ($\mathcal{F}$) . . . 140
6.3 Interferometr Fabry'ego-Perota jako analizator widma . . . 141
6.3.1 Zdolność rozdzielcza . . . 141
6.3.2 Przedział spektralny . . . 143
6.3.3 Dobroć wnęki i czas życia fotonu . . . 144
6.4 Interferencja wiązki rozbieżnej . . . 145
6.5 Ograniczenia interferometru Fabry'ego-Perota . . . 146
6.6 Konfokalny interferometr Fabry'ego-Perota . . . 147
6.7 Zastosowania interferencji wielowiązkowej . . . 147
6.7.1 Analiza widmowa . . . 147
6.7.2 Filtry spektralne . . . 148
6.7.3 Lustra dielektryczne . . . 149
6.7.4 Pokrycia antyrefleksyjne . . . 151
6.7.5 Lasery i stabilizacja częstotliwości . . . 151
6.7.6 Macierze ABCD i stabilność wnęki optycznej . . . 151
6.8 Podsumowanie rozdziału . . . 152

7 Dyfrakcja **155**
7.1 Strefy Fresnela . . . 155
7.1.1 Zasada Huygensa-Fresnela . . . 155
7.1.2 Strefy Fresnela . . . 155
7.1.3 Plamka Poissona . . . 158
7.1.4 Ugięcie na krawędzi nieprzepuszczalnej półpłaszczyzny . . . 161
7.1.5 Ograniczenia metody stref Fresnela . . . 164

7.2 Teoria dyfrakcji Kirchhoffa . . . 165
7.2.1 Twierdzenie całkowe Helmholtza-Kirchhoffa . . . 165
7.2.2 Amplituda fali w punkcie za przesłoną . . . 167
7.2.3 Apertura oświetlana falą płaską . . . 169
7.2.4 Apertura oświetlana falą sferyczną . . . 171
7.2.5 Problem obliczeń dyfrakcji Kirchhoffa . . . 172
7.3 Reżimy dyfrakcyjne . . . 172
7.3.1 Przybliżenie Fresnela (Pole bliskie) . . . 172
7.3.2 Przybliżenie Fraunhofera (Pole dalekie) . . . 173
7.3.3 Ewolucja obrazu dyfrakcyjnego . . . 174
7.4 Dyfrakcja światła na różnych obiektach . . . 174
7.4.1 Ugięcie światła na szczelinie jednowymiarowej . . . 174
7.4.2 Ugięcie światła na otworze prostokątnym . . . 177
7.4.3 Ugięcie światła na otworze okrągłym . . . 178
7.5 Rozdzielczość układów optycznych . . . 180
7.6 Dyfrakcja na dwóch szczelinach . . . 182
7.7 Układ N_s szczelin czyli siatka dyfrakcyjna . . . 183
7.7.1 Rozkład natężenia światła . . . 183
7.7.2 Równanie siatki i dyspersja . . . 185
7.7.3 Padanie pod kątem i siatki profilowane . . . 187
7.7.4 Zdolność rozdzielcza siatki . . . 188
7.8 Zasada Babineta . . . 190
7.9 Dyfrakcja na fali akustycznej . . . 191
7.9.1 Ugięcie jako oddziaływanie fotonu z fononem . . . 192
7.9.2 Reżimy ugięcia: Raman-Nath oraz Bragg . . . 193
7.10 Dyfrakcja światła białego . . . 193
7.11 Optyka fourierowska . . . 194
7.12 Podsumowanie rozdziału . . . 196

8 Polaryzacja światła **200**
8.1 Formalizm opisu stanu polaryzacji . . . 200
8.1.1 Formalizm Jonesa . . . 200
8.1.2 Formalizm Stokesa . . . 203
8.1.3 Sfera Poincarégo . . . 204
8.2 Sposoby polaryzacji światła . . . 206
8.2.1 Polaryzacja przez odbicie . . . 206
8.2.2 Polaryzacja przez rozproszenie . . . 207
8.2.3 Polaryzacja przez selektywną absorpcję . . . 208
8.2.4 Polaryzacja przez dwójłomność . . . 209
8.3 Prawo Malusa . . . 213
8.4 Płytki fazowe . . . 213
8.5 Aktywność optyczna . . . 216
8.6 Wymuszone efekty optyczne . . . 218
8.6.1 Fotoelastyczność . . . 218
8.6.2 Efekty elektrooptyczne: Kerra i Pockelsa . . . 219
8.6.3 Efekt Faradaya . . . 220

8.7 Podsumowanie rozdziału 222

9 Model Lorentza **226**
9.1 Wprowadzenie 226
9.2 Emisja fali przez ładunek elektryczny 226
9.2.1 Ładunek w spoczynku lub poruszający się ruchem jednostajnym prostoliniowym 226
9.2.2 Emisja fali przez ładunek rozpędzany liniowo 227
9.2.3 Promieniowanie synchrotronowe 228
9.2.4 Dipol elektryczny 230
9.3 Model Lorentza 231
9.4 Propagacja fali w ośrodku materialnym 238
9.5 Wiele linii rezonansowych 239
9.6 Dyspersja normalna i anomalna 240
9.6.1 Nadświetlność i wolne światło 240
9.7 Zależność pomiędzy absorpcją i dyspersją 241
9.8 Podsumowanie rozdziału 243

10 Optyka nieliniowa **246**
10.1 Wprowadzenie 246
10.2 Alternatywna postać równania falowego 246
10.3 Zjawiska nieliniowe 247
10.3.1 Fizyczne źródła nieliniowej polaryzacji 247
10.3.2 Równanie Lorentza w ośrodku nieliniowym 248
10.4 Optyczne zjawiska nieliniowe 253
10.5 Podsumowanie rozdziału 255

11 Źródła Światła **258**
11.1 Wprowadzenie 258
11.2 Termiczne źródła światła i promieniowanie ciała doskonale czarnego 259
11.2.1 Katastrofa w nadfiolecie 260
11.2.2 Kwantowanie pola 260
11.2.3 Rzeczywiste termiczne źródła światła 261
11.2.4 Prawo Stefana-Boltzmanna i prawo Wiena 263
11.3 Niespójne źródła atomowe i półprzewodnikowe 264
11.3.1 Emisja spontaniczna 264
11.3.2 Lampy wyładowcze i diody elektroluminescencyjne 267
11.4 Źródła koherentne – lasery 270
11.4.1 Inwersja obsadzeń i emisja wymuszona 270
11.4.2 Rezonatory optyczne i własności wiązki 273
11.4.3 Przegląd rodzajów laserów 276
11.4.4 Lasery impulsowe 278
11.5 Źródła światła nieklasycznego 281

Indeks alfabetyczny **284**

Lista oznaczeń

α_a	kąt akceptacji
α_B	kąt Brewstera
α_g	kąt graniczny
α_o	kąta odbicia
α_p	kąta padania
α_z	kąta załamania
β	„wektor falowy” fali zanikającej
β_q	stała tłumienia w modelu Lorentza
$\Delta\lambda_{\mathrm{FSR}}$	wolny przedział spektralny w długościach fali
$\Delta\nu_{\mathrm{FSR}}$	wolny przedział spektralny w częstotliwościach
$\delta\theta_{\mathrm{min}}$	rozdzielczość kątowa
δ_w	głębokość wnikania
η	zespolony współczynnik załamania
η	zespolony współczynnika załamania
γ_e	ułamkowy współczynnik wzmocnienia
γ_L	czynnik Lorentza
γ_q	stała tłumienia drgań dipola w modelu Lorentza
$\hat{\mathbf{d}}_{12}$	operator momentu dipolowego
$\hat{\mathbf{r}}$	operator położenia
λ	długość fali
λ_0	długość fali w próżni

$\mathcal{F}$	finezja interferometru
$\mathcal{M}$	powiększenie soczewki/układu optycznego
NA	apertura numeryczna
ε	tensor przenikalności elektrycznej
J	macierz Jonesa
M_M	macierz Müllera
ω	częstość kołowa fali
ω_p	częstość plazmowa
ω_{ak}	częstość fali akustycznej
ψ	funkcja falowa
ρ	gęstość ładunków elektrycznych
$\rho_p(\omega)$	spektralna gęstość mocy promieniowania
θ_0	kąt rozbieżności wiązki gaussowskiej
θ_m	kąt pod którym obserwowane jest m-te maksimum dyfrakcyjne
θ_m	kąt ugięcia m-tego prążka dyfrakcyjnego
$\tilde{\varepsilon}$	zespolona przenikalność elektryczna
φ	faza początkowa fali
$\mathbf{S}_S$	wektor Stokesa
$\mathbf{a}$	przyspieszenie cząstki naładowanej
$\mathbf{B}$	indukcja pola magnetycznego
$\mathbf{D}$	indukcja pola elektrycznego
$\mathbf{E}$	natężenie pola elektrycznego
$\mathbf{e}_B$	zespolony wersor pola magnetycznego fali elektromagnetycznej
$\mathbf{e}_E$	zespolony wersor polaryzacji światła
$\mathbf{F}$	siła
$\mathbf{H}$	natężenie pola magnetycznego
$\mathbf{J}$	moment pędu cząstki
$\mathbf{j}$	gęstość prądu

$\mathbf{k}_{ak}$ wektor falowy fali akustycznej

$\mathbf{o}_q$ moment dipolowy

a odległość przedmiotu od zwierciadła (lub płaszczyzny głównej przedmiotowej)

b odległość obrazu od zwierciadła (lub płaszczyzny głównej obrazowej)

B_o indukcja pola składowej magnetycznej fali odbitej

B_p indukcja składowej magnetycznej fali padającej

B_z indukcja pola składowej magnetycznej fali załamanej

B_{12} współczynnik absorpcji Einsteina

B_{21} współczynnik emisji wymuszonej Einsteina

C stała elektrooptyczna / stała Brewstera

c prędkość światła w próżni

D szerokość apertury

D szerokość szczeliny na której dochodzi do dyfrakcji

d odległość między dwoma punktami

D_a średnica apertury kołowej

D_k dyspersja kątowa

d_u zdolność rozdzielcza

e ładunek cząstki

$E_{l,m}^{(+)}(\mathbf{r})$ pole elektryczne wiązki Hermita-Gaussa

$E_0^{(+)}$ zespolona amplituda pola elektrycznego fali elektromagnetycznej

E_{pw} szerokość przerwy wzbronionej

E_e energia ładunku elektrycznego

E_o natężenie składowej elektrycznej światła odbitego

E_p natężenie składowej elektrycznej światła padającego

E_z natężenie składowej elektrycznej światła załamanego

$E_{l,m}^{(+)}(\mathbf{r})$ mod p

F ognisko układu

F	współczynnik finezji
f	ogniskowa układu
h_i	odległość płaszczyzny głównej od wierzchołka układu
H_o	natężenie pola składowej magnetycznej fali odbitej
H_p	natężenie pola składowej magnetycznej fali padającej
H_z	natężenie pola składowej magnetycznej fali załamanej
I	natężenie światła
K	stała Kerra
k_0	wektor falowy w próżni
k_0	wektor falowy w próżni
k_q	stała sprężystości w modelu Lorentza
l_{koh}	długość spójności/droga koherencji
M_i	macierz ABCD transformacji promienia
n	współczynnik załamania światła
N_d	koncentracja elektrycznych momentów dipolowych
n_e	współczynnik załamania promienia nadzwyczajnego
n_g	grupowy współczynnik załamania
n_o	współczynnika załamania promienia zwyczajnego
n_p	współczynnik załamania płaszcza światłowodu
n_r	współczynnik załamania rdzenia światłowodu
Q	dobroć wnęki
R	reflektancja
$r_{\parallel}$	amplitudowy współczynnik odbicia światła o polaryzacji p
$r_{\perp}$	amplitudowy współczynnik odbicia światła o polaryzacji s
r_a	promień apertury kołowej
r_s	amplitudowy współczynnik odbicia fali o polaryzacji s
r_{ij}	stała elektrooptyczna
T	transmitancja

$t_{\parallel}$	amplitudowy współczynnik transmisji światła o polaryzacji p
$t_{\perp}$	amplitudowy współczynnik transmisji światła o polaryzacji s
V	stała Verdeta
V	widzialność prążdków w interferometrze
v	prędkość fazowa fali
v_{ψ}	prędkość fali mechanicznej
V_E	potencjał elektronu w polu elektrycznym
v_e	prędkość ładunku
V_p	stopień polaryzacji światła
v_q	prędkość ładunku elektrycznego
v_{gr}	prędkość grupowa
$w(z)$	promień wiązki gaussowskiej
w_0	promień przewężenia wiązki gaussowskiej w ognisku
z_R	długość RAyleigha
μ	przenikalność magnetyczna
ε	przenikalność elektryczna
NA	apertura numeryczna

Od autora

Niniejszy podręcznik z optyki został stworzony z myślą o studentach fizyki i kierunków pokrewnych. Powstał on w oparciu o notatki i doświadczenie zdobyte podczas wieloletnich wykładów jakie prowadziłem dla II roku studiów fizyki, astronomii oraz studiów matematyczno-przyrodniczych, choć nie wykluczam, że słuchaczami wykładu byli także studenci informatyki, fizyki dla firm czy zaawansowanych materiałów i nanotechnologii. Celem podręcznika jest dostarczenie klarownej i przystępnej wiedzy na temat podstaw optyki, które są kluczowe dla zrozumienia wielu aspektów fizyki oraz jej zastosowań praktycznych. W swoim zamierzeniu stanowić ma on jednak przede wszystkim wprowadzenie do optyki nieliniowej i kwantowej, fotoniki, a nawet fizyki atomowej, czyli fascynujących działów fizyki współczesnej.

Na przestrzeni lat szereg osób pomagało mi w stworzeniu tego podręcznika. Kilkukrotnie swoich uwag dostarczali mi studenci. Jednak moje serdeczne podziękowania kieruję do Pani Izabeli Grochowskiej. Niezmiernie doceniam jej ogromny wysiłek włożony w przeczytanie tego podręcznika, szereg cennych uwag merytorycznych oraz ogromną liczbę poprawek edytorskich. To właśnie jej ciężka praca sprawiła, że zdecydowałem się w końcu udostępnić ten podręcznik szerszej grupie odbiorców.

Chciałem również podziękować swoim Kolegom. Moje szczególne podziękowania kieruję do Prof. Krzysztofa Dzierżęgi, z którym wielokrotnie wielokrotnie dyskutowaliśmy o różnych aspektach materiału prezentowanego w tym podręczniku, nierzadko odkrywając przy tej okazji rzeczy, z których istnienia nie zdawałem sobie sprawy. Chciałbym również podziękować Prof. Wojciechowi Gawlikowi, który wiele lat temu wprowadzał mnie w tajniki optyki, który był moim mentorem i przewodnikiem. Jego nieustająca, mimo upływu lat pasja, oraz ogromna wiedza, witalność i kreatywność są dla mnie niedoścignionym wzorem.

Szymon Pustelny

Wstęp

Światło jest jedną z najbardziej intrygujących, ale także inspirujących wielkości fizycznych, z którymi mamy na co dzień do czynienia. Niesie ono ze sobą nie tylko „jasność", ale również ciepło, co sprawiło, że już od najdawniejszych czasów kojarzone było ono z dobrem, pięknem i szczęściem. Nieprzypadkowo w większości religii bóg był utożsamiany ze światłością, a często najwyższym w niebieskim panteonie był bóg światła lub słońca.

Z czasem pytanie o to czym jest światło stało się obiektem dociekań myślicieli. Pytanie o jego naturę było jednym z tych podstawowych, które zadawali sobie starożytni filozofowie. Już wtedy ścierały się ze sobą dwa poglądy dotyczące natury światła. Pierwszy, reprezentowany przez Pitagorejczyków, stanowił, że światło jest strumieniem mikroskopijnych cząsteczek, które niosą ze sobą energię. Uczniowie szkoły arystotelejskiej uważali natomiast, że światło to fala, która podobna jest w swej naturze do fal na wodzie, z tą jednak różnicą, że dotyczy zdecydowanie mniejszych obiektów. Oczywiście, zarówno jedni, jak i drudzy prowadzili swoje rozważania na gruncie czysto filozoficznym bez oparcia się o badania empiryczne. Zatem ani jedni ani drudzy nie mogli podać empirycznych argumentów, które dowodziłyby poprawności ich teorii.

Pytanie o naturę światła nabrało nowego charakteru w momencie powstania nowożytnej nauki, tj. na przełomie XVII i XVIII wieku. Wtedy to holenderski matematyk, fizyk i astronom Christiaan Huygens (1629-1695) zaproponował pogląd, że światło jest falą. Jako punkt wyjścia do tego twierdzenia przyjął szereg zjawisk obserwowanych dla fal mechanicznych (odbicie, ugięcie), które obserwowane były również dla światła. Z tym poglądem nie zgadzał się jednak współczesny Huygensowi Izaak Newton (1643-1727), który uważał, że światło jest zbiorem cząstek oddziałujących ze sobą siłami poprzez pola i to te siły odpowiedzialne są za takie zjawiska jak odbicie czy załamanie. Choć za poglądem Newtona zdawała się przemawiać mniejsza liczba argumentów, to jednak jego niepodważalny autorytet sprawił, że powszechnie uznano hipotezę o korpuskularnej naturze światła. Zahamowało to rozwój optyki na przeszło sto lat.

Mimo, że eksperymenty Gian Domenico Romagnosi (1761-1835), Hansa Christiana Oersteda (1777-1851), Andre Ampere'a (1775-1836) czy Micheala Faradaya (1791-1867) wskazywały na bliski związek pomiędzy elektrycznością i magnetyzmem, to do połowy XIX wieku badania nad obiema dziedzinami prowadzone były niezależnie. Trzeba było dopiero geniuszu Jamesa Clarka Maxwella

(1831-1879), aby wykazać, że elektryczność i magnetyzm są jedynie dwiema postaciami tej samej rzeczywistości, jaką jest **pole elektromagnetyczne**. Definicja i opis fal elektromagnetycznych, którego dokonał Maxwell, utorowały drogę rozwoju optyki. Okazało się bowiem, że prędkość z jaką rozchodzą się fale elektromagnetyczne i prędkość propagacji światła, którą zmierzył Ole Romer (1644-1710), obserwując zaćmienia księżyców Jowisza, mają tę samą wartość. Ponieważ wartość ta była nieporównywalna z prędkością jakiegokolwiek innego obiektu znanego wówczas w przyrodzie, skłoniło to uczonych do stwierdzenia, że światło jest falą elektromagnetyczną. Dziś wiemy, że rzeczywiście tak jest, ale fale z zakresu widzialnego stanowią tylko niewielką część całego zakresu spektralnego fal elektromagnetycznych.

Kolejny skok jakościowy w optyce nastąpił na przełomie XIX i XX wieku kiedy niemiecki fizyk Max Planck (1858-1947) zaproponował wytłumaczenie problemu promieniowania ciała doskonale czarnego poprzez dyskretyzację energii i sformułowanie pojęcia **kwantu promieniowania**. Kilka lat później Albert Einstein (1879-1955), wykorzystując pojęcie kwantu promieniowania, zaproponował wytłumaczenie efektu fotoelektrycznego zewnętrznego. Dziś teoria Plancka jest umowną cezurą, która wyznacza narodziny nowej teorii, jaką jest **mechanika kwantowa**.

Hipoteza Plancka rozpoczęła badania nad mechaniką kwantową. Następne kilkadziesiąt lat charakteryzowały się gwałtownym rozwojem szeregu nowych hipotez. Przykładowo, 20 lat po wytłumaczeniu efektu fotoelektrycznego, Louis de Broglie (1892-1987) zaproponował jedną z najbardziej zaskakujących hipotez dotyczących właściwości natury tj. jej **dualizm korpuskularno-falowy**. Mówi ona, że światło, co z czasem zostało uogólnione na całą materię, posiada własności, które pozwalają mu się zachować w niektórych zjawiskach jak fala, w innych zaś jak zbiór cząsteczek.

Wiek XIX i XX to triumf optyki i to zarówno w jej formie falowej, jak i kwantowej. Badania prowadzone w ramach optyki falowej doprowadziły do powstania i rozwoju m.in. takich dziedzin jak optyka nieliniowa, telekomunikacja, holografia, mikroskopia czy interferometria. Z czasem, prace w ramach tzw. optyki kwantowej otworzyły drzwi do badań w ramach fizyki atomowej, sensoryki, kwantowej teorii informacji czy kryptografii kwantowej. Wydaje się, że optyka kwantowa wkrótce znajdzie również zastosowanie praktyczne w **komputerach kwantowych**.

W dzisiejszych czasach optyka zajmuje się nie tylko falami elektromagnetycznymi z zakresu widzialnego, ale obejmuje ogół badań nad falami elektromagnetycznymi w zakresie od promieniowania X do dalekiej podczerwieni (a czasem w jeszcze szerszym zakresie). O tym, że jest ona niezmiernie istotną dziedziną nauki niech świadczy liczba Nagród Nobla „przyznanych optyce":

- 1901, Wilhelm Röntgen, Promieniowania X
- 1907, Albert Michelson, Interferometr Michelsona
- 1908, Gabriel Lippmann, Fotografia kolorowa
- 1914, Max von Laue, Dyfrakcja promieniowania X

- 1915, William Henry Bragg, William Lawrence Bragg, Analiza dyfraktogramów promieniowania X
- 1917, Charles Barkla, Właściwości promieniowania X
- 1918, Max Planck, Rozwiązanie problemu promieniowania ciała doskonale czarnego i wprowadzenia pojęcia kwantu promieniowania
- 1919, Johannes Stark, Efekt Dopplera dla światła
- 1921, Albert Einstein, Wyjaśnienie efektu fotoelektrycznego
- 1925, Karl Siegbahn, Spektroskopia promieniowania X
- 1927, Arthur Compton, Efekt Comptona
- 1929, Louis de Broglie, Falowa natura elektronów
- 1930, Chandrasekhara Raman, Efekt Ramana
- 1936, Peter Debey, Dyfrakcja promieniowania X
- 1937, Clinton Davisson, George Thomson, Dyfrakcja elektronów na kryształach
- 1951, Cecil Powell, Rozwój technik fotograficznych do badaniach mezonów
- 1952, Archer Martin, Richard Synge, Chromatografia
- 1953, Frits Zernike, Mikroskopia z kontrastem fazowym
- 1955, Willis Lamb, Odkrycie rozszczepienia Lamba
- 1958, Pavel Cherenkov, Ilja Franck, Igor Tamm, Promieniowanie Cherenkova
- 1959, Jaroslav Heyrovsky, Polarymetria
- 1964, Charles Townes, Nikolay Basov, Aleksandr Prokhorov, Maser
- 1964, Dorothy Hodgkin, Spektrometria promieniowania X w badaniach biochemicznych
- 1966, Alfred Kastler, Pompowanie optyczne
- 1971, Denis Gabor, Holografia
- 1981, Nicolaas Bloembergen, Arthur Schawlow, Spektroskopia laserowa
- 1986, Ernest Ruska, Gerd Binnig, Heinrich Rohrer, Mikroskopia elektronowa
- 1989, Norman Ramsey, Hans Dehmelt, Wolfgang Paul, Pułapkowanie jonów i spektroskopia ramseyowska

- 1998, Steven Chu, Claude Cohen-Tannoudji, William Phillips, Chłodzenie i pułapkowanie atomowe
- 2000, Zhores Alferov, Herbert Krömer, Rozwój struktur półprzewodnikowych dla optoelektroniki i telekomunikacji
- 2001, Eric Cornell, Wolfgang Ketterle, Karl Weinman, Kondensacja Bosego-Einsteina
- 2005, Roy Glauber, John Hall, Theodor Hänsch, Grzebień optyczny
- 2009, Charles Kao, William Boyle, Goerge Smith, Telekomunikacja światłowodowa i kamery CCD
- 2012, Serge Haroche, David Wineland, Techniki manipulacji układami kwantowymi
- 2014, Isama Akasaki, Hiroshi Amano, Shuji Nakamura, Niebieskie diody elektroluminescencyjne
- 2014, Eric Betzig, Stefan Hell, William Möerner, Mikroskopia superzdolności rozdzielczej
- 2017, Rainer Weiss, Barry Barish, Kip Thorne, Fale grawitacyjne
- 2018, Arthur Ashkin, Gerhard Mourou, Donna Strickland, Spektroskopia laserowa z ultrakrótkimi impulsami światła
- 2022, Alain Aspect, John Clauser, Anton Zeilinger, Za eksperymenty ze splątanymi fotonami i pionierską informatykę kwantową
- 2023, Pierre Agostini, Ferenc Krausz i Anne L'Huillier, Za eksperymentalne metody generowania attosekundowych impulsów światła dla celów badania dynamiki elektronów w materii

Jak widać szczególnie na przestrzeni ostatnich dwóch dekad pojawiło się wiele „optycznych" Nagród Nobla. Co ciekawe, są to zarówno nagrody za badania podstawowe, jak w przypadku chłodzenia i pułapkowania atomowego, ale także takie, które są dużo bliższe zastosowaniom (telekomunikacja światłowodowa czy niebieskie diody elektroluminescencyjne), co zresztą lepiej oddaje ideę, która przyświecała Alfredowi Noblowi fundującemu tę nagrodę.

Możliwości, których dostarcza obecnie światło są olbrzymie. Rozwój w telekomunikacji, który dokonał się na przestrzeni ostatnich dwudziestu lat, był możliwy tylko dzięki opracowaniu technik wytwarzania i detekcji światła, ale również technik jego transmisji. Wykorzystanie światła, którego częstotliwość jest rzędu 100 THz, pozwoliło na zakodowanie sygnału poprzez jego modulację z częstościami dochodzącymi do setek GHz. Propagacja światła w światłowodach nie jest silnie związana z procesem rozpraszania energii (nie ma np. wydzielania ciepła jak w przypadku przepływu prądu), a jedyne straty pochodzą z rozproszeń czy absorpcji, które mogą być we włóknach optycznych ograniczone

do poziomu 0,3 dB/km. Dzięki temu możliwa jest transmisja informacji na odcinkach nawet kilkudziesięciu kilometrów bez konieczności jej wzmacniania. W końcu w jednym światłowodzie mogą zostać przesłane setki długości fali, które w pierwszym przybliżeniu są od siebie zupełnie niezależne.[1] Innym przykładem, w którym rozwój optyki dostarczył nowych możliwości jest mikroskopia. Zrozumienie źródeł ograniczeń zdolności rozdzielczej pozwoliło na opracowanie technik mikroskopowych umożliwiających obrazowanie obiektów wielokrotnie mniejszych od długości fali światła wykorzystywanych do obserwacji obiektu. Dzięki temu udało się zobaczyć obiekty o rozmiarach pojedynczego wirusa, co otworzyło nowe pola badań w naukach biologicznych. Już te dwa przykłady, ale też szereg innych nieprzywołanych tutaj pokazują, że optyka jest w chwili obecnej jedną z najdynamiczniej rozwijających się gałęzi fizyki.

Niniejszy podręcznik prezentuje podstawy optyki, które doprowadziły do rozwoju tej dziedziny oraz szeregu ciekawych i dziś powszechnie wykorzystywanych zastosowań. Podstawy te często zostały opracowane jeszcze w XIX wieku lub nawet wcześniej. Tym niemniej za każdym razem, gdy omawiane będą nowe zagadnienia pojawiać się będą odnośniki do współczesnych badań, gdzie są one wykorzystane. Ma to ukazać, że nawet proste idee mogą prowadzić do dalece nietrywialnych obserwacji i wynalazków.

[1] Badaniem procesów, które są źródłem ograniczeń tego twierdzenia zajmuje się optyka nieliniowa.

Rozdział 1

Równanie falowe i fale elektromagnetyczne

1.1 Równania Maxwella

Jednym z najważniejszych osiągnięć fizyki XIX wieku było sformułowanie **równań Maxwella**. W rzeczywistości równania te stanowią kompilację wcześniej znanych praw, ale połączenie ich w spójną całość jednocześnie opisującą pole elektryczne i magnetyczne pozwoliło na wykazanie istnienia **fal elektromagnetycznych** oraz określenie sposobu ich propagacji.

Równania Maxwella w swojej różniczkowej formie można zapisać jako:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}\,, \tag{1.1a}$$

$$\nabla \cdot \mathbf{D} = \rho\,, \tag{1.1b}$$

$$\nabla \cdot \mathbf{B} = 0\,, \tag{1.1c}$$

$$\nabla \times \mathbf{H} = \mathbf{j} + \frac{\partial \mathbf{D}}{\partial t}\,, \tag{1.1d}$$

gdzie ρ jest gęstością ładunków, $\mathbf{j}$ jest gęstością prądu, $\mathbf{E}$ i $\mathbf{D}$ są odpowiednio **natężeniem** i **indukcją pola elektrycznego**, zaś $\mathbf{H}$ oraz $\mathbf{B}$ są **natężeniem** i **indukcją pola magnetycznego**. O ile natężenie jest wielkością, która charakteryzuje pole niezależnie od ośrodka, w którym to pole istnieje, o tyle indukcja jest wielkością fizyczną, która uwzględnia również odpowiedź ośrodka na zewnętrzne pole. Dlatego pomiędzy natężeniem i indukcją istnieją określone relacje zwane **równaniami materiałowymi**:

$$\mathbf{D} = \varepsilon \mathbf{E}\,, \tag{1.2a}$$

$$\mathbf{B} = \mu \mathbf{H}\,, \tag{1.2b}$$

gdzie ε i μ są odpowiednio **przenikalnością elektryczną** i **przenikalnością magnetyczną** ośrodków. W ogólności wielkości te są tensorami, z czym, na

przykład, mamy do czynienia w opisie ośrodków anizotropowych. Jednocześnie w przypadku ośrodków izotropowych, a z takimi będziemy mieli do czynienia w większości przypadków rozważanych w tym podręczniku, ε i μ są skalarami. W takim przypadku przenikalność elektryczna jest często nazywana **stałą dielektryczną**.

Z uwagi na fakt, że światło widzialne nie rozchodzi się w metalach, bardzo często na potrzeby opisu propagacji światła rozpatruje się **ośrodki dielektryczne**, w których nie istnieją swobodne ładunki elektryczne ($\rho = 0$). Korzystając z równań materiałowych, możemy równania Maxwella zapisać w takim przypadku w postaci:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}\,, \tag{1.3a}$$

$$\nabla \cdot \mathbf{E} = 0\,, \tag{1.3b}$$

$$\nabla \cdot \mathbf{B} = 0\,, \tag{1.3c}$$

$$\nabla \times \mathbf{B} = \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t}\,. \tag{1.3d}$$

Najprostszym przykładem dielektryka jest próżnia, w której przenikalność elektryczna i magnetyczna przyjmują charakterystyczne wartości:

$$\varepsilon_0 = 8,85 \times 10^{-12}\ \frac{\mathrm{F}}{\mathrm{m}}\,, \tag{1.4a}$$

$$\mu_0 = 4\pi \times 10^{-7}\ \frac{\mathrm{N}}{\mathrm{A}^2}\,. \tag{1.4b}$$

Dlaczego E i B, a nie D i H?

W literaturze naukowej, a zwłaszcza w optyce, do opisu fali elektromagnetycznej powszechnie używa się pary wektorów: **natężenia pola elektrycznego E** oraz **indukcji pola magnetycznego B**. Może to budzić zdziwienie, gdyż nazwy sugerują pewną asymetrię (jeden wektor to „natężenie", a drugi to „indukcja"). Dlaczego nie używamy pary **E** i **H**?

Odpowiedź kryje się w **sile Lorentza**:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

Równanie to opisuje rzeczywistą siłę, jaką pole elektromagnetyczne wywiera na ładunek elektryczny. Jak widać, wielkościami fizycznymi, które bezpośrednio „popychają" ładunki, są właśnie **E** oraz **B**. Wielkości **D** i **H** są tzw. polami pomocniczymi, wprowadzonymi w celu ułatwienia obliczeń w ośrodkach materialnych (uwzględniają one uśrednioną odpowiedź ośrodka), ale to **E** i **B** są polami fundamentalnymi.

1.2 Równanie fali elektromagnetycznej

Rozpatrzmy teraz równania Maxwella w ośrodku dielektrycznym. Korzystając z prawa Faradaya i obkładając opisujące je równanie (1.3a) lewostronnie operatorem rotacji, otrzymujemy:

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \frac{\partial \mathbf{B}}{\partial t}. \tag{1.5}$$

Dygresja matematyczna

Dla przeprowadzenia dalszych analiz konieczne jest skorzystanie z tożsamości wektorowej:

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{A}\,(\mathbf{B} \cdot \mathbf{C}) - (\mathbf{A} \cdot \mathbf{B})\,\mathbf{C}. \tag{1.6}$$

Korzystając z równania (1.6), lewą stronę równania (1.5) można przepisać do postaci:

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}\,. \tag{1.7}$$

Pierwszy z wyrazów po prawej stronie znika na mocy prawa Gaussa dla ośrodków dielektrycznych [równanie (1.3b)].

W równaniu (1.5) możemy zamienić kolejność różniczkowania po przestrzeni i czasie. Pozwala to skorzystać z równania Ampère'a dla dielektryków [równanie (1.3d)] i ostatecznie zapisać:

$$\boxed{\nabla^2 \mathbf{E} - \mu\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.} \tag{1.8}$$

W pełni analogiczną postać równania falowego można wyprowadzić dla pola magnetycznego:

$$\boxed{\nabla^2 \mathbf{B} - \mu\varepsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0.} \tag{1.9}$$

Dygresja

W pierwszej połowie XVIII wieku Jean d'Alembert (1717–1783) opisał ruch struny wprawionej w drgania równaniem zwanym dziś równaniem falowym:

$$\frac{\partial^2 \psi}{\partial z^2} - \frac{1}{v_\psi^2} \frac{\partial^2 \psi}{\partial t^2} = 0. \tag{1.10}$$

W powyższym równaniu ψ jest funkcją skalarną spełniającą równanie falowe, zaś v_ψ jest prędkością, z jaką porusza się fala. Pomimo że funkcja falowa ψ zależy zarówno od czasu, jak i od przestrzeni, $\psi = \psi(z, t)$, z

uwagi na rozchodzenie się fali ze skończoną prędkością obie wielkości łączy zależność $\Delta z = v_\psi \Delta t$, gdzie Δz i Δt oznaczają odpowiednio: zmianę położenia fali oraz upływ czasu. Zależność ta pozwala zapisać funkcję falową w postaci:

$$\psi(z,t) = \phi_+(z - v_\psi t). \tag{1.11}$$

Funkcja (1.11) opisuje falę poruszającą się w dodatnim kierunku osi z. Warto jednak zauważyć, że w równaniu falowym prędkość v_ψ występuje w kwadracie. Dlatego też równorzędnym rozwiązaniem równania falowego jest funkcja falowa opisująca falę poruszającą się w przeciwnym kierunku z tą samą prędkością:

$$\psi(z,t) = \phi_-(z + v_\psi t). \tag{1.12}$$

Ponieważ równanie falowe jest równaniem różniczkowym liniowym, to fale opisywane są nie tylko przez funkcje spełniające równanie falowe, ale również przez ich kombinacje (sumy). W ten sposób manifestuje się jeszcze jedna z podstawowych właściwości fali zwana **zasadą superpozycji**. Wykorzystując zasadę superpozycji, można zapisać najbardziej ogólną postać rozwiązania równania falowego:

$$\psi(z,t) = \phi_+(z - v_\psi t) + \phi_-(z + v_\psi t). \tag{1.13}$$

Porównując postać równania (1.8) z postacią równania falowego (1.10) wnioskujemy, że równanie (1.8) jest równaniem falowym. Okazuje się zatem, że w oparciu o analizy równań Maxwella można wykazać istnienie fal elektromagnetycznych, które rozchodzą się w przestrzeni z prędkością :

$$v = \frac{1}{\sqrt{\varepsilon\mu}}. \tag{1.14}$$

Sformułowanie równania falowego dla fal elektromagnetycznych od razu zaowocowało identyfikacją światła jako fali elektromagnetycznej. Było to możliwe dzięki obserwacjom Ole Rømera (1644–1710) i wyznaczeniu prędkości światła na podstawie obrotu księżyców Jowisza oraz znajomości stałej dielektrycznej próżni ε_0 oraz przenikalności magnetycznej próżni μ_0. Prędkość rozchodzenia się światła w próżni jest dziś stałą przyrody, której wartość została ustalona jako :

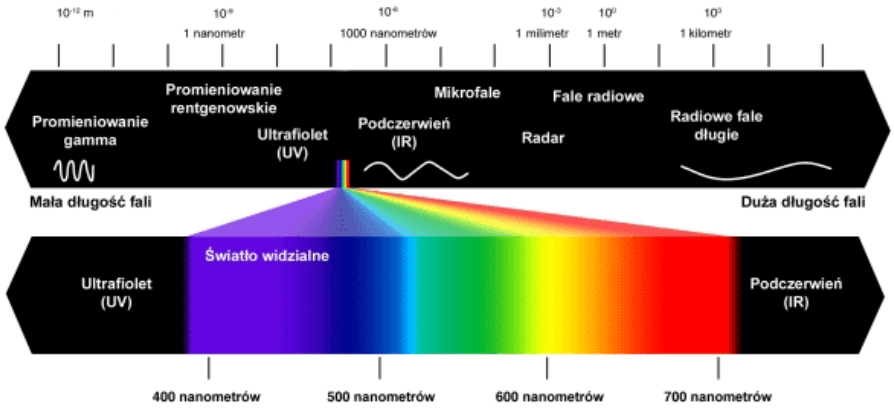
$$c = \frac{1}{\sqrt{\varepsilon_0\mu_0}} = 299\ 792\ 458\ \frac{\mathrm{m}}{\mathrm{s}}. \tag{1.15}$$

Porównując równania (1.14) i (1.15), możemy zdefiniować wielkość, która określać nam będzie, ile razy wolniej porusza się światło w danym ośrodku, niż poruszałoby się w próżni:

$$n = \frac{c}{v} = \sqrt{\frac{\varepsilon\mu}{\varepsilon_0\mu_0}}. \tag{1.16}$$

Wielkość tę będziemy nazywać **współczynnikiem załamania światła**.

Ponieważ zarówno w dielektrykach, jak i w przewodnikach oddziaływanie magnetyczne jest słabe, tzn. $\mu \approx \mu_0$, to równanie (1.16) może zostać przepisane



Rysunek 1.1: Widmo fal elektromagnetycznych

do postaci:

$$n \approx \sqrt{\frac{\varepsilon}{\varepsilon_0}} = \sqrt{\varepsilon_r}, \tag{1.17}$$

gdzie ε_r jest tzw. **względną przenikalnością elektryczną ośrodka**.

Dziś wiadomo, że spektrum widzialne jest tylko bardzo niewielką częścią całego widma fal elektromagnetycznych. Na rysunku 1.1 przedstawiono znane obecnie fale elektromagnetyczne wraz z odpowiadającymi im zakresami długości fali i częstości. I choć oryginalnie nazwa optyka zarezerwowana była dla opisu fal z zakresu widzialnego, dziś często tego pojęcia używa się do opisu fal elektromagnetycznych z szerokiego zakresu spektralnego: od promieniowania X do dalekiej podczerwieni (a czasami nawet jeszcze szerzej).

1.3 Fale monochromatyczne

Zgodnie z przeprowadzoną powyżej dyskusją, rozwiązaniem równania falowego dla składowej elektrycznej fali elektromagnetycznej jest dowolna funkcja w postaci $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r} \pm \mathbf{v}t)$. Szczególnym przykładem rozwiązania równania falowego jest funkcja opisująca **falę monochromatyczną**, czyli falę o pojedynczej częstości ω :

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}) \cos(\omega t - \varphi)\,, \tag{1.18}$$

gdzie φ jest fazą początkową fali. W równaniu (1.18) posłużono się metodą separacji zmiennych, która pozwoliła rozdzielić zależność przestrzenną pola elektrycznego fali elektromagnetycznej od jej zależności czasowej. Szczególnym, ale nie jedynym, rozwiązaniem części czasowej równania falowego jest funkcja cosinus.

Warto zwrócić uwagę, że zgodnie z transformacją Fouriera dowolny przebieg czasowy, czy to periodyczny, czy też nieokresowy, można przedstawić jako

złożenie fal monochromatycznych o odpowiednich amplitudach. W tym kontekście, przedstawienie zależności czasowej równania falowego w postaci fal monochromatycznych jest rozwiązaniem ogólnym i dlatego będzie ono powszechnie wykorzystywane w niniejszym podręczniku.

1.4 Zapis zespolony fali elektromagnetycznej

Z matematycznego punktu widzenia, opis fal przy pomocy funkcji trygonometrycznych nie jest najwygodniejszy z uwagi na konieczność przekształcania przy różniczkowaniu funkcji sinus i cosinus przy jednoczesnym uwzględnieniu znaku. Z drugiej jednak strony, korzystając ze związków pomiędzy funkcjami trygonometrycznymi i wykładniczymi, możemy przepisać te równania do postaci wykładniczej:

$$\begin{aligned}
\mathbf{E}(\mathbf{r},t) =& \mathbf{E}(\mathbf{r})\frac{e^{i\varphi}}{2}e^{-i\omega t} + \mathbf{E}(\mathbf{r})\frac{e^{-i\varphi}}{2}e^{i\omega t} \\
=& \mathbf{E}^{(+)}(\mathbf{r})e^{-i\omega t} + \mathbf{E}^{(-)}(\mathbf{r})e^{i\omega t} \\
=& \mathbf{E}^{(+)}(\mathbf{r})e^{-i\omega t} + \text{c.c.} \\
=& 2\text{Re}\left[\mathbf{E}^{(+)}(\mathbf{r})e^{-i\omega t}\right], \\
\mathbf{B}(\mathbf{r},t) =& \mathbf{B}^{(+)}(\mathbf{r})e^{-i\omega t} + \text{c.c.} \\
=& 2\text{Re}\left[\mathbf{B}^{(+)}(\mathbf{r})e^{-i\omega t}\right],
\end{aligned} \tag{1.19}$$

gdzie $\mathbf{E}^{(+)}(\mathbf{r}) = \mathbf{E}(\mathbf{r})e^{i\varphi}/2$ oraz $\mathbf{B}^{(+)}(\mathbf{r}) = \mathbf{B}(\mathbf{r})e^{i\varphi}/2$ są zespolonymi amplitudami wektorowymi pola elektrycznego i magnetycznego. Z uwagi na zasadę superpozycji fale elektromagnetyczne często zapisuje się wprost w postaci wykładniczej porzucając człon sprzężony:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}^{(+)}(\mathbf{r})e^{-i\omega t}, \tag{1.20a}$$

$$\mathbf{B}(\mathbf{r},t) = \mathbf{B}^{(+)}(\mathbf{r})e^{-i\omega t}. \tag{1.20b}$$

W takim przypadku należy jednak pamiętać, że po obliczeniu rozwiązania interesująca jest jedynie jego część rzeczywista.

Przedstawienie pola elektrycznego i magnetycznego w postaci zespolonej znacząco ułatwia obliczenia. W szczególności, różniczkowanie po czasie zwraca człon $-i\omega$, co sprawia, że równania Maxwella dla ośrodków dielektrycznych mogą zostać zapisane w postaci zespolonej:

$$\nabla \cdot \mathbf{E}^{(+)} = 0, \tag{1.21a}$$

$$\nabla \cdot \mathbf{B}^{(+)} = 0, \tag{1.21b}$$

$$\nabla \times \mathbf{E}^{(+)} = i\omega\mathbf{B}^{(+)}, \tag{1.21c}$$

$$\nabla \times \mathbf{B}^{(+)} = -i\omega\varepsilon\mu\mathbf{E}^{(+)}, \tag{1.21d}$$

które uzupełnione są zmodyfikowanymi wersjami równań materiałowych:

$$\mathbf{D}^{(+)} = \varepsilon \mathbf{E}^{(+)}\,, \tag{1.22a}$$

$$\mathbf{B}^{(+)} = \mu \mathbf{H}^{(+)}\,, \tag{1.22b}$$

dając pełen opis układu.

1.5 Parametry opisujące fale elektromagnetyczne

1.5.1 Podstawowe wielkości opisujące fale elektromagnetyczne

Fale elektromagnetyczne są rozchodzącym się w przestrzeni zaburzeniem pola elektrycznego i magnetycznego. Z tego też powodu opisuje je szereg podstawowych wielkości. Pierwszą jest **częstość kołowa** ω, zdefiniowana jako:

$$\omega = \frac{2\pi}{T}, \tag{1.23}$$

gdzie T jest okresem oscylacji fali. Częstość ω wiąże się z **częstotliwością fali** ν zależnością:

$$\omega = 2\pi\nu. \tag{1.24}$$

Podczas gdy częstość kołowa wyrażana jest w radianach na sekundę [rad/s], częstotliwość wyrażona jest w hercach [Hz].

Inną wielkością charakteryzującą falę jest **długość fali** λ, będąca najmniejszą odległością pomiędzy dwoma punktami fali drgającymi w tej samej fazie. Długość fali i jej częstotliwość związane są zależnością:

$$v = \lambda\nu, \tag{1.25}$$

gdzie v jest prędkością propagacji fali.[1] W przypadku fal elektromagnetycznych z zakresu widzialnego (światło) długość fali mieści się w zakresie od 400 do 700 nanometrów [nm].

Kolejną wielkością charakteryzującą fale jest jej amplituda. W przypadku fali elektromagnetycznej amplitudą fali jest amplituda pola elektrycznego fali świetlnej E_0 (lub $E_0^{(+)}$, gdy korzystać będziemy z zapisu zespolonego). Jej jednostką jest wolt na metr [V/m].

Zwróćmy teraz uwagę na fakt, że fala elektromagnetyczna jest wielkością wektorową. Możliwe jest zatem rozważenie dwóch fal, które drgają w kierunkach prostopadłych:

$$\mathbf{E}_1(\mathbf{r}, t) = \mathbf{x} E_{0x}^{(+)}(\mathbf{r}) e^{-i\omega t}, \tag{1.26a}$$

$$\mathbf{E}_2(\mathbf{r}, t) = \mathbf{y} E_{0y}^{(+)}(\mathbf{r}) e^{-i(\omega t+\varphi)}, \tag{1.26b}$$

[1] W dalszej części podręcznika wielkość ta będzie nazywana prędkością fazową fali.

gdzie $E_{0x}^{(+)}$ i $E_{0y}^{(+)}$ są rzeczywistymi amplitudami pola, podczas gdy różnica faz pomiędzy drganiami pola elektrycznego fali została zdeterminowana przez fazę φ. Zgodnie z zasadą superpozycji, również suma dwóch fal drgających w prostopadłych kierunkach spełnia równanie falowe:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_1(\mathbf{r},t) + \mathbf{E}_2(\mathbf{r},t) = (\mathbf{x}a_x + \mathbf{y}a_y e^{i\varphi})\tilde{E}_0^{(+)} e^{-i\omega t}, \tag{1.27}$$

gdzie $a_x = E_{0x}/\tilde{E}_0$ i $a_y = E_{0y}/\tilde{E}_0$ przy czym $\tilde{E}_0 = \sqrt{E_{0x}^2 + E_{0y}^2}$. Pole elektryczne dane równaniem (1.27) opisuje pole elektromagnetyczne, którego amplituda zorientowana jest w płaszczyźnie xy, a pomiędzy składowymi x i y istnieje określona relacja fazowa. Wyrażenie $\mathbf{x}a_x + \mathbf{y}a_y \exp(i\varphi)$ jest wektorem zespolonym w płaszczyźnie xy i opisuje ono **polaryzację fali elektromagnetycznej**.

Analizując możliwe stany polaryzacji światła, rozpatrzmy kilka szczególnych przypadków:

$a_x = 0$ **lub** $a_y = 0$

W najprostszym przypadku amplituda jednej ze składowych liniowych pola $\mathbf{E}(\mathbf{r})$ jest równa zero, co sprawia, że wektor elektryczny światła drga albo wzdłuż x ($a_y = 0$), albo wzdłuż y ($a_x = 0$). Taki rodzaj polaryzacji światła nazywany jest **polaryzacją liniową**.

$\varphi = 0$ **lub** $\varphi = \pi$

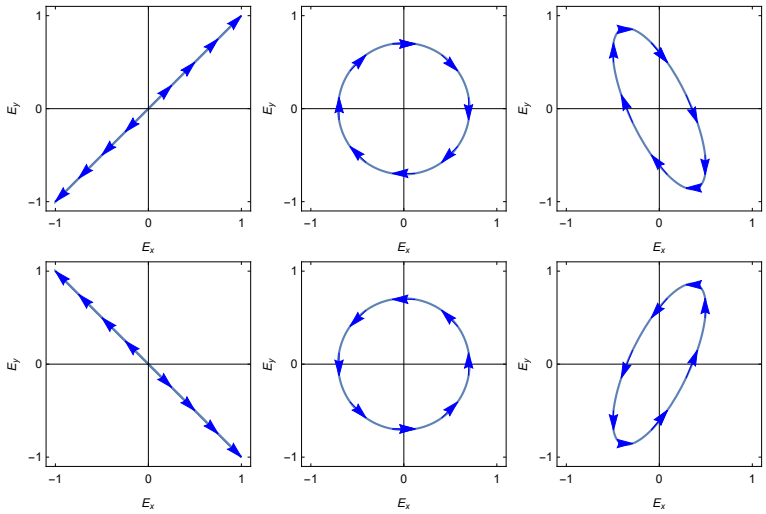
Choć w takim przypadku drganie nie odbywa się wyłącznie wzdłuż x lub y, to, co można pokazać wykorzystując dodawanie wektorowe składowych x i y, wektor pola elektrycznego drga wzdłuż określonego kierunku, a sama fala spolaryzowana jest liniowo. Kierunek drgań pola elektrycznego wypadkowej fali $\mathbf{E}(\mathbf{r})$ jest zorientowany względem osi x pod kątem α_E:

$$\alpha_E = \operatorname{arctg}\left(\frac{a_y}{a_x}\right). \tag{1.28}$$

Warto zauważyć, że w przypadku gdy różnica faz $\varphi = 0$ wektor tworzy kąt dodatni z osią x, podczas gdy dla $\varphi = \pi$ jest on ujemny.

$\varphi = \pi/2$ **lub** $\varphi = 3\pi/2$ **oraz** $a_x = a_y$

Innym szczególnym przypadkiem jest sytuacja, w której $E_{0x} = E_{0y}$, a między składowymi istnieje przesunięcie fazy $\pi/2$ lub $3\pi/2$. W tym pierwszym przypadku drganie w kierunku y wyprzedza drganie w kierunku x o ćwierć okresu i nie tylko w żadnej chwili długość wektora $\mathbf{E}$ nie jest zerowa, ale jest ona stała. Oznacza to, że koniec wektora pola elektrycznego $\mathbf{E}$ porusza się po okręgu w kierunku zgodnym z kierunkiem ruchu wskazówek zegara. Z tego też powodu polaryzację tę nazywamy **polaryzacją kołową prawoskrętną**. Analogicznie w przypadku gdy różnica faz wynosi $3\pi/2$, drganie w kierunku y jest opóźnione względem drgania w kierunku x i wypadkowy wektor pola elektrycznego porusza się po okręgu w kierunku przeciwnym do ruchu wskazówek zegara, a samą polaryzację nazywamy **polaryzacją kołową lewoskrętną**.



Rysunek 1.2: Różne stany polaryzacji. Pierwsza kolumna: polaryzacja liniowa, dwie polaryzacje prostopadłe. Druga kolumna: polaryzacja kołowa prawo- i lewoskrętna. Trzecia kolumna: polaryzacja eliptyczna prawo- i lewoskrętna.

Uwaga!

Rodzaj skrętności polaryzacji determinuje nie kierunek wektora falowego, ale kierunek osi układu odniesienia. Oznacza to, że niezależnie czy fala porusza się w dodatnim czy w ujemnym kierunku osi z, jeśli wektor pola elektrycznego zatacza ten sam okrąg, mówimy o tej samej skrętności.

$a_x \neq a_y$ **oraz** $\varphi \notin \{0, \pi\}$

W sytuacji, w której długości składowych x i y nie są równe, a pomiędzy nimi istnieje różnica faz inna niż 0 lub π, wektor pola elektrycznego również nie przyjmuje wartości zerowej w żadnym momencie, ale jego długość zmienia się w czasie. Dlatego też taki typ polaryzacji nazywamy **polaryzacją eliptyczną**. Podobnie jak w przypadku polaryzacji kołowej, polaryzacja eliptyczna może być zarówno prawoskrętna, jak i lewoskrętna. Co więcej, różna może być również orientacja tej elipsy w płaszczyźnie xy. Kąt, jaki tworzy wielka oś elipsy polaryzacji z osią x dany jest zależnością:

$$\alpha_E = \frac{1}{2}\mathrm{arctg}\left(\frac{2E_{0x}^{(+)}E_{0y}^{(+)}\cos\Delta\varphi}{\left[E_{0x}^{(+)}\right]^2 + \left[E_{0y}^{(+)}\right]^2}\right). \tag{1.29}$$

Rysunek 1.2 przedstawia przykłady różnych stanów polaryzacji światła opisanych powyżej.

Dygresja

W przypadku oddziaływania światła z ośrodkami materialnymi, naturalną bazą polaryzacji nie jest baza dwóch ortogonalnych polaryzacji liniowych, ale baza, którą tworzy polaryzacja prawo- i lewoskrętna. Dzieje się tak dlatego, że z punktu widzenia mikroskopowego wielkością odpowiadającą polaryzacji jest moment pędu, zaś polaryzacja kołowa prawo- i lewoskrętna odpowiada fotonowi, którego rzut wewnętrznego momentu pędu (spinu) na oś kwantyzacji jest równy odpowiednio 1 lub -1. Ponieważ przy absorpcji światła spełnione muszą być nie tylko zasady zachowania energii i pędu, ale również zasada zachowania momentu pędu, przy absorpcji dochodzi do zmiany momentu pędu atomu o określoną wartość. Prowadzi to do tzw. **reguł wyboru** związanych z odpowiednią polaryzacją. Światło liniowo spolaryzowane, które jest superpozycją dwóch fal kołowo spolaryzowanych o odpowiednich fazach, nie ma jednej reguły wyboru, a jest to kombinacja reguł obu polaryzacji. Prowadzi to do bardziej skomplikowanego, z punktu widzenia mikroskopowego, oddziaływania między światłem i atomami.

1.5.2 Gęstość energii i wektor Poyntinga

Fala elektromagnetyczna, rozchodząc się w przestrzeni, transportuje energię. Zanim jednak przejdziemy do opisu jej transportu, zastanówmy się, ile energii zgromadzone jest w samym polu. Z teorii elektromagnetyzmu wiemy, że w jednostce objętości ośrodka liniowego [Równania (1.2)] pole elektryczne i magnetyczne „magazynują" energię o gęstości odpowiednio u_E i u_B:

$$u_E = \frac{1}{2}\mathbf{E}\cdot\mathbf{D} = \frac{1}{2}\varepsilon E^2, \tag{1.30a}$$

$$u_B = \frac{1}{2}\mathbf{B}\cdot\mathbf{H} = \frac{1}{2\mu}B^2. \tag{1.30b}$$

Całkowita, chwilowa gęstość energii u jest sumą obu tych wkładów:

$$u = u_E + u_B = \frac{1}{2}\varepsilon E^2 + \frac{1}{2\mu}B^2. \tag{1.31}$$

Analizując sprzężone równania Maxwella, które prowadzą do równania falowego, można pokazać, że dla dowolnej fali elektromagnetycznej wędrującej w dielektryku, wartości chwilowe pola elektrycznego i magnetycznego są ze sobą ściśle powiązane zależnością:

$$E = vB = \frac{1}{\sqrt{\varepsilon\mu}}B. \tag{1.32}$$

Podstawiając tę zależność do wzoru na gęstość energii magnetycznej, otrzymujemy:

$$u_B = \frac{1}{2\mu}\left(E\sqrt{\varepsilon\mu}\right)^2 = \frac{1}{2}\varepsilon E^2 = u_E. \tag{1.33}$$

Oznacza to, że w fali elektromagnetycznej energia jest rozłożona równomiernie pomiędzy pole elektryczne i magnetyczne ($u_E = u_B$). Pozwala to zapisać całkowitą chwilową gęstość energii fali, korzystając wyłącznie z natężenia pola elektrycznego:

$$u = \varepsilon E^2. \tag{1.34}$$

W przypadku fal monochromatycznych (harmonicznych), którymi zajmujemy się w tym rozdziale, pole E oscyluje w czasie sinusoidalnie. W praktyce zazwyczaj interesuje nas średnia gęstość energii $\langle u \rangle$. Ponieważ średnia wartość kwadratu funkcji sinus lub cosinus po okresie wynosi $1/2$, otrzymujemy ważną zależność:

$$\boxed{\langle u \rangle = \frac{1}{2}\varepsilon E_0^2,} \tag{1.35}$$

gdzie E_0 jest amplitudą pola elektrycznego fali.

Doświadczenie Obserwacja mocy panelu fotowoltaicznego oświetlanego pod różnym kątem

Wielkością opisującą kierunkowy przepływ tej energii jest **wektor Poyntinga**:

$$\boxed{\mathbf{S} = \mathbf{E} \times \mathbf{H} = \frac{1}{\mu}\mathbf{E} \times \mathbf{B}.} \tag{1.36}$$

Wektor ten wyznacza kierunek propagacji energii, a jego zwrot jest zgodny z kierunkiem rozchodzenia się fali. Wartość chwilowa wektora Poyntinga dla fali wędrującej wiąże się z gęstością energii prostą relacją (analogiczną do iloczynu gęstości i prędkości w mechanice płynów):

$$S = vu. \tag{1.37}$$

Przepisując równanie (1.36) do postaci zespolonej (dla fal harmonicznych), otrzymujemy:

$$\begin{aligned}\mathbf{S} =& (\mathbf{E}^{(+)}e^{-i\omega t} + \mathbf{E}^{(-)}e^{i\omega t}) \times (\mathbf{H}^{(+)}e^{-i\omega t} + \mathbf{H}^{(-)}e^{i\omega t}) \\ =& \mathbf{E}^{(+)} \times \mathbf{H}^{(-)} + \mathbf{E}^{(-)} \times \mathbf{H}^{(+)} + \mathbf{E}^{(+)} \times \mathbf{H}^{(+)}e^{-2i\omega t} + \mathbf{E}^{(-)} \times \mathbf{H}^{(-)}e^{2i\omega t}.\end{aligned} \tag{1.38}$$

Wielkością fizyczną rejestrowaną przez detektory (np. oko ludzkie, matryce CCD) jest **natężenie światła**, zdefiniowane jako średnia wartość strumienia energii:

$$\boxed{I = |\langle \mathbf{S} \rangle| = v\langle u \rangle,} \tag{1.39}$$

gdzie uśrednianie odbywa się po czasie znacznie dłuższym niż okres drgań fali. Wykorzystując postać zespoloną, szybkozmienne człony z czynnikami $e^{\pm 2i\omega t}$ uśredniają się do zera, co daje:

$$I = \left|\mathbf{E}^{(+)} \times \mathbf{H}^{(-)} + \text{c.c.}\right| = 2\left|\text{Re}\left(\mathbf{E}^{(+)} \times \mathbf{H}^{(-)}\right)\right|. \tag{1.40}$$

Łącząc równania (1.35) i (1.39), dla fali rozchodzącej się w próżni możemy zapisać użyteczny wzór praktyczny:

$$I = \frac{1}{2} c \varepsilon_0 E_0^2. \tag{1.41}$$

Z analizy wymiarowej wynika, że jednostką natężenia jest wat na metr kwadratowy [W/m^2].

Doświadczenie Obserwacja mocy panelu fotowoltaicznego oświetlanego pod różnym kątem

Dygresja

Do wytłumaczenia problemu widma promieniowania ciała doskonale czarnego Max Planck wprowadził pojęcie kwantu promieniowania, którego energia dana jest zależnością:

$$E_f = \hbar \omega, \tag{1.42}$$

gdzie ω jest częstością promieniowania, zaś $\hbar \cong 1,05 \times 10^{-34}$ J s jest stałą, która nosi nazwę **zredukowanej stałej Plancka**. W oparciu o tę definicję można przepisać równanie na natężenie światła do postaci:

$$I = \frac{d\Phi_f}{dt} \hbar \omega, \tag{1.43}$$

gdzie $\dot{\Phi}_f$ jest nazywany **strumieniem fotonów**, tzn. liczbą fotonów padającą na jednostkową powierzchnię.

1.5.3 Ciśnienie światła i pęd fali elektromagnetycznej

Doświadczenie Obserwacja obrotu łopatek w młynku Crookesa pod wpływem oświetlenia

Już od czasów Johannesa Keplera (1571–1630) postulowano, że promieniowanie wywiera ciśnienie na obiekty. Dowodem na poparcie tej tezy było orientowanie się ogonów komet w stronę od Słońca podczas ich obiegania tej gwiazdy. Bezpośredni dowód na to został jednak podany dopiero w 1900 roku w eksperymencie zrealizowanym przez Piotra Lebiediewa (1866–1912). W swoim eksperymencie pokazał on, że lekkie skrzydełka zawieszone w próżni wychylają się pod wpływem oświetlenia.

Zjawisko ciśnienia światła można w pełni wyjaśnić na gruncie klasycznej elektrodynamiki, analizując siły działające na ładunki elektryczne w materii, która pochłania promieniowanie. Załóżmy, że fala elektromagnetyczna pada prostopadle na powierzchnię materiału. Fala ta składa się z oscylującego pola elektrycznego $\mathbf{E}$ i magnetycznego $\mathbf{B}$. Gdy fala wnika w materiał, pole elektryczne wprawia w ruch znajdujące się tam ładunki q (np. elektrony), nadając im pewną chwilową prędkość, którą oznaczymy jako $\mathbf{v}_q$.

Na każdy taki poruszający się ładunek działa całkowita siła Lorentza:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v}_q \times \mathbf{B}). \tag{1.44}$$

Pierwszy człon ($q\mathbf{E}$) jest siłą elektryczną. Jak pokażemy w rozdziale 3, w ośrodkach jednorodnych pole $\mathbf{E}$ oscyluje poprzecznie do kierunku rozchodzenia się fali.

W konsekwencji w tym samym kierunku działa na ładunek siła, wykonując nad nim pracę. Szybkość wykonywania tej pracy, tj. moc chwilowa przekazywana ładunkowi, wynosi:

$$\frac{dW_q}{dt} = \mathbf{F}_E \cdot \mathbf{v}_q = q\mathbf{E} \cdot \mathbf{v}_q. \tag{1.45}$$

Drugi człon $(q\mathbf{v}_q \times \mathbf{B})$ to siła magnetyczna. Ponieważ prędkość ładunku $\mathbf{v}_q$ jest wymuszana przez pole $\mathbf{E}$, jest ona do niego równoległa. W rozdziale 3 pokazano, że w ośrodkach jednorodnych wektor $\mathbf{B}$ jest prostopadły do $\mathbf{E}$, a więc jest też prostopadły do wektora prędkości ładunku $\mathbf{v}_q$. Zgodnie z regułą iloczynu wektorowego, siła magnetyczna $\mathbf{F}_B = q\mathbf{v}_q \times \mathbf{B}$ jest skierowana prostopadle do obu tych wektorów, czyli dokładnie w kierunku rozchodzenia się fali. To właśnie ta siła odpowiada za ciśnienie promieniowania.

Obliczmy wartość siły $\mathbf{F}_B$. Ponieważ $\mathbf{v}_q \perp \mathbf{B}$, wartość siły to $F_B = qv_qB$. Korzystając ze związku między polami w fali ($B = E/c$ w próżni), otrzymujemy:

$$F_B = qv_q\left(\frac{E}{c}\right) = \frac{1}{c}(qv_qE). \tag{1.46}$$

Zauważmy, że wyrażenie w nawiasie (qv_qE) jest dokładnie równe mocy przekazywanej ładunkowi [z równania (1.45)]. Zatem siła parcia wywierana na pojedynczy ładunek wiąże się z absorbowaną przez niego mocą relacją:

$$F_B = \frac{1}{c}\frac{dW_q}{dt}. \tag{1.47}$$

Sumując siły działające na wszystkie ładunki w oświetlanej powierzchni, otrzymujemy całkowitą siłę F działającą na materiał, która zależy od całkowitej mocy pochłanianej P_{abs}:

$$F_{calk} = \sum F_B = \frac{1}{c}\sum \frac{dW_q}{dt} = \frac{P_{abs}}{c}. \tag{1.48}$$

Definiując ciśnienie $\mathcal{P}$ jako siłę na jednostkę powierzchni (F/A) oraz pamiętając, że moc na jednostkę powierzchni to natężenie światła $(I = P/A)$, otrzymujemy ostateczny wzór na **ciśnienie promieniowania** dla ciała doskonale absorbującego:

$$\boxed{\mathcal{P} = \frac{F_{calk}}{A} = \frac{P_{abs}}{Ac} = \frac{I}{c}.} \tag{1.49}$$

Ponieważ wcześniej wykazaliśmy, że I/c jest równe gęstości energii u, dochodzimy do wniosku, że ciśnienie światła jest liczbowo równe gęstości energii pola przy powierzchni $(\mathcal{P} = u)$.

W przypadku całkowitego odbicia (np. od idealnego lustra), zmiana pędu jest dwukrotnie większa, co prowadzi do dwukrotnie wyższego ciśnienia:

$$\mathcal{P}_{odb} = \frac{2I}{c}. \tag{1.50}$$

Warto na koniec zauważyć, że kierunek ciśnienie światła jest zgodny z kierunkiem rozchodzenia się energii, co pozwala nam zapisać zależność

$$\mathcal{P} = \frac{\langle \mathbf{S} \rangle}{c}. \tag{1.51}$$

Dygresja

Choć powyższe wyprowadzenie jest czysto klasyczne, wynik ten jest w pełni zgodny z mechaniką kwantową. Jeśli potraktujemy światło jako strumień fotonów, z których każdy niesie energię $E_f = \hbar\omega$, to zgodnie z teorią względności musi on również nieść pęd:

$$p_f = \frac{E_f}{c} = \frac{\hbar\omega}{c} = \hbar k, \tag{1.52}$$

gdzie k jest wartością wektora falowego. Sumaryczny pęd przekazywany przez ogromną liczbę fotonów uderzających w powierzchnię daje dokładnie to samo ciśnienie makroskopowe, co siła Lorentza działająca na ładunki.

1.5.4 Moment pędu

Weźmy polaryzację kołową prawoskrętną daną równaniem:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{x}E^{(+)}(\mathbf{r})e^{-i\omega t} + \mathbf{y}E^{(+)}(\mathbf{r})e^{-i(\omega t + \pi/2)}. \tag{1.53}$$

Załóżmy teraz, że pole elektryczne tej fali działa na ładunek elektryczny (elektronu) e wprawiając go w ruch obrotowy. Posłużymy się tu klasycznym, poglądowym modelem elektronu poruszającego się po orbicie kołowej pod wpływem siły pola elektrycznego. Z II zasady dynamiki Newtona wiemy, że zmiana momentu pędu $\mathbf{J}$ ładunku w czasie zależy od momentu siły, jaki działa na ten ładunek:

$$\frac{d\mathbf{J}}{dt} = \mathbf{r} \times \mathbf{F}. \tag{1.54}$$

Zakładając, że siła działająca na ładunek jest siłą elektrostatyczną, otrzymujemy:

$$\frac{dJ}{dt} = rE_0 e, \tag{1.55}$$

gdzie uwzględniono fakt, że $\mathbf{J} \| (\mathbf{r} \times \mathbf{E_0})$. Praca, jaką wykonuje pole elektryczne nad ładunkiem, dana jest zależnością:

$$W = e\mathbf{E_0} \cdot \mathbf{s}, \tag{1.56}$$

skąd też z zasady zachowania energii można pokazać, że zmiana energii w czasie dE_e/dt jest równa (zakładamy, że siła działa wzdłuż drogi, tj. $\mathbf{E_0} \| d\mathbf{s}$):

$$\frac{dE_e}{dt} = eE_0 \frac{ds}{dt} = eE_0 v_e, \tag{1.57}$$

gdzie v_e jest prędkością elektronu. Przekształcając równanie (1.57), otrzymujemy:

$$v_e = \frac{dE_e/dt}{eE_0}. \tag{1.58}$$

Uwzględniając, że $r = v_e/\omega$, równanie (1.58) można przepisać do postaci:

$$rE_0 e = \frac{1}{\omega}\frac{dE_e}{dt}, \tag{1.59}$$

co po wstawieniu do równania (1.55) pozwala zapisać:

$$\frac{dJ}{dt} = \frac{1}{\omega}\frac{dE_e}{dt}. \tag{1.60}$$

Po wycałkowaniu tego równania po czasie otrzymujemy:

$$\boxed{J = \frac{E_e}{\omega}}. \tag{1.61}$$

Dygresja

Ponieważ zgodnie z mechaniką kwantową energia jest skwantowana $E_e = N\hbar\omega$, równanie (1.61) można przepisać do postaci:

$$J = N\hbar. \tag{1.62}$$

Z powyższej zależności widać, że **moment pędu fotonu** jest równy $\hbar$.

1.5.5 Oddziaływanie światła z materią a zasady zachowania

Wyprowadzone powyżej pojęcia energii, pędu i momentu pędu fali elektromagnetycznej mają fundamentalne znaczenie dla jej oddziaływania z ośrodkami materialnymi, bowiem przy oddziaływaniu muszą być spełnione trzy zasady zachowania: **zasada zachowania energii**, **zasada zachowania pędu** oraz **zasada zachowania momentu pędu**. O ile spełnienie zasad zachowania pędu jest dość proste z uwagi na brak kwantowania tej wielkości, o tyle w myśl zasady zachowania energii, energia niesiona przez pojedynczy foton musi być możliwa do przyjęcia przez ośrodek. Innymi słowy, w ośrodku muszą istnieć poziomy energetyczne, na które atom może przejść po przyjęciu energii. Przejście to będzie jednak możliwe tylko wtedy, gdy spełniona zostanie ostatnia ze wspomnianych wyżej zasad — zasada zachowania momentu pędu. Okazuje się bowiem, że z uwagi na określoną wartość momentu pędu niesioną przez foton i istnienie maksymalnej wartości momentu pędu, jaką może posiadać atom, atomy mogą zaabsorbować tylko skończoną liczbę spolaryzowanych fotonów.

W konsekwencji może się okazać, że nawet rezonansowe światło może nie być w ośrodku absorbowane, ponieważ budujące go atomy nie są w stanie przyjąć od niego momentu pędu. Prowadzi to do wielu ciekawych efektów, takich jak pompowanie optyczne, elektromagnetycznie indukowana przeźroczystość, „wolne" i „szybkie" światło itp., które intensywnie badane są w ramach optyki nieliniowej.

1.6 Podsumowanie rozdziału

W niniejszym rozdziale sformułowano teoretyczne fundamenty optyki klasycznej, wychodząc od makroskopowych równań Maxwella. Wykazano, że ich spójny układ różniczkowy prowadzi bezpośrednio do falowego opisu propagacji pola elektromagnetycznego, pozwalając na wyznaczenie prędkości światła oraz zdefiniowanie współczynnika załamania ośrodków materialnych. Wprowadzono również aparat matematyczny, w tym formalizm zapisu zespolonego fal monochromatycznych oraz wektorowy opis stanów polaryzacji. Zdefiniowano również wielkości fizyczne kluczowe dla opisu transportu pola elektromagnetycznego: gęstość energii, natężenie światła (opisane wektorem Poyntinga), a także ciśnienie promieniowania i moment pędu fali.

Zwrócono ponadto uwagę na fundamentalne zasady zachowania (energii, pędu i momentu pędu) rządzące oddziaływaniem promieniowania z materią, co buduje naturalny pomost pomiędzy klasyczną falą a kwantowym charakterem absorpcji manifestującym się w regułach wyboru. Tak skonstruowany, kompletny obraz fali elektromagnetycznej stanowi bazę dla dalszych rozważań podręcznika, zarówno dla optyki geometrycznej, jak i zjawisk dyfrakcji, interferencji czy optyki nieliniowej.

Zadania

1.1. Proszę sprawdzić, które z podanych poniżej funkcji opisują falę:

(a) $f_1(z,t) = Ae^{-b(z-vt)^2}$,

(b) $f_2(z,t) = A\sin[b(z-vt)]$,

(c) $f_3(z,t) = \frac{A}{b(z-vt)^2+1}$,

(d) $f_4(z,t) = Ae^{b(bz^2+vt)}$,

(e) $f_5(z,t) = A\sin(bz)\cos^3(bvt)$.

1.2. Dana jest fala elektromagnetyczna w postaci:

$$\mathbf{E}(z,t) = \mathbf{y}(30\mathrm{V/m})\cos[(6,28\mathrm{rad/m})z - (20\mathrm{rad/s})t + \pi/4].$$

Proszę wyznaczyć:

(a) częstotliwość,

(b) długość fali,

(c) okres,

(d) amplitudę,

(e) prędkość rozchodzenia się fali,

(f) kierunek rozchodzenia się fali,

(g) kierunek polaryzacji fali.

1.3. Średnia moc promieniowania emitowanego przez Słońce równa jest $3,9 \cdot 10^{26}$ W. Określ średnią amplitudę pola elektrycznego fali świetlnej docierającej do górnej warstwy atmosfery Ziemi (odległej od Słońca o $1,5 \cdot 10^{11}$ m).

1.4. Typowy czerwony wskaźnik laserowy emituje wiązkę światła o mocy $P = 5$ mW i promieniu przekroju poprzecznego $r = 1$ mm.

(a) Oblicz średnią wartość natężenia światła I (wektora Poyntinga) tej wiązki.

(b) Korzystając z relacji między natężeniem światła a amplitudą pola elektrycznego, oblicz E_0.

(c) Wiedząc, że w próżni (i w powietrzu) $E_0 = cB_0$, oblicz amplitudę indukcji pola magnetycznego B_0 tej fali świetlnej.

(d) Porównaj otrzymaną wartość B_0 ze średnim natężeniem ziemskiego pola magnetycznego (ok. 50 μT). Co ten wynik mówi o tym, dlaczego w optyce oddziaływanie światła z materią opisuje się głównie poprzez pole elektryczne?

1.5. Proszę obliczyć ciśnienie, jakie na powierzchni Ziemi wywiera promieniowanie emitowane ze Słońca. Dla uproszczenia proszę przyjąć, że natężenie tego światła, zwane stałą słoneczną, jest równe 1400 W/m^2.

1.6. Wiązka atomów sodu emitowana jest kierunkowo z pieca. Prędkość atomów opuszczających piec jest równa 500 m/s. Atomy te oświetlane są rezonansową wiązką światła o długości fali 590 nm. Po wzbudzeniu atom powraca do stanu podstawowego emitując foton w przypadkowym kierunku. W ten sposób średni przekaz pędu przy absorpcji jednego fotonu równy jest pędowi fotonu. Proszę obliczyć:

(a) Ile fotonów musi być zaabsorbowanych, żeby całkowicie zatrzymać atom?

(b) Jaką drogę w tym czasie przebędą spowalniane atomy? Dla obliczeń proszę przyjąć, że fotony absorbowane są natychmiast po tym jak atom znajdzie się w stanie podstawowym, podczas gdy reemisja zaabsorbowanych fotonów odbywa się po czasie $T = 16$ ns.

1.7. W czytnikach DVD typu Blu-ray zainstalowany był laser niebieski o długości fali 405 nm i mocy 30 mW. Zakładając, że promień laserowy skupiany jest na powierzchni okręgu o średnicy 400 nm, proszę obliczyć liczbę fotonów padającą na powierzchnię płyty w ciągu sekundy.

1.8. Światło lampy błyskowej (550 nm) zasilanej napięciem 3 V, przez którą w czasie błysku przepływa prąd o natężeniu 0,25 A, emitowane jest kierunkowo w walec o polu poprzecznym 10 cm^2. Zakładając, że efektywność konwersji energii na światło wynosi 1%, proszę obliczyć liczbę fotonów emitowanych z lampy w ciągu błysku.

1.9. Laser femtosekundowy emituje światło w postaci impulsów o czasie trwania 10 fs i częstotliwości repetycji 1 kHz. Średnia moc światła emitowanego przez laser wynosi 5 W, a średnica wiązki światła wynosi 2 mm. Zakładając prostokątny kształt impulsu światła emitowanego przez laser, oblicz średnią wartość natężenia pola elektrycznego w impulsie. Porównaj otrzymaną wartość z:

(a) wartością pola elektrycznego powstałego na powierzchni kuli ($r =$ 30 cm) generatora Van de Graaffa naładowanego do potencjału 100 kV,

(b) wartością pola elektrycznego, jakie odczuwa elektron na powłoce walencyjnej w atomie uranu.

1.10. Wiązka światła lasera o mocy $P = 1$ W i długości fali $\lambda = 532$ nm jest całkowicie spolaryzowana kołowo (prawoskrętnie). Wiązka ta pada prostopadle na idealnie pochłaniającą, zawieszoną swobodnie małą tarczę i zostaje przez nią w 100% zaabsorbowana.

(a) Ile fotonów uderza w tarczę w ciągu jednej sekundy?

(b) Pamiętając, że każdy foton fali spolaryzowanej kołowo niesie moment pędu równy $\hbar$, oblicz całkowity moment pędu przekazywany tarczy w ciągu jednej sekundy.

1.11. Krótki impuls światła lasera zostaje wprowadzony do rdzenia światłowodu ze szkła kwarcowego o długości $L = 15$ km. Względna przenikalność elektryczna tego szkła dla użytej długości fali wynosi $\varepsilon_r \approx 2,16$, a przenikalność magnetyczną można uznać za równą przenikalności próżni ($\mu_r \approx 1$).

(a) Oblicz współczynnik załamania n tego szkła.

(b) Oblicz czas, jakiego potrzebuje impuls światła na pokonanie całego światłowodu, i porównaj go z czasem, jakiego potrzebowałby na pokonanie tego samego dystansu w próżni.

Rozdział 2

Optyka geometryczna

2.1 Równanie Helmholtza

Rozważmy trójwymiarowe równanie falowe

$$\left(\nabla^2 - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E}(\mathbf{r},t) = 0, \tag{2.1}$$

do którego podstawiamy rozwiązanie w postaci fali monochromatycznej

$$\mathbf{E}^{(+)}(\mathbf{r},t) = \mathbf{E}^{(+)}(\mathbf{r})e^{-i\omega t}. \tag{2.2}$$

Podwójne różniczkowanie po czasie pozwala na zapisanie równania (2.1) w postaci

$$\left(\nabla^2 + k^2\right)\mathbf{E}^{(+)}(\mathbf{r}) = 0, \tag{2.3}$$

gdzie $k = \omega/v = \omega\sqrt{\varepsilon\mu}$ jest **liczbą falową** opisującą propagację fali w ośrodku o stałej dielektrycznej ε oraz przenikalności magnetycznej μ. Równanie to, zwane **równaniem Helmholtza**, zależy tylko od składowych przestrzennych, co sprawia, że pozwala ono określić rozkład pola elektrycznego światła w przestrzeni. W tej sytuacji, korzystając z definicji współczynnika załamania n, możemy nadać równaniu (2.3) postać:

$$\nabla^2 E^{(+)}(\mathbf{r}) + n^2 k_0^2 E^{(+)}(\mathbf{r}) = 0, \tag{2.4}$$

gdzie $k_0 = \omega/c$ jest liczbą falową opisującą propagację fali w próżni.

Warto zwrócić uwagę, że równanie (2.4) pozwala określić propagację światła w ośrodku, nie tylko w przypadku gdy stałe przenikalności elektrycznej i magnetycznej nie zależą od położenia, ale również w przypadku gdy są one funkcjami położenia, tj. gdy $\varepsilon(\mathbf{r}) \neq \text{const}$ i $\mu(\mathbf{r}) \neq \text{const}$. W takim przypadku równanie Helmholtza należy albo rozwiązać niezależnie w każdej podprzestrzeni opisanej stałymi wartościami ε oraz μ, a następnie narzucić na rozwiązania warunki ciągłości i w ten sposób dobrać wartości stałych w rozwiązaniu, albo sparametryzować równanie i rozwiązać je analitycznie (o ile to możliwe) lub numerycznie.

2.2 Równanie eikonału i przybliżenie optyki geometrycznej

Załóżmy, że propagację światła w przestrzeni (np. w ośrodku o zmiennym współczynniku załamania $n(\mathbf{r})$) opisuje równanie falowe, które dla fali monochromatycznej sprowadza się do równania Helmholtza [równanie (2.4)]. W ogólności pole elektryczne jest wielkością wektorową, jednakże w wielu zagadnieniach optyki można uprościć opis, analizując każdą z jego składowych kartezjańskich osobno jako wielkość skalarną. Oznaczmy taką wybraną składową zespolonego pola elektrycznego przez $E^{(+)}(\mathbf{r})$.

Zaproponujmy rozwiązanie równania Helmholtza w postaci fali, w której wyraźnie rozdzielimy część amplitudową od czynnika fazowego:

$$E^{(+)}(\mathbf{r}) = E_0^{(+)}(\mathbf{r}) e^{ik_0 R(\mathbf{r})}, \tag{2.5}$$

gdzie $E_0^{(+)}(\mathbf{r})$ jest rzeczywistą amplitudą przestrzenną, $k_0 = 2\pi/\lambda_0$ jest liczbą falową w próżni, a $R(\mathbf{r})$ jest poszukiwaną rzeczywistą funkcją skalarną, która określa rozkład fazy w przestrzeni. Powierzchnie zdefiniowane równaniem $R(\mathbf{r}) =$ const stanowią powierzchnie stałej fazy, czyli tzw. **czoła fali**.

Aby sprawdzić, pod jakimi warunkami funkcja (2.5) spełnia skalarne równanie Helmholtza $\nabla^2 E^{(+)} + k_0^2 n^2(\mathbf{r}) E^{(+)} = 0$, musimy wyznaczyć jej laplasjan.

Korzystając ze wzoru na pochodną iloczynu, w pierwszym kroku obliczamy gradient pola:

$$\nabla E^{(+)}(\mathbf{r}) = \left[\nabla E_0^{(+)}(\mathbf{r}) + ik_0 E_0^{(+)}(\mathbf{r}) \nabla R(\mathbf{r})\right] e^{ik_0 R(\mathbf{r})}. \tag{2.6}$$

Następnie, stosując operator dywergencji $(\nabla\cdot)$ do powyższego wektora, po uporządkowaniu wyrazów otrzymujemy laplasjan:

$$\begin{aligned}\nabla^2 E^{(+)}(\mathbf{r}) = \Big\{ &\nabla^2 E_0^{(+)}(\mathbf{r}) - k_0^2 E_0^{(+)}(\mathbf{r}) [\nabla R(\mathbf{r})]^2 \\ &+ ik_0 \left(E_0^{(+)}(\mathbf{r}) \nabla^2 R(\mathbf{r}) + 2 \nabla E_0^{(+)}(\mathbf{r}) \cdot \nabla R(\mathbf{r}) \right) \Big\} e^{ik_0 R(\mathbf{r})}.\end{aligned} \tag{2.7}$$

Wstawmy teraz wyznaczony laplasjan (2.7) z powrotem do równania Helmholtza. Czynnik wykładniczy $e^{ik_0 R(\mathbf{r})}$ występuje w każdym wyrazie, więc ulega skróceniu, co prowadzi do równania:

$$\nabla^2 E_0^{(+)} - k_0^2 E_0^{(+)} (\nabla R)^2 + ik_0 \left(E_0^{(+)} \nabla^2 R + 2 \nabla E_0^{(+)} \cdot \nabla R \right) + k_0^2 n^2(\mathbf{r}) E_0^{(+)} = 0. \tag{2.8}$$

Dla przejrzystości podzielmy obie strony tego równania przez wyraz $k_0^2 E_0^{(+)}$ i pogrupujmy człony:

$$\left\{ n^2(\mathbf{r}) - [\nabla R(\mathbf{r})]^2 \right\} + \frac{\nabla^2 E_0^{(+)}(\mathbf{r})}{k_0^2 E_0^{(+)}(\mathbf{r})} + \frac{i}{k_0} \left(\nabla^2 R(\mathbf{r}) + \frac{2}{E_0^{(+)}(\mathbf{r})} \nabla E_0^{(+)}(\mathbf{r}) \cdot \nabla R(\mathbf{r}) \right) = 0. \tag{2.9}$$

Powyższe równanie musi być spełnione jednocześnie dla części rzeczywistej i urojonej. W tym miejscu wprowadzamy kluczowe założenie, które pozwala nam przejść od falowego opisu światła do **optyki geometrycznej**. Reżim ten obowiązuje, gdy długość fali świetlnej jest zaniedbywalnie mała w porównaniu z rozmiarami elementów układu optycznego i niejednorodnościami ośrodka, co matematycznie zapisujemy jako granicę $\lambda_0 \to 0$ (a zatem $k_0 \to \infty$). W tej granicy człony zawierające w mianowniku k_0 oraz k_0^2 dążą do zera:

$$\frac{1}{k_0} \to 0 \quad \text{oraz} \quad \frac{1}{k_0^2} \to 0. \tag{2.10}$$

Pozwala to na całkowite pominięcie drugiego i trzeciego wyrazu w równaniu (2.9). W rezultacie, z części rzeczywistej pozostaje jedynie bardzo proste i eleganckie równanie różniczkowe:

$$[\nabla R(\mathbf{r})]^2 = n^2(\mathbf{r}). \tag{2.11}$$

Z równania $[\nabla R(\mathbf{r})]^2 = n^2(\mathbf{r})$ wynika, że kwadrat długości wektora gradientu eikonału jest równy kwadratowi współczynnika załamania. Ponieważ długość wektora z definicji jest nieujemna, a współczynnik załamania $n(\mathbf{r}) > 0$, wnioskujemy, że wartość (długość) wektora gradientu eikonału w każdym punkcie wynosi:

$$|\nabla R(\mathbf{r})| = n(\mathbf{r}). \tag{2.12}$$

Z matematyki wiemy, że każdy wektor możemy zapisać jako iloczyn jego długości oraz wektora jednostkowego (wersora) wskazującego jego zwrot. Oznaczmy ten kierunkowy wersor jako $\mathbf{s}(\mathbf{r})$ (gdzie $|\mathbf{s}| = 1$). Pozwala to na wektorowy zapis powyższej zależności, co prowadzi do fundamentalnego dla całej optyki geometrycznej **równania eikonału**:

$$\boxed{\nabla R(\mathbf{r}) = n(\mathbf{r})\mathbf{s}(\mathbf{r}),} \tag{2.13}$$

gdzie funkcja skalarna $R(\mathbf{r})$ nosi nazwę **eikonału** (od greckiego słowa *eikōn* oznaczającego obraz lub podobiznę).

Równanie to niesie ze sobą głęboką fizyczną interpretację zachowania fali świetlnej, którą możemy rozbić na dwa aspekty:

1. Z definicji matematycznej wektor gradientu funkcji skalarnej (∇R) jest zawsze prostopadły do powierzchni jej stałej wartości. Ponieważ powierzchnie $R(\mathbf{r}) = \text{const}$ reprezentują czoła fali, wersor $\mathbf{s}(\mathbf{r})$ wyznacza lokalny kierunek prostopadły do czoła fali. Kierunek ten definiuje tor, wzdłuż którego transportowana jest energia świetlna.

2. Długość gradientu informuje nas, jak szybko narasta wartość funkcji. Równanie eikonału stwierdza, że faza fali zmienia się w przestrzeni tym gwałtowniej, im większy jest współczynnik załamania $n(\mathbf{r})$ w danym punkcie. Jest to w pełni zgodne z intuicją fizyczną: w ośrodku gęstszym

optycznie (większe n) światło spowalnia, a jego długość fali ulega skróceniu ($\lambda = \lambda_0/n$). W konsekwencji, na tym samym odcinku drogi fala "upakowuje"więcej cykli oscylacji, co oznacza szybszy przyrost fazy przestrzennej.

W ten sposób równanie eikonału stanowi elegancki i bezpośredni pomost pomiędzy fizyką falową (czoła fali) a optyką geometryczną. Te abstrakcyjne krzywe przestrzenne, do których wektor $\mathbf{s}$ jest w każdym punkcie styczny, nazywamy w fizyce **promieniami świetlnymi**, a badanie ich zachowania stanowi fundament optyki geometrycznej.

2.3 Zasada Fermata

Doświadczenie Propagacja światła w akwarium wypełnionym wodnym roztworem soli

Równanie eikonału (2.13) wyznacza drogę, po której rozchodzą się promienie świetlne w danym ośrodku, przy czym może to być ośrodek cechujący się niejednorodnym współczynnikiem załamania, $n(\mathbf{r}) \neq \text{const}$. Z fenomenologicznego punktu widzenia bieg promienia pomiędzy dwoma punktami może być także wyznaczony w oparciu o **zasadę Fermata**, w myśl której *promień świetlny porusza się pomiędzy dwoma punktami w przestrzeni w taki sposób, że czas przejścia pomiędzy tymi punktami jest minimalny.* Warto podkreślić, że w ośrodku może istnieć więcej niż jeden promień, który minimalizuje czas przejścia pomiędzy dwoma punktami.

Jak wspomniano wyżej, zasada Fermata jest prawem fenomenologicznym sformułowanym po raz pierwszy przez Pierra Fermata (1601-1665) w 1658 roku. Jak się jednak okazuje, może ona zostać wyprowadzona z równania eikonału. Aby to wykazać, obliczmy rotację z równania eikonału

$$\nabla \times [n(\mathbf{r})\mathbf{s}(\mathbf{r})] = \nabla \times [\nabla R(\mathbf{r})]. \tag{2.14}$$

Z własności operatorów różniczkowych wynika, że rotacja z gradientu jest równa zero. Dlatego też prawa strona równania (2.14) znika. Scałkujmy teraz lewą stronę równania (2.14) po powierzchni zamkniętej i skorzystajmy z twierdzenia Stokesa. W takim przypadku otrzymujemy

$$\int_A \nabla \times [n(\mathbf{r})\mathbf{s}(\mathbf{r})]da = \oint_C n(\mathbf{r})\mathbf{s}(\mathbf{r}) \cdot d\mathbf{l} = 0. \tag{2.15}$$

Ponieważ całka po powierzchni zamkniętej C jest równa zero, oznacza to, że nie zależy ona od drogi. Wybierzmy teraz w przestrzeni dwa punkty A i B i załóżmy, że droga łącząca je ze sobą przebiega wzdłuż krzywej, która w każdym punkcie jest styczna do wersora $\mathbf{s}(\mathbf{r})$. W takim przypadku iloczyn skalarny $\mathbf{s}(\mathbf{r})$ i $d\mathbf{l}$ można zapisać jako $\mathbf{s}(\mathbf{r}) \cdot d\mathbf{l} = dl$, gdyż kąt między tymi wektorami jest równy zero w każdym miejscu wzdłuż drogi. Dla każdej innej drogi kąt pomiędzy $\mathbf{s}(\mathbf{r})$ i $d\mathbf{l}$ jest niezerowy, co sprawia, że cosinus kąta między tymi wektorami jest mniejszy niż 1. Z drugiej jednak strony, na mocy naszej wcześniejszej dyskusji wiemy, że sama całka (2.15) przyjmuje tę samą wartość dla wszystkich dróg. Oznacza to, że iloczyn $n(\mathbf{r})dl$ wycałkowany wzdłuż innej drogi musi być większy

niż ten przebiegający wzdłuż $\mathbf{s}(\mathbf{r})$ (żeby zrekompensować mniejszą od jedności wartość cosinusa). Pozwala to zatem zapisać

$$\int_A^B n(\mathbf{r})\mathbf{s}(\mathbf{r}) \cdot d\mathbf{l} = \min\left\{\int_A^B n(\mathbf{r})dl\right\}. \tag{2.16}$$

Całka w argumencie funkcji minimum (min) nazywana jest **drogą optyczną**. Warto zwrócić uwagę, że droga optyczna różni się od drogi geometrycznej tym, że uwzględnia ona wolniejsze rozchodzenie się fali w ośrodku związane z większym od jedności współczynnikiem załamania. Oznacza to, że najkrótsza droga geometryczna nie musi być tożsama z najkrótszą drogą optyczną (dzieje się tak tylko w ośrodkach jednorodnych). Zauważmy na koniec, że drogę optyczną można w prosty sposób związać z czasem przejścia pomiędzy dwoma punktami $\Delta t|_A^B$. Przekształcając równanie

$$dl = v(\mathbf{r})dt = \frac{cdt}{n(\mathbf{r})}, \tag{2.17}$$

poprzez wymnożenie go przez współczynnik załamania $n(\mathbf{r})$ oraz scałkowanie od punktu A do punktu B otrzymujemy, że czas przejścia między tymi punktami dany jest zależnością

$$\Delta t|_A^B = \frac{\int_A^B n(\mathbf{r})dl}{c}. \tag{2.18}$$

Łącząc ze sobą równania (2.18) oraz (2.16) można stwierdzić, że promień świetlny porusza się po drodze, dla której minimalizowany jest czas przejścia, co jest treścią zasady Fermata.

2.4 Prawa optyki geometrycznej

Doświadczenie Demonstracja zachowania promieni przy użyciu tablicy z płaskimi elementami optycznymi

Jak wcześniej wspomniano z optyką geometryczną mamy do czynienia, w przypadku gdy długość fali światła, a co za tym idzie wszystkie zjawiska związane z jego falową naturą, mogą być zaniedbane. W takim przypadku prawa optyki redukują się do prostych praw geometrycznych, którym podlega promień świetlny. Co ciekawe, prawa optyki systematycznie badane były już w starożytności przez Empedoklesa (494-434 p.n.e.) i Euklidesa (365-295 p.n.e), zaś samo prawo załamania opisane zostało na początku XVI wieku przez holenderskiego uczonego Willebrorda Snella (1580-1626).

2.4.1 Prawo odbicia

Pierwszym prawem optyki geometrycznej jest **prawo odbicia**. Zgodnie z nim promień świetlny odbija się od granicy dwóch ośrodków w taki sposób, że kąt padania α_p równy jest kątowi odbicia α_o

$$\boxed{\alpha_p = \alpha_o}, \tag{2.19}$$

przy czym oba kąty liczone są względem normalnej do powierzchni, od której odbija się światło.

2.4.2 Prawo załamania

Drugim prawem optyki geometrycznej jest **prawo załamania**, znane również pod nazwą **prawa Snella**. Zgodnie z nim stosunek sinusa kąta padania α_p do sinusa kąta załamania α_z ma się tak, jak stosunek współczynników załamania ośrodka, do którego załamuje się światło do ośrodka, z którego światło się załamało

$$n_1 \sin \alpha_p = n_2 \sin \alpha_z, \tag{2.20}$$

gdzie n_1 i n_2 to odpowiednio współczynniki załamania ośrodków, z którego i do którego załamało się światło, zaś α_p i α_z, podobnie jak w przypadku prawa odbicia, liczone są względem normalnej do powierzchni.

2.4.3 Zjawisko całkowitego wewnętrznego odbicia

Z ciekawym zjawiskiem mamy do czynienia w przypadku wyjścia z ośrodka gęstszego optycznie, tj. ośrodka o większym współczynniku załamania, do ośrodka rzadszego optycznie, tj. do ośrodka o mniejszym współczynniku załamania. W takiej sytuacji istnieje pewien kąt, zwany **kątem granicznym** α_{gr}, dla którego światło załamane porusza się wzdłuż granicy ośrodków, $\alpha_z = \pi/2$. By spełnione było prawo Snella dla kątów padania większych od kąta granicznego, $\alpha_p > \alpha_{gr}$, sinus kąta załamania musiałby przyjmować wartość większą od jedności. Ponieważ nie jest to możliwe, zmienić się musi charakter oddziaływania światła. Konsekwencją tego jest odbicie od, a nie załamanie na granicy dwóch ośrodków. Zjawisko to nosi nazwę **zjawiska całkowitego wewnętrznego odbicia**.

Warto w tym miejscu zauważyć, że z punktu widzenia optyki falowej podczas całkowitego wewnętrznego odbicia rzeczywiście dochodzi do odbicia światła od powierzchni, ale sam proces i zachowanie światła na granicy dwóch ośrodków jest złożony. Przyjrzymy się temu zagadnieniu bliżej podczas omawiania praw rządzących zachowaniem światła w ramach optyki falowej.

2.4.4 Prawo odbicia i załamania a zasada Fermata

Zarówno prawo odbicia, jak i prawo załamania można wyprowadzić z zasady Fermata. W szczególności dokonując minimalizacji drogi optycznej przy odbiciu od zwierciadła płaskiego, można łatwo pokazać, że czas przejścia pomiędzy dwoma punktami jest minimalny, gdy kąt padania jest równy kątowi odbicia. Analogicznie można pokazać, że czas przejścia między dwoma punktami znajdującymi się w ośrodkach o różnych współczynnikach załamania jest najmniejszy wtedy, kiedy stosunek sinusów kątów padania i załamania spełnia relację daną prawem załamania.

2.5 Przybliżenie przyosiowe i macierze ABCD

Opis zachowania światła przy pomocy praw geometrycznych otwiera możliwość wykorzystania algebry do opisu propagacji promienia światła. Punktem wyjścia do tego typu opisu jest operowanie w ramach tzw. **przybliżenia przyosiowego**. Oznacza ono, że ograniczamy swoje rozważania jedynie do promieni prawie równoległych do **osi optycznej** układu. Operowanie w ramach przybliżenia przyosiowego ma konkretną konsekwencję. Pomiędzy kątem θ, pod którym promień porusza się względem osi optycznej, a jego funkcjami trygonometrycznymi zachodzi związek $\theta \approx \sin\theta \approx \mathrm{tg}\theta$. Na przybliżenie przyosiowe można zatem patrzeć jak na konsekwencję rozwinięcia funkcji trygonometrycznych w szereg Taylora i późniejsze uwzględnienie jedynie liniowych elementów rozwinięcia tych funkcji.

2.5.1 Konwencja znaków (Konwencja kartezjańska)

Zanim przejdziemy do algebraicznego opisu układów optycznych za pomocą macierzy, konieczne jest rygorystyczne zdefiniowanie reguł określających znaki poszczególnych parametrów geometrycznych, takich jak promienie krzywizn, kąty propagacji czy odległości. W optyce geometrycznej standardowo przyjmuje się tzw. **konwencję kartezjańską**. Opiera się ona na umieszczeniu wierzchołka elementu optycznego w początku układu współrzędnych i analizie propagacji względem osi optycznej (oś z) oraz osi poprzecznej (oś y).

Aby zapewnić pełną spójność wyprowadzanych w dalszej części równań (m.in. macierzy ABCD, wzoru szlifierzy soczewek czy równania soczewkowego), w całej książce będziemy stosować reguły zebrane w Tabeli 2.1.

Przyjęcie powyższej konwencji tłumaczy bezpośrednio reguły stosowane w kolejnych podrozdziałach – określa na przykład dlaczego po odbiciu od zwierciadła odległość przebyta przez promień przyjmuje wartość ujemną ($d \to -d$), a promienie krzywizn powierzchni wypukłych i wklęsłych mają przeciwne znaki.

2.5.2 Macierz ABCD dla pustej przestrzeni

Rysunek 2.1 przedstawia rozchodzenie się promienia świetlnego w ośrodku jednorodnym. Analizując go, można łatwo zauważyć, że znając takie parametry promienia jak jego wysokość nad osią optyczną y_1 oraz kąt, jaki z nią tworzy θ_1 w punkcie z_1 (wartość kąta θ mieści się w przedziale od $-\pi/2$ do $\pi/2$), jesteśmy w stanie określić parametry tego promienia w punkcie z_2 oddalonym wzdłuż osi optycznej o d

$$y_2 = y_1 + d\mathrm{tg}\theta_1 \approx y_1 + d\theta_1, \tag{2.21}$$

$$\theta_2 = \theta_1. \tag{2.22}$$

Równania te można w zgrabny sposób zapisać w postaci macierzowej

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}. \tag{2.23}$$

Tablica 2.1: Kartezjańska konwencja znaków stosowana w optyce geometrycznej i formalizmie macierzowym ABCD.

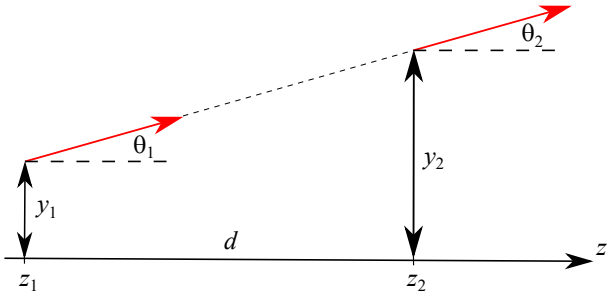
Wielkość fizyczna	**Znak dodatni (+)**	**Znak ujemny (-)**
Kierunek padania światła	Światło propaguje się początkowo z lewej strony w prawo, tj. wzdłuż dodatniej osi z.	
Wysokość promienia (y)	Promień znajduje się **powyżej** osi optycznej.	Promień znajduje się **poniżej** osi optycznej.
Kąt propagacji promienia (θ)	Promień odchyla się w górę (rośnie wartość y wzdłuż osi z).	Promień odchyla się w dół (maleje wartość y wzdłuż osi z).
Promień krzywizny (R)	Środek krzywizny powierzchni (C) leży **po prawej** stronie wierzchołka (powierzchnia wypukła dla padającego światła).	Środek krzywizny powierzchni (C) leży **po lewej** stronie wierzchołka (powierzchnia wklęsła dla padającego światła).
Odległość wzdłuż osi (d, a, b)	Mierzona w prawo (zgodnie z początkowym kierunkiem światła).	Mierzona w lewo (pod prąd, np. droga przebyta po odbiciu od zwierciadła).
Ogniskowa (f)	Układ **skupiający** (np. soczewka dwuwypukła, zwierciadło wklęsłe).	Układ **rozpraszający** (np. soczewka dwuwklęsła, zwierciadło wypukłe).

Macierz transformacji promienia przy propagacji w ośrodku jednorodnym można zapisać jako

$$M_p = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix}. \tag{2.24}$$

Warto podkreślić, że dla propagacji w kierunku przeciwnym do osi optycznej odległość d przyjmuje wartość ujemną, tj. $d \to -d$.

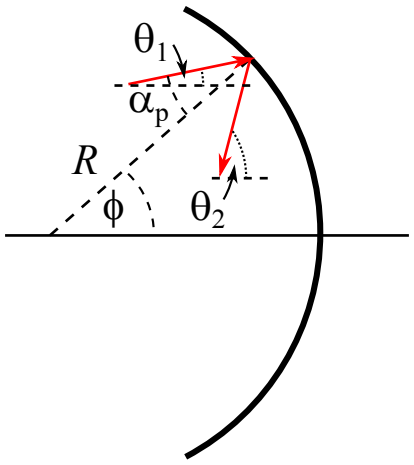
Zapis propagacji światła w postaci macierzy nosi nazwę **formalizmu macierzowego ABCD** i pozwala on w bardzo prosty sposób wyznaczać propagację światła w układach optycznych. By było to jednak możliwe, musimy do opisu macierzowego wprowadzić prawo odbicia i załamania.



Rysunek 2.1: Rozchodzenie się promienia świetlnego w ośrodku jednorodnym. Znajomość wysokości y_1 promienia nad osią optyczną w punkcie z_1 oraz kąta θ_1 w tym punkcie pozwala na obliczenie wysokości y_2 i kąta θ_2 w punkcie z_2 oddalonym od punktu z_1 o d.

2.5.3 Macierz ABCD dla odbicia

Rozważmy najpierw odbicie od powierzchni sferycznej. Należy podkreślić, że o ile na rys. 2.2 przedstawiono jako przykład odbicie od powierzchni wklęsłej, prowadzone rozważania są ogólne i obowiązują również dla powierzchni wypukłej.



Rysunek 2.2: Odbicie od powierzchni sferycznej.

Przy odbiciu od powierzchni nie zmienia się pozycja (wysokość) promienia świetlnego, tzn. $y_2 = y_1$. Zastanówmy się teraz nad transformacją kątów. Z zależności geometrycznych widać, że kąt propagacji światła po odbiciu jest równy

$$\theta_2 = -(\theta_1 + \alpha_p + \alpha_o) = -(\theta_1 + 2\alpha_p), \tag{2.25}$$

gdzie skorzystaliśmy z prawa odbicia, $\alpha_p = \alpha_o$. Warto również zwrócić uwagę,

że w równaniu pojawił się znak minus. Jest on skutkiem przyjętej przez nas konwencji, w myśl której zmiana kierunku propagacji światła na przeciwny powoduje zmianę znaku (w tym przypadku, zmianę znaku kąta).

Zauważmy teraz, że kąt padania może być wyrażony poprzez kąt ϕ, jaki odcinek łączący środek krzywizny zwierciadła z punktem padania promienia tworzy z osią optyczną:

$$\alpha_p = \phi - \theta_1. \tag{2.26}$$

Podstawiając równanie (2.26) do równania (2.25), otrzymujemy:

$$\theta_2 = \theta_1 - 2\phi. \tag{2.27}$$

Wykorzystując teraz przybliżenie przyosiowe, kąt ϕ można powiązać z wysokością padania promienia y_1 oraz promieniem krzywizny R. Pamiętając o naszej kartezjańskiej konwencji znaków, środek krzywizny zwierciadła wklęsłego leży w ujemnej części osi z, zatem jego promień R przyjmuje wartość ujemną ($R < 0$). Geometrycznie kąt ϕ wynosi $y_1/|R|$, co algebraicznie pozwala nam zapisać $\phi \approx -y_1/R$. Podstawiając to do powyższego równania, otrzymujemy:

$$\theta_2 \approx \theta_1 - 2\left(-\frac{y_1}{R}\right) = \theta_1 + 2\frac{y_1}{R}. \tag{2.28}$$

Powyższe równanie sprzężone z warunkiem braku zmiany wysokości ($y_2 = y_1$) może być zapisane w zwartej postaci macierzowej:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{2}{R} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}, \tag{2.29}$$

co pozwala na bezpośrednie wyznaczenie macierzy ABCD opisującej odbicie od powierzchni sferycznej:

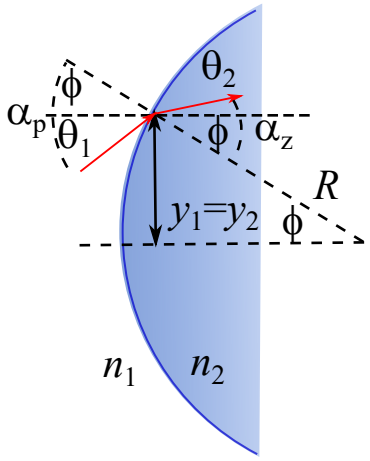
$$\boxed{M_o = \begin{bmatrix} 1 & 0 \\ \frac{2}{R} & 1 \end{bmatrix}}. \tag{2.30}$$

Zwróćmy uwagę na wewnętrzną spójność z naszą tabelą znaków: dla zwierciadła wklęsłego (skupiającego) promień krzywizny R jest ujemny. Wówczas element $C = 2/R$ staje się ujemny, co oznacza, że promień biegnący początkowo równolegle do osi optycznej ($y_1 > 0$, $\theta_1 = 0$) zostanie odchylony w dół ($\theta_2 < 0$), czyli poprawnie zogniskowany w jej stronę.

Należy podkreślić, że macierz (2.30) jest macierzą uniwersalną dla każdej sferycznej powierzchni odbijającej. Aby opisać odbicie od powierzchni wypukłej, zgodnie z Tabelą 2.1 wystarczy podstawić wartość promienia $R > 0$. Na koniec zauważmy, że dla odbicia od zwierciadła płaskiego promień krzywizny dąży do nieskończoności ($R \to \infty$), przez co element macierzowy C znika, pozostawiając macierz jednostkową.

2.5.4 Macierz ABCD dla załamania

Analogiczne rozważania do tych przedstawionych powyżej można przeprowadzić dla załamania promienia na granicy dwóch ośrodków o współczynnikach załamania n_1 i n_2. Podobnie jak w poprzednim podrozdziale, choć na rys. 2.3 przedstawiony został przypadek załamania na powierzchni wypukłej, to poniższe rozważania są ogólne.



Rysunek 2.3: Załamanie na granicy sferycznej dwóch ośrodków o współczynnikach załamania n_1 i n_2.

Podobnie jak w przypadku odbicia, również w przypadku załamania pozycja promienia (wysokość) bezpośrednio na granicy ośrodków nie ulega zmianie, tj. $y_2 = y_1$. Dodatkowo, korzystając z relacji geometrycznych, możemy związać kąty θ_1 i θ_2 z kątami padania α_p, załamania α_z oraz kątem ϕ, jaki odcinek łączący środek krzywizny powierzchni z punktem padania tworzy z osią optyczną:

$$\begin{aligned} \alpha_p &= \phi + \theta_1, \\ \alpha_z &= \phi + \theta_2. \end{aligned} \tag{2.31}$$

Należy w tym miejscu zauważyć, że w przeciwieństwie do odbicia, przy załamaniu promień światła wciąż propaguje się w dodatnią stronę osi optycznej i w konsekwencji znak θ_2 (wynikający z konwencji opisu kierunku) nie ulega odwróceniu.

Możemy teraz skorzystać z prawa Snella w przybliżeniu przyosiowym

$$n_1 \alpha_p \approx n_2 \alpha_z, \tag{2.32}$$

które po podstawieniu do układu równań (2.31) i przekształceniu daje

$$\theta_2 = \frac{n_1 - n_2}{n_2}\phi + \frac{n_1}{n_2}\theta_1. \tag{2.33}$$

Po podstawieniu zależności trygonometrycznych w przybliżeniu małych kątów, $\phi \approx \sin\phi = y_1/R$, otrzymujemy

$$\theta_2 = \frac{n_1 - n_2}{n_2}\frac{y_1}{R} + \frac{n_1}{n_2}\theta_1. \tag{2.34}$$

Równanie (2.34) pozwala ostatecznie zapisać równanie dla załamania na powierzchni wypukłej w formie macierzowej

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2 R} & \frac{n_1}{n_2} \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}. \tag{2.35}$$

Macierz załamania dana jest zatem zależnością

$$M_z = \begin{bmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2 R} & \frac{n_1}{n_2} \end{bmatrix}. \tag{2.36}$$

Zaznaczmy na koniec, że podobnie jak wcześniej obowiązuje konwencja dotycząca znaku promieni krzywizny.

2.5.5 Propagacja promienia w przestrzeni a macierze ABCD

Jedną z podstawowych zalet podejścia macierzowego jest możliwość opisu propagacji światła w złożonych układach optycznych poprzez mnożenie ze sobą macierzy odpowiadających za zachowanie promienia czy to w pustej przestrzeni, czy przy odbiciu od lub załamaniu na granicy dwóch ośrodków

$$\begin{aligned} M_{12\ldots n} = \begin{bmatrix} A_{12\ldots n} & B_{12\ldots n} \\ C_{12\ldots n} & D_{12\ldots n} \end{bmatrix} &= M_n \ldots M_2 M_1 \\ &= \begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} \ldots \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}. \end{aligned} \tag{2.37}$$

2.5.6 Właściwości macierzy ABCD

Jak się okazuje, elementy macierzy ABCD nie są dowolne, ale wykazują bardzo użyteczne właściwości matematyczne. Łatwo pokazać w oparciu o macierze propagacji w pustej przestrzeni (2.24) oraz odbicia (2.30), że wyznacznik tych dwóch macierzy jest dokładnie równy jeden. Inaczej zachowuje się jednak macierz załamania (2.36) na granicy dwóch ośrodków, dla której wyznacznik wynosi:

$$\det M_z = \frac{n_1}{n_2}. \tag{2.38}$$

Z własności algebry wiemy, że wyznacznik iloczynu macierzy jest równy iloczynowi ich wyznaczników. Jeśli rozważymy promień przechodzący przez układ optyczny składający się z wielu granic ośrodków (np. $n_1 \to n_2 \to n_3$), to wewnętrzne współczynniki w iloczynie wyznaczników ulegną skróceniu (np.

$\frac{n_1}{n_2}\frac{n_2}{n_3} = \frac{n_1}{n_3}$). Oznacza to, że dla dowolnego złożonego układu optycznego, w którym promień zaczyna swój bieg w ośrodku o współczynniku załamania n_p, a kończy w ośrodku o współczynniku n_k, wyznacznik wypadkowej macierzy M wynosi:

$$\det M = \frac{n_p}{n_k}. \tag{2.39}$$

W najczęściej spotykanym w praktyce przypadku, wtedy kiedy promień zaczyna i kończy swój bieg w tym samym ośrodku (np. światło przechodzące przez szklaną soczewkę lub złożony obiektyw umieszczony w całości w powietrzu), stosunek ten wynosi 1. Wtedy i tylko wtedy wyznacznik macierzy całego układu optycznego jest jednostkowy:

$$\det M = 1. \tag{2.40}$$

Stanowi to potężne narzędzie diagnostyczne – obliczając macierz ABCD skomplikowanego układu optycznego znajdującego się w powietrzu, zawsze warto na koniec sprawdzić, czy $AD - BC = 1$, co pozwala szybko wykryć ewentualne błędy rachunkowe.

2.6 Odbicie od zwierciadła

2.6.1 Równanie soczewkowe dla zwierciadła

Rozważmy przypadek, w którym przyosiowy promień świetlny, opisany położeniem y_1 i kątem θ_1, propaguje się w przestrzeni na odległość a , po czym odbija się od lustra sferycznego i rozchodzi się w przeciwnym kierunku na odległość $-b$ (zgodnie z przyjętą przez nas konwencją kartezjańską). Macierz opisującą taki przypadek można zapisać w postaci:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} 1 & -b \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{2}{R} & 1 \end{bmatrix} \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 - \frac{2b}{R} & a - b - \frac{2ab}{R} \\ \frac{2}{R} & 1 + \frac{2a}{R} \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}. \tag{2.41}$$

Załóżmy teraz, że punkt y_1, znajdujący się w odległości a, jest przedmiotem, który chcemy zobrazować. Aby powstał ostry obraz tego punktu w płaszczyźnie odległej o $-b$, wszystkie promienie opuszczające ten punkt (tj. poruszające się pod różnymi, małymi kątami θ_1) muszą przeciąć się w tym samym miejscu w przestrzeni. Oznacza to, że dla ogólnego układu opisanego macierzą ABCD:

$$\begin{bmatrix} y_2 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}, \tag{2.42}$$

położenie obrazu y_2 nie może zależeć od początkowego kąta emisji θ_1. Warunek ten matematycznie sprowadza się do tego, że wyraz B macierzy układu musi być tożsamościowo równy zero ($B = 0$). Wykorzystując ten warunek dla wyprowadzonej wcześniej macierzy (2.41), otrzymujemy:

$$a - b - \frac{2ab}{R} = 0, \tag{2.43}$$

skąd po podzieleniu obustronnie przez ab i przeniesieniu wyrazów otrzymujemy rygorystyczną relację między położeniem przedmiotu i obrazu:

$$\frac{1}{a} - \frac{1}{b} = -\frac{2}{R}. \tag{2.44}$$

Załóżmy teraz, że przedmiot umieszczony jest w nieskończoności ($a \to \infty$). Zgodnie z równaniem (2.44), obraz powstanie w odległości $b = R/2$. Pamiętając, że w macierzy wejściowej założyliśmy propagację wsteczną na odległość $-b$, fizyczna współrzędna punktu przecięcia osi wynosi $-R/2$. Jest to miejsce charakterystyczne, gdyż wszystkie promienie propagujące się początkowo równolegle do osi optycznej skupiać się będą w tym konkretnym punkcie leżącym na osi. Z uwagi na tę cechę punkt ten nazywany jest **ogniskiem** F, zaś odległość zdefiniowana jako $f = -R/2$ nazywana jest **ogniskową** f. Podstawiając ogniskową do równania (2.44), otrzymujemy spójną z naszą konwencją postać:

$$\boxed{\frac{1}{a} - \frac{1}{b} = \frac{1}{f}.} \tag{2.45}$$

Wielkość $1/f$ nazywamy często **zdolnością skupiającą** elementu lub układu optycznego (patrz rozdział 2.13) i mierzymy ją w dioptriach (zakładając, że ogniskowa wyrażona jest w metrach). Równanie to, rygorystycznie zachowujące ujęcie kartezjańskie, w którym odpowiednie współrzędne w lewej i prawej półprzestrzeni mają przypisane znaki, nazywane jest **równaniem soczewkowym** i obowiązuje dla wszystkich układów optycznych.

Co ciekawe, nie tylko promienie równoległe do osi optycznej ogniskują się w jednym punkcie. W ogólności wszystkie wzajemnie równoległe promienie padające na zwierciadło pod dowolnym, małym kątem skupiać się będą w jednym miejscu, choć punkt ten nie będzie już leżał bezpośrednio na osi optycznej. W ramach przybliżenia przyosiowego promienie równoległe padające ze wszystkich dozwolonych kierunków tworzą w miejscu przecięcia płaszczyznę. Płaszczyzna ta, zawierająca ognisko i prostopadła do osi optycznej układu, zwana jest **płaszczyzną ogniskową**.

Warto na koniec zauważyć, że w wybranym przez nas formalizmie element C, pojawiający się w macierzy ABCD dla zwierciadła, wiąże się z ogniskową następującą zależnością:

$$f = -\frac{1}{C}. \tag{2.46}$$

Zależność ta może być z powodzeniem wykorzystywana do szybkiego obliczania ogniskowej zarówno pojedynczych elementów, jak i całych, złożonych układów optycznych.

2.6.2 Powiększenie układu optycznego

Ważną cechą układu optycznego jest jego **powiększenie poprzeczne** $\mathcal{M}$, które definiuje się jako stosunek rozmiaru (wysokości) obrazu do rozmiaru przedmiotu:

$$\mathcal{M} \equiv \frac{y_2}{y_1}. \tag{2.47}$$

Warto w tym miejscu zwrócić uwagę na to, że $\mathcal{M}$ może przyjmować wszystkie wartości rzeczywiste. Gdy $\mathcal{M} > 0$, oznacza to, że wysokości mają ten sam znak, a zatem obraz jest prosty (nieodwrócony). Z kolei dla $\mathcal{M} < 0$ obraz jest odwrócony względem osi optycznej. Dodatkowo, gdy $|\mathcal{M}| > 1$ obraz jest powiększony, podczas gdy dla $|\mathcal{M}| < 1$ jest pomniejszony. W konwencji odległości bezwzględnych powiększenie to wyraża się często relacją $\mathcal{M} = -b/a$.

2.7 Soczewki

2.7.1 Propagacja światła w soczewce

Szczególnym przypadkiem układu, dla którego możemy wykorzystać opis macierzowy, jest **soczewka**. Soczewka to element optyczny wykonany z materiału dielektrycznego (najczęściej szkła), który załamuje promień światła na co najmniej jednej powierzchni, a samo światło propaguje się również wewnątrz tego materiału.

Jako przykład rozważmy **soczewkę dwuwypukłą** o grubości Δ (liczonej wzdłuż osi optycznej), której dwie sferyczne powierzchnie łamiące mają promienie krzywizn R_1 oraz R_2. Bezwzględnie obowiązuje nas tutaj rygorystyczna konwencja kartezjańska: znak promienia krzywizny zależy od tego, po której stronie wierzchołka leży środek krzywizny powierzchni. Stanowi to klasyczną pułapkę koncepcyjną – dla soczewki dwuwypukłej pierwsza powierzchnia ma środek krzywizny po prawej stronie, zatem $R_1 > 0$. Jednak środek krzywizny drugiej (zamykającej) powierzchni leży po lewej stronie, w związku z czym musi ona przyjąć ujemny promień krzywizny ($R_2 < 0$).

Korzystając z równania (2.37), możemy zapisać propagację światła w takiej grubej soczewce jako:

$$\begin{aligned} M_s &= \begin{bmatrix} 1 & 0 \\ \frac{n_2-n_1}{n_1 R_2} & \frac{n_2}{n_1} \end{bmatrix} \left(\begin{bmatrix} 1 & \Delta \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{n_1-n_2}{n_2 R_1} & \frac{n_1}{n_2} \end{bmatrix} \right) \\ &= \begin{bmatrix} 1 - \frac{(n_2-n_1)\Delta}{n_2 R_1} & \frac{n_1}{n_2}\Delta \\ -\left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2} + \frac{n_2-n_1}{n_2}\frac{\Delta}{R_1 R_2}\right) & 1 + \frac{(n_2-n_1)\Delta}{n_2 R_2} \end{bmatrix}, \end{aligned} \tag{2.48}$$

gdzie nawias okrągły w pierwszej linii określa kolejność wykonywania działań (mnożenie macierzy nie jest przemienne). Jak widać, macierz charakteryzująca przejście przez taką soczewkę jest dość skomplikowana, ale precyzyjnie uwzględnia ona propagację światła wewnątrz jej struktury.

2.7.2 Równanie soczewkowe dla soczewki

Analogicznie jak w przypadku zwierciadła, analizując padanie na układ promieni równoległych do osi optycznej, możemy wyznaczyć ogniskową soczewki. Skorzystamy z wyprowadzonego już związku pomiędzy ogniskową f a elementem C macierzy ABCD:

$$\frac{1}{f} = \left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2} + \frac{n_2 - n_1}{n_2}\frac{\Delta}{R_1 R_2}\right). \tag{2.49}$$

Powyższe równanie nosi nazwę **wzoru szlifierzy soczewek**. Wiąże ono geometryczne promienie krzywizn powierzchni soczewki oraz współczynnik załamania jej materiału z jej podstawową właściwością optyczną, czyli ogniskową.

2.7.3 Soczewka cienka i wzór szlifierzy

Rozpatrzmy najpierw uproszczony przypadek, w którym grubość soczewki jest pomijalnie mała, tzn. $\Delta \to 0$. W takiej sytuacji wypadkowa macierz przejścia M_s przyjmuje znacznie prostszą postać:

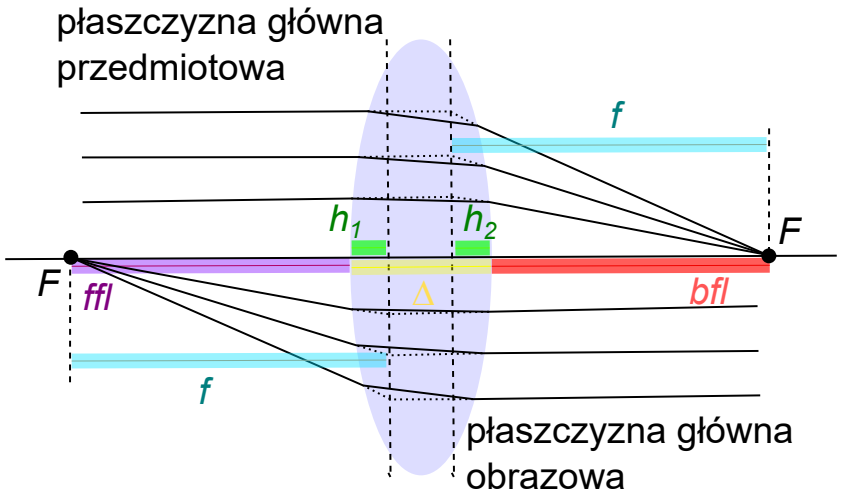
$$M_s = \begin{bmatrix} 1 & 0 \\ -\left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) & 1 \end{bmatrix}. \tag{2.50}$$

Ponieważ element C macierzy dla soczewki wynosi $-1/f$, wzór szlifierzy (zakładając bezwzględną zgodność konwencji kartezjańskiej dla obu promieni krzywizn) można zapisać w jego klasycznej i fundamentalnej postaci:

$$\frac{1}{f} = \left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right). \tag{2.51}$$

2.7.4 Płaszczyzny główne

Rozpatrzmy teraz przypadek, w którym grubość soczewki nie może być zaniedbana. Na rys. 2.4 pokazano zachowanie promieni równoległych padających i wychodzących z takiej soczewki. Dla soczewki grubej zarówno na pierwszej, jak i na drugiej powierzchni łamiącej dochodzi do zmiany kierunku propagacji światła. Przedłużmy teraz linie biegu promieni: padającego oraz wychodzącego z soczewki, w taki sposób, aby przecięły się one wewnątrz soczewki. Co ciekawe, wszystkie promienie równoległe do osi optycznej padające na układ przetną się w tej samej płaszczyźnie, określając miejsce efektywnego załamania. W pełni analogiczne rozważania można przeprowadzić dla promieni, które po wyjściu z soczewki poruszają się równolegle do osi optycznej. Również ich przedłużenia przetną się, tworząc płaszczyznę. Obie te płaszczyzny nazywamy **płaszczyznami głównymi**, przy czym ta związana z promieniami wychodzącymi



Rysunek 2.4: Zachowanie promieni w soczewce grubej. Równoległe promienie padające na soczewkę uginają się dwukrotnie na płaszczyźnie wejściowej i wyjściowej. Przedłużenia tych promieni (oraz promieni załamanych skupiających się w ognisku) tworzą płaszczyznę główną obrazową. Podobnie, promienie równoległe wychodzące z soczewki również uginają się dwukrotnie, a przedłużenia promieni padających z ogniska i ugiętych równolegle tworzą płaszczyznę główną przedmiotową. Na wykresie zaznaczono również odległość obu płaszczyzn od wierzchołków powierzchni soczewki (odpowiednio h_1 i h_2), grubość soczewki liczoną wzdłuż osi Δ oraz odległości ognisk od wierzchołków soczewki, odpowiednio FFL i BFL. Oprócz tych wielkości zaznaczono także efektywną ogniskową f oraz ognisko F.

równolegle jest **płaszczyzną główną przedmiotową**, zaś ta związana z promieniami wpadającymi równolegle nosi nazwę **płaszczyzny głównej obrazowej**. Punkty przecięcia tych płaszczyzn z osią optyczną nazywamy **punktami głównymi** (H_1, H_2). To właśnie względem nich (a nie względem fizycznych powierzchni soczewki) odmierzane są efektywne ogniskowe f.

Wyznaczmy teraz algebraicznie położenie płaszczyzn głównych. Rozważmy propagację światła od pierwszej płaszczyzny głównej do powierzchni soczewki (przesunięcie o odległość h_1), następnie transformację w samej soczewce, którą dla zachowania ogólności opiszemy generyczną macierzą ABCD, i w końcu propagację od tylnej powierzchni soczewki do drugiej płaszczyzny głównej (przesunięcie o h_2):

$$M_{obr} = \begin{bmatrix} 1 & h_2 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} 1 & h_1 \\ 0 & 1 \end{bmatrix} \right) = \begin{bmatrix} A + Ch_2 & h_1 A + B + h_1 h_2 C + h_2 D \\ C & Ch_1 + D \end{bmatrix}. \tag{2.52}$$

Załóżmy teraz, że układ taki chcemy zastąpić soczewką cienką (umieszczoną w płaszczyznach głównych), której macierz ABCD z definicji na przekątnej posiada

jedynki, co dane jest równaniem (2.50). Z porównania obu macierzy możemy zapisać równości:

$$\begin{aligned} A + Ch_2 &= 1, \\ Ch_1 + D &= 1. \end{aligned} \tag{2.53}$$

Pozwala to wyznaczyć zależności pomiędzy odległościami płaszczyzn głównych od powierzchni granicznych układu optycznego :

$$\begin{aligned} h_1 &= \frac{1-D}{C}, \\ h_2 &= \frac{1-A}{C}. \end{aligned} \tag{2.54}$$

Zastosujmy teraz zależności (2.54) do obliczenia pozycji płaszczyzn głównych soczewki grubej. Podstawiając wartości A, B, C i D z wyprowadzonej wcześniej macierzy (2.48), otrzymujemy:

$$\boxed{\begin{aligned} h_1 &= -\frac{n_1 R_1 \Delta}{n_2(R_1 - R_2 - \Delta) + n_1 \Delta}, \\ h_2 &= \frac{n_1 R_2 \Delta}{n_2(R_1 - R_2 - \Delta) + n_1 \Delta}. \end{aligned}} \tag{2.55}$$

Warto w tym miejscu zwrócić uwagę, że płaszczyzny główne pełnią rolę efektywnych płaszczyzn, względem których liczona jest efektywna ogniskowa soczewki f. Ma to daleko idące konsekwencje. Okazuje się bowiem, że o ile ogniskowe f liczone względem płaszczyzn głównych są stałe, o tyle odległości ognisk od samych wierzchołków fizycznych powierzchni soczewki – odpowiednio **przednia ogniskowa** FFL (ang. *Front Focal Length*) oraz **tylna ogniskowa** BFL (ang. *Back Focal Length*) – mogą przyjmować zupełnie inne wartości:

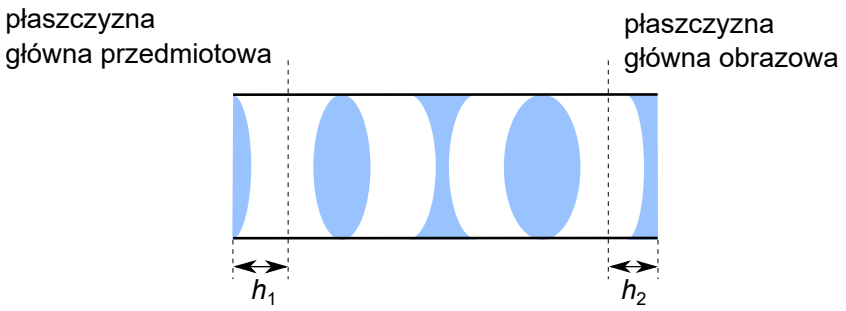
$$\begin{aligned} \text{FFL} &= f + h_1, \\ \text{BFL} &= f + h_2. \end{aligned} \tag{2.56}$$

(Uwaga: dla standardowej soczewki dwuwypukłej wartości h_1 i h_2 są z reguły algebraicznie ujemne, gdyż płaszczyzny leżą wewnątrz soczewki. Dodanie tych ujemnych przesunięć poprawnie skraca ogniskowe wierzchołkowe względem ogniskowej efektywnej).

W przypadku skomplikowanych soczewek grubych położenia płaszczyzn głównych nie tylko nie muszą być symetryczne względem geometrycznego środka soczewki, ale mogą nawet znajdować się całkowicie poza samą soczewką (np. w soczewkach meniskowych).

O ile powyższe rozważania przeprowadzone zostały dla pojedynczej soczewki grubej, o tyle już sama generyczna postać macierzy ABCD, którą wykorzystaliśmy do ich wyprowadzenia, sugeruje, że koncepcję tę można uogólnić. I

rzeczywiście, zależności (2.54) wyprowadzone w powyższym rozdziale są bezwzględnie prawdziwe dla dowolnego złożonego układu optycznego i mogą być wykorzystywane do zastąpienia wielosoczewkowego obiektywu pojedynczą macierzą (reprezentatywny przykład takiego układu pokazano na rys. 2.5).



Rysunek 2.5: Przykładowy układ optyczny składający się z szeregu elementów (soczewek) wraz z zaznaczonymi płaszczyznami głównymi całego, wypadkowego układu.

Punkty węzłowe

Do pełnego opisu układu optycznego, oprócz ognisk i punktów głównych, wprowadza się parę **punktów węzłowych** (N_1, N_2). Definiuje się je jako taką parę punktów sprzężonych na osi optycznej, dla której promień wchodzący do układu i skierowany do pierwszego punktu węzłowego (N_1) wychodzi z układu z drugiego punktu węzłowego (N_2) dokładnie równolegle do kierunku pierwotnego. Oznacza to, że dla punktów węzłowych powiększenie kątowe wynosi dokładnie jedność $(\gamma = 1)$.

W szczególnym (i najczęstszym) przypadku, gdy ośrodek przed i za soczewką lub układem optycznym jest identyczny (np. układ zanurzony jest z obu stron w powietrzu, $n_1 = n_2$), punkty węzłowe idealnie pokrywają się z punktami głównymi ($N_1 = H_1$ oraz $N_2 = H_2$).

Rozróżnienie to staje się jednak kluczowe w przypadku **ludzkiego oka**, które wewnątrz wypełnione jest cieczą (ciałem szklistym) o współczynniku załamania wyraźnie różnym od powietrza ($n_1 \neq n_2$). W oku punkty węzłowe są znacząco przesunięte względem punktów głównych. To właśnie przez punkt węzłowy oka przechodzą promienie, które nie zmieniają swojego kierunku, co bezpośrednio determinuje kątowy rozmiar obrazu powstającego na siatkówce.Poniżej załączam również kod TeX-a trzeciego rozdziału.

2.7.5 Równanie soczewkowe Newtona

Ciekawą alternatywą do klasycznego opisu powstawania obrazu jest **równanie soczewkowe Newtona**. Zgodnie z jego treścią, odległość przedmiotu

(oznaczmy ją jako x_p) i odległość obrazu (x_o) mierzone nie od płaszczyzn głównych, lecz **od ognisk układu**, wiążą się z ogniskową zależnością:

$$x_p x_o = f^2. \tag{2.57}$$

2.7.6 Soczewki gradientowe

Analizując zależność (2.49) można zauważyć, że równanie szlifierzy soczewek obowiązuje tylko dla materiałów wykonanych z jednorodnego materiału. Jednak już zasada Fermata podpowiada nam, że zmienność przestrzenna współczynnika załamania sprawia, że promienie nie poruszają się w ogólności po liniach prostych, a jedynie po trajektoriach odpowiadających najkrótszemu czasowi przejścia pomiędzy dwoma punktami.

Szczególnym przykładem elementu wykorzystującego tę właśnie właściwość do zmiany kierunku rozchodzenia światła są **soczewki gradientowe**. Soczewki te cechują się tym, że współczynnik załamania zmienia się wraz z odległością od osi soczewki. Co ciekawe, soczewka taka w ogóle nie musi mieć kształtu obłego i może być np. płytką płaskorównoległą. Z uwagi na modulację współczynnika załamania, droga optyczna pokonywana przez promień w środku soczewki jest znacząco dłuższa od tej, jaką pokonuje promień bliżej jej krawędzi.

Współczynnik załamania światła w soczewce gradientowej musi być opisany zależnością kwadratową, aby ogniskowanie wiązki odpowiadało temu w jednorodnej soczewce sferycznej:

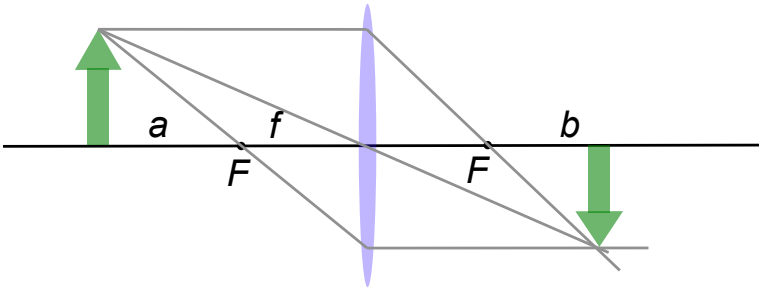
$$n = n_{max} - \frac{r^2}{2f\Delta_r}, \tag{2.58}$$

gdzie Δ_r jest grubością soczewki zaś n_{max} największą wartością współczynnika załamania.

Szczególnym przykładem zastosowania soczewek gradientowych jest telekomunikacja światłowodowa. Dzieje się tak dlatego, że cylindryczna geometria soczewki pozwala na jej stosunkowo łatwe przymocowanie do włókna światłowodowego, zaś odpowiedni dobór długości sprawia, że może ona albo skupiać wiązkę do punktu albo kolimować wiązkę rozbieżną wychodzącą z włókna. Możliwe są też przypadki, w których wiązka rozbieżna jest z powrotem skupiana.

2.8 Diagramy promieni

Alternatywą do podejścia macierzowego, która pozwala na znalezienie obrazu powstałego przy pomocy układu optycznego, jest metoda **diagramów promieni**. W jej przypadku obraz wyznaczany jest w oparciu o istnienie pewnych charakterystycznych promieni i graficzne analizy ich zachowania. Przykładowo, fakt przechodzenia przez ognisko obrazowe promieni optycznych padających na soczewkę równolegle do osi optycznej jest pierwszym krokiem pozwalającym na



Rysunek 2.6: Diagram promieni.

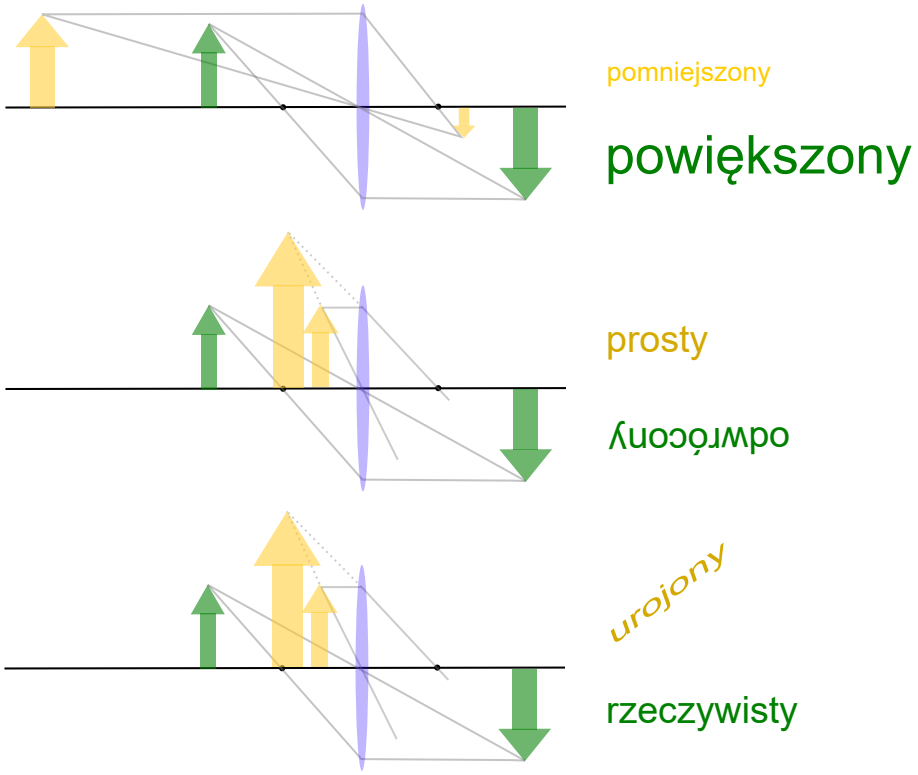
tworzenie obrazów w metodzie diagramów promieni. Drugim charakterystycznym promieniem, które umożliwia znalezienie obrazu jest promień, który po wyjściu z przedmiotu przechodzi przez ognisko przedmiotowe, a następne biegnie równolegle do osi optycznej. Trzecim promieniem charakterystycznym jest promień padający na element optyczny (zwierciadło lub soczewkę) w miejscu jego przecięcia z osią optyczną. W takim przypadku promień albo przechodzi przez element bez zmiany kierunku (soczewka), albo odbija się symetrycznie pod tym samym kątem (zwierciadło), ale zmienia kierunek propagacji. Przecięcie co najmniej dwóch z trzech promieni charakterystycznych wyznacza położenie obrazu. Przykład zastosowania diagramu promieni przedstawiony jest na rys. 2.6.

Warto w tym miejscu jednak zwrócić uwagę, że metoda diagramów promieni, o ile intuicyjna i szybka w stosowaniu, o tyle jest ona jedynie metodą przybliżoną, podczas gdy metoda macierzowa pozwala na osiągnięcie dokładniejszych wyników.

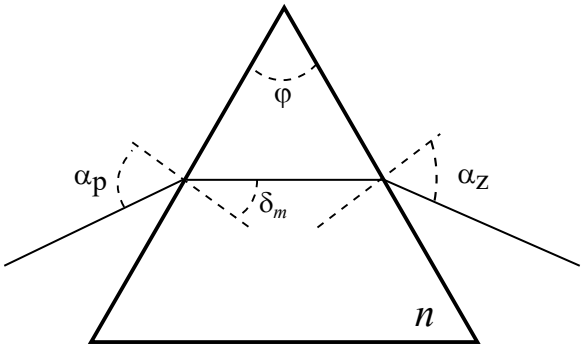
2.9 Rodzaje obrazów optycznych

Przeanalizujmy teraz cechy obrazów powstających w układach optycznych. Jak wspomniano wyżej, obraz taki może być rzeczywisty albo pozorny. W **obrazie rzeczywistym** mamy do czynienia z sytuacją, w której obraz powstaje na przecięciu rzeczywistych promieni, które przeszły przez układ optyczny. **Obraz pozorny** natomiast powstaje wtedy, gdy promienie przechodzące przez układ się ze sobą nie przecinają. Z taką sytuacją mamy np. do czynienia w przypadku zwierciadeł wypukłych. Korzystając z diagramów promieni, możemy jednak wyznaczyć obraz poprzez przecięcie przedłużeń promieni.

Innymi właściwościami obrazu jest jego orientacja przestrzenna. W szczególności obraz przedmiotu może być **prosty** wtedy, kiedy zorientowany jest on w tę samą stronę co przedmiot lub **obrócony**, kiedy orientacja ta jest przeciwna.



Rysunek 2.7: Rożne typy obrazów powstające w układach optycznych.



Rysunek 2.8: Minimal kąt ugięcia promieniowania w pryzmacie równoramiennym.

2.10 Inne elementy optyczne

2.10.1 Pryzmaty

Pryzmat jest elementem optycznym, którego kształt jest tak dobrany, że realizuje jeden z dwóch celów: rozpraszanie światła lub zmiana kierunku światła.

2.10.2 Pryzmaty rozpraszające

Doświadczenie Ugięcia światła w pryzmacie

W **pryzmatach rozpraszających** wykorzystuje się zjawisko dyspersji, polegające na zależności współczynnika załamania światła od długości fali (rozdział 9.6). Sprawia to, że różne składowe widmowe światła padającego na pryzmat załamywane są pod różnymi kątami. Pozwala to na przestrzenne rozdzielenie widma promieniowania, a w konsekwencji na analizę natężenia jego poszczególnych składowych spektralnych. W typowym pryzmacie przyjmuje się, że istnieje monotoniczna zależność (najczęściej bliska liniowej) współczynnika załamania od długości fali (wiele materiałów dielektrycznych ma właśnie taką zależność w zakresie światła widzialnego).

Z uwagi na zależność współczynnika załamania od długości fali, jednym z zastosowań pryzmatów jest określanie współczynnika załamania szkła, z którego wykonany jest pryzmat oraz współczynnika załamania ośrodka. Aby przeanalizować to zastosowanie rozważmy pryzmat równoramienny przedstawiony na rys. 2.8. Można pokazać, że dla pomiarów w takim pryzmacie minimalny kąt ugięcia wiązki osiągany jest wtedy, gdy promień załamany rozchodzi się w pryzmacie równolegle do jego podstawy. W takim przypadku współczynnik załamania dany jest zależnością

$$n = \frac{\sin\left[\left(\delta_m + \varphi\right)/2\right]}{\sin(\varphi/2)}, \tag{2.59}$$

gdzie δ_m jest minimalnym kątem ugięcia pryzmatu.

Pryzmaty wykorzystuje się w urządzeniach spektralnych, które rozdzielają fale kątowo w zależności od długości fali. Z uwagi na tę cechę można wprowadzić pojęcie **dyspersji kątowej**

$$D_k = \frac{\partial \delta}{\partial \lambda} = \frac{\partial \delta}{\partial n}\frac{\partial n}{\partial \lambda}. \tag{2.60}$$

Równanie to można przepisać korzystając z zależności (2.59)

$$D_k = \frac{\sin(\varphi/2)}{\sqrt{1 - n^2 \sin^2 \varphi/2}}\frac{\partial n}{\partial \lambda}. \tag{2.61}$$

Warto na koniec zwrócić uwagę, że dla uniknięcia wpływu wad tego układu, w tego typu badaniach wybiera się zazwyczaj kąt padania bliski kątowi minimalnemu.

2.11 Pryzmaty odbijające

Drugim obszarem zastosowań pryzmatów jest zmiana kierunku propagacji światła oraz przekształcenie obrazu. W takim przypadku do zmiany kierunku wykorzystuje się zjawisko całkowitego wewnętrznego odbicia wewnątrz pryzmatu.

Ponieważ przy odbiciu światła w pryzmacie może dojść do zmiany orientacji obrazu, **pryzmaty odbijające** dają dużą swobodę manipulacji obrazem. Pozwala to na dużą miniaturyzację układów optycznych i jest często wykorzystywane w tego typu układach.

Niepożądaną cechą pryzmatów odbijających jest ich dyspersja. Aby uniknąć problemu z tym zjawiskiem, pryzmaty buduje się w taki sposób, że wiązka światła pada prostopadle do ściany pryzmatu (kąt padania wynosi zero). W konsekwencji wszystkie składowe spektralne wnikają do pryzmatu bez zmiany kierunku i problem dyspersji przestaje mieć znaczenie.

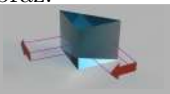
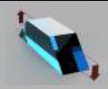
Tabela 2.2 przedstawia różne typu pryzmatów stosowane do zmiany kierunku promieni światła. Większość z tych pryzmatów pozwala na zmianę kierunku propagacji światła i obrócenie obrazu. Wyjątkiem jest w tym przypadku pryzmat Dovego, który obraca obraz bez zmiany kierunku propagacji światła. Szczególnie ciekawymi przypadkami pryzmatów są retroreflektory (kocie oko). W tym przypadku promień odbity po wyjściu z pryzmatu zawsze porusza się równolegle do promienia padającego. Sprawia to, że patrząc na taki pryzmat, zawsze będziemy widzieć swoje oko i to nie zależnie od tego pod jakim kątem patrzymy (to różni ten pryzmat od zwierciadła płaskiego). Właśnie tego typu pryzmat został zainstalowany przez misję Apollo na Księżycu. Dzięki niemu promień laserowy kierowany z Ziemi na pryzmat powraca do źródła. Wykorzystywane jest to m.in. w pomiarach odległości Ziemia–Księżyc, które bazują na zjawisku interferencji (patrz rozdział 5).

2.12 Światłowody

Doświadczenie
Propagacja światła w stróżce wody

Propagacja światła w pęczku światłowodów

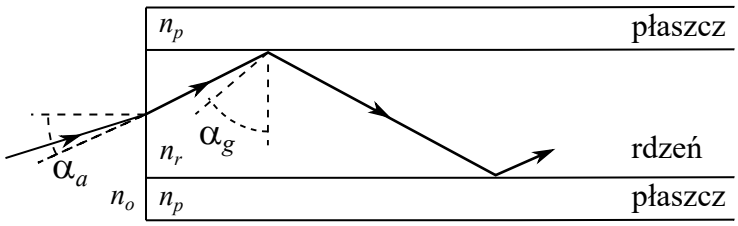
Tablica 2.2: Różne rodzaje pryzmatów i przekształcenia którym poddają one obraz.

	Pryzmat prostokątny Pryzmat taki działa jak lustro płaskie
	Pryzmat Porro Pryzmat Porro zachowuje orientację przedmiotów i z tego też powodu stosowany jest w lornetkach
	Pryzmat Dovego Nie zmienia kierunku, ale obraca obraz
	Pryzmat pentagonalny Odchyla wiązkę światła o 90° bez obracania jej.
	Retroreflektor/Kocie oko Narożnik sześcianu, który odbija wiązkę dokładnie przeciwbieżnie do kierunku padania. Tego typu elementy zostały zostawione na powierzchni Księżyca podczas misji Apollo i do dziś są wykorzystywane do badania odległości Ziemia–Księżyc

Jednym z bardzo ciekawych elementów optycznych pozwalających na prowadzenie światła w często bardzo skomplikowanej geometrii są **światłowody**. Pozwalają one na prowadzenie promieniowania elektromagnetycznego z zakresu optycznego, ale również bliskiej podczerwieni oraz ultrafioletu. Konwencjonalne światłowody mają kształt włókna zbudowanego z co najmniej dwóch dielektryków o różnych współczynnikach załamania. Wewnętrzny dielektryk, zwany **rdzeniem światłowodu** charakteryzuje się wyższym współczynnikiem załamania od otaczającego go dielektryka zwanego **płaszczem światłowodu**.

W najprostszym podejściu propagację światła w światłowodzie można wyjaśnić w oparciu o zjawisko całkowitego wewnętrznego odbicia. W tym podejściu promień prowadzony w światłowodzie pada na granicę pomiędzy płaszczem a rdzeniem, pod kątem większym od kąta granicznego α_g, odbijając się od tej granicy bezstratnie (rys. 2.9). Łatwo można pokazać, że w myśl tej zasady sam światłowód może być wyginany a światło i tak będzie w jego wnętrzu prowadzone. Oczywiście dla zbyt dużych kątów wygięcia światłowodu kąt padania może być mniejszy od kąta granicznego, a samo światło może wnikać do płaszcza, gdzie podlega stratom. Ten mechanizm utraty energii światła prowadzonego w światłowodzie zwany jest **stratami zgięciowymi**. Minimalny promień zgięcia światłowodu, przy którym wciąż prowadzi on światło nosi nazwę **promienia gięcia światłowodu**.

Pomiędzy promieniem padającym na granicę pomiędzy płaszczem i rdze-



Rysunek 2.9: Geometryczne podstawy propagacji światła w światłowodach.

niem, a kątem pomiędzy czołem (przodem) światłowodu i ośrodkiem zewnętrznym istnieją relacje geometryczne. Z uwagi na zjawisko całkowitego wewnętrznego odbicia istnieje pewien kąt graniczny zwany **kątem akceptacji** α_a, pod którym wchodzące promienie ciągle prowadzone są w światłowodzie. Dla kątów większych od kąta akceptacji wprowadzane promienie wnikają do płaszcza, gdzie ulegają tłumieniu.

Kąt akceptacji związany jest z inną wielkością charakterystyczną dla układów optycznych zwaną **aperturą numeryczną NA**. W przypadku światłowodów relacja pomiędzy obiema wielkościami dana jest zależnością

$$\mathrm{NA} = n_o \sin \alpha_a = \sqrt{n_r^2 - n_o^2}, \tag{2.62}$$

gdzie n_o, n_r oraz n_p są odpowiednio współczynnikami załamania ośrodka otaczającego światłowód, rdzenia światłowodu oraz płaszcza światłowodu, zaś α_a zwane jest **kątem akceptacji**.

Sama idea światłowodu jako elementu optycznego prowadzącego światło narodziła się jeszcze w XIX wieku. Dopiero jednak w latach 70. XX w. nastąpił prawdziwy wystrzał zainteresowania tego typu elementami. W tym właśnie okresie firma Corning wyprodukowała pierwszy **światłowód włóknisty**, zwany też **włóknem światłowodowym**, który charakteryzował się stratami rzędu 10 dB/km (po kilometrze natężenie spada do 10% początkowej wartości). Przez następne dekady technologia produkcji światłowodów była udoskonalana, tak że współczesne światłowody charakteryzują się stratami na poziomie 0.13 dB/km. Taki niski poziom strat pozwala przesyłać sygnały optyczne na odległości do 80 km bez konieczności ich wzmacniania.

Choć współczesne światłowody wykorzystywane są w oświetleniu i analogowej transmisji obrazu, to ich głównym obszarem zastosowań jest telekomunikacja i przesył informacji cyfrowej. W tych pierwszych zastosowaniach główną zaletą światłowodu jest możliwość doprowadzenia i wyprowadzenia światła z miejsc niedostępnych dla typowych rozwiązań. To właśnie ta zaleta jest głównym powodem zastosowań laserów w badaniach endoskopowych, ale również ich wykorzystania w medycynie inwazyjnej (np. przy wykorzystaniu światłowodów do niszczenia/wycinania tkanki nowotworowej). To ostatnie zastosowanie ma szereg zalet w stosunku do stosowanych konwencjonalnie technik (szybsze gojenie ran, brak blizn itp.).

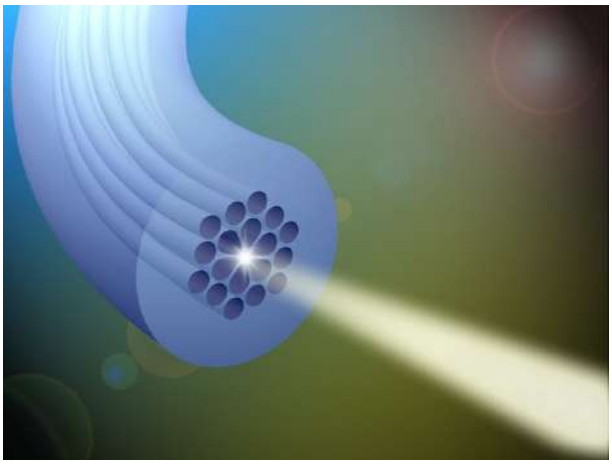
W przypadku zastosowania światłowodów w telekomunikacji zaleta świa-

tłowodów jest nieco inna. Z jednej strony zasadza się ona na możliwość bardzo szybkiego modulowania światła prowadzonego w światłowodach. W takim przypadku możliwe jest bardzo efektywne kodowanie informacji cyfrowej w postaci impulsów światła. Należy zaznaczyć, że tak szybka modulacja nie jest możliwa w przypadku kabli elektrycznych, gdzie bardzo szybka modulacja prądu prowadzonego w przewodzie prowadzi do emisji fali elektromagnetycznej i związanych z tym strat energii. W przypadku światłowodów problem ten nie istnieje. Dodatkową zaletą światłowodów jest tzw. możliwość **multipleksowania sygnałów z podziałem długości fali**, tj. jednoczesnego przesyłania informacji przez światłowód na wielu długościach fali. Ponieważ w pierwszym przybliżeniu (w przybliżeniu optyki liniowej) fale propagują się w światłowodzie niezależnie, możliwe jest duże bardziej efektywne przesyłanie informacji przez taki światłowód. Obecnie, w zależności od techniki multipleksowania, możliwe jest jednoczesne przesyłanie informacji przy pomocy światła w od kilkunastu do stu niezależnych kanałach. Pozwala to na zwiększenie przepustowości już istniejących sieci światłowodowych nawet o dwa rzędy wielkości.

Warto zaznaczyć, że współcześnie wzmocnienie sygnałów optycznych rozchodzących się w sieciach światłowodowych nie wymaga ich konwersji na sygnały elektryczne, które następnie byłyby wzmacniane i wykorzystywane do generacji sygnałów optycznych. Role tego typu układów pełnią tzw. **lasery światłowodowe**, którymi są światłowody domieszkowane jonami pierwiastków ziem rzadkich. Poddania tak domieszkowanych światłowodów oświetleniu światłem o odpowiednio dobranej długości fali prowadzi do powstania tzw. **inwersji obsadzeń** (więcej atomów znajduje się w stanie o wyższej energii niż w stanie o energii niższej), co sprawia, że światło przechodzące przez taki układ podlega wzmocnieniu. Ostatnim aspektem wielkiej popularności światłowodów jest ich cena, która czyni je niezmiernie atrakcyjnymi rozwiązaniem.

Światłowody fotoniczne

Na przestrzeni ostatnich dwóch dekad rozwinął się jeszcze jeden typ światłowodów zwany **światłowodami fotonicznymi**. Budowa światłowodów tego typu znacząco różni się od budowy światłowodów konwencjonalnych, gdyż nie tyle składają się one z dwóch współosiowych cylindrów wykonanych z materiałów o różnych współczynnikach załamania, ale jeden z materiałów tworzy w drugim specyficzną strukturę zwaną **kryształem fotonicznym**. Właściwości tego kryształu sprawiają, że zmianie ulega sposób propagacji światła. Nie bazuje on już na zjawisku całkowitego wewnętrznego odbicia, gdyż nie ma wyraźnej granicy pomiędzy płaszczem i rdzeniem, ale światło podlega rozproszeniom i interferencji. Interferencja ta staje się konstruktywna dla pewnego zakresu częstotliwości, „pułapkując" światło w obszarze rdzenia. Zakres częstotliwości, który podlega pułapkowaniu w rdzeniu światłowodu fotonicznego zwany jest **fotoniczną przerwą wzbronioną**.



Rysunek 2.10: Model światłowodu fotonicznego z otaczającą szklany rdzeń strukturą (kryształem fotonicznym) tworzoną przez przez układ współosiowych kanałów powietrznych. Rysunek ze strony Phys.org.

2.13 Układy optyczne

Jedną z podstawowych zalet optyki geometrycznej jest możliwość tworzenia, opisywania i analizowania złożonych układów optycznych.

2.13.1 Ogniskowa układu optycznego

Wyobraźmy sobie układ optyczny składający się z dwóch soczewek o ogniskowych f_1 i f_2 umieszczonych względem siebie w odległości d, która liczona jest między płaszczyzną główną obrazową pierwszej soczewki i płaszczyzną główną przedmiotową drugiej soczewki. Wykorzystując formalizm macierzy ABCD, możemy zapisać

$$M_s = \begin{bmatrix} A_2 & B_2 \\ -\frac{1}{f_2} & D_2 \end{bmatrix} \left(\begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ \frac{1}{f_1} & D_2 \end{bmatrix} \right) = \begin{bmatrix} A_1 & B_1 \\ -\frac{1}{f_{ef}} & D_2 \end{bmatrix}. \tag{2.63}$$

Łatwo pokazać, że wypadkowa ogniskowa f_{ef} takiego układu dana jest zależnością

$$\boxed{\frac{1}{f_{ef}} = \frac{D_2}{f_1} + \frac{A_1}{f_2} - \frac{d}{f_1 f_2}.} \tag{2.64}$$

Analizując równanie (2.64) można zauważyć, że w przypadku soczewek cienkich, dla których $A_1 = A_2 = D_1 = D_2 = 1$ wzór na ogniskową dwóch soczewek można zapisać w postaci

$$\frac{1}{f_{ef}} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}. \tag{2.65}$$

Co więcej, w przypadku gdy odległości między soczewkami są zaniedbywalne, można wprowadzić proste wyrażenie na efektywną ogniskową układu soczewek składających się z n soczewek cienkich, która dana jest zależnością

$$\frac{1}{f} = \sum_{i=1}^{n} \frac{1}{f_i}. \tag{2.66}$$

Pozwala ona na łatwe określenie wypadkowej ogniskowej układu optycznego

Dla soczewek grubych wzór na efektywną ogniskową soczewki jest bardziej złożony, ale można go stosunkowo łatwo wyznaczyć, korzystając z równania (2.64) oraz wyrażenia na macierz ABCD soczewki grubej [równanie (2.48)].

2.13.2 Ograniczenie wiązek: Przysłony i źrenice

W każdym rzeczywistym układzie optycznym istnieją fizyczne ograniczenia rozmiarów wiązek światła (oprawy soczewek, diafragmy). Elementy te pełnią kluczową rolę w determinowaniu jasności obrazu oraz pola widzenia.

Przysłona aperturowa i polowa

Element optyczny (lub jego oprawa), który najbardziej ogranicza pęk promieni wychodzących z osiowego punktu przedmiotu, nazywany jest **przysłoną aperturową**. To ona decyduje o tym, jak dużo światła dociera do obrazu, a zatem determinuje jego jasność (natężenie oświetlenia).

Z kolei element, który ogranicza kątowy rozmiar wiązki (czyli to, jak duży fragment przedmiotu widzimy), nazywany jest **przysłoną polową**. W mikroskopach i teleskopach jest to zazwyczaj pierścień umieszczony w płaszczyźnie obrazu pośredniego, co zapewnia ostre krawędzie pola widzenia.

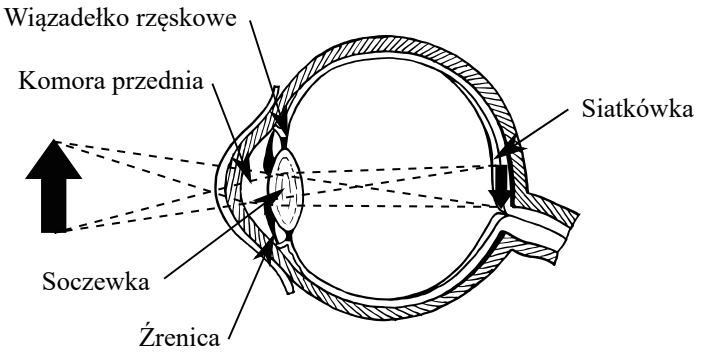
Źrenice układu

Przysłona aperturowa widziana przez soczewki znajdujące się przed nią (od strony przedmiotu) tworzy obraz zwany **źrenicą wejściową**. Analogicznie, obraz przysłony aperturowej widziany przez soczewki znajdujące się za nią (od strony obrazu) nazywamy **źrenicą wyjściową**.

Znajomość położenia źrenic jest kluczowa przy łączeniu układów optycznych. Przykładowo dla układu składającego się z teleskopu i oka, aby nie tracić światła, źrenica wyjściowa teleskopu powinna pokrywać się ze źrenicą wejściową oka. Odległość źrenicy wyjściowej od ostatniej soczewki okularu nazywana jest **odstępem źrenicy**. Jest to parametr krytyczny dla osób noszących okulary, gdyż zbyt mały odstęp uniemożliwia objęcie wzrokiem całego pola widzenia.

2.13.3 Oko

Oko jest niezmiernie złożonym, ale i bardzo ciekawym układem optycznym. Z punktu widzenia optycznego jest to układ dwóch soczewek zwanych **komorą przednią** i **soczewką**. Oprócz tego, w układzie tym mamy również przesłonę



Rysunek 2.11: Uproszczony schemat budowy ludzkiego oka wraz z przykładowym powstawaniem w nim obrazu.

(źrenica), której rolę pełni źrenica regulująca ilość światła wnikającą do wnętrza oka. Sam obraz powstaje na siatkówce, która pokryta jest czopkami i pręcikami. Dużo liczniejsze z nich (ok. 25 mln) pręciki odpowiadają za wykrycie ruchu, podczas gdy mniej liczne (5 mln), ale większe czopki odpowiedzialne są za widzenie barwne. Zdolność akomodacyjna oka polega na możliwości obrazowania różnoodległych przedmiotów na siatkówce. Umożliwia to mięsień zwany wiązadełkiem rzęskowym, który przez fizyczny ruch soczewki zmienia ogniskową oka. Z wiekiem mięsień ten wiotczeje, co prowadzi do pogorszenia wzroku. Ponieważ degraduje się możliwość efektywnego przesuwania soczewki, oko nie jest w stanie skupić się na obiektach znajdujących się blisko soczewki, co prowadzi do dalekowzroczności.

2.13.4 Mikroskop

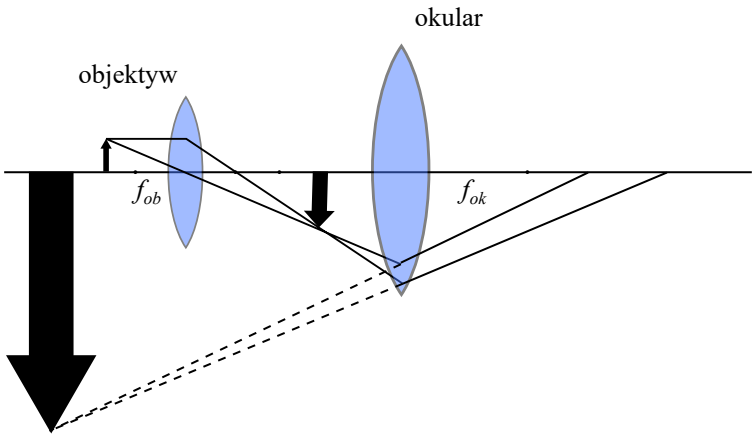
Mikroskop jest urządzeniem optycznym, którego zadaniem jest powiększanie obiektów znajdujących się w bliskiej odległości. W ogólności składa się on z dwóch elementów (soczewek), z których pierwszy, zwany **obiektywem**, tworzy obraz, podczas gdy drugi, zwany **okularem**, rzutuje go na oku.

Na rys. 2.12 przedstawiono diagram promieni pozwalający na wyznaczenie obrazu powstałego przy pomocy mikroskopu. Obraz rzeczywisty przedmiotu powstaje w obszarze między obiektywem a okularem i jest skutkiem przejścia promieni przez obiektyw. Obraz ten jest następnie przedmiotem dla okularu. W tubusie odległości soczewek są tak dobrane, że ogniskowa okularu przewyższa odległość między obrazem utworzonym przez obiektyw a okularem. Dlatego okular przekształca obraz, czyniąc go pozornym i powiększonym, co pozwala na rejestrację tego obrazu przez oko.

Powiększenie mikroskopu dane jest jako iloczyn powiększenia okularu i obiektywu

$$M_m = M_{ob} M_{ok}. \tag{2.67}$$

Zdolność rozdzielcza mikroskopu tj. zdolność rozdzielenia dwóch obiek-



Rysunek 2.12: Diagram promieni umożliwiający wyznaczenie obrazu powstałego przy pomocy mikroskopu.

tów znajdujących się blisko siebie jest ograniczona ilością światła, która wchodzi do obiektywu. Dlatego tak ważne jest wykorzystanie obiektywów o wysokiej aperturze numerycznej NA. Co więcej, aby uzyskać wysoką rozdzielczość ważne jest także kompensowanie wad optycznych (aberracji) soczewek.

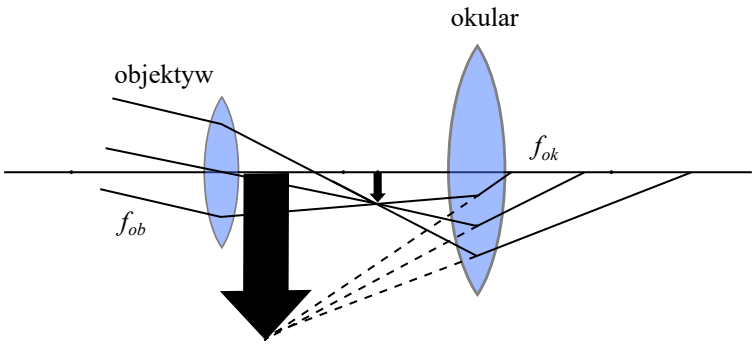
2.13.5 Teleskop

W przeciwieństwie do mikroskopów **teleskopy** używane są do powiększania odległych obiektów, z których docierające promienie światła są równoległe. W tradycyjnej swojej formie teleskop jest układem dwóch soczewek (**obiektywu** i **okularu**), których celem jest przekształcenie promieni tak, aby przy ich pomocy powstał powiększony obraz pozorny. Ponieważ przy pomocy teleskopu obserwacji dokonuje człowiek, to rozmiar wiązki wychodzącej z teleskopu powinien być dopasowany do źrenicy oka lub powinien być od niej mniejszy (jeśli byłby większy, to część fotonów byłaby tracona, co skutkowałoby mniejszą jasnością rejestrowanego obrazu).

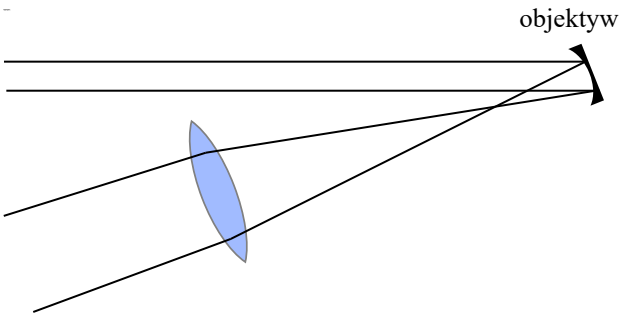
Diagram promieni dla teleskopu przedstawiony został na rys. 2.13. Widać, że urządzenie to jest tak skonstruowane, że przekrywają się ze sobą ogniska obiektywu i okularu. Sprawia to, że powiększenie teleskopu, które w ogólności jest iloczynem powiększenia obiektywu i okularu, można zapisać jako stosunek ogniskowych obiektywu i okularu

$$M_{tel} = \frac{f_{ob}}{f_{ok}}. \tag{2.68}$$

Z uwagi na zależność pomiędzy rozdzielczością i ilością docierającego do oka światłą, ważne jest, aby budowa teleskopu pozwalała na dostarczanie do obiektywu jak największej liczby promieni, tzn. obrazowanie jak największej



Rysunek 2.13: Diagram promieni umożliwiający wyznaczenie obrazu powstałego przy pomocy teleskopu.

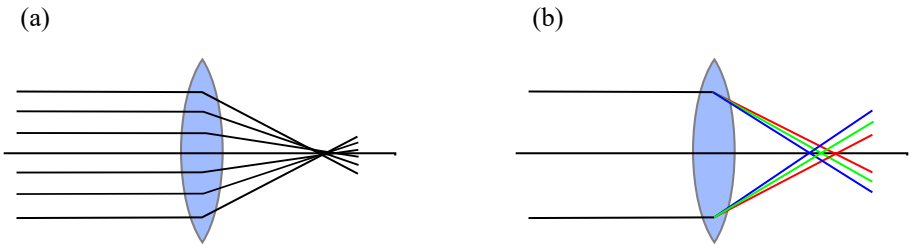


Rysunek 2.14: Diagram promieni w teleskopie zwierciadlanym.

powierzchni. Z drugiej jednak strony wykonanie dużych soczewek nastręcza poważnych problemów technicznych z jednorodnością materiału i ich masą (problem mocowania). W tym kontekście zastąpienie soczewek lustrami (rys. 2.14) pozwala na ominięcie tych problemów i konstrukcję **teleskopów zwierciadlanych**. Jasnym jest, że zwierciadła nie są wolne od wszystkich wad optycznych (rozdział 2.14), z których najpoważniejszą jest aberracja sferyczna.

Optyka adaptatywna

Jednym z poważnych problemów obserwacji kosmosu z Ziemi są fluktuacje atmosfery, które prowadzą do zaburzeń obserwowanego obrazu. Okazuje się jednak, że w pewnym ograniczonym obszarze światło przechodzące przez atmosferę podlega niewielkim zaburzeniom. Chcąc wykorzystać ten fakt, ludzie opracowali układy optycznej zwane układami **optyki adaptatywnej**. Układy takie są zbudowane z wieloelementowych luster, których poszczególne części mogą się niezależnie poruszać. W ten sposób przesuwając element lustra, jesteśmy w stanie korygować zmiany wynikające z fluktuacji



Rysunek 2.15: (a) Aberracja sferyczna i (b) aberracja chromatyczna.

atmosfery.

Aby korekcja z uwagi na zmiany gęstości atmosfery była możliwa, konieczne jest zmierzenie ich. Robi się to, wytwarzając w górnych warstwa atmosfery (w mezosferze), referencyjne źródło światła. Źródło to powstaje w oparciu o fluorescencyjne wzbudzenie sodu znajdującego się w mezosferze. Oświetlenie ich światłem o odpowiedniej (rezonansowej) długości fali spowoduje powstanie „sztucznej gwiazdy". To właśnie obserwacja światła tej gwiazdy pozwala dokonać korekcji fluktuacji gęstości atmosfery i odpowiednio ustawić elementy zwierciadła adaptatywnego.

2.14 Wady soczewek

2.14.1 Aberracje

Zgodnie z definicją, obraz jest to takie miejsce w przestrzeni, w którym wszystkie promienie wychodzące z danego punktu przedmiotu przecinają się. Dodatkowym warunkiem jest, aby każdy tak zobrazowany punkt przedmiotu był wzmocniony w ten sam sposób. Jak się jednak okazuje warunki te nie są spełniony dokładnie, a wszystkie odstępstwa od przywołanego wcześniej zachowania zwane są **aberracjami**.

Dla promieni pozaosiowych, odległość od osi optycznej układu sprawia, że kąt padania na powierzchnię soczewki staje się znaczący. W konsekwencji rozwinięcie funkcji trygonometrycznych w szereg Taylora i uwzględnienie tylko pierwszych wyrazów rozwinięcia jest niewystarczające do poprawnego opisu propagacji promienia. W przypadku elementów o symetrii sferycznej, tzn. luster i soczewek, których powierzchnie odbijające mają kształt fragmentu sfery, problem promieni pozaosiowych manifestuje się jako inne położenie ogniska dla różnych odległości od osi optycznej [rys. 2.15(a)]. Zjawisko to nazywane jest **aberracją sferyczną** i jest ono jedną z podstawowych wad soczewek i zwierciadeł.

Z matematycznego punktu widzenia, jednym ze sposobów usunięcia problemu promieni pozaosiowych jest uwzględnienie wyższych rzędów rozwinięcia funkcji sinus i tangens w zależnościach propagacyjnych. Takie rozwiązanie nie pozwala jednak na przedstawienie propagacji promienia jako prostych przekształceń algebraicznych. Alternatywnie problem ten może zostać usunięty

przez zastąpienie kąta pojawiającego się w wektorze propagacji promienia świetlnego przez jego sinus. Choć jest to w ogólności możliwe, to jednak prowadzi to do dużo bardziej skomplikowanych zależności pomiędzy różnymi wielkościami przy przejściu przez poszczególne elementy optyczne. Z doświadczalnego punktu widzenia aberrację sferyczną można ograniczyć poprzez wykonanie soczewki, której powierzchnie łamiące mają kształt paraboloidy. Taki kształt pozwala na skupienie wszystkich promieni w tym samym miejscu.

Drugim źródłem aberracji jest dyspersja ośrodka, z którego wykonane są elementy optyczne. Ponieważ współczynnik załamania materiałów cechuje się zależnością od długości fali,dlatego fale o różnej długości będą miały ogniska w innym miejscu w przestrzeni [rys. 2.15(b)]. Wada ta nosi nazwę **aberracji chromatycznej**.

Aberracje chromatyczną można ograniczyć poprzez zastosowanie materiałów o niskiej dyspersji lub układów soczewek wykonanych z materiałów o różnych współczynnikach załamania, tak aby zjawisko dyspersji kompensowane było dla określonych długości fali. W przypadku kompensacji dyspersji mówimy o **soczewkach achromatycznych** lub o **apochromatach**.

2.14.2 Koma

Jeśli jasny punkt, który jest źródłem szerokiej wiązki światła znajduje się poza osią optyczną, to jego obraz zamiast tworzyć jasne koło w płaszczyźnie obrazowej, pojawia się jako dość skomplikowany asymetryczny kształt przypominający przecinek. To właśnie ten kształt daje nazwę powyższej wadzie zwanej **komą** (przecinkiem).

Eliminacja aberracji sferycznej dla punktu osiowego nie gwarantuje, że punkty leżące blisko osi również zostaną odwzorowane jako punkty. Aby układ optyczny był wolny zarówno od aberracji sferycznej, jak i od komy (taki układ nazywamy **aplanatycznym**), musi zostać spełniony **warunek sinusów Abbego**:

$$n_1 y_1 \sin\theta_1 = n_2 y_2 \sin\theta_2, \tag{2.69}$$

gdzie y_1 i y_2 to odległości punktów od osi optycznej (wysokość przedmiotu i obrazu), a θ_1 i θ_2 to kąty, jakie tworzy dowolny promień wychodzący z punktu osiowego z osią optyczną.

Warunek ten implikuje, że powiększenie poprzeczne układu musi być stałe dla wszystkich stref soczewki (dla promieni przyosiowych i skrajnych). Jest on fundamentalnym kryterium przy projektowaniu wysokiej jakości obiektywów mikroskopowych, które muszą odwzorowywać ostre obrazy nie tylko w centrum, ale i w szerszym polu widzenia.

2.14.3 Astygmatyzm

Jeżeli na soczewkę pada wiązka światła, to promienie leżące w płaszczyźnie południkowej będą uginać się inaczej niż te, które leżą w płaszczyźnie równi-

kowej. W konsekwencji po przejściu przez soczewkę promienie nie będą do siebie równoległe i zmieniać się będzie przekrój poprzeczny przechodzącej wiązki. Przykładowo, gdy wiązką o kołowym profilu poprzecznym oświetlimy pod kątem soczewkę to w pewnej odległości od soczewki wiązka ta będzie widoczna jako odcinek, następnie jako elipsa, później koło, jeszcze dalej znów jako elipsa, jednak o osi długiej obróconej o kąt 90° w stosunku do pierwszej elipsy, aż będzie widoczna jako pionowy odcinek (dalej wiązka zacznie się rozbiegać do nieskończoności). Przykład ten pokazuje, że ogniskowa dla promieni leżących w płaszczyźnie południkowej jest inna od tej dla promieni równikowych. Tego typu aberracja nosi nazwę **astygmatyzmu**.

Warto w tym miejscu podkreślić, że ognisko dla promieni leżących pod kątem do osi optycznej nie znajduje się na tej osi, ale poza nią. Co więcej, nie jest ono nawet odległe od soczewki o długość ogniskowej. W rzeczywistości ugięciu ulega płaszczyzna obrazowa, co jest jedną z wad układów optycznych. W konsekwencji obrazy powstałe w układach optycznych podlegają **zniekształceniu**, a powiększenie układu w ogólności zależy od kąta padania promieni względem osi optycznej.

2.15 Podsumowanie rozdziału

W niniejszym rozdziale dokonano przejścia od falowego opisu światła do optyki geometrycznej. Wykazano, że w granicy krótkofalowej równanie Helmholtza redukuje się do równania eikonału, co stanowi rygorystyczne uzasadnienie dla koncepcji promieni świetlnych oraz fenomenologicznej zasady Fermata. Następnie wprowadzono formalizm macierzy ABCD, który ujednolica i upraszcza analizę propagacji światła w przybliżeniu przyosiowym. Pozwala on na elegancki, algebraiczny opis zarówno prostych zjawisk odbicia i załamania, jak i zachowania światła w grubych soczewkach czy złożonych systemach wieloelementowych.

Teoretyczne rozważania uzupełniono analizą rzeczywistych instrumentów optycznych (oka, mikroskopów, teleskopów) oraz elementów prowadzących światło (światłowodów i pryzmatów). Zwrócono jednocześnie uwagę na fizyczne i geometryczne ograniczenia układów rzeczywistych, manifestujące się w postaci aberracji sferycznej, chromatycznej czy komy. Zrozumienie tych wad – oraz zjawisk zachodzących w reżimie, gdy rozmiary przeszkód stają się porównywalne z długością fali świetlnej – wymaga jednak porzucenia przybliżenia promieni na rzecz pełnego opisu falowego, który jest punktem wyjścia do rozważań prowadzonych w następnych rozdziałach.

Zadania

2.1. Korzystając z zasady Fermata wyprowadź

(a) prawo odbicia,

(b) prawo załamania,

(c) odbicia od powierzchni eliptycznej.

2.2. Załóżmy, że chcemy zaprojektować zwierciadło, które idealnie skupi wszystkie promienie wiązki równoległej (pochodzącej od gwiazdy) dokładnie w jednym punkcie (ognisku), eliminując całkowicie aberrację sferyczną. Wykorzystując zasadę Fermata (wymagając, aby droga optyczna OPL każdego promienia docierającego do ogniska była identyczna), wykaż analitycznie, że powierzchnia takiego zwierciadła musi być paraboloidą obrotową.

2.3. Moneta leży w wodzie na głębokości h. Na jakiej głębokości będzie ją widział obserwator patrzący pionowo z góry przez powierzchnię wody o współczynniku załamania $n_w = 1,33$?

2.4. Obiekt znajduje się przed zwierciadłem wklęsłym. Proszę znaleźć położenie obrazu d_i i wartość powiększenia M dla $d_\mathrm{o} = R$, $d_\mathrm{o} = R/2$, $d_\mathrm{o} = R/4$ i $d_\mathrm{o} = -R/2$ (przedmiot pozorny). Dla każdego z tych przypadków proszę narysować promienie biegnące z pojedynczego punktu (znajdującego się poza osią optyczną) przedmiotu do odpowiedniego punktu obrazu. Proszę zauważyć, że można narysować promień, który początkowo biegnie równolegle do osi optycznej i promień, który początkowo porusza się w kierunku przecinającym oś w ognisku $R/2$.

2.5. Kulę o promieniu $R = 2$ cm wykonaną ze szkła o współczynniku załamania $n = 1,5$ wstawiono w skolimowaną wiązkę światła. W jakiej odległości od środka kuli znajduje się ognisko wiązki przechodzącej przez środek kuli? Rozmiary poprzeczne wiązki są niewielkie w porównaniu z promieniem kuli.

2.6. Na szklaną półkulę o promieniu R, leżącą płaską powierzchnią na stole, pada pionowo od góry wiązka światła o przekroju kołowym z promieniem, $r < R$. Oś wiązki światła przechodzi centralnie przez półkulę. Jaki będzie rozmiar plamki światła na stole?

2.7. Soczewka dwuwypukła wykonana ze szkła o współczynniku załamania $n = 1,5$ ma grubość 1 cm. Promienie krzywizn obu płaszczyzn soczewki mają odpowiednio wartości 10 cm i -15 cm. Proszę policzyć różnicę ogniskowych oraz ich położenia na osi zakładając, że soczewka ta jest soczewką grubą oraz, że jest ona traktowana jako soczewka cienka, aby określić błąd wprowadzany przez przyjęcie takiego przybliżenia.

2.8. Wyznacz odległość obrazu b dwuwypukłej soczewki grubej wykonanej ze szkła o współczynniku załamania 1,5, promieniach krzywizn 20 cm i -40 cm oraz grubości 1,5 cm.

2.9. Wychodząc z równania soczewkowego, wyprowadź równanie soczewkowe Newtona. Następnie wyprowadź to samo równanie posługując się diagramem promieni.

2.10. Proszę wyznaczyć postać macierzy ABCD dla płytki płaskorównoległej o grubości Δ i współczynniku załamania n umieszczonej w powietrzu. Następnie proszę obliczyć o ile zmienią się parametry promienia padającego na płytkę pod kątem 5° na wysokości 1 cm, w przypadku gdy płytka wykonana jest ze szkła o współczynniku załamania $n = 1,5$ i ma ona grubość 5 mm.

2.11. Proszę rozważyć soczewkę o grubości $d = 5$ cm, $R_1 = 5$ cm, $R_2 = -10$ cm, $n = 1,5$. Proszę wyliczyć elementy macierzy ABCD dla takiej soczewki.

2.12. Soczewka płasko-wypukła zrobiona jest ze szkła ($n = 1,50$) i znajduje się w powietrzu. Ma ona promień krzywizny $r = 10$ cm i grubość 1 cm. Proszę wyznaczyć macierz ABCD tej soczewki oraz pokazać, że jej wyznacznik jest równy 1. Pod jakim kątem (mierzonym powyżej osi optycznej) powinien padać na soczewkę promień światła tak, aby zarówno wejście jak i wyjście z soczewki znajdowały się na wysokości 2 cm?

2.13. Proszę wyznaczyć postać macierzy ABCD, która opisuje bieg promienia (tam i z powrotem) w prostej wnęce lasera, złożonej ze zwierciadła płaskiego i zwierciadła wklęsłego o promieniu krzywizny R, które oddalone są od siebie o odległość L. Rozważania warto rozpocząć od powierzchni zwierciadła płaskiego.

2.14. Wnęka optyczna lasera (rezonator) składa się z dwóch zwierciadeł sferycznych o promieniach krzywizn R_1 i R_2, oddalonych od siebie o L. Światło wielokrotnie odbija się między nimi. Aby promień nie „uciekł" z wnęki po wielu obiegach (co jest warunkiem stabilnej pracy lasera), ślad macierzy opisującej pełen obieg zamknięty (w jedną i z powrotem) musi spełniać tzw. warunek stabilności: $-1 \leq \frac{A+D}{2} \leq 1$.

(a) Wyznacz pełną macierz ABCD dla jednego obiegu (od zwierciadła 1, przez odległość L, odbicie od zwierciadła 2, powrót na odległość L i odbicie od zwierciadła 1).

(b) Korzystając z elementów A i D wyznaczonej macierzy, wyprowadź warunek stabilności rezonatora w postaci: $0 \leq g_1 g_2 \leq 1$, gdzie parametry rezonatora to $g_1 = 1 - \frac{L}{R_1}$ oraz $g_2 = 1 - \frac{L}{R_2}$.

(c) Określ, czy rezonator konfokalny (gdzie $R_1 = R_2 = L$) jest stabilny.

2.15. Proszę obliczyć wartość kąta akceptacji dla światłowodu, którego rdzeń ma współczynnik załamania światła równy 1,48, a płaszcz 1,46.

2.16. Rozważmy lunetę o ogniskowej obiektywu lunety $f_1 = 60$ cm, okularu $f_2 = 4$ cm i współczynniku załamania szkła obiektywu i lunety $n = 1,5$. Do środka lunety nieznośny Dyzio nalał wody ($n = 1,33$). Jakim obiektywem z tego samego gatunku szkła trzeba zastąpić poprzedni obiekty lunety, aby można było używać lunety

(a) w powietrzu,

(b) pod wodą?

2.17. Obiektyw teleskopu składa się z dwóch soczewek, tj. soczewki skupiającej o ogniskowej 20 cm oraz soczewki rozpraszającej o ogniskowej 5 cm, umieszczonej za pierwszą soczewką w odległości 15,5 cm. Teleskop skierowany jest na wieżę o wysokości 30 m, znajdującą się w odległości 3 km. Jakie będzie powiększenie obrazu wieży i jaka będzie odległość między pierwszą soczewką a obrazem?

2.18. Soczewka zbieżna o ogniskowej 15 cm i średnicy 4 cm używana jest do skupiania równoległych promieni światła. Określ aberrację sferyczną dla tej soczewki, zakładając, że soczewka ma kształt sfery.

2.19. W fotografii i projektowaniu okularów powszechnie stosuje się układy wielosoczewkowe. Rozważmy układ dwóch soczewek cienkich o ogniskowych $f_1 = f_2 = 10$ cm, umieszczonych w odległości $d = 5$ cm od siebie (w powietrzu).

(a) Oblicz wypadkową (efektywną) ogniskową f_{ef} tego układu.

(b) Skonstruuj macierz ABCD całego układu (od soczewki 1 do soczewki 2).

(c) Wykorzystując odpowiednie wzory na płaszczyzny główne z tekstu głównego podręcznika, znajdź odległości płaszczyzny głównej przedmiotowej od pierwszej soczewki (h_1) oraz płaszczyzny obrazowej od drugiej soczewki (h_2). Skomentuj fizyczne znaczenie znaku tych odległości.

Rozdział 3

Fale płaskie, sferyczne i gaussowskie

3.1 Wprowadzenie

Doświadczenie
Falownica generująca fale różnego typu

Jak pokazano w poprzednim rozdziale, separacja zmiennych przestrzennych i czasowych oraz zapisanie części czasowej pola w postaci fali monochromatycznej pozwala na sformułowanie równania Helmholtza [równanie (2.3)]. Jego rozwiązanie umożliwia wyznaczanie przestrzennego rozkładu pola, czym zajmiemy się poniżej dla różnych geometrii problemu.

3.2 Fale płaskie

Rozważmy propagację światła w ośrodku jednorodnym. W takim przypadku pojawiający się w równaniu Helmholtza operator Laplace'a najłatwiej zapisać w postaci

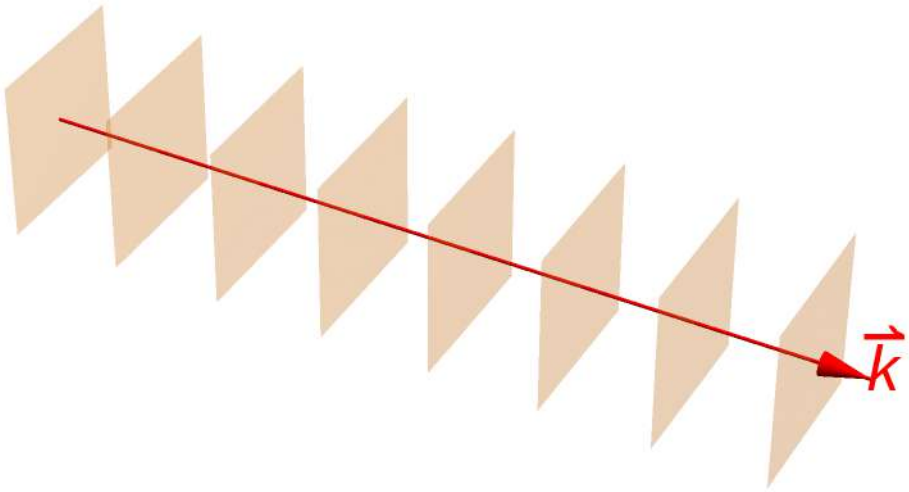
$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \tag{3.1}$$

Szczególnym przykładem rozwiązania równania Helmholtza z takim laplasjanem jest tzw. **fala płaska**

$$E^{(+)} = E_0^{(+)} e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{3.2}$$

gdzie $E_0^{(+)}$ jest zespoloną amplitudą fali, zaś $\mathbf{k}$ jest jej wektorem falowym. Fale płaskie to takie fale, których **fronty falowe**, tj. powierzchnie łączące punkty o tej samej fazie, są płaszczyznami. Przykład fali płaskiej został przedstawiony na rysunku 3.1.

Łącząc postać funkcji falowej w ośrodku jednorodnym [równanie (3.2)] z równaniem fali monochromatycznej [równanie (1.18)] oraz uwzględniając wektorowy



Rysunek 3.1: Fala płaska.

charakter pola elektrycznego możemy zapisać równanie na falę elektromagnetyczną jako

$$\mathbf{E}^{(+)}(\mathbf{r},t) = \mathbf{e}_E E_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{3.3}$$

gdzie $\mathbf{e}_E$ jest wersorem polaryzacji pola elektrycznego $\mathbf{E}$ (dla fal eliptycznych i kołowych wersor ten może przyjmować postać zespoloną).

W pełni analogiczne rozważania można przeprowadzić dla pola magnetycznego, uzyskując wyrażenie na składową pola magnetycznego fali elektromagnetycznej

$$\mathbf{B}^{(+)}(\mathbf{r},t) = \mathbf{e}_B B_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{3.4}$$

gdzie $B_0^{(+)}$ jest składową magnetyczną fali, zaś $\mathbf{e}_B$ jest wersorem pola magnetycznego fali. Pojawiające się w równaniu (3.4) wielkości ω i $\mathbf{k}$ są dokładnie tymi samymi wielkościami, które pojawiają się w równaniu (3.3).

Korzystając z równań Maxwella w ośrodkach dielektrycznych dla fali opisanej zależnościami (3.3) oraz (3.4), utożsamiając operator ∇ z wielkością $i\mathbf{k}$ otrzymujemy

$$\begin{aligned} \mathbf{k}\times\mathbf{B}_0 &= -\frac{\omega}{v^2}\mathbf{E}_0, \\ \mathbf{k}\times\mathbf{E}_0 &= \omega\mathbf{B}_0. \end{aligned} \tag{3.5}$$

gdzie $\mathbf{E}_0 = E_0\mathbf{e}_E$ i $\mathbf{B}_0 = B_0\mathbf{e}_B$. Zależności (3.5) wskazują, że dla fali płaskiej

$$\mathbf{E} \perp \mathbf{B} \perp \mathbf{k}, \tag{3.6}$$

co oznacza, że fala płaska jest **poprzeczną falą elektromagnetyczną**. Dodatkowo, ponieważ wektory te są do siebie prostopadłe, mamy następującą relację pomiędzy amplitudami pola elektrycznego i magnetycznego światła

$$B_0^{(+)} = \frac{E_0^{(+)}}{v}. \tag{3.7}$$

Warto w tym miejscu zauważyć, że z uwagi na fakt, że v jest wielkością wymiarową, relacja (3.7) nie mówi, że pole magnetyczne jest v-razy mniejsze od pola elektrycznego. Ponieważ są to inne wielkości, to ich bezpośrednie porównanie nie jest możliwe.

Zwróćmy na koniec uwagę, że równanie (3.7) pozwala obliczyć natężenie (średnią wartość wektora Poytinga) fali płaskiej

$$I = \langle |\mathbf{S}| \rangle = \frac{v\varepsilon}{2} \left| E_0^{(+)} \right|^2. \tag{3.8}$$

3.3 Fale sferyczne

Rozważmy teraz nieco inny przypadek, w którym fala rozchodzi się w ośrodku o symetrii sferycznej, a źródłem fali jest punkt. Laplasjan we współrzędnych sferycznych dany jest zależnością

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2}. \tag{3.9}$$

Ponieważ układ ma mieć symetrię sferyczną, to różniczkowanie po θ i ϕ daje zero, a sama funkcja falowa dana jest zależnością

$$E(\mathbf{r}) = E(r, \theta, \phi) = E(r). \tag{3.10}$$

Powyższe równanie można sprowadzić do postaci:

$$\nabla^2 E(r) = \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2 \frac{\partial E(r)}{\partial r}\right]. \tag{3.11}$$

Równanie (3.11) można przepisać do postaci

$$\nabla^2 E(\mathbf{r}) = \frac{1}{r}\frac{\partial^2}{\partial r^2}\left[rE(r)\right]. \tag{3.12}$$

Podstawiając równanie (3.12) do równania Helmholtza (2.3) i mnożąc to równanie obustronnie przez r otrzymujemy

$$\frac{\partial^2}{\partial r^2}\left[rE(r,t)\right] = -k^2 rE(r,t). \tag{3.13}$$

Rozwiązaniem tego równania są funkcje w postaci

$$E(r) = \frac{E_0^{(+)} e^{ikr}}{r}. \tag{3.14}$$

Warto zwrócić uwagę, że fronty falowe tej fali mają kształt sfer, których środek jest w źródle fali. Z tego też powodu fale takie nazywamy **falami sferycznymi**.

Po uwzględnieniu zależności czasowej w równaniu fali sferycznej, otrzymujemy ogólne wyrażenie na pole elektryczne fali elektromagnetycznej emitowanej ze źródła punktowego umieszczonego w początku układu współrzędnych

$$E(r,t) = \frac{E_0^{(+)} e^{i(kr-\omega t)}}{r}. \tag{3.15}$$

Analogiczną zależność można wprowadzić dla składowej magnetycznej

$$B(r,t) = \frac{B_0^{(+)} e^{i(kr-\omega t)}}{r}. \tag{3.16}$$

Warto zauważyć, że w wyrażeniach na pole elektryczne i magnetyczne fali sferycznej pominięto wektorowy charakter pola. Zrobiono tak, ponieważ oba pola drgają prostopadle do kierunku propagacji, tak więc dla każdego z promieni ten kierunek jest inny, przy czym oba pola wciąż pozostają do siebie prostopadłe.

Dlaczego nie istnieje idealnie izotropowe źródło światła?

Światło jest falą elektromagnetyczną poprzeczną, co oznacza, że wektory pola elektrycznego drgają w płaszczyźnie prostopadłej do kierunku rozchodzenia się fali. Wyobraźmy sobie punktowe źródło, które emituje falę sferyczną we wszystkich możliwych kierunkach z równą mocą. W takim przypadku fronty falowe mają kształt rozszerzających się sfer, a propagacja następuje wzdłuż promieni. Ponieważ pole musi być poprzeczne, wektor pola elektrycznego w każdym punkcie musi leżeć na powierzchni tej sfery (musi być do niej styczny).

W tym miejscu fizyka zderza się jednak z topologią, a konkretnie z **twierdzeniem o zaczesywaniu sfery** (ang. *Hairy Ball Theorem*). Twierdzenie to najprościej wyobrazić sobie na przykładzie tenisowej piłki gęsto pokrytej włosami. Jeśli spróbujemy gładko zaczesać wszystkie włosy tak, aby leżały całkowicie płasko na powierzchni piłki, szybko zdamy sobie sprawę, że to niemożliwe. Zawsze powstanie przynajmniej jeden „wicherek" – punkt, w którym włosy sterczą do góry, albo miejsce, gdzie po prostu ich nie ma. Z matematycznego punktu widzenia oznacza to, że na powierzchni sfery nie może istnieć ciągłe pole wektorów stycznych, które nigdzie nie

przyjmuje wartości zero.

W kontekście optyki owymi „włosami" są wektory pola elektrycznego. Skoro nie da się ich ciągle i niezerowo rozłożyć na całej sferze, to znaczy, że na froncie falowym każdej wektorowej fali sferycznej muszą istnieć kierunki, w których amplituda pola bezwzględnie spada do zera. Z tego powodu fizycznie niemożliwe jest stworzenie źródła (np. atomu czy anteny dipolowej), które promieniowałoby falę poprzeczną z idealnie taką samą, niezerową amplitudą we wszystkich kierunkach w przestrzeni. Zawsze znajdą się osie („bieguny"), wzdłuż których źródło w ogóle nie emituje promieniowania. Z tego względu izotropowa wektorowa fala sferyczna jest jedynie wygodną, niefizyczną matematyczną abstrakcją.

Z równania (3.15) widać, że wraz z oddalaniem się od źródła promieniowania spada amplituda pola elektrycznego fali sferycznej. Z tego też powodu średnia wartość wektora Poytinga, czyli natężenie światła, dane jest zależnością

$$I_s = |\langle S \rangle| = \frac{v\varepsilon}{2} \frac{|E_0^{(+)}|^2}{r^2} = \frac{I_0}{r^2}, \tag{3.17}$$

gdzie I_0 jest natężeniem światła emitowanego ze źródła.

Fale sferyczne są przypadkami wyidealizowanymi, gdyż ośrodki, w których rozchodzą się fale albo nie są idealnie izotropowe, co zaburza propagację fali, albo źródło nie jest punktowe, a ma pewną rozciągłość przestrzenną. Tym niemniej w szeregu przypadków fale sferyczne są dobrymi przybliżeniami sytuacji rzeczywistej.

Zauważmy na koniec, że obserwując jedynie niewielki wycinek fali sferycznej, np. obserwując ją z odpowiednio dużej odległości, fronty falowe są niemal płaskie, a sama fala sferyczna w takim przypadku może być z dobrym przybliżeniem aproksymowana falą płaską (rysunek 3.2).

3.4 Fale cylindryczne

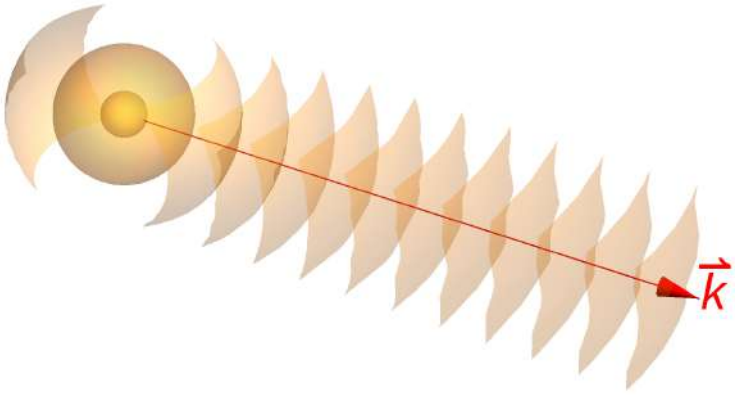
Jeszcze inny wyidealizowany przypadek rozchodzenia się fal dotyczy ośrodków z symetrią cylindryczną. Z tego typu sytuacją mamy do czynienia, gdy źródłem fali elektromagnetycznej jest przewodnik prostoliniowy, przez który płynie prąd przemienny.

Operator Laplace'a we współrzędnych cylindrycznych dany jest zależnością

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\partial^2}{\partial z^2}, \tag{3.18}$$

przy czym, podobnie jak w przypadku układów z symetrią sferyczną, w układach z symetrią cylindryczną wektor pola elektrycznego nie powinien zależeć od współrzędnej θ oraz z

$$E(\mathbf{r}) = E(r, \theta, z) = E(r). \tag{3.19}$$



Rysunek 3.2: Fala sferyczna emitowana z punktowego źródła. Dobrym przybliżeniem fali sferycznej obserwowanej w dużej odległości od źródła jest fala płaska.

Dzięki temu równanie Helmholtza przyjmuje postać

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dE(r)}{dr}\right) + k^2 E(r) = \frac{d^2E(r)}{dr^2} + \frac{1}{r}\frac{dE(r)}{dr} + k^2 E(r) = 0 \tag{3.20}$$

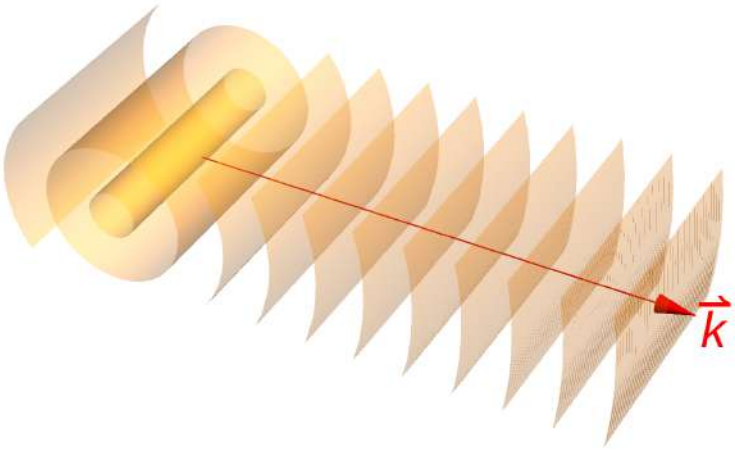
Powyższe równanie jest specjalnym przykładem równania Bessela zerowego rzędu, które można rozwiązać rozwijając funkcję $E(r)$ w szereg potęgowy, a następnie różniczkując kolejne rzędy rozwinięcia, by na koniec porównać je z odpowiednimi potęgami w wyrażeniu nieróżniczkowanym zależnym od k^2. Otrzymana w ten sposób relacja rekurencyjna jest sumą dwóch funkcji Bessela

$$E(r) = C_1 J_0(kr) + C_2 Y_0(kr), \tag{3.21}$$

gdzie J_0 jest funkcją Bessela pierwszego rodzaju, Y_0 jest funkcją Neumanna (zwaną też funkcją Bessela drugiego rodzaju), zaś współczynniki C_1 i C_2 dobierane są w oparciu o warunki brzegowe.

W kontekście fal rozbiegających się od źródła (propagujących na zewnątrz), znacznie wygodniej jest zastosować inną bazę rozwiązań równania Bessela, tzw. funkcje Hankela. Funkcje Hankela pierwszego i drugiego rodzaju definiuje się jako zespolone kombinacje liniowe J_0 i Y_0: $H_0^{(1)}(kr) = J_0(kr) + iY_0(kr)$ oraz $H_0^{(2)}(kr) = J_0(kr) - iY_0(kr)$. Wybór funkcji $H_0^{(1)}(kr)$ w naturalny sposób opisuje falę wybiegającą promieniście od osi (przy przyjętej konwencji czasowej $e^{-i\omega t}$).

Dla dużych odległości od źródła ($kr \gg 1$), funkcję Hankela pierwszego rodzaju można przybliżyć asymptotycznie oscylującą funkcją zespoloną zanikającą



Rysunek 3.3: Fala cylindryczna.

proporcjonalnie do $1/\sqrt{r}$, co pozwala nam zapisać rozwiązanie w postaci

$$E(r) = \frac{E_0^{(+)}}{\sqrt{r}} e^{ikr}. \tag{3.22}$$

Ponieważ równanie to opisuje cylindry, mówimy, że w takim przypadku mamy do czynienia z **falą cylindryczną**.

Po połączeniu wyrażenia na falę cylindryczną (3.22) z funkcją harmoniczną otrzymujemy pełne wyrażenie na amplitudę pola elektrycznego

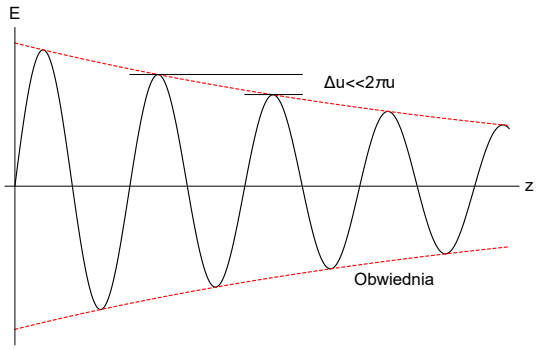
$$E(r,t) = \frac{E_0^{(+)}}{\sqrt{r}} e^{i(kr-\omega t)}. \tag{3.23}$$

Analogiczne rozwiązanie otrzymujemy dla składowej magnetycznej pola

$$B(r,t) = \frac{B_0^{(+)}}{\sqrt{r}} e^{i(kr-\omega t)}. \tag{3.24}$$

Jak pokazuje równanie (3.23), rozwiązaniem równania falowego w takim przypadku są fale o kształcie współosiowych walców (rysunek 3.3), zaś zależność natężenia światła od odległości dana jest przez

$$I_c = |\langle S_c \rangle| = \frac{c\varepsilon}{2} \frac{|E_0|^2}{r} = \frac{I_0}{r}, \tag{3.25}$$



Rysunek 3.4: Zależność przestrzenna pola elektrycznego od położenia w przybliżeniu przyosiowym.

czyli, jak można się było spodziewać natężenie spada liniowo z odległością od źródła.

Podobnie jak w przypadku fali sferycznej, również i w tym przypadku, dla obserwatora oddalonego odpowiednio daleko od źródła fale te mogą być przybliżane falami płaskimi (rysunek 3.3).

3.5 Wiązka gaussowska

3.5.1 Fale przyosiowe i przybliżenie wolnej obwiedni

Jak pokazano na rysunkach 3.2 oraz 3.3, dla odpowiednio dużych odległości od źródła fale sferyczne i cylindryczne mogą być przybliżane przez fale płaskie. Uogólniając to podejście, możemy wprowadzić pojęcie **fal przyosiowych**, których wektory falowe są równoległe do osi optycznej układu (kierunek z). W takim przypadku pole elektryczne fali można zapisać w postaci

$$E^{(+)}(\mathbf{r}) = u(\mathbf{r})e^{ikz}, \tag{3.26}$$

gdzie dla uproszenia pole elektryczne zostało zapisane w postaci skalarnej. W równaniu (3.26) $u(\mathbf{r})$ jest **wolnozmienną obwiednią fali**, podczas gdy e^{ikz} jest nośną tej fali. Wprowadzenie pojęcia wolnozmiennej obwiedni oznacza, że zmiany $u(\mathbf{r})$ na odległościach rzędu długości fali są małe. Opisany przypadek został przedstawiony na rysunku 3.4.

Rozdzielając zmienne równoległe i prostopadłe do kierunku propagacji fali, równanie Helmholtza można zapisać w postaci

$$\left(\nabla_\perp^2 + \frac{\partial^2}{\partial z^2} + i2k\frac{\partial}{\partial z} - k^2 + k^2\right) u(\mathbf{r}) = 0, \tag{3.27}$$

gdzie $\nabla_\perp$ jest poprzecznym operatorem Laplaca danym we współrzędnych kartezjańskich w postaci, $\nabla_\perp = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. Ponieważ zmiany funkcji $u(\mathbf{r})$ są niewielkie na odcinku rzędu długości fali można przyjąć, że $\partial^2 u/\partial z^2 \ll k\partial u/\partial z$,

co jest treścią **przybliżenia wolnozmiennej obwiedni**. W takim przypadku równanie (3.26) przyjmuje postać **równania falowego dla fal przyosiowych**

$$\left(\nabla_\perp^2 + i2k\frac{\partial}{\partial z}\right) u(\mathbf{r}) = 0 \tag{3.28}$$

Fale o parabolicznym froncie falowym

Jednym z najprostszych rozwiązań równania falowego w przybliżeniu przyosiowym (3.28) jest wiązka opisująca paraboliczny front falowy

$$E_p^{(+)}(\mathbf{r}) = E_p^{(+)}(r_\perp, z) = \frac{E_0}{z}\exp\left(ik\frac{r_\perp^2}{2z}\right)\exp(ikz). \tag{3.29}$$

W najprostszy sposób można dowieść tego, że fala paraboliczna rzeczywiście spełnia równanie falowe w przybliżeniu przyosiowym, podstawiając funkcję (3.29) do równania (3.28) i sprawdzając, czy równanie to jest spełnione.

Uwaga!

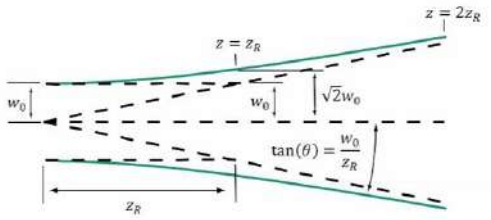
Warto przy tym pamiętać, że najwygodniejszą geometrią dla rozwiązania powyższego problemu jest geometria cylindryczna z pełną symetrią wokół kierunku z. W takim przypadku laplasjan $\nabla_\perp^2$ przyjmuje postać

$$\nabla_\perp^2 = \frac{d^2}{dr_\perp^2} + \frac{1}{r_\perp}\frac{d}{dr_\perp}. \tag{3.30}$$

Wiązka gaussowska

Inne rozwiązanie równania falowego w przybliżeniu przyosiowym można otrzymać korzystając z własności równań różniczkowych, zgodnie z którą jeżeli jakaś funkcja jest rozwiązaniem równania różniczkowego, to również ta sama funkcja przesunięta o wartość stałą jest rozwiązaniem tego równania. Własność tę możemy wykorzystać przesuwając zmienną z o wartość ζ, tj. $z \to z + \zeta$. W przypadku gdy ζ jest liczbą urojoną, $\zeta = -iz_R$, gdzie z_R nazywane jest **długością Rayleigha**, rozwiązanie równania falowego w przybliżeniu przyosiowym nazywane jest **wiązką gaussowską**

$$E_g^{(+)}(\mathbf{r}) = E^{(+)}(r_\perp, z) = \frac{E_0}{z - iz_R}\exp\left[ik\frac{r_\perp^2}{2(z - iz_R)}\right]\exp(ikz). \tag{3.31}$$



Rysunek 3.5: Wiazka gaussowska z zaznaczonymi charakterystycznymi wielkościami: z_R długość Rayleigha, w_0 przewężenie.

Warto zaznaczyć, że ostatni człon powyższego wyrażenia nie został przesunięty o wartość ζ. Stało się tak dlatego, że jest to człon, który nie wynika z rozwiązania równania (3.28), ale pochodzi z nośnej fali wyseparowanej w równaniu (3.26).

Przeanalizujmy teraz rozwiązanie (3.31). Dla $z = 0$ równanie (3.31) przyjmuje postać

$$E_g^{(+)}(\mathbf{r}) = \frac{E_0}{z_R} \exp\left(-\frac{kr_\perp^2}{2z_R}\right) \exp\left[i\left(kz + \frac{\pi}{2}\right)\right]. \tag{3.32}$$

Powyższa funkcja opisuje falą płaską, która, w przeciwieństwie do wcześniejszego przypadku, nie posiada stałej amplitudy (amplituda fali zmniejsza się wraz z oddalaniem się od osi optycznej układu). Analogiczne rozważania możemy przeprowadzić dla $z \gg z_R$. W takim przypadku równanie (3.31) przyjmuje postać

$$E_g^{(+)}(\mathbf{r}) \approx \frac{E_0}{z} \exp(ikz), \tag{3.33}$$

co jest równaniem fali sferycznej. Jak widać, w oparciu o te dwa przypadki, rozwiązanie (3.31) jest rozwiązaniem hybrydowym opisującym jednocześnie dwa reżimy: fale płaskie dla małych z oraz fale sferyczne dla dużych z. Na rysunku 3.5 przedstawiono wygląd wiązki światła w pobliżu $z = 0$.

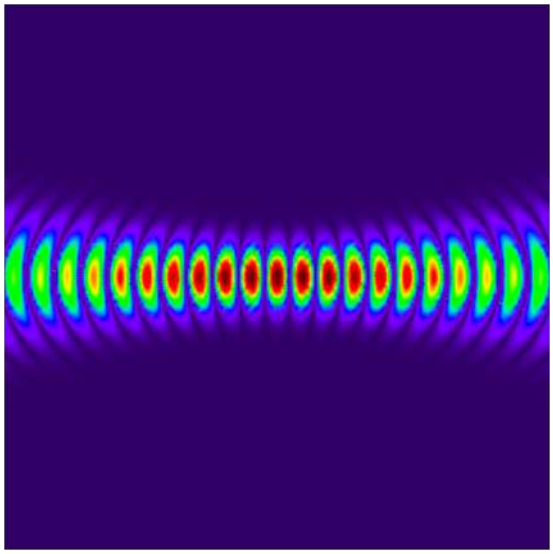
Przekształćmy teraz wyrażenie (3.31) używając podstawienia

$$\frac{1}{z - iz_R} = \frac{1}{R(z)} + i\frac{\lambda}{\pi w^2(z)}, \tag{3.34}$$

gdzie

$$\begin{aligned} R(z) =& z\left[1 + \left(\frac{z_R}{z}\right)^2\right], \\ w^2(z) =& w_0^2\left[1 + \left(\frac{z}{z_R}\right)^2\right]. \end{aligned} \tag{3.35}$$

gdzie $w_0 = \sqrt{\lambda z_R/\pi}$ nazywane jest **promieniem przewężenia wiązki** (rysunek 3.5). Analizując definicję przewężenia, można zauważyć, że dla zaniedbywalnej długości fali, $\lambda \to 0$, przewężenie zmierza do zera, tzn. mamy do



Rysunek 3.6: Chwilowe pole elektryczne wiązki gaussowskiej.

czynienia z punktowym ogniskiem. Jest to zatem reżim optyki geometrycznej. Jednakże kiedy długość fali nie może być zaniedbana, tzn. gdy w grę wchodzą zjawiska falowe, to ognisko staje się przewężeniem o skończonych rozmiarach. W ten sposób manifestuje się zjawisko dyfrakcji, które opisane zostanie później.

Podstawienie zależności (3.35) do równania (3.31) pozwala zapisać je w postaci

$$E_g^{(+)}(\mathbf{r}) = E_0 \frac{w_0}{w(z)} \exp\left(-\frac{r_\perp^2}{w^2(z)}\right) \exp\left[i\left(kz + k\frac{r_\perp^2}{2R(z)} + \xi(z) + \frac{\pi}{2}\right)\right], \tag{3.36}$$

gdzie

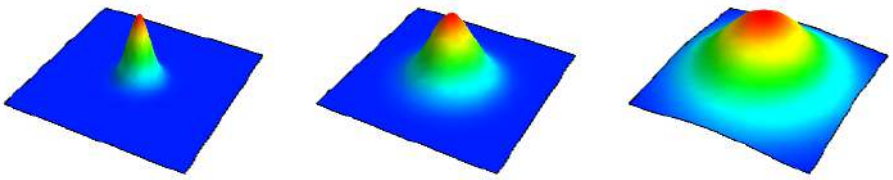
$$\xi(z) = \operatorname{arctg}\left(\frac{z}{z_R}\right) \tag{3.37}$$

nosi nazwę fazy Gouya. Przy wyprowadzaniu tego wzoru skorzystano z zależności $e^{i\operatorname{arctg}(\xi)} = \sqrt{(1+i\xi)/(1-i\xi)}$. Chwilowe pole elektryczne wiązki gaussowskiej pokazane jest na rysunku 3.6.

Analizując równanie (3.36), można w nim zidentyfikować dwa czynniki: czynnik amplitudowy

$$E_0(\mathbf{r}) = E_0(r_\perp, z) = E_0 \frac{w_0}{w(z)} \exp\left(-\frac{r_\perp^2}{w^2(z)}\right), \tag{3.38}$$

który determinuje zachowanie amplitudy wiązki gaussowskiej, oraz czynnik fazowy odpowiadający za fazę fali. W tym drugim można dodatkowo wyróżnić



Rysunek 3.7: Kształt wiązki gaussowskiej dla 3 różnych wartości przewężenia w_0: 0.5, 1, 2 (od lewej).

część radialną $\varphi_r(r_\perp)$ i część podłużną $\varphi_z(z)$ fazy

$$\begin{aligned} \varphi_r(r_\perp, z) &= \exp\left(ik\frac{r_\perp^2}{2R(z)}\right), \\ \varphi_z(z) &= \exp\left[ikz + i\text{arctg}\left(\frac{z}{z_0}\right)\right]. \end{aligned} \tag{3.39}$$

Natężenie wiązki gaussowskiej

Korzystając z postaci amplitudy pola elektrycznego wiązki gaussowskiej [równanie (3.38)] można obliczyć natężenie światła takiej wiązki

$$I_g(\mathbf{r}) = I_g(r_\perp, z) = \frac{c\varepsilon}{2}|E_0(r_\perp, z)|^2 = I_0\left[\frac{w_0}{w(z)}\right]^2 \exp\left(-\frac{2r_\perp^2}{w^2(z)}\right). \tag{3.40}$$

Rysunek 3.7 przedstawia profile natężeniowe wiązki gaussowskej, dla trzech różnych wartości przewężenia w_0. Jak widać, natężenie spada wraz z oddalaniem się od osi optycznej ($r_\perp = 0$) zgodnie z funkcją Gaussa, co notabene jest powodem, dla którego ten typ wiązki jest nazywany wiązka gaussowską.

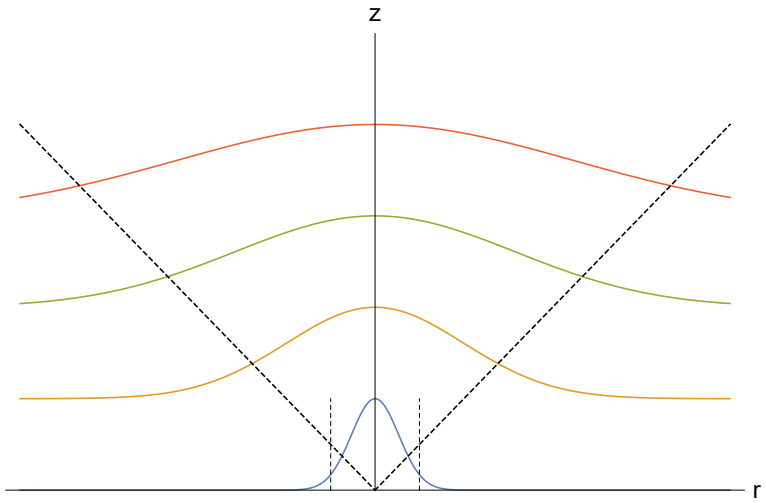
Korzystając z zależności (3.40), można obliczyć całkowitą moc jaką niesie ze sobą taka fala gaussowska

$$P_g = \int_0^\infty I_g(r_\perp, z) 2\pi r_\perp dr_\perp = \frac{I_0\pi}{2}\frac{w_0^2}{w^2(z)}w^2(z) = \frac{I_0}{2}\pi w_0^2. \tag{3.41}$$

Zgodnie z zasadą zachowania energii, jeżeli podczas propagacji nie ma procesów rozpraszania energii, moc całkowita niesiona przez falę nie zmienia się. Dlatego też wraz z oddalaniem się od przewężenia spada maksymalne natężenie fali.

Faza podłużna wiązki gaussowskiej

Jak pokazano w równaniu (3.39), przy propagacji wzdłuż kierunku z, faza zależy od dwóch czynników. Pierwszy z nich, tj. człon $\exp(ikr)$, jest identyczny jak w



Rysunek 3.8: Dyfrakcja wiązki gaussowskiej.

przypadku fali płaskiej i zależy liniowo od współrzędnej z. Drugi z czynników, dany zależnością $\exp[i\mathrm{arctg}(z/z_R)]$ i zwany **opóźnieniem fazy Gouya**, określa ilościowo odstępstwo frontu falowego wiązki gaussowskiej od płaszczyzny. Ponieważ wraz z oddalaniem się od punktu $z = 0$ fala staje się coraz bardziej sferyczna, to zależność ta jest monotoniczna. Efekt opóźnienia Gouya występuje np. dla wiązek które zostały skupione i staje się on szczególnie ważny w przypadku obliczenia częstotliwości rezonansowych w rezonatorach.

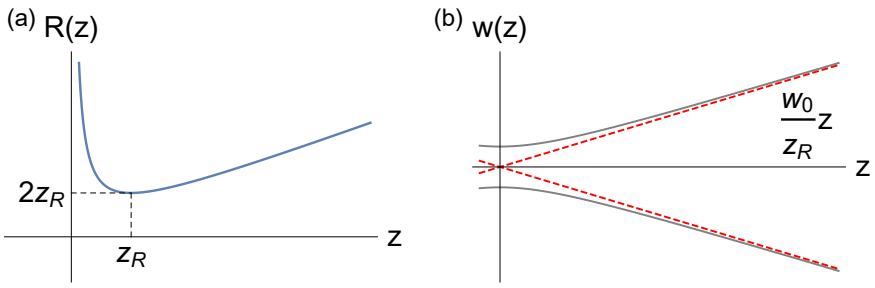
Faza radialna wiązki gaussowskiej

W równaniu (3.39) pokazano, że faza fali gaussowskiej zależy również od współrzędnej poprzecznej $r_\perp$, $\exp[ikr_\perp^2/R(z)]$. Podobnie jak poprzednio, człon ten jest związany z odstępstwem frontu falowego wiązki gaussowsiej od płaszczyzny, a ten jest funkcją położenia, dlatego też czynnik ten jest również zależny od położenia z.

Krzywizna frontu falowego i szerokość wiązki

Do tej pory nie zidentyfikowaliśmy znaczenia parametrów $R(z)$ i $w(z)$ zdefiniowanych równaniami (3.35). Jak się okazuje, czynnik $R(z)$ jest niczym innym jak promieniem krzywizny frontu falowego, która przedstawiona została na rysunku 3.9. W dwóch skrajnych przypadkach krzywizna ta przyjmuje wartości:

$$\begin{aligned} R(z) &\approx \frac{z_R^2}{z} \text{ dla } |z| \ll z_R, \\ R(z) &\approx z \text{ dla } |z| \gg z_R, \end{aligned} \tag{3.42}$$



Rysunek 3.9: Krzywizna frontu falowego (a) oraz szerokość (b) wiązki gaussowskiej.

które odpowiadają fali płaskiej w obszarze przewężenia i fali sferycznej dla odległości znacząco większych od długości Rayleigha. Analizując równanie (3.35), możemy znaleźć minimum krzywizny frontu falowego $R(z_R) = 2z_R$, które osiągane jest w odległości Rayleigha od przewężenia.

Podobne rozważania można przeprowadzić dla szerokości wiązki $w(z)$. Korzystając z równania (3.35), można pokazać, że

$$\begin{aligned} w(z) &= w_0 \text{ dla } z = 0, \\ w(z) &\approx \frac{w_0}{z_0} z = \frac{\theta_0}{z} \text{ dla } |z| \gg z_R, \end{aligned} \tag{3.43}$$

gdzie θ_0 jest kątem rozbieżności wiązki. Zależność szerokości od z pokazana została na rysunku 3.9(b).

Przeprowadzona powyżej analiza dowodzi jednej z najważniejszych własności wiązek gaussowskich – podczas ich propagacji w przestrzeni nie zmienia się profil przestrzenny wiązki, a jedynie zmianie mogą ulec pewne jego parametry.

3.6 Macierze ABCD dla wiązki gaussowskiej

Zdefiniujmy zespolony parametr $q(z) = z - iz_R$ wiązki gaussowskiej, który w pełni opisuje jej stan w danym punkcie z:

$$\frac{1}{q(z)} = \frac{1}{R(z)}_i \frac{\lambda}{\pi w^2(z)}. \tag{3.44}$$

Wielkość ta pozwala na eleganckie opisanie ewolucji wiązki w układach optycznych. Zgodnie z tzw. **prawem ABCD**, jeśli układ optyczny jest opisany macierzą o elementach A, B, C oraz D, to parametry wiązki na wejściu (q_1) i wyjściu (q_2) wiąże zależność:

$$q_2(z) = \frac{Aq_1(z) + B}{Cq_1(z) + D}. \tag{3.45}$$

Prawo to ukazuje niezwykłą użyteczność macierzy ABCD również poza reżimem optyki geometrycznej.

Korzystając z parametru $q(z)$, możemy przepisać równanie wiązki gaussowskiej (3.31) do postaci

$$E^{(+)}(\mathbf{r}) = E_0^{(+)} \frac{q_0}{q(z)} \exp\left[ik \frac{r_\perp^2}{2q(z)}\right] \exp(ikz), \tag{3.46}$$

gdzie $q_0 = q(0) = -iz_R$.

Określmy teraz zmianę parametrów wiązki gaussowskiej w dwóch prostych przypadkach. Zacznijmy od propagacji światła w pustej przestrzeni. Wykorzystując macierz propagacji światła w przestrzeni [równanie (2.24)], dla której $A = 1, B = d, C = 0, D = 1$, możemy łatwo pokazać, że zmiana parametru q, przy propagacji na odległość d, dana jest zależnością

$$q_2(z + d) = q_1(z) + d. \tag{3.47}$$

Rozpisując tę zależność w sposób jawny można pokazać, że

$$z_2 - iz_{R2} = (z_1 + d) - iz_{R1}, \tag{3.48}$$

tak więc zmianie nie ulega długość Rayleigha, a jedynie przesunięciu ulega współrzędna określająca promień krzywizny wiązki, co jest w pełni zgodne z naszą intuicją.

Sytuacja jest odmienna w przypadku przejścia przez soczewkę. W przypadku cienkiej soczewki skupiającej opisanej macierzą (2.50) (gdzie $A = 1, B = 0, C = -1/f, D = 1$), współczynnik q transformuje się dużo łatwiej, gdy analizujemy jego odwrotność:

$$\frac{1}{q_2(z)} = \frac{Cq_1(z) + D}{Aq_1(z) + B} = \frac{-\frac{1}{f}q_1(z) + 1}{q_1(z)} = \frac{1}{q_1(z)} - \frac{1}{f}. \tag{3.49}$$

Rozpisując tę zależność na część rzeczywistą i urojoną, zgodnie z definicją parametru q, natychmiastowo widać, że:

$$\begin{aligned} \frac{1}{R_2(z)} &= \frac{1}{R_1(z)} - \frac{1}{f}, \\ w_2(z) &= w_1(z). \end{aligned} \tag{3.50}$$

Powyższy wynik dowodzi bardzo ważnego faktu: cienka soczewka natychmiastowo zmienia promień krzywizny frontu falowego (zmienia fazę radialną wiązki), ale nie wpływa na samą szerokość wiązki (jej obwiednię) w płaszczyźnie soczewki.

Transformację dla bardzo złożonych układów optycznych można przeprowadzić korzystając z reguły mnożenia macierzy ABCD i podstawienia parametrów wynikowej macierzy bezpośrednio do równania (3.45).

3.7 Wiązki Hermita-Gaussa

Podczas wyprowadzania równania (3.31) opisującego wiązkę gaussowską założono, że rozważany układ cechuje się symetrią walcową. Jednakże w ogólności założenie takie nie musi być prawdziwe, co prowadzi do innej formy wyrażenia na pole elektryczne fali

$$E^{(+)}_{l,m}(\mathbf{r}) = E^{(+)}_g(\mathbf{r}) H_l\left[\frac{\sqrt{2}x}{w(z)}\right] H_m\left[\frac{\sqrt{2}y}{w(z)}\right] \exp[-i(m+n)\xi(z)] \tag{3.51}$$

gdzie $E^{(+)}_g(\mathbf{r})$ jest zależnością pola elektrycznego funkcji gaussowskiej daną równaniem (3.31). Jak widać istnieją dwie podstawowe różnice względem funkcji gaussowskiej. Pierwszą jest obecność wielomianów Hermita n-tego stopnia H_n, drugą człon normalizacyjny pojawiający się przed nimi, zapewniający stałość mocy niezależnie od l i m. Trzecią różnicą jest dodatkowy człon fazowy pojawiający się w opóźnieniu fazowym Gouya. Należy tu podkreślić, że całkowita faza Gouya dla modu wyższego rzędu wynosi w istocie $(l+m+1)\text{arctg}(z/z_R)$, jednak ponieważ funkcja bazowa $E^{(+)}_g(\mathbf{r})$ zawiera już w sobie fundamentalną fazę Gouya (odpowiadającą wartości 1), we wzorze (3.53) dodajemy do niej jedynie dodatkowe opóźnienie fazowe wynikające bezpośrednio z rzędów l oraz m.

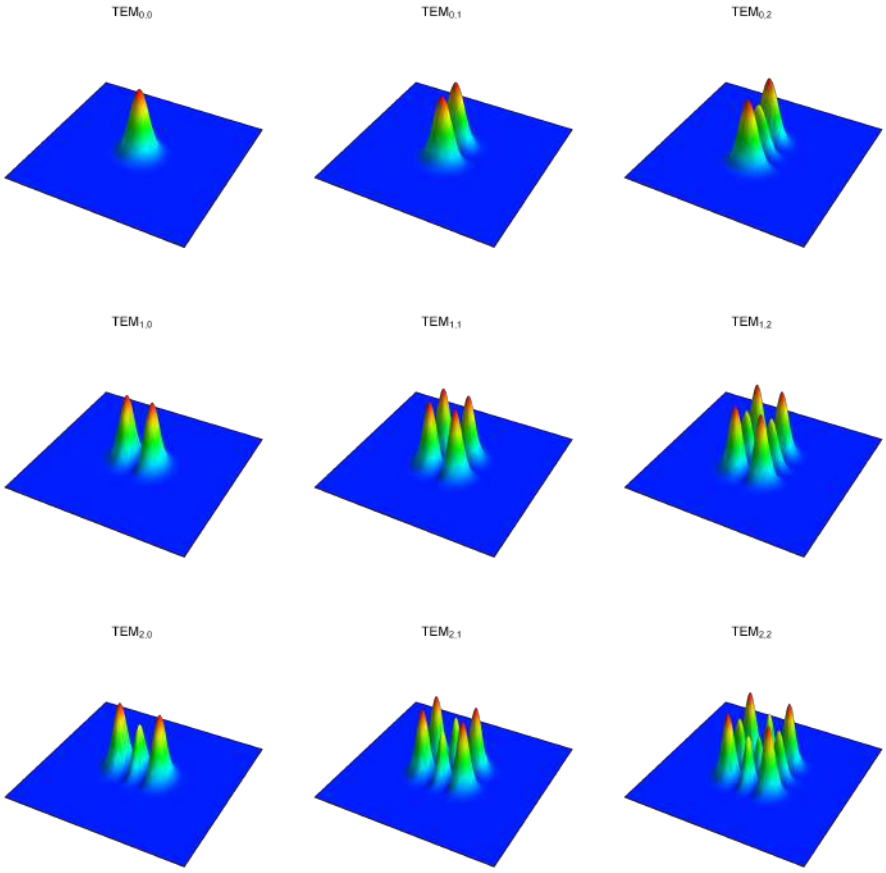
Rozwiązanie (3.53) zwane jest **poprzecznym modem pola elektromagnetycznego TEM**$_{l,m}$. Mod TEM odpowiada funkcji własnej równania Helmholtza w przybliżeniu przyosiowym bez symetrii obrotowej i, podobnie jak w przypadku wiązki gaussowskiej, fala opisana takim modem porusza się w przestrzeni nie zmieniając kształtu, a co najwyżej zmieniając swoje parametry.

Kluczową właściwością modów TEM jest fakt, że zarówno pole elektryczne, jak i magnetyczne tego promieniowania, nie posiadają składowych równoległych do kierunku propagacji. Choć warunek ten wydaje się być oczywisty, tak w końcu jest w przypadku fali płaskiej, sferycznej i cylindrycznej, to okazuje się, że jest on tylko spełniony w kilku bardzo szczególnych przypadkach.

Pojawiające się w równaniu (3.53) wielomiany Hermita n-tego stopnia H_n są stablicowane, a ich postać może zostać podana wprost

$$\begin{aligned}
H_0(x) &= 1, \\
H_1(x) &= 2x, \\
H_2(x) &= 2(2x^2-1), \\
H_n(x) &= (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2} \qquad \text{(jawna relacja)}. \\
H_{n+1}(x) &= 2xH_n(x) - 2nH_{n-1}(x). \qquad \text{(relacja rekurencyjna)}.
\end{aligned} \tag{3.52}$$

Wielomiany Hermita n-tego rzędu charakteryzują się posiadaniem n zer funkcji. W oparciu o tę własność można pokazać, że funkcja TEM$_{0,0}$ ($l=m=0$) opisuje wiązkę gaussowską. Analogicznie TEM$_{l,m}$ odpowiada funkcji posiadającej $l+1$ jasnych maksimów w kierunku x oraz $m+1$ jasnych maksimów w kierunku y.



Rysunek 3.10: Przestrzenny rozkład natężenia kilku pierwszych modów poprzecznych pola elektromagnetycznego TEM.

Warto w końcu zauważyć, że pomimo tego, że wielomiany Hermita rozbiegają się dla dużych wartości $|x|$ lub $|y|$ to obecność funkcji Gaussa w wyrażeniu $E_g(\mathbf{r})$ sprawia, że całe wyrażenie zbiega do zera. Przykłady kilku pierwszych modów TEM zostały przedstawione na rysunku 3.10.

3.8 Wiązki Laguerre'a-Gaussa

Skorzystajmy jeszcze raz z zasady superpozycji i wyznaczmy superpozycję dwóch modów TEM: $\mathrm{TEM}_{l,m}$ oraz $\mathrm{TEM}_{m,l}$. Superpozycja takich modów opisana jest

polem danym w postaci

$$E^{(+)}_{l,m}(\mathbf{r}) = E^{(+)}_{g}(\mathbf{r}) L^{|l|}_{p}\left[\frac{\sqrt{2}x}{w(z)}\right] \exp(il\phi) \exp[(2p+|l|)\xi(z)] \tag{3.53}$$

gdzie l jest całkowite i nieujemne, a m może przyjmować dowolne wartości całkowite (w tym zero). Pojawiające się w równaniu wyrażenie $L^{|m|}_{l}(x)$ jest tzw. stowarzyszonym wielomianem Laguerre'a rzędu l-tego, którego postać dana jest zależnościami

$$\begin{aligned} L^{m}_{0}(x) &= 1, \\ L^{m}_{1}(x) &= -x + m + 1, \\ L^{m}_{2}(x) &= \frac{x^2}{2} - (m+2)x + \frac{(m+2)(m+1)}{2}, \\ L^{m}_{n}(x) &= \frac{z^{-m}e^{z}}{n!}\frac{d^n}{dx^n}\left(e^{-x}x^{n+m}\right), \end{aligned} \tag{3.54}$$

zdefiniowanymi dla $m \geq 0$. W praktyce liczba l determinuje liczbę zer wielomianu Laguerre'a, a co za tym idzie liczbę ciemnych pierścieni jakie posiada wiązka. Podobnie jak w przypadku wiązek kwadratowych, w modach cylindrycznych funkcja bazowa $E^{(+)}_{g}(\mathbf{r})$ niesie ze sobą ładunek fundamentalnej fazy Gouya, dlatego we wzorze (3.53) człon eksponencjalny uzupełnia jedynie fazę do poprawnej wielkości $(2l + |m| + 1)\text{arctg}(z/z_R)$.

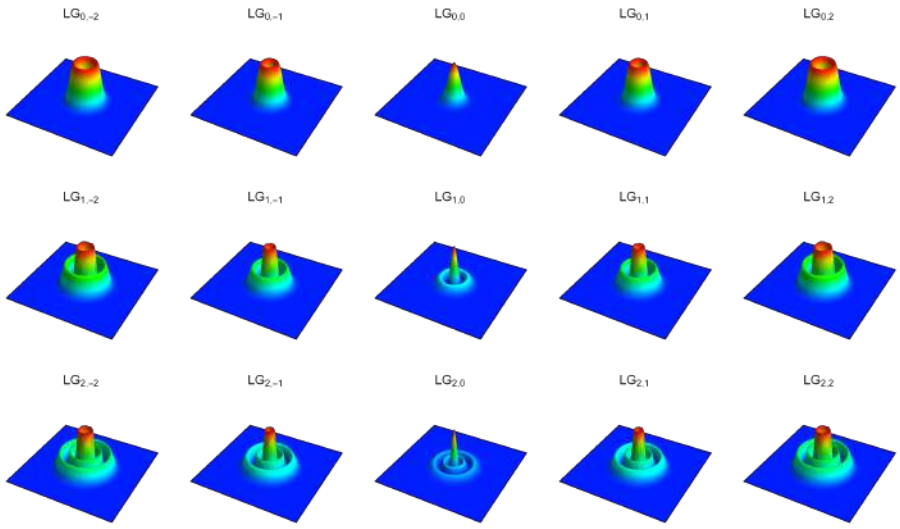
Amplituda pola modu Laguerre'a-Gaussa zależy zarówno od współrzędnej radialnej $r_\perp$, jak i współrzędnej z, ale jest niezależna od współrzędnej azymutalnej ϕ. Od tej współrzędnej zależy tylko faza, która zmienia się wraz z obrotem wokół osi z. W ten sposób mody Laguerre'a-Gaussa posiadają tzw. **ładunek topologiczny**, czyli posiadają tzw. **wir optyczny**, co oznacza, że oprócz momentu pędu wynikającego z polaryzacji posiadają też dodatkowy, tzw. **orbitalny moment pędu**. Posiadanie orbitalnego momentu pędu przez światło powoduje, że zmianie ulegać może jego oddziaływanie z atomami. W szczególności możliwe stają się przejścia pomiędzy poziomami energetycznymi, które nie były możliwe dla wiązek bez wiru.

Ponieważ wiązki Laguerre'-Gaussa są superpozycjami modów TEM, również te wiązki charakteryzują się brakiem składowej podłużnej do kierunku propagacji zarówno dla pola elektrycznego, jak i magnetycznego światła.

Kilka przykładowych rozkładów natężenia światłą wiązki Laguerre'a-Gaussa pokazanych jest na rysunku 3.11.

Wykorzystanie wiązek światła z wirem

Wiązki z niezerowym ładunkiem topologicznym znalazły w ostatnich dekadach szereg fascynujących zastosowań, rewolucjonizując wiele dziedzin nauki i techniki. Jednym z pierwszych przełomów była zaawansowana mikromanipulacja optyczna. Podczas gdy tradycyjne **szczypce optyczne** (wykorzystujące skupione wiązki gaussowskie) pozwalają na przestrzenne



Rysunek 3.11: Przestrzenny rozkład natężenia światła kilku pierwszych modów Laguerre'a-Gaussa.

uwięzienie i przesuwanie mikroskopijnych obiektów, użycie wiązek z orbitalnym momentem pędu pozwala na przekazanie tego momentu materii. W efekcie uwięziona cząstka zostaje wprawiona w precyzyjny ruch obrotowy. Uzyskuje się dzięki temu narzędzie w biofizyce, pozwalające na badanie właściwości mechanicznych DNA czy napędzanie mikrosilników w układach mikroprzepływowych.

Innym, niezwykle obiecującym kierunkiem rozwoju jest telekomunikacja światłowodowa i satelitarna. Ponieważ mody o różnych wartościach ładunku topologicznego są do siebie ściśle ortogonalne, mogą one propagować się w tym samym ośrodku bez wzajemnego zakłócania się (interferencji). Właściwość ta stała się fundamentem tzw. multipleksowania z podziałem przestrzennym. Technika ta pozwala na jednoczesne przesyłanie wielu niezależnych strumieni danych na tej samej długości fali świetlnej, co otwiera drogę do drastycznego i teoretycznie nieograniczonego zwiększenia przepustowości globalnych sieci teleinformatycznych.

Zjawisko wirów optycznych z powodzeniem eksploatuje się również w astronomii i optyce kwantowej. W teleskopach nowej generacji stosuje się tzw. koronagrafy wirowe, tj. instrumenty, które wykorzystują osobliwość fazową w centrum wiązki do precyzyjnego wygaszenia oślepiającego światła odległej gwiazdy, co umożliwia bezpośrednią obserwację krążących wokół niej, słabo świecących egzoplanet. Z kolei w zaawansowanej kryptografii kwantowej fotony niosące orbitalny moment pędu pozwalają na kodowanie informacji w układach wielowymiarowych, co radykalnie zwiększa bezpie-

czeństwo i gęstość przesyłania kwantowych kluczy szyfrujących.

3.9 Podsumowanie rozdziału

W niniejszym rozdziale zaprezentowano przegląd rozwiązań równania falowego dla pola elektromagnetycznego, stanowiących fundament współczesnej optyki klasycznej i fotoniki. Wychodząc od ogólnego ujęcia równania Helmholtza w różnych układach współrzędnych, wprowadzono idealizacje w postaci fal płaskich, sferycznych i cylindrycznych. Szczególną uwagę poświęcono zdefiniowaniu pojęć wektora falowego, frontów falowych oraz związku amplitud pól elektrycznego i magnetycznego, wykazując na ich podstawie jednoznacznie poprzeczny charakter propagującego się zaburzenia.

Kluczowym rozwinięciem tego zagadnienia było wprowadzenie przybliżenia przyosiowego i przybliżenia wolnozmiennej obwiedni, które pozwoliło na matematyczny opis fizycznie realizowalnych wiązek promieniowania. Rozwiązanie zmodyfikowanego równania Helmholtza zaowocowało sformułowaniem pojęcia wiązki gaussowskiej. Omówiono jej fundamentalne parametry przestrzenne – odległość Rayleigha, przewężenie oraz fazę Gouya – które definiują strukturę wiązek emitowanych przez współczesne układy laserowe. Ukazano ponadto, że baza rozwiązań może być skutecznie rozszerzana o wyższe mody poprzeczne, takie jak wiązki Hermita-Gaussa i Laguerre'a-Gaussa. Wprowadzenie tych ostatnich, związanych z występowaniem orbitalnego momentu pędu i zjawiskiem wiru optycznego, sygnalizuje istnienie zaawansowanych stopni swobody pola promieniowania, które otwierają zupełnie nowe perspektywy w kontekście kontroli materii i kodowania informacji. Całość omówionego tu formalizmu falowego, oprócz opisania samej formy propagacji, stanowi niezbędne podglebie do późniejszego, rygorystycznego zrozumienia zjawisk dyfrakcji i propagacji w reżimach, w których przybliżenie optyki promieni zawodzi, o czym przeczytasz w dalszych rozdziałach podręcznika.

Zadania

3.1. Porównaj falę monochromatyczną opisaną zespoloną amplitudą $E_0^{(+)}$ z falą monochromatyczną o tej samej częstotliwości, ale sprzężonej amplitudzie, biorąc pod uwagę natężenie, kształt i kierunek wektora falowego. Do obliczeń użyj:

(a) fali płaskiej $E(\mathbf{r}) = E_0^{(+)} \exp[ik(x+y)/\sqrt{2}]$,

(b) fali sferycznej $E(\mathbf{r}) = \frac{E_0^{(+)}}{r} \exp(ikr)$.

3.2. Zakładając układ o symetrii cylindrycznej, wyprowadź równanie fali cylindrycznej. Przy wyprowadzaniu zależności skorzystaj z operatora Laplace'a zapisanego we współrzędnych cylindrycznych.

3.3. Udowodnij, że funkcja paraboliczna spełnia równanie falowe w przybliżeniu przyosiowym.

3.4. Wyprowadź funkcję opisującą wiązkę gaussowską wykorzystującą pojęcie przewężenia i fazę Gouya [równanie (3.36)] wychodząc z zależności jawnie wykorzystującej pojęcie długości Rayleigha [równanie (3.31)].

3.5. Przyjmując, że laser emituje wiązkę światła o profilu gaussowskim, oblicz rozmiar wiązki w odległości 10 km od przewężenia, zakładając, że średnica wiązki w przewężeniu wynosi 0,5 cm. Do obliczeń przyjmij długość fali $\lambda = 700$ nm.

3.6. Laser emituje światło o profilu gaussowskim i długości fali λ, gdzie przewężenie w_0 obserwowane jest na wyjściu z lasera, które oświetla ekran umieszczony w odległości d od lasera.

(a) W jakiej odległości należy ustawić ekran, by krzywizna docierającego do niego frontu falowego była minimalna, i jaki promień (liczony jako odległość od środka dla punktu, w którym natężenie spada $1/e^2$ razy) będzie mieć wtedy ta wiązka?

(b) Jaką wartość liczbową przyjmuje ten promień dla światła o długości fali $\lambda = 200\pi$ nm, odległości ekranu $d = 10$ cm oraz przewężeniu $w_0 = 0,1$ mm.

3.7. Korzystając z zależności

$$\frac{1}{q(z)} = \frac{1}{R(z)} + i\frac{\lambda}{\pi w^2(z)}$$

wyprowadź wzory na $R(z)$ i $w^2(z)$.

3.8. Promień lasera ma profil gaussowski o średnicy 2 mm i całkowitą moc 10 mW. Oblicz powierzchniową gęstość mocy (natężenie) i amplitudę fali elektrycznej w centrum wiązki oraz w punkcie oddalonym od centrum o 1 mm.

3.9. Powszechnie stosowana technika pomiaru promienia wiązki gaussowskiej polega na zasłanianiu jej przesłoną z ostrą krawędzią (np. żyletką) prostopadle do kierunku rozchodzenia się wiązki oraz na pomiarze mocy pozostałej (niezasłoniętej) wiązki. Zgodnie z zasadą 10-90, pomiaru położenia żyletki dokonuje się, gdy transmitowana moc równa jest 10% i 90% całkowitej mocy wiązki. W takim przypadku promień wiązki $w(z)$ w punkcie pomiaru dany jest zależnością:

$$w(z) = \alpha|x_{10\%} - x_{90\%}|,$$

gdzie α jest pewnym współczynnikiem. Wyznacz wartość współczynnika α.

3.10. Wyznacz rozkład natężenia światła powstałego przez nałożenie na siebie dwóch wiązek Hermita-Gaussa $\mathrm{TEM}_{1,0}$ i $\mathrm{TEM}_{0,1}$ o jednakowych natężeniach (wiązki te mają niezależne i losowe fazy, więc ich natężenia dodają się bez interferencji). Wyznacz całkowitą moc, jaką niesie ze sobą ta wiązka.

3.11. Wszystkie mody Hermite'a-Gaussa charakteryzowane tym samym przewężeniem w_0 tworzą bazę zupełną. Oznacza to, że dowolny rozkład pola $E^{(+)}$ może być przedstawiony jako nieskończona suma modów Hermite'a-Gaussa:

$$E^{(+)}(r_\perp, 0) = \sum_{l,m=0}^{\infty} c_{l,m} E^{(+)}_{l,m}(r_\perp, 0),$$

gdzie $E^{(+)}_{l,m}$ oznacza rozkład odpowiadający modowi $\mathrm{TEM}_{l,m}$. Przyjmując, że mody Hermite'a-Gaussa są ortogonalne, zaproponuj sposób obliczenia współczynników $c_{l,m}$.

3.12. Rozważ wiązki Hermite'a-Gaussa $\mathrm{TEM}_{l,m}$ o długości Rayleigha $z_R = 30$ cm w próżni. Określ częstotliwości w zakresie $\nu = 10^{14} \pm 2 \times 10^9$ Hz, dla których opóźnienia fazowe między punktami $z = -z_R$ oraz $z = z_R$ umieszczonymi na osi wiązki są całkowitą wielokrotnością π. Co stanie się z tymi punktami wraz z oddalaniem się od osi wiązki?

Rozdział 4

Wzory Fresnela

4.1 Zachowanie pola elektrycznego i magnetycznego na granicy ośrodków

Rozważmy dwa ośrodki dielektryczne opisane stałymi dielektrycznymi ε_1 i ε_2 oraz przenikalnościami magnetycznymi μ_1 i μ_2. Na granicy pomiędzy tymi ośrodkami dochodzi do transformacji składowych pola elektrycznego i magnetycznego zgodnie z tzw. **warunkami brzegowymi**. Aby wyprowadzić te warunki, możemy przywołać całkowe postacie prawa Gaussa dla pola elektrycznego i magnetycznego w dielektrykach:

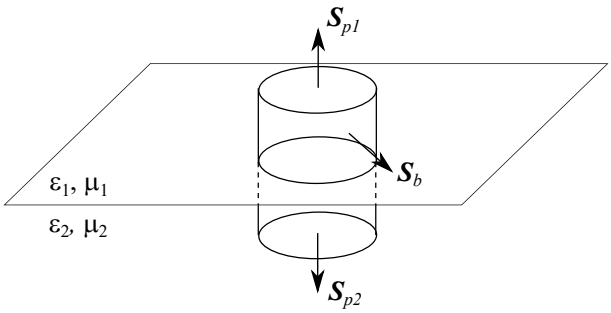
$$\begin{aligned} \oint_S \mathbf{B} \cdot \mathrm{d}\mathbf{S} &= 0, \\ \oint_S \mathbf{D} \cdot \mathrm{d}\mathbf{S} &= 0, \end{aligned} \tag{4.1}$$

gdzie całkowanie odbywa się po powierzchni zamkniętej.

Wyobraźmy sobie walec, którego podstawy o polu powierzchni S_p znajdują się w dwóch ośrodkach, zaś powierzchnia boczna ma pole S_b (Rys. 4.1). W takim przypadku równanie (4.1) można przepisać do postaci:

$$\begin{aligned} \mathbf{B}_1 \cdot \mathbf{S}_{p1} + \mathbf{B}_{12} \cdot \mathbf{S}_b + \mathbf{B}_2 \cdot \mathbf{S}_{p2} &= 0, \\ \mathbf{D}_1 \cdot \mathbf{S}_{p1} + \mathbf{D}_{12} \cdot \mathbf{S}_b + \mathbf{D}_2 \cdot \mathbf{S}_{p2} &= 0. \end{aligned} \tag{4.2}$$

Załóżmy teraz, że wysokość tego walca dąży do zera, tak by dwie podstawy znajdowały się nieskończenie blisko siebie, ale jednak w różnych ośrodkach. W takim przypadku powierzchnia boczna walca jest równa zero, $S_b = 0$, i drugi wyraz w obu wyrażeniach znika. Zauważmy ponadto, że obie podstawy walca mają takie same pola, ale ich normalne skierowane są w przeciwnych kierunkach, tzn. $\mathbf{S}_{p1} = -\mathbf{S}_{p2}$. Podstawiając tę zależność do równań (4.2), otrzymujemy warunki transformacji dla składowej normalnej indukcji pola elektrycznego i



Rysunek 4.1: Walec rozciągający się pomiędzy dwoma ośrodkami dielektrycznymi, pozwalający na wyprowadzenie transformacji składowej normalnej indukcji pola elektrycznego i magnetycznego.

magnetycznego:

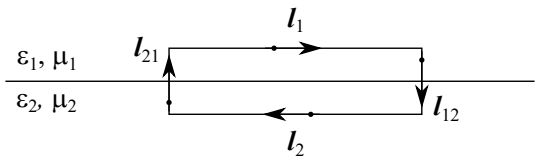
$$\begin{aligned} D_{1n} &= D_{2n}, \\ B_{1n} &= B_{2n}, \end{aligned} \tag{4.3}$$

gdzie indeks n oznacza składową normalną do granicy. Należy w tym miejscu wyraźnie podkreślić, że ciągłość składowej normalnej wektora indukcji pola elektrycznego ($D_{1n} = D_{2n}$) wynika wprost z założenia o braku ładunków swobodnych na granicy obu dielektryków (powierzchniowa gęstość ładunku $\sigma_{\text{swob}} = 0$). Gdyby na granicy ośrodków celowo zgromadzono ładunek elektryczny, prawo Gaussa wymagałoby uwzględnienia go w całkowitym strumieniu. Doprowadziłoby to do nieciągłości, a sam skok wartości składowej normalnej byłby proporcjonalny do gęstości tego ładunku ($D_{2n} - D_{1n} = \sigma_{\text{swob}}$). W przypadku standardowych zagadnień optycznych, gdzie badamy propagację fali w obojętnych elektrycznie dielektrykach, przyjmuje się jednak, że $\sigma_{\text{swob}} = 0$.

W celu wyprowadzenia warunków brzegowych dla składowych stycznych pola elektrycznego i magnetycznego skorzystamy z całkowej postaci prawa Faradaya i prawa Ampère'a:

$$\begin{aligned} \oint_L \mathbf{E} \cdot \mathrm{d}\mathbf{l} &= 0, \\ \oint_L \mathbf{H} \cdot \mathrm{d}\mathbf{l} &= 0, \end{aligned} \tag{4.4}$$

gdzie po raz kolejny założyliśmy, że w ośrodku nie ma swobodnych ładunków, zaś same pola nie zmieniają się w czasie. W analogii do poprzednich rozważań tworzymy krzywą zamkniętą w postaci prostokąta, którego dwa boki są równoległe do granicy ośrodków, dwa zaś są do niej prostopadłe (Rys. 4.2). Podobnie jak w przypadku powyższych obliczeń, rozbijamy powyższy prostokąt na cztery boki i zakładamy, że długość boków prostopadłych do granicy między ośrodkami zmierza do zera, $l_{12} \to 0$. Uwzględniając, że kierunki wektorów związanych z



Rysunek 4.2: Prostokątny obrys wykorzystany jako krzywa zamknięta rozciągnięta między dwoma ośrodkami dielektrycznymi do wyprowadzenia warunków brzegowych dla natężenia pola elektrycznego i magnetycznego.

odcinkami równoległymi do granicy ośrodków są przeciwne, ale ich długości są identyczne, $\mathbf{l}_1 = -\mathbf{l}_2$, otrzymujemy

$$\begin{aligned} \mathbf{E}_1 \cdot \mathbf{l}_1 - \mathbf{E}_2 \cdot \mathbf{l}_1 = 0, \\ \mathbf{H}_1 \cdot \mathbf{l}_1 - \mathbf{H}_2 \cdot \mathbf{l}_1 = 0, \end{aligned} \tag{4.5}$$

co po wykonaniu iloczynu skalarnego daje wzory na składowe styczne natężenia pola elektrycznego i magnetycznego:

$$\boxed{\begin{aligned} E_{1t} = E_{2t}, \\ H_{1t} = H_{2t}, \end{aligned}} \tag{4.6}$$

gdzie indeks t oznacza składową styczną do granicy.

Równania (4.3) oraz (4.6) są równaniami brzegowymi dla odpowiednich składowych natężenia i indukcji pola elektrycznego i magnetycznego, które pozwolą nam na analizę zachowania fali elektromagnetycznej na granicy dwóch ośrodków.

4.2 Transformacja pola elektrycznego światła na granicy dwóch ośrodków

Załóżmy, że na granicę pomiędzy dwoma dielektrykami pada światło opisane polem E_p, które podlega odbiciu i załamaniu, dzieląc się na dwie wiązki opisane odpowiednio polami E_o i E_z. Fale te możemy zapisać jako fale płaskie dane w ogólnej postaci równaniami (dla uproszczenia pominięto fazę początkową):

$$\begin{aligned} E_p &= E_{p,0}^{(+)} e^{i(\mathbf{k}_p \cdot \mathbf{r} - \omega_p t)}, \\ E_z &= E_{z,0}^{(+)} e^{i(\mathbf{k}_z \cdot \mathbf{r} - \omega_z t)}, \\ E_o &= E_{o,0}^{(+)} e^{i(\mathbf{k}_o \cdot \mathbf{r} - \omega_o t)}. \end{aligned} \tag{4.7}$$

gdzie

$$\begin{aligned} k_p &= \frac{\omega_p}{v_1}, \\ k_o &= \frac{\omega_o}{v_1}, \\ k_z &= \frac{\omega_z}{v_2}, \end{aligned} \tag{4.8}$$

oraz $E_{i,0}^{(+)}$ jest amplitudą zespoloną pola elektrycznego odpowiedniej wiązki światła, przy czym $i \in \{p, z, o\}$.

4.3 Stałość częstotliwości fali

Z równań brzegowych składowej stycznej natężenia pola elektrycznego możemy zapisać dla światła następujące zależności:

$$\left[E_{p,0}^{(+)} e^{i(\mathbf{k}_p \cdot \mathbf{r} - \omega_p t)} - E_{o,0}^{(+)} e^{i(\mathbf{k}_o \cdot \mathbf{r} - \omega_o t)}\right]_t = \left[E_{z,0}^{(+)} e^{i(\mathbf{k}_z \cdot \mathbf{r} - \omega_z t)}\right]_t . \tag{4.9}$$

Powyższe równanie prawdziwe jest dla dowolnej chwili czasu t. Aby warunek ten był spełniony, musi zachodzić zależność

$$\omega_p t = \omega_o t = \omega_z t, \tag{4.10}$$

skąd ostatecznie mamy, że

$$\boxed{\omega_p = \omega_o = \omega_z \equiv \omega.} \tag{4.11}$$

Z równania (4.11) wynika, że zarówno fala padająca, odbita, jak i fala załamana mają tę samą częstotliwość. Zależność ta dowodzi stałości częstotliwości światła w każdym ośrodku, w którym rozchodzi się światło.

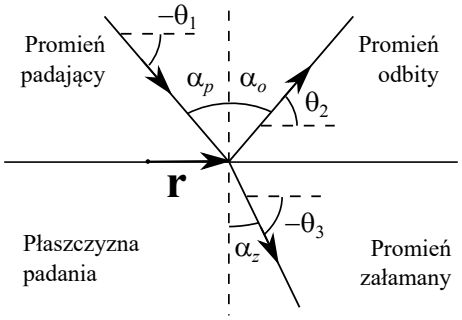
4.4 Przestrzenna orientacja promienia padającego, odbitego i załamanego

Zajmijmy się teraz kierunkami rozchodzenia się wiązki padającej, odbitej i załamanej. Prowadząc analizę analogiczną do tej przeprowadzonej powyżej dla częstości, można pokazać, że pomiędzy określonymi wyrażeniami w wykładniku opisującym fale zachodzą relacje

$$\mathbf{k}_p \cdot \mathbf{r} = \mathbf{k}_o \cdot \mathbf{r} = \mathbf{k}_z \cdot \mathbf{r}. \tag{4.12}$$

Na potrzeby rozważań wybierzmy układ współrzędnych w taki sposób, aby wektor wodzący $\mathbf{r}$ leżał w płaszczyźnie granicy ośrodków i był on prostopadły do wektora falowego fali padającej $\mathbf{k}_p$, tzn.

$$\mathbf{k}_p \cdot \mathbf{r} = 0. \tag{4.13}$$



Rysunek 4.3: Odbicie od granicy dwóch dielektryków z zaznaczoną falą padającą, odbitą i załamaną.

Łącząc tę zależność z równaniem (4.12), otrzymujemy również, że dla fali odbitej i załamanej spełniona jest zależność

$$\mathbf{k}_o \cdot \mathbf{r} = \mathbf{k}_z \cdot \mathbf{r} = 0. \tag{4.14}$$

Równanie (4.14) dowodzi, że fala padająca, załamana i odbita leżą w tej samej płaszczyźnie zwanej **płaszczyzną padania**.

4.5 Prawo odbicia i załamania

Wybierzmy teraz układ współrzędnych w taki sposób, by jego środek należał do płaszczyzny padania, zaś sam wektor $\mathbf{r}$ był styczny do granicy ośrodków $\mathbf{r} = r\hat{t}$ (Rys. 4.3). Analizując zależności geometryczne, można zapisać

$$\begin{aligned}
\mathbf{k}_p \cdot \mathbf{r} &= \mathbf{k}_p \cdot \hat{t}r = k_p r\cos(-\theta_1) = k_p r\cos\theta_1 = k_p r\cos(\pi/2-\alpha_p) = k_p r\sin\alpha_p,\\
\mathbf{k}_o \cdot \mathbf{r} &= \mathbf{k}_o \cdot \hat{t}r = k_o r\cos\theta_2 = k_o r\cos(\pi/2-\alpha_o) = k_o r\sin\alpha_o,\\
\mathbf{k}_z \cdot \mathbf{r} &= \mathbf{k}_z \cdot \hat{t}r = k_z r\cos(-\theta_3) = k_z r\cos\theta_3 = k_z r\cos(\pi/2-\alpha_z) = k_z r\sin\alpha_z,
\end{aligned} \tag{4.15}$$

gdzie przyjęto konwencję, że kąty dane są w wartościach dodatnich (stąd znak minus przy kątach θ_1 oraz θ_3).

Z przytoczonej już wcześniej zależności mamy

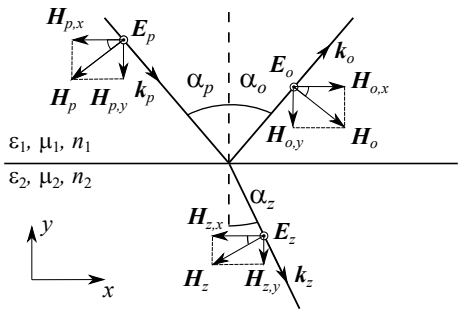
$$k_p r\sin\alpha_p = k_o r\sin\alpha_o = k_z r\sin\alpha_z, \tag{4.16}$$

co po wstawieniu $k = \omega/v$ daje

$$\frac{\sin\alpha_p}{v_1} = \frac{\sin\alpha_o}{v_1} = \frac{\sin\alpha_z}{v_2}. \tag{4.17}$$

Z równania (4.17) otrzymujemy zarówno prawo odbicia

$$\boxed{\alpha_p = \alpha_o,} \tag{4.18}$$



Rysunek 4.4: Zachowanie składowych elektrycznych i magnetycznych fali na granicy dwóch ośrodków w sytuacji, w której wektor pola elektrycznego prostopadły jest do płaszczyzny padania (polaryzacja s).

jak również prawo załamania

$$\frac{\sin\alpha_p}{\sin\alpha_z} = \frac{v_1}{v_2}, \tag{4.19}$$

które można także zapisać w postaci

$$n_1 \sin\alpha_p = n_2 \sin\alpha_z. \tag{4.20}$$

4.6 Wzory Fresnela

Dowolny stan polaryzacji światła może być przedstawiony jako superpozycja fal liniowych drgających w kierunkach prostopadłych z odpowiednimi fazami. Dlatego też niezależnie od liczby możliwości, z jakimi mamy do czynienia przy analizie procesu padania na granicę dwóch ośrodków, wystarczy rozważyć tylko dwa przypadki, w których wektor pola elektrycznego (polaryzacja liniowa światła) jest prostopadły (**polaryzacja** s) albo równoległy (**polaryzacja** p) do płaszczyzny padania. Poniżej przedstawiamy analizę tych dwóch przypadków.

4.6.1 Wektor E prostopadły do płaszczyzny padania (polaryzacja s)

Rysunek 4.4 przedstawia orientacje przestrzenne wektorów pól elektrycznego i magnetycznego fali padającej, odbitej i załamanej w sytuacji, w której kierunek wektora pola elektrycznego prostopadły jest do płaszczyzny padania. Z warunków ciągłości dla składowej stycznej (w rozważanej sytuacji skierowanej wzdłuż osi x) pola magnetycznego mamy

$$-H_{p,x} + H_{o,x} = -H_{z,x}. \tag{4.21}$$

Korzystając z zależności pomiędzy natężeniem H i indukcją B pola magnetycznego w ośrodkach izotropowych ($B = \mu_i H$, gdzie μ_i jest bezwzględną przenikalnością magnetyczną) oraz z relacji trygonometrycznych, można zapisać

$$-\frac{B_p}{\mu_1}\cos\alpha_p + \frac{B_o}{\mu_1}\cos\alpha_o = -\frac{B_z}{\mu_2}\cos\alpha_z. \tag{4.22}$$

Wiedząc, że $B = E/v$ oraz korzystając z prawa odbicia, $\alpha_p = \alpha_o$, równanie (4.22) można zapisać jako

$$\frac{E_p - E_o}{\mu_1 v_1}\cos\alpha_p = \frac{E_z}{\mu_2 v_2}\cos\alpha_z, \tag{4.23}$$

co po wstawieniu $v = c/n$ daje

$$\frac{n_1(E_p - E_o)}{\mu_1}\cos\alpha_p = \frac{n_2 E_z}{\mu_2}\cos\alpha_z. \tag{4.24}$$

Wykorzystując przywołane już kilkukrotnie równanie ciągłości dla natężenia pola elektrycznego trzech fal ($E_p + E_o = E_z$), można z równania (4.23) wyrugować albo pole elektryczne fali załamanej E_z, albo odbitej E_o.

Zdefiniujmy teraz współczynniki, zwane **amplitudowymi współczynnikami Fresnela**, które charakteryzować będą stosunki amplitud fal odbitej lub załamanej do fali padającej

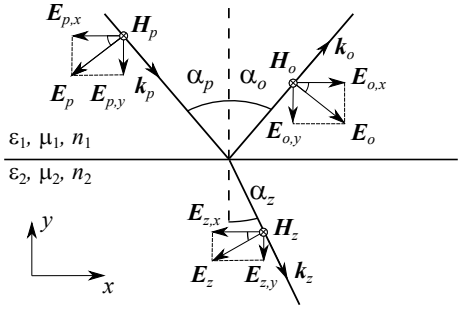
$$\begin{aligned} r &= \frac{E_o}{E_p}, \\ t &= \frac{E_z}{E_p}. \end{aligned} \tag{4.25}$$

Korzystając z tych definicji, można w rozważanym przypadku zdefiniować **amplitudowe współczynniki Fresnela dla składowej prostopadłej** fali odbicia i załamania:

$$\boxed{\begin{aligned} r_\perp &= \left(\frac{E_o}{E_p}\right)_\perp = \frac{\frac{n_1}{\mu_1}\cos\alpha_p - \frac{n_2}{\mu_2}\cos\alpha_z}{\frac{n_1}{\mu_1}\cos\alpha_p + \frac{n_2}{\mu_2}\cos\alpha_z}, \\ t_\perp &= \left(\frac{E_z}{E_p}\right)_\perp = \frac{2\frac{n_1}{\mu_1}\cos\alpha_p}{\frac{n_1}{\mu_1}\cos\alpha_p + \frac{n_2}{\mu_2}\cos\alpha_z}. \end{aligned}} \tag{4.26}$$

W ośrodkach niemagnetycznych, tj. w sytuacji, w której $\mu_1 \approx \mu_2 \approx \mu_0$, amplitudowe współczynniki Fresnela dla składowej prostopadłej przyjmują postać

$$\boxed{\begin{aligned} r_\perp &= \frac{n_1\cos\alpha_p - n_2\cos\alpha_z}{n_1\cos\alpha_p + n_2\cos\alpha_z}, \\ t_\perp &= \frac{2n_1\cos\alpha_p}{n_1\cos\alpha_p + n_2\cos\alpha_z}. \end{aligned}} \tag{4.27}$$



Rysunek 4.5: Zachowanie wektorów pola elektrycznego i magnetycznego na granicy dwóch ośrodków w sytuacji, w której wektor pola elektrycznego leży w płaszczyźnie padania (polaryzacja p).

4.6.2 Wektor E równoległy do płaszczyzny padania (polaryzacja p)

Rozważmy teraz sytuację, w której wektor pola elektrycznego fali świetlnej leży w płaszczyźnie padania. Aby jednoznacznie zdefiniować amplitudowe współczynniki Fresnela, musimy przyjąć określoną konwencję znaków. W niniejszym opracowaniu stosujemy układ spopularyzowany m.in. przez E. Hechta, w którym dodatnie zwroty wektorów pola elektrycznego fali padającej i odbitej są określone tak, by przy normalnym padaniu (kąt 0°) ich rzuty na płaszczyznę granicy pokrywały się z konwencją dla polaryzacji s. Powoduje to, że w granicy normalnego padania, z powodów geometrycznych, zachodzi relacja $r_{\parallel} = -r_{\perp}$.

Warto w tym miejscu odnotować, że w starszej, klasycznej literaturze (np. M. Born, E. Wolf *Principles of Optics*) stosowana jest odmienna konwencja, w której wektor fali odbitej jest odwrócony. Skutkuje to odwrotnym znakiem w liczniku współczynnika odbicia. Należy mieć świadomość tych różnic podczas korzystania z różnych źródeł.

Z warunków brzegowych dla stycznej składowej pola elektrycznego można w naszej konwencji zapisać

$$-E_p \cos\alpha_p + E_o \cos\alpha_p = -E_z \cos\alpha_z. \tag{4.28}$$

Dla składowej magnetycznej światła otrzymujemy zaś zależność

$$\frac{B_p}{\mu_1} + \frac{B_o}{\mu_1} = \frac{B_z}{\mu_2}. \tag{4.29}$$

Na podstawie tych równań, w sposób analogiczny do opisanego powyżej można

wyprowadzić **amplitudowe wzory Fresnela dla składowej równoległej**

$$\begin{aligned} r_{\parallel} &= \left(\frac{E_o}{E_p}\right)_{\parallel} = \frac{\frac{n_2}{\mu_2}\cos\alpha_p - \frac{n_1}{\mu_1}\cos\alpha_z}{\frac{n_2}{\mu_2}\cos\alpha_p + \frac{n_1}{\mu_1}\cos\alpha_z}, \\ t_{\parallel} &= \left(\frac{E_z}{E_p}\right)_{\parallel} = \frac{2\frac{n_1}{\mu_1}\cos\alpha_p}{\frac{n_2}{\mu_2}\cos\alpha_p + \frac{n_1}{\mu_1}\cos\alpha_z}, \end{aligned} \tag{4.30}$$

które w ośrodkach niemagnetycznych ($\mu_i \approx \mu_0$) przyjmują postać

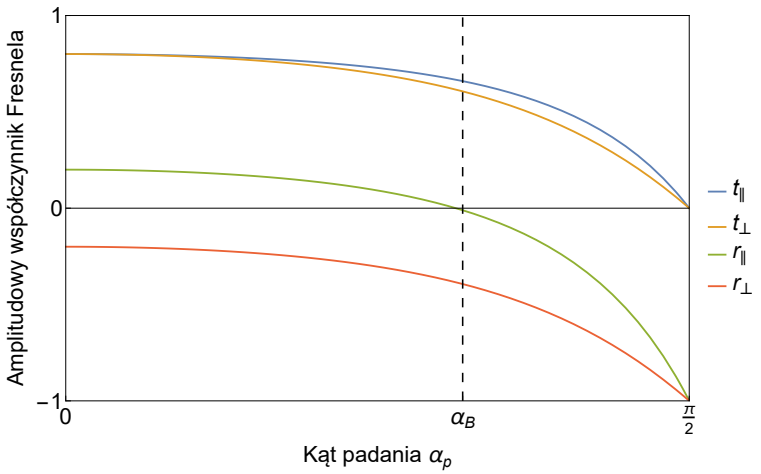
$$\begin{aligned} r_{\parallel} &= \frac{n_2\cos\alpha_p - n_1\cos\alpha_z}{n_2\cos\alpha_p + n_1\cos\alpha_z}, \\ t_{\parallel} &= \frac{2n_1\cos\alpha_p}{n_2\cos\alpha_p + n_1\cos\alpha_z}. \end{aligned} \tag{4.31}$$

4.6.3 Zależność współczynników Fresnela od kąta padania

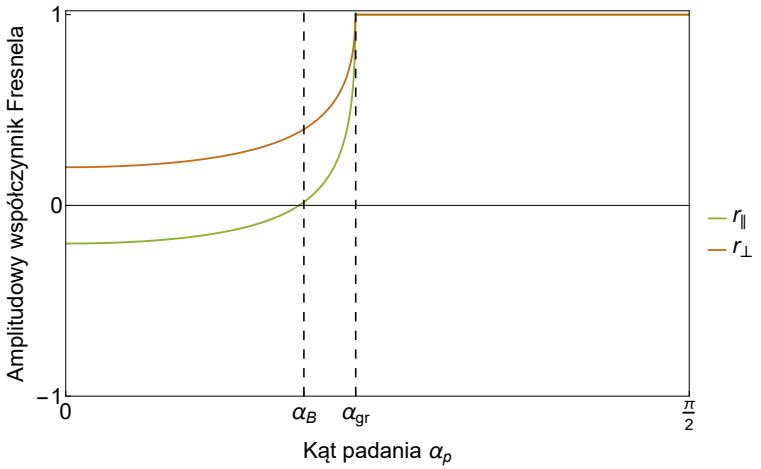
W oparciu o wyprowadzone powyżej zależności możliwe jest wyznaczenie amplitudowych i fazowych relacji dla fali odbitej i załamanej na granicy dwóch ośrodków. W praktyce można rozpatrywać zarówno przypadek, w którym światło załamane wychodzi z ośrodka rzadszego optycznie (o niższym współczynniku załamania n_1) do ośrodka gęstszego optycznie (o wyższym współczynniku załamania n_2), jak i odwrotnie.

Rysunek 4.6 przedstawia zależność amplitudowych współczynników Fresnela w funkcji kąta padania α_p, przy czym współczynnik załamania pierwszego ośrodka n_1 jest mniejszy od współczynnika załamania drugiego ośrodka n_2 ($n_2 > n_1$). Jak pokazano na wykresach, w takim przypadku zarówno współczynniki transmisji, jak również współczynniki odbicia zmieniają się wraz ze wzrostem kąta padania w sposób monotoniczny, przy czym wartość współczynnika odbicia dla składowej prostopadłej jest ujemna dla wszystkich wartości kąta padania. Oznacza to, że w przypadku odbicia od ośrodka gęstszego optycznie następuje skok fazy fali o π. W przypadku współczynnika dla fali odbitej o polaryzacji równoległej $r_{\parallel}$, współczynnik ma początkowo wartość dodatnią i maleje wraz ze wzrostem kąta. Dla pewnego specjalnego kąta α_B, zwanego **kątem Brewstera**, współczynnik ten osiąga wartość równą zero. Oznacza to, że dla tego kąta prostopadła składowa światła nie jest odbijana od granicy ośrodków. Jeśli chodzi o światło transmitowane widać, że istnieją dwa szczególne kąty dla których amplitudowe współczynniki transmisji składowej prostopadłej $t_{\perp}$ i równoległej $t_{\parallel}$ są identyczne. Kątami tymi są 0 oraz $\pi/2$.

Podobne zależności jak dla odbicia od ośrodka gęstszego optycznie można wyprowadzić dla odbicia od ośrodka rzadszego optycznie. Zależność amplitudowych współczynników odbicia od kąta padania przedstawia Rys. 4.7. W takim przypadku współczynniki odbicia $r_{\perp}$ i $r_{\parallel}$ cechują się podobną monotoniczną zależnością jak we wcześniejszym przypadku. Pierwszy z tych współczynników



Rysunek 4.6: Zależność amplitudowych współczynników Fresnela od kąta dla fali odbitej i załamanej. Do odbicia dochodzi na granicy pomiędzy powietrzem ($n_1 = 1$) i szkłem ($n_2 = 1.5$).



Rysunek 4.7: Zależność amplitudowych współczynników Fresnela od kąta dla fali odbitej. Do odbicia dochodzi na granicy pomiędzy szkłem ($n_1 = 1,5$) i powietrzem ($n_2 = 1$).

$r_\perp$ rośnie wraz ze wzrostem kąta, podczas gdy drugi, początkowo osiągający wartości ujemne, rośnie osiągając w końcu wartość zerową dla α_B. Jak widać również w takim przypadku światło odbite od granicy dwóch ośrodków może być całkowicie spolaryzowane. Dla jeszcze większych kątów współczynnik $r_\parallel$ rośnie wraz z kątem osiągając wartość 1 dla kąta granicznego α_{gr}. Dla kątów większych od kąta granicznego dochodzi do zjawiska całkowitego wewnętrznego odbicia.

4.6.4 Kąt Brewstera

Powróćmy jeszcze do sytuacji, w której kąt padania jest kątem Brewstera, w którym to przypadku współczynnik odbicia składowej równoległej równy jest 0. Oznacza to, że dla kąta Brewstera fala odbita od granicy ośrodków jest całkowicie spolaryzowana w kierunku prostopadłym do płaszczyzny padania. Warto w tym miejscu jednak zaznaczyć, że polaryzacja fali odbitej nie oznacza, że fala transmitowana do ośrodka jest spolaryzowana. Owszem, zmienia się jej polaryzacja, ale wciąż jest w niej obecna składowa prostopadła, jak i równoległa.

Używając definicji współczynnika odbicia dla składowej równoległej $r_\parallel$ oraz korzystając z prawa Snella, równanie na współczynnik $r_\parallel$ można przepisać do postaci

$$r_\parallel = \frac{\mathrm{tg}(\alpha_p - \alpha_z)}{\mathrm{tg}(\alpha_p + \alpha_z)}. \tag{4.32}$$

Z powyższej postaci równania można pokazać, że $r_\parallel = 0$ osiągane jest dla $\alpha_p + \alpha_z = \pi/2$. Oznacza to, że dla kąta Brewstera kąt pomiędzy wiązką odbitą i załamaną jest prosty. Pozwala nam to wyznaczyć wartość kąta Brewstera, korzystając z równania Snella:

$$\frac{\sin\alpha_B}{\sin(\pi/2 - \alpha_B)} = \frac{\sin\alpha_B}{\cos\alpha_B} = \mathrm{tg}\alpha_B = \frac{n_2}{n_1}, \tag{4.33}$$

która dana jest przez

$$\alpha_B = \mathrm{arctg}\left(\frac{n_2}{n_1}\right). \tag{4.34}$$

Co ciekawe, z polaryzacją fali na skutek odbicia pod kątem Brewstera mamy do czynienia zarówno w przypadku, w którym padanie odbywa się od ośrodka rzadszego optycznie, jak i w przypadku padania od strony ośrodka gęstszego optycznie.

Fizyczny powód tego zjawiska ujawnia się niezwykle elegancko w ujęciu mikroskopowym. Kiedy fala elektromagnetyczna wnika do drugiego ośrodka, wymusza ona oscylacje dipoli elektrycznych (chmur elektronowych atomów i cząsteczek) w tym materiale. Oscylujące dipole stają się wtórnymi źródłami promieniowania. Zależności geometryczne sprawiają, że dla polaryzacji p przy padaniu pod kątem Brewstera, oś oscylacji dipoli wewnątrz drugiego ośrodka pokrywa się dokładnie z kierunkiem propagacji hipotetycznej fali odbitej. Ponieważ klasyczny dipol nie wypromieniowuje energii wzdłuż własnej osi drgań,

fala odbita o tej polaryzacji po prostu nie może powstać. Cała energia składowej p jest w tym specyficznym ujęciu transmitowana.

Możliwość polaryzacji światła poprzez odbicie ma szereg ważnych zastosowań, stanowiąc doskonały przykład pomostu między teorią elektromagnetyzmu a praktyczną inżynierią optyczną. Jednym z najważniejszych jest możliwość wykorzystania tego typu polaryzacji do wytwarzania światła spolaryzowanego w laserach.

Dygresja

Znakomitym praktycznym wykorzystaniem kąta Brewstera są tzw. okienka Brewstera, montowane na końcach rur wyładowczych w laserach gazowych (np. w popularnych laserach hel-neonowych). Umieszczenie tych szklanych okienek pod kątem Brewstera względem osi optycznej rezonatora sprawia, że składowa promieniowania spolaryzowana równolegle do płaszczyzny padania (polaryzacja p) przechodzi przez nie bez najmniejszych strat na odbicie ($R_{\parallel} = 0$). Z kolei składowa prostopadła (polaryzacja s) jest przy każdym przejściu przez okienko częściowo odbijana i wyprowadzana poza rezonator. W efekcie, po wielu obiegach światła wewnątrz wnęki rezonansowej, straty dla polaryzacji s stają się na tyle duże, że wzmacniana jest wyłącznie bezstratna składowa p. Dzięki temu laser emituje silnie spolaryzowaną liniowo wiązkę światła, unikając jednocześnie konieczności stosowania dodatkowych, stratnych polaryzatorów wewnątrz układu.

4.6.5 Całkowite wewnętrzne odbicie i fala zanikająca

Przeprowadzone już wcześniej rozważania dotyczące zjawiska Snella wykazały, że przy padaniu na granicę dwóch ośrodków od strony ośrodka gęstszego optycznie istnieje pewien kąt graniczny α_{gr}, dla którego promień ulega całkowitemu wewnętrznemu odbiciu. Okazuje się jednak, że choć fala odbija się od granicy tych ośrodków, to w drugim ośrodku, tuż przy granicy ośrodków, pojawia się pewne niezerowe pole elektryczne. Żeby tego dowieść rozważmy falę padającą na granicę dwóch dielektryków o współczynnikach załamania n_1 i n_2 (Rys. 4.8). Ogólna postać wektora falowego w drugim ośrodku dana jest zależnością

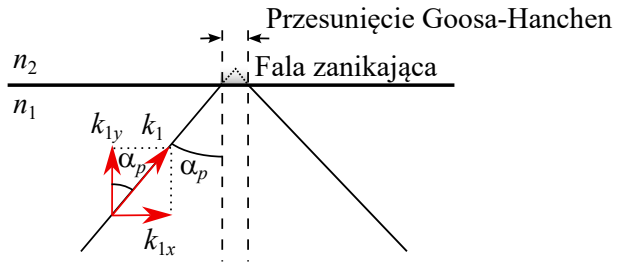
$$k_2 = \sqrt{k_{2x}^2 + k_{2y}^2}. \tag{4.35}$$

Korzystając z faktu, że $k = \omega n/c$ możemy zapisać związek pomiędzy wektorami falowymi oraz współczynnikami załamania w określonych ośrodkach

$$\frac{k_1}{n_1} = \frac{k_2}{n_2}. \tag{4.36}$$

Wykorzystując tę samą zależność możemy także zapisać prawo Snella w nieco innej postaci

$$k_1 \sin \alpha_p = k_2 \sin \alpha_z. \tag{4.37}$$



Rysunek 4.8: Zależności geometryczne pól dla fali zanikającej.

W końcu posługując się Rys. 4.8 możemy pokazać geometryczne relacje

$$k_{1,2,x} = k_{1,2} \sin \alpha_{p,z}. \tag{4.38}$$

Łącząc te relacje możemy przepisać równanie (4.35) do postaci

$$\frac{n_2}{n_1} k_1 = \sqrt{k_{1x}^2 + k_{2y}^2}. \tag{4.39}$$

Proste operacje algebraiczne pozwalają przepisać powyższe równanie do postaci

$$k_{2y}^2 = \left(\frac{n_2}{n_1}\right)^2 k_1^2 - k_{1x}^2, \tag{4.40}$$

skąd dla kątów $\alpha_p > \alpha_{gr}$ można zapisać wyrażenie na poprzeczną do granicy ośrodków składową wektora falowego w drugim ośrodku k_{2y}

$$k_{2y}^2 = -\left(\frac{n_2}{n_1}\right)^2 k_1^2 \left(\frac{n_1^2}{n_2^2} \sin^2 \alpha_p - 1\right). \tag{4.41}$$

Warto w tym miejscu zaznaczyć, że w rozważanych warunkach kwadrat składowej wektora falowego jest wielkością ujemną. Dlatego sama składowa może zostać zapisana w postaci

$$k_{2y} = \pm i\beta, \tag{4.42}$$

gdzie

$$\beta = n_2 \frac{\omega}{c} \sqrt{\frac{n_1^2}{n_2^2} \sin^2 \alpha_p - 1}. \tag{4.43}$$

Dygresja

W przypadku fali zanikającej, obecność trzeciego ośrodka w zasięgu fali zanikającej może spowodować zmianę jej zachowania. W szczególności może dojść do przenikania promieniowania między tymi ośrodkami, mimo że kąt padania na granicę pierwszego z nich i obszaru otaczającego go jest większy

od kąta całkowitego wewnętrznego odbicia. Zjawisko to nosi nazwę **udaremnionego całkowitego wewnętrznego odbicia**. Co więcej, transmisja energii przez tę przerwę jest optycznym odpowiednikiem znanego z mechaniki kwantowej zjawiska tunelowania przez barierę potencjału. Warto zauważyć, że tunelowanie to odbywa się niezwykle szybko.

Doświadczenie Transmisja światła pomiędzy dwoma pryzmatami mikrofalowymi rozdzielonymi szczeliną powietrzną.

Jak widać istnieją dwa rozwiązania na składową k_{2y}. Pierwsze z nich, tj. rozwiązanie proporcjonalne do wyrażenia $e^{\beta y}$, należy jednak odrzucić jako niefizyczne, gdyż w tym przypadku pole elektryczne musiałoby rosnąć do nieskończoności, co wiązałoby się z nieskończoną energią pojawiającą się w drugim ośrodku. Drugie z rozwiązań jest fizyczne i przyjmuje postać

$$E_2 = E_{20} e^{-\beta y} e^{i(k_{2x}x - \omega t)}. \tag{4.44}$$

Rozwiązanie to opisuje falę biegnącą wzdłuż granicy ośrodków (kierunek x), która, choć penetruje do drugiego ośrodka, to jest to wnikanie z eksponencjalnie zanikającą amplitudą wraz z odległością od granicy. Fala taka nazywana jest **falą zanikającą**, a do drugiego ośrodka wnika ona na głębokość porównywalną z długością fali.

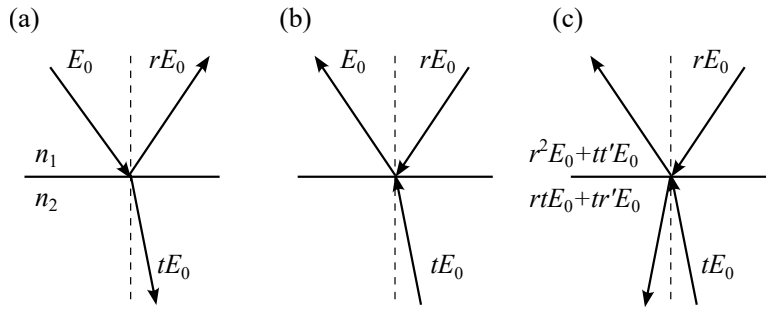
Warto zauważyć, że w rozpatrywanym układzie nie ma transportu energii netto przez granicę ośrodków, tzn. w kierunku osi y uśredniony wektor Poyntinga równy jest zeru, a strumień energii skierowany jest docelowo równolegle do granicy ośrodków (wzdłuż osi x).

Jedną z konsekwencji propagacji wzdłuż powierzchni granicznej jest istnienie tzw. **przesunięcia Goosa-Hänchena** (Herman Goos 1883-1968, Hilda Hänchen 1919-2013). Polega ono na mikroskopijnym, poprzecznym rozsunięciu punktu padania i punktu odbicia na granicy ośrodków. Zjawisko to wynika bezpośrednio z faktu, że choć uśredniony transport energii w głąb drugiego ośrodka wynosi zero, pole elektromagnetyczne przedostaje się i transportuje energię wzdłuż granicy w obszarze fali zanikającej, zanim ostatecznie powróci do pierwszego ośrodka jako promień odbity.

Dygresja

Zjawisko fali zanikającej znalazło genialne zastosowanie we współczesnej biologii molekularnej i medycynie pod postacią mikroskopii TIRF (*Total Internal Reflection Fluorescence*). Kiedy badamy komórki pod standardowym mikroskopem fluorescencyjnym, światło lasera wzbudza świecenie cząsteczek w całej objętości preparatu, tworząc tło rozmywające szczegóły.

W mikroskopie TIRF promień lasera kierowany jest na szkiełko pod kątem większym niż kąt całkowitego wewnętrznego odbicia. Światło ulega odbiciu, ale na styku szkiełka i badanej komórki powstaje fala zanikająca, która wnika do próbki na głębokość zaledwie 100–200 nanometrów. Dzięki temu wzbudzane są wyłącznie te fluorofory, które znajdują się tuż przy samej błonie komórkowej. Pozwala to na niezwykle ostre, pozbawione szumów z tła obrazowanie procesów życiowych, takich jak uwalnianie neuro-



Rysunek 4.9: Zachowanie promienia świetlnego na granicy dwóch ośrodków dla przypadku rzeczywistego (a), przypadku „odwróconego w czasie" (b) oraz przypadku, w którym wiązki padają tak jak w przypadku „odwróconym w czasie", ale rozważane są wszystkie przypadki odbicia i załamania (c). Na rysunkach zaznaczono amplitudy wszystkich rozpatrywanych wiązek.

przekaźników w synapsach czy interakcje wirusów z powierzchnią komórki. Zjawisko opisane przez Fresnela na początku XIX wieku pozwala dziś dosłownie podglądać molekularne sekrety życia w czasie rzeczywistym.

4.6.6 Relacje Stokesa

W równaniu falowym zależność od czasu pojawia się w formie drugiej pochodnej po funkcji falowej. Oznacza to, że odbicie w czasie, tzn. zmiana $t \to -t$, nie powoduje zmiany równania falowego. W ten sposób manifestuje się jedna z podstawowych zasad Przyrody mówiąca o tym, że prawa optyki, a szerzej prawa elektromagnetyzmu, są **niezmiennicze względem symetrii odbicia w czasie**.

Jednym z przykładów manifestacji symetrii odbicia w czasie są tzw. **równania Stokesa**. Żeby je wyprowadzić, rozważmy wiązkę światła padającą na granicę dwóch dielektryków (Rys. 4.9). Zgodnie z prowadzonymi do tej pory rozważaniami, część wiązki ulegnie odbiciu od tej granicy, a część wniknie do drugiego ośrodka (bez utraty ogólności zakładamy, że wnikamy z ośrodka o niższym współczynniku załamania do ośrodka gęstszego optycznie, $n_1 < n_2$). Amplituda światła odbitego i załamanego determinowana jest przez amplitudowe współczynniki odbicia r i załamania t, tzn. $E_z = tE_0$ oraz $E_o = rE_0$. Zgodnie z niezmienniczością czasową, analogiczne zachowanie zachodzić powinno w odwrotnym przypadku. Wówczas dwie fale o amplitudach E_z i E_o, padając na granicę dwóch ośrodków, łączą się w pojedynczą falę o amplitudzie E_0 obserwowaną tylko w ośrodku pierwszym (kluczowe w tym przypadku jest zjawisko konstruktywnej i destruktywnej interferencji opisane w następnym rozdziale).

Rozważmy teraz odbicie i załamanie każdej z wiązek w „obróconym czasowo" przypadku. Rozchodząca się w ośrodku fala o amplitudzie rE_0 ulega odbiciu i załamaniu, tworząc wiązki o amplitudach r^2E_0 oraz rtE_0. Od drugiej strony na granicę pada fala o amplitudzie tE_0, która rozdziela się na dwie wiązki o

amplitudach $tr'E_0$ oraz $tt'E_0$, gdzie prim oznacza zachowanie fali padającej od strony ośrodka drugiego. Łącząc ten przypadek z przypadkiem opisanym wcześniej, możemy zapisać relacje pomiędzy amplitudami fal powstałych w dwóch ośrodkach

$$\begin{aligned} r^2 E_0 + tt' E_0 &= E_0, \\ rtE_0 + tr' E_0 &= 0. \end{aligned} \tag{4.45}$$

Wydzielając pierwsze równanie przez pole E_0 oraz drugie równanie przez tE_0 otrzymujemy następujące zależności

$$\boxed{\begin{aligned} r^2 + tt' &= 1, \\ r &= -r', \end{aligned}} \tag{4.46}$$

zwane wzorami Stokesa. Relacje te stanowić mogą ciekawe uzupełnienie wyprowadzonych wcześniej wzorów Fresnela. Można na ich podstawie np. pokazać, że o ile amplituda odbitej od granicy ośrodków fali nie zależy od tego czy padanie zachodzi od strony ośrodka optycznie gęstszego czy rzadszego, o tyle gdy odbicie zachodzi od ośrodka o wyższym współczynniku załamania n, to dochodzi do zmiany fazy fali, podczas gdy w drugim przypadku takiej zmiany nie ma.

4.6.7 Reflektancja i transmitancja

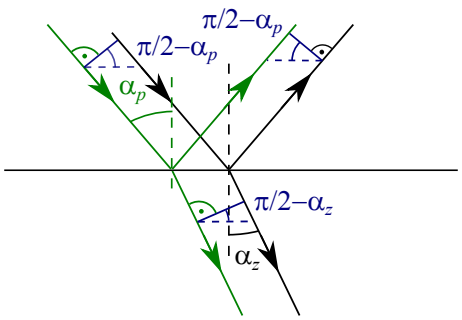
Zgodnie ze wcześniejszymi rozważaniami, oscylacja pola elektrycznego światła jest tak szybka, że wielkością, która jest mierzona w praktyce, jest natężenie światła. Dlatego też warto wprowadzić wielkości, które opisują stosunek mocy wiązki transmitowanej i odbitej od granicy dwóch ośrodków do mocy wiązki padającej.

Z zasady zachowania energii wiadomo, że suma mocy wiązki transmitowanej i odbijanej musi być równa mocy wiązki padającej. Przy założeniu braku przepływu energii między dwiema polaryzacjami, zasadę tę można rozpatrywać niezależnie dla składowej prostopadłej (polaryzacja s) i równoległej (polaryzacja p) do płaszczyzny padania

$$\begin{aligned} P_{p,\perp} &= P_{o,\perp} + P_{z,\perp}, \\ P_{p,\parallel} &= P_{o,\parallel} + P_{z,\parallel}. \end{aligned} \tag{4.47}$$

Współczynnik R, który opisuje stosunek mocy fali odbitej do padającej, zwany jest **reflektancją** i może on być powiązany z wyprowadzonymi wcześniej amplitudowymi współczynnikami Fresnela

$$\begin{aligned} R_\parallel &= \left(\frac{P_o}{P_p}\right)_\parallel = \left(\frac{A_o I_o}{A_p I_p}\right)_\parallel = \left(\frac{I_o}{I_p}\right)_\parallel = \left(\frac{|E_o|^2}{|E_p|^2}\right)_\parallel, \\ R_\perp &= \left(\frac{P_o}{P_p}\right)_\perp = \left(\frac{A_o I_o}{A_p I_p}\right)_\perp = \left(\frac{I_o}{I_p}\right)_\perp = \left(\frac{|E_o|^2}{|E_p|^2}\right)_\perp, \end{aligned} \tag{4.48}$$



Rysunek 4.10: Zmiana kierunku propagacji światła na skutek ugięcia fali i związana z nią zmiana pola powierzchni wiązki.

gdzie wykorzystaliśmy fakt, że pola powierzchni wiązki padającej i odbitej są takie same ($A_o = A_p$). Wykorzystując powyższe równanie można ostatecznie zapisać

$$\boxed{\begin{aligned} R_{\parallel} &= |r_{\parallel}|^2, \\ R_{\perp} &= |r_{\perp}|^2. \end{aligned}} \tag{4.49}$$

W oparciu o równania (4.48) można obliczyć całkowite natężenie fali odbitej, które równe będzie

$$I_o = I_{o,\parallel} + I_{o,\perp} = (R_{\parallel} + R_{\perp}) I_p = R I_p, \tag{4.50}$$

gdzie $R = R_{\parallel} + R_{\perp}$.

Prowadząc analogiczne rozważania możemy obliczyć współczynniki określające stosunek mocy wiązki załamanej i padającej, zwane **transmitancją**. W tym przypadku należy jednak zwrócić uwagę, że na skutek zmiany kierunku propagacji światła w drugim ośrodku, przy załamaniu zmienia się powierzchnia przekroju wiązki. Zmianę tę można obliczyć prowadząc proste analizy geometryczne (Rys. 4.10):

$$d_p \sin\left(\frac{\pi}{2} - \alpha_p\right) = d_z \sin\left(\frac{\pi}{2} - \alpha_z\right), \tag{4.51}$$

gdzie d_p i d_z są wymiarami poprzecznymi odpowiednio wiązki padającej i załamanej. W oparciu o tę zależność można pokazać, że podczas załamania rozmiar wiązki transmitowanej zmienia się $\cos\alpha_z / \cos\alpha_p$-razy.

Ostatni brakujący element konieczny do wyliczenia transmitancji związany jest z natężeniem światła w ośrodku dielektrycznym niebędącym próżnią. Przywołując definicję natężenia światła dla współrzędnych zespolonych [równanie (1.40)], korzystając z równań materiałowych pomiędzy indukcją i natężeniem pola magnetycznego oraz z faktu, że $B_0 = E_0/v$, gdzie v jest prędkością światła w danym ośrodku, możemy przepisać natężenie światła do postaci

$$I = \frac{\varepsilon v}{2} |E_0|^2 = \frac{\varepsilon_0 \varepsilon_r c}{n} |E_0|^2, \tag{4.52}$$

co po wstawieniu $n = \sqrt{\varepsilon_r \mu_r} \approx \sqrt{\varepsilon_r}$ (dla ośrodków niemagnetycznych) daje

$$I = \frac{n\varepsilon_0 c}{2}|E_0|^2. \tag{4.53}$$

Prowadząc obliczenia analogiczne do tych przeprowadzonych dla reflektancji oraz uwzględniając zmianę parametrów współczynnika załamania przy zmianie ośrodków otrzymujemy

$$\begin{aligned} T_\parallel &= \left(\frac{P_z}{P_p}\right)_\parallel = \left(\frac{A_z I_z}{A_p I_p}\right)_\parallel = \left(\frac{\cos\alpha_z I_z}{\cos\alpha_p I_p}\right)_\parallel = \left(\frac{n_2 \cos\alpha_z |E_z|^2}{n_1 \cos\alpha_p |E_p|^2}\right)_\parallel, \\ T_\perp &= \left(\frac{P_z}{P_p}\right)_\perp = \left(\frac{A_z I_z}{A_p I_p}\right)_\perp = \left(\frac{\cos\alpha_z I_z}{\cos\alpha_p I_p}\right)_\perp = \left(\frac{n_2 \cos\alpha_z |E_z|^2}{n_1 \cos\alpha_p |E_p|^2}\right)_\perp, \end{aligned} \tag{4.54}$$

gdzie A_z i A_p są powierzchniami prostopadłymi do kierunku propagacji odpowiednio dla wiązki załamanej i padającej światła. Korzystając z amplitudowych współczynników Fresnela, powyższe równania można przepisać do ostatecznej postaci:

$$\boxed{\begin{aligned} T_\parallel &= \frac{n_2 \cos\alpha_z}{n_1 \cos\alpha_p}|t_\parallel|^2, \\ T_\perp &= \frac{n_2 \cos\alpha_z}{n_1 \cos\alpha_p}|t_\perp|^2. \end{aligned}} \tag{4.55}$$

Z postaci równania widać, że transmitancja nie jest równa kwadratowi modułu amplitudowego współczynnika transmisji, ale uwzględnia również stosunek impedancji ośrodków oraz przekrojów geometrycznych obu wiązek.

Podobnie jak w przypadku wiązki odbitej, możemy zapisać, że całkowite natężenie wiązki załamanej jest równe

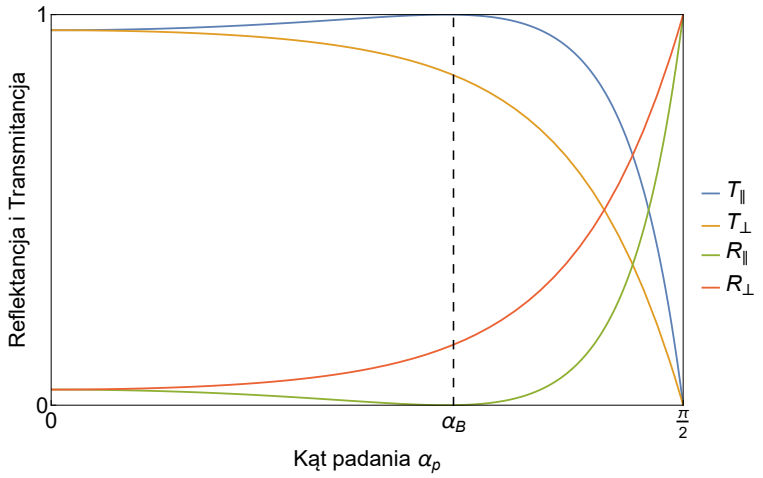
$$I_z = T_\parallel I_p + T_\perp I_p = (T_\parallel + T_\perp) I_p = T I_p, \tag{4.56}$$

gdzie $T = T_\parallel + T_\perp$.

Co ciekawe, transmitancja może być również obliczona korzystając z zasady zachowania energii. W przypadku gdy w ośrodku nie ma strat, zasada ta przyjmuje postać:

$$\boxed{\begin{aligned} R_\parallel + T_\parallel &= 1, \\ R_\perp + T_\perp &= 1. \end{aligned}} \tag{4.57}$$

Rysunek 4.11 przedstawia zależność reflektancji i transmitancji dla polaryzacji równoległej i prostopadłej do płaszczyzny padania w funkcji kąta, dla padania na granicę dwóch ośrodków od strony ośrodka rzadszego optycznie ($n_2 > n_1$). Jak można się było spodziewać, składowa równoległa do płaszczyzny padania ma pewną niezerową wartość dla padania wzdłuż normalnej do powierzchni granicznej i monotonicznie maleje osiągając zero dla kąta Brewstera. W takim przypadku, fala o tej polaryzacji jest całkowicie transmitowana, co manifestuje



Rysunek 4.11: Zależność reflektancji i transmitancji fali dwóch składowych polaryzacyjnych w funkcji kąta padania.

się jako $T_\parallel = 1$. Dla kątów większych od kąta Brewstera współczynnik odbicia fali rośnie i osiąga wartość równą jedności dla kątów bliskich padaniu ślizgowemu (wzdłuż granicy). Dla porównania, reflektancja składowej prostopadłej wiązki odbitej rośnie monotonicznie wraz z kątem, osiągając wartość maksymalną równą jedności. W takim też przypadku zależność transmitancji jest odwrotna.

4.7 Odbicie na granicy dielektryk-przewodnik

4.7.1 Równanie falowe w przewodnikach

O ile do tej pory rozważaliśmy falę elektromagnetyczną w dielektrykach, w których nie ma swobodnych ładunków, o tyle warto się teraz zastanowić, jak będzie wyglądało zachowanie fali na granicy pomiędzy dielektrykiem a przewodnikiem. Aby to zrobić, musimy się odwołać do pełnej formy równań Maxwella. Istnienie swobodnych ładunków sprawia, że wykorzystywane do tej pory równanie Gaussa dla pól elektrycznych musi zostać uzupełnione o niezerową gęstość ładunków ρ, zaś prawo Ampère'a musi zostać zmodyfikowane przez dodanie do prawej strony gęstości prądu $\mathbf{j}$ pomnożoną przez przenikalność magnetyczną ośrodka μ. Na potrzeby rozważań będziemy musieli również skorzystać z prawa Ohma w postaci

$$\mathbf{j} = \sigma \mathbf{E}, \tag{4.58}$$

gdzie σ jest konduktywnością materiału.

Z rozważań elektrostatycznych łatwo wywnioskować, że skutkiem oddziaływania pola elektrycznego na swobodne ładunki jest ruch tych ostatnich. W

podobny sposób odpowiadać będą te ładunki na pole zmienne, przy czym w przypadku fal optycznych o częstościach setek teraherców, pomiędzy ruchem harmonicznym ładunków a oscylującym polem elektrycznym pojawia się przesunięcie fazy. W zapisie zespolonym manifestuje się to przez fakt, że konduktywność σ staje się wielkością zespoloną, tzn. posiada ona niezerową część urojoną.

Z prawa zachowania ładunku elektrycznego można zapisać, że zmiana gęstości ładunków w czasie musi być związana z prądem, który wpływa lub wypływa z rozważanego obszaru

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{j}. \tag{4.59}$$

Korzystając z równania (4.58), powyższe równanie możemy zapisać w postaci

$$\frac{\partial \rho}{\partial t} = -\sigma \nabla \cdot \mathbf{E} = -\frac{\sigma}{\varepsilon}\rho, \tag{4.60}$$

gdzie skorzystaliśmy z równania Gaussa dla pola elektrycznego i równania materiałowego wiążącego indukcję i natężenie pola elektrycznego.

Rozwiązaniem równania (4.60) jest funkcja

$$\rho(t) = \rho_0 \exp\left(-\frac{\sigma}{\varepsilon}t\right), \tag{4.61}$$

gdzie ρ_0 jest początkową gęstością ładunków.

W przewodnikach, gdzie konduktywność jest bardzo duża, stała tłumienia σ/ε przyjmuje duże wartości, tak że na dłuższych przedziałach czasu można założyć, iż makroskopowa gęstość ładunków $\rho = 0$. Założenie to jest treścią tzw. **przybliżenia adiabatycznego**. W ramach tego przybliżenia prawo Gaussa dla pola elektrycznego można zapisać w postaci identycznej z tą znaną dla dielektryków:

$$\varepsilon \nabla \cdot \mathbf{E} = 0. \tag{4.62}$$

Z powyższego wzoru wynika, że jedynym źródłem pola wewnątrz materiału są prądy przewodnictwa.

Postać równań Maxwella uwzględniająca swobodne ładunki elektryczne ma znaczący wpływ na propagację fali w ośrodku. Możemy tego dowieść, postępując podobnie jak w przypadku ośrodków dielektrycznych, tzn. obkładając obie strony prawa Faradaya operatorem $\nabla\times$, a następnie przekształcając to równanie do postaci równania falowego. W przewodnikach równanie to przyjmuje postać

$$\nabla^2 \mathbf{E} - \mu\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{E}}{\partial t} = 0, \tag{4.63}$$

gdzie skorzystaliśmy z postaci rotacji pola $\mathbf{B}$ dla ośrodków zawierających ładunki swobodne oraz z prawa Ohma. Jak widać z powyższej zależności, równanie falowe w przewodnikach różni się od równania falowego w dielektrykach pojawieniem się członu $-\mu\sigma\frac{\partial \mathbf{E}}{\partial t}$, który związany jest z dysypacyjnym przepływem prądu w materiale.

Podobnie jak w przypadku dielektryków, falę elektromagnetyczną w przewodnikach możemy zapisać w postaci zespolonej

$$\mathbf{E}^{(+)}(\mathbf{r},t) = \mathbf{e}_E E_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}. \tag{4.64}$$

Podstawiając tę postać fali do równania falowego, otrzymujemy że

$$k^2 = \mu\varepsilon\omega^2 + i\mu\sigma\omega = \omega^2\mu\left(\varepsilon + i\frac{\sigma}{\omega}\right). \tag{4.65}$$

Zdefiniujmy teraz zespoloną przenikalność elektryczną $\tilde{\varepsilon}$

$$\tilde{\varepsilon} = \varepsilon + i\frac{\sigma}{\omega}, \tag{4.66}$$

co pozwala nam zapisać **równanie dyspersji w przewodnikach**

$$\boxed{k = \sqrt{\mu\tilde{\varepsilon}}\omega.} \tag{4.67}$$

Podobnie jak w przypadku dielektryków, również dla przewodników możemy wprowadzić pojęcie współczynnika załamania. Należy jednak zaznaczyć, że wielkość ta jest zespolona i w literaturze oznaczana często symbolem $\tilde{n}$ lub η

$$\eta = \sqrt{\frac{\tilde{\varepsilon}}{\varepsilon_0}} = n + i\kappa, \tag{4.68}$$

gdzie część urojona κ nazywana jest **współczynnikiem ekstynkcji** lub absorpcji.

Wprowadzenie zespolonego współczynnika załamania η pozwala nam stosować identyczne podejście do propagacji fali elektromagnetycznej w przewodniku do tego, jakie stosowaliśmy wcześniej w dielektrykach. Jedyna różnica polega na tym, że z uwagi na fakt, że współczynnik załamania jest zespolony, liczba falowa k jest wielkością zespoloną:

$$k = k_r + ik_i. \tag{4.69}$$

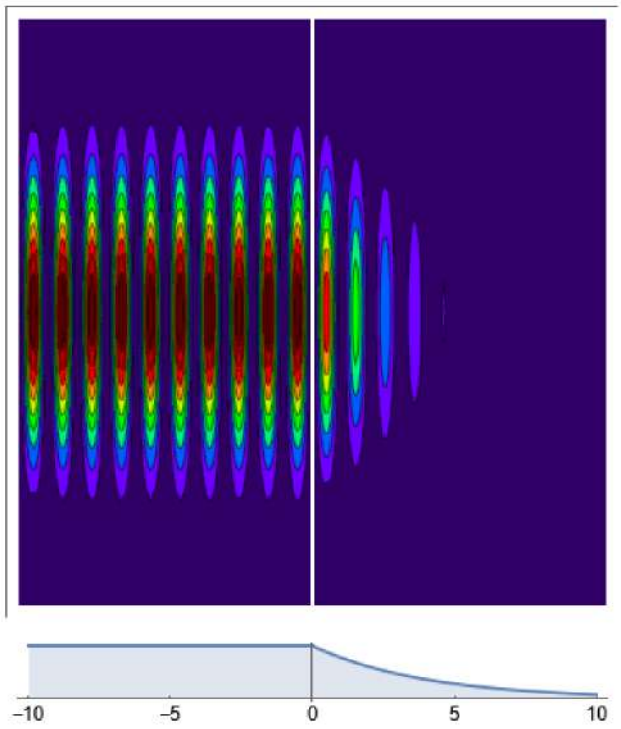
Podstawiając taką postać liczby falowej do równania (4.65) otrzymujemy

$$k_r^2 + 2ik_rk_i - k_i^2 = \omega^2\mu\left(\varepsilon + i\frac{\sigma}{\omega}\right), \tag{4.70}$$

skąd można wyprowadzić jawne zależności na współczynniki k_r oraz k_i

$$k_r^2 = \frac{\omega^2\mu\varepsilon}{2}\left[\sqrt{1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2}+1\right], \tag{4.71a}$$

$$k_i^2 = \frac{\omega^2\mu\varepsilon}{2}\left[\sqrt{1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2}-1\right]. \tag{4.71b}$$



Rysunek 4.12: Efekt naskórkowy. Górny wykres przedstawia pole elektryczne fali padającej na granicę dielektryk (lewa strona) – przewodnik (prawa strona) od strony dielektryka prostopadle do powierzchni granicznej. Dolny wykres pokazuje natężenie fali w obu ośrodkach. Widać, że fala, która wnika do przewodnika, podlega absorpcji, a jej natężenie spada wraz z odległością od granicy ośrodków.

4.7.2 Efekt naskórkowy i głębokość wnikania

Rozważmy teraz falę rozchodzącą się wzdłuż kierunku z, opisaną zespoloną liczbą falową k. W takim przypadku funkcja opisująca falę dana jest równaniem

$$\mathbf{E}(z,t) = \mathbf{e}_E E_0^{(+)} e^{i(kz-\omega t)} = \mathbf{e}_E E_0^{(+)} e^{i(k_r z-\omega t)} e^{-k_i z}. \tag{4.72}$$

Z postaci równania (4.72) widać, że fala wnikająca do metalu podlega silnemu tłumieniu i jej amplituda spada eksponencjalnie wraz z odległością wnikania do przewodnika [czynnik $e^{-k_i z}$]. Bazując na tej obserwacji, do opisu propagacji światła w przewodnikach można wprowadzić pojęcie **głębokości wnikania**, (Rys. 4.12), która jest odległością, na jakiej amplituda fali maleje e-krotnie:

$$\delta_w = \frac{1}{k_i}. \tag{4.73}$$

Łącząc ze sobą równanie (4.68), (4.69) oraz związek pomiędzy wektorem falowym w ośrodku materialnym i w próżni ($k = \eta k_0 = 2\pi\eta/\lambda_0$, gdzie λ_0 jest długością fali w próżni), otrzymujemy wyrażenie na głębokość wnikania:

$$\boxed{\delta_w = \frac{\lambda_0}{2\pi\kappa}.} \tag{4.74}$$

Samo zjawisko silnej absorpcji i zanikania pola w przewodnikach nazywane jest **efektem naskórkowym**.

Ponieważ w przewodnikach tłumienność fal z zakresu widzialnego jest duża, $\kappa \gtrsim 1$, głębokość wnikania światła w tego typu materiałach jest zdecydowanie krótsza od długości fali. Przykładowo, dla srebra, charakteryzującego się współczynnikiem załamania $\eta = 0.27 + 4.47i$, głębokość wnikania dla światła o długości fali 780 nm jest równa ok. 28 nm. Co więcej, z uwagi na zależność głębokości wnikania od długości fali, wielkość ta może przyjmować skrajnie różne wartości w odmiennych zakresach spektralnych.

Dla niższych częstości (np. fale radiowe) $\sigma/(\varepsilon\omega) \gg 1$ i równanie (4.71) może zostać znacząco uproszczone:

$$k_i \approx \sqrt{\frac{\sigma\omega\mu}{2}}. \tag{4.75}$$

W konsekwencji równanie na głębokość wnikania można zapisać w postaci:

$$\boxed{\delta_w = \sqrt{\frac{2}{\omega\sigma\mu}}.} \tag{4.76}$$

Ponieważ fala niesie ze sobą energię, można zadać sobie pytanie, co dzieje się z energią traconą przez światło. Tłumienie ruchu elektronów w przewodnikach związane jest z rozpraszaniem nośników na sieci krystalicznej metalu. Ponieważ proces ten jest tożsamy z mechanizmem odpowiadającym za opór elektryczny w przewodnikach (wydzielanie ciepła Joule'a), łatwo można się domyśleć, że jego skutkiem jest dysypacja energii na ciepło.

Rozważmy na koniec przypadek słabych przewodników, tj. ośrodków o znikomym przewodnictwie. Ponieważ w takich ośrodkach $\sigma \to 0$, to $\sigma/(\varepsilon\omega) \to 0$. Oznacza to, że wszystkie wyrażenia, w których pojawia się stosunek $\sigma/(\varepsilon\omega)$, należy rozwinąć w szereg Taylora, a następnie uwzględnić tylko najniższe rzędy tego rozwinięcia. Przykładowo wektor falowy k_r dany jest zależnością:

$$k_r = \frac{\omega\sqrt{\mu\varepsilon}}{\sqrt{2}}\sqrt{\sqrt{1+\left(\frac{\sigma}{\varepsilon\omega}\right)^2}+1} \approx \omega\sqrt{\mu\varepsilon}\left[1+\frac{1}{8}\left(\frac{\sigma}{\varepsilon\omega}\right)^2\right]. \tag{4.77}$$

Z równania (4.77) widać, że dla $\sigma \to 0$, wektor falowy przyjmuje postać standardową dla idealnych dielektryków.

4.7.3 Wzory Fresnela w przewodnikach

Analogicznie jak to zostało zrobione w przypadku dielektryków, również dla metali można wprowadzić amplitudowe współczynniki odbicia. Ich postać matematyczna jest strukturalnie zbieżna z tą znaną z dielektryków:

$$\begin{aligned} r_\perp &= \frac{n_1 \cos\alpha_p - \eta_2 \cos\alpha_z}{n_1 \cos\alpha_p + \eta_2 \cos\alpha_z}, \\ r_\parallel &= \frac{\eta_2 \cos\alpha_p - n_1 \cos\alpha_z}{\eta_2 \cos\alpha_p + n_1 \cos\alpha_z}. \end{aligned} \tag{4.78}$$

Dokładniejsza analiza wykazuje jednak, że w kontekście interpretacji fizycznej istnieją istotne różnice. Ich źródłem jest fakt, że współczynnik załamania przewodnika η_2 jest wielkością zespoloną, co z kolei powoduje, że również kąt załamania α_z musi zostać ujęty jako formalna wartość zespolona. Zależności tej można dowieść w oparciu o **zespolone prawo Snella**:

$$n_1 \sin\alpha_p = \eta_2 \sin\alpha_z. \tag{4.79}$$

W konsekwencji przy odbiciu od przewodników dochodzi do istotnej zmiany fazy światła odbitego i to zarówno dla składowej równoległej (polaryzacja p), jak i prostopadłej (polaryzacja s) względem granicy ośrodków. Zmiana ta nabiera charakteru płynnego i zależy od kąta padania światła na granicę tych materiałów. Stanowi to wyrazistą różnicę w stosunku do dielektryków, gdzie podczas odbicia od ośrodka gęstszego optycznie dochodziło do stałego skoku fazy o π, natomiast przy odbiciu od granicy rzadszej faza fali w ogóle nie ulegała zmianie.

Podobnie jak w przypadku dielektryków możemy również obliczyć moc fali, która odbijana jest od granicy. Przez analogię można wykazać, że:

$$R_{\perp,\parallel} = |r_{\perp,\parallel}|^2. \tag{4.80}$$

Z uwagi na fakt, że współczynnik r jest wielkością zespoloną, której moduł w optyce widzialnej jest na ogół mniejszy niż 1, natężenie fali odbitej jest mniejsze od natężenia fali padającej. Przywołując raz jeszcze przykład srebra i padania światła wzdłuż normalnej do granicy ośrodków, możemy pokazać, że współczynnik odbicia energetycznego wynosi w przybliżeniu 95%. Pozostałe 5% mocy strumienia falowego jest trwale absorbowane w metalu w warstwie przypowierzchniowej. Straty wymuszane przez pochłanianie radiacji mają istotne znaczenie w technologiach wymagających światła o bardzo dużym natężeniu i mocno ograniczają użycie standardowych metalicznych powłok lustrzanych w niektórych zaawansowanych aplikacjach fotonicznych.

Warto przy tym odnotować, że ze względu na niejednakowe zespolone współczynniki odbicia dla obu składowych polaryzacyjnych, fale o polaryzacji p oraz s doznają podczas odbicia ukośnego różnych przesunięć fazowych. W konsekwencji światło padające np. jako spolaryzowane liniowo pod pewnym kątem przekształca się na wyjściu z lustra w światło **spolaryzowane eliptycznie**.

Zjawisko to, polegające na zróżnicowanej transformacji amplitudowo-fazowej odbitych składowych falowych, dało początek wysoce precyzyjnej i powszechnie dziś używanej technice badawczej znanej jako **elipsometria**, która stanowi potężne narzędzie w inżynierii optycznej do bezinwazyjnego charakteryzowania cienkich warstw i nanostruktur.

4.8 Podsumowanie rozdziału

W niniejszym rozdziale dokonano rygorystycznej analizy zachowania fali elektromagnetycznej na granicy dwóch ośrodków materialnych. Wychodząc z całkowych postaci równań Maxwella, sformułowano fundamentalne warunki ciągłości dla składowych stycznych i normalnych pól, które stały się punktem wyjścia do wyprowadzenia praw optyki geometrycznej (prawa odbicia i załamania) bezpośrednio na gruncie elektrodynamiki.

Kluczowym osiągnięciem było analityczne wyznaczenie amplitudowych współczynników Fresnela dla składowych polaryzacji stycznej i prostopadłej do płaszczyzny padania. Analiza tych wzorów ujawniła bogactwo zjawisk fizycznych, wykraczających poza prosty rachunek wektorowy: od polaryzacji światła przez odbicie pod kątem Brewstera, przez zjawisko fali zanikającej i subtelne przesunięcie Goosa-Hänchena przy całkowitym wewnętrznym odbiciu, aż po analog tunelowania kwantowego. Rozdział domknięto rozszerzeniem formalizmu na ośrodki stratne (przewodniki) za pomocą zespolonego współczynnika załamania, co pozwoliło opisać efekt naskórkowy, dyssypację energii na ciepło Joule'a oraz pojawienie się eliptyczności liniowo spolaryzowanego światła po odbiciu od granicy dielektryk–metal. Przedstawiony aparat matematyczny dostarcza niezbędnych narzędzi do projektowania zaawansowanych układów fotonicznych, przygotowując jednocześnie grunt do dyskusji o superpozycji wielu fal, co stanowić będzie przedmiot rozważań w kolejnych częściach podręcznika.

Zadania

4.1. Wykorzystując warunki ciągłości pól elektromagnetycznych na płaskiej granicy dwóch dielektryków, wyprowadź:

(a) amplitudowe współczynniki odbicia (r_s, r_p) oraz transmisji (t_s, t_p) dla fali spolaryzowanej prostopadle i równolegle do płaszczyzny padania,

(b) natężeniowe współczynniki odbicia (reflektancję R_s, R_p) oraz transmisji (transmitancję T_s, T_p).

Wskazówka: Przy wyprowadzaniu transmitancji pamiętaj, że na skutek załamania zmianie ulega pole przekroju poprzecznego transmitowanej wiązki.

4.2. Wychodząc z zasady odwracalności biegu promieni świetlnych, udowodnij

prawdziwość formuł Stokesa:

$$r_{12} + r_{21} = 0, \qquad t_{12}t_{21} - r_{12}r_{21} = 1,$$

gdzie r_{12} i t_{12} to amplitudowe współczynniki odbicia i transmisji przy padaniu światła z ośrodka 1 do 2, a r_{21} i t_{21} – przy propagacji z ośrodka 2 do 1.

4.3. Wiązka laserowa pada z powietrza na płaską powierzchnię plastiku (o względnej przenikalności dielektrycznej $\varepsilon_r = 2$) pod kątem 20°. Zakładając, że amplituda składowej poprzecznej pola elektrycznego w powietrzu wynosi 20 V/m, a składowej podłużnej 10 V/m, oblicz wartości obu tych składowych tuż pod powierzchnią plastiku.

4.4. Oblicz kąt Brewstera dla zewnętrznego odbicia na granicy powietrze–woda ($n = 1,3$). Narysuj schemat biegu promieni na granicy tych ośrodków, ze szczególnym uwzględnieniem kąta pomiędzy promieniem odbitym a załamanym.

4.5. Niespolaryzowane światło pada na szklaną płytkę ($n = 1,5$) pod kątem Brewstera. Stopień polaryzacji światła definiujemy jako:

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}},$$

gdzie $I_{\parallel}$ oraz $I_{\perp}$ to natężenia światła o polaryzacji odpowiednio równoległej i prostopadłej do płaszczyzny padania.

(a) Oblicz stopień polaryzacji światła po przejściu przez pojedynczą płytkę. Pamiętaj, że do załamania dochodzi dwukrotnie: na granicy powietrze–szkło oraz szkło–powietrze.

(b) Zmodyfikuj wzór dla stosu składającego się z N identycznych płytek.

(c) Sporządź wykres zależności stopnia polaryzacji światła transmitowanego od liczby płytek N dla zakresu $N \in \{1, 2, \ldots, 20\}$.

4.6. Wyprowadź wzory na całkowitą reflektancję R i transmitancję T płaskiej fali o długości λ, padającej pod kątem α na dielektryczną płytkę o grubości d i współczynniku załamania n, umieszczoną w próżni. Znajdź minimalne i maksymalne wartości tych współczynników.
Wskazówka: Rozważ nieskończony szereg wielokrotnych odbić wewnątrz płytki i oblicz jego sumę.

4.7. Rozważ spolaryzowane mikrofale ($\lambda_0 = 3$ cm) padające na układ dwóch parafinowych pryzmatów (trójkąty prostokątne równoramienne, $n = 1,5$), ułożonych przeciwprostokątnymi do siebie i oddzielonych szczeliną powietrzną o grubości $d = 5$ mm. Zakładając padanie prostopadłe na pierwszą ścianę (i kąt padania na szczelinę równy 45°), oblicz transmitancję mikrofal przez ten układ. Pomiń straty absorpcyjne w parafinie.

4.8. Światło o długości fali $\lambda = 780$ nm pada prostopadle z próżni na wypolerowane powierzchnie trzech różnych metali.

(a) Oblicz współczynnik odbicia R (reflektancję) oraz głębokość wnikania δ_w dla srebra ($\eta = 0,27 + 4,47i$), złota ($\eta = 0,08 + 4,60i$) oraz miedzi ($\eta = 0,24 + 4,80i$).

(b) Jak zmienią się wartości R_s, R_p oraz δ_w dla srebra, jeśli światło będzie padać pod kątem 45°?

4.9. Rozważ światło padające pod bardzo dużym kątem $\theta_i = 80°$ na powierzchnię srebra ($n = 0,13$, $\kappa = 4,0$). Oblicz natężeniowe współczynniki odbicia $R_\perp$ (polaryzacja s) oraz $R_\parallel$ (polaryzacja p). Skomentuj uzyskane wyniki.

4.10. Zmienne pole elektromagnetyczne o częstotliwości $f = 200$ kHz pada prostopadle na powierzchnię tytanu, którego konduktywność wynosi $\sigma = 2 \times 10^6$ S/m.

(a) Oszacuj głębokość wnikania tego promieniowania.

(b) Rozważ tytanowy dysk o kształcie i rozmiarze monety 1 zł, umieszczony wewnątrz długiej cewki zasilanej prądem o tej samej częstotliwości (dysk ułożony prostopadle do osi cewki). Zakładając jednorodność pola na zewnątrz, opisz jakościowo i ilościowo przestrzenny rozkład absorpcji energii wewnątrz dysku.

Rozdział 5

Interferencja

5.1 Superpozycja fal

Jedną z podstawowych właściwości równania falowego jest jego addytywność i związana z nią zasada superpozycji fal. Zgodnie z nią, jeżeli dwie fale spełniają równanie falowe, to ich suma również je spełnia:

$$\left.\begin{array}{l}\nabla^2\psi_1 - \frac{1}{v^2}\frac{\partial^2\psi_1}{\partial t^2} = 0 \\ \nabla^2\psi_2 - \frac{1}{v^2}\frac{\partial^2\psi_2}{\partial t^2} = 0\end{array}\right\} \quad \nabla^2(\psi_1+\psi_2) - \frac{1}{v^2}\frac{\partial^2(\psi_1+\psi_2)}{\partial t^2} = \nabla^2\psi - \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} = 0. \tag{5.1}$$

Zasada superpozycji pozwala na analizę przebiegów falowych powstałych w wyniku nałożenia się fal o różnych kierunkach propagacji, amplitudach czy częstotliwościach. W szczególności umożliwia ona analizę propagacji dowolnego przebiegu czasowego, który da się przedstawić jako sumę przebiegów monochromatycznych o odpowiednich amplitudach. Pozwala to zatem na uogólnienie prowadzonych dotychczas rozważań.

Uwaga!

Omawiając zasadę superpozycji, należy zwrócić uwagę, że obowiązuje ona wyłącznie w tzw. **ośrodkach liniowych**, czyli takich, których właściwości nie zależą od natężenia propagującej się fali. W wielu przypadkach założenie to nie jest jednak spełnione. Przykładowo intensywne światło może powodować zmianę współczynnika załamania ośrodka, co w konsekwencji prowadzi do uzależnienia prędkości rozchodzenia się fali od natężenia światła. W ośrodkach tego typu, nazywanych **ośrodkami nieliniowymi**, może wręcz dochodzić do mieszania się przyczynków pochodzących od różnych fal. Skutkuje to powstawaniem nowych fal o częstotliwościach równych na przykład sumie lub różnicy częstotliwości fal składowych. Wówczas zasada superpozycji przestaje obowiązywać. Opis propagacji światła w ośrodkach nieliniowych jest zagadnieniem złożonym i zostanie pokrótce przedstawiony w rozdziale 10.

5.2 Przykłady nakładania się fal

5.2.1 Dwie fale o jednakowych częstotliwościach i wektorach falowych

Doświadczenie Obserwacja sygnału sumy dwóch periodycznych przebiegów elektrycznych o regulowanych fazach na oscyloskopie.

Warto przeanalizować przypadek dwóch skalarnych fal płaskich o identycznych amplitudach $E_0^{(+)}$, wektorach falowych $\mathbf{k}$ oraz częstościach ω, różniących się jedynie fazami początkowymi. Wykorzystując opis funkcji falowych za pomocą funkcji trygonometrycznych (zapis zespolony):

$$\begin{aligned} E_1(\mathbf{r},t) &= E_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\varphi_1)}, \\ E_2(\mathbf{r},t) &= E_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\varphi_2)}. \end{aligned} \tag{5.2}$$

Superpozycję tych fal można zapisać jako:

$$\begin{aligned} E(\mathbf{r},t) &= E_1(\mathbf{r},t) + E_2(\mathbf{r},t) \\ &= E_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \left(e^{i\varphi_1} + e^{i\varphi_2}\right) \\ &= E_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\bar{\varphi})} \left(e^{i\Delta\varphi/2} + e^{-i\Delta\varphi/2}\right) \\ &= 2E_0^{(+)} \cos(\Delta\varphi/2) e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\bar{\varphi})}, \end{aligned} \tag{5.3}$$

gdzie $\Delta\varphi = \varphi_1 - \varphi_2$ oraz $\bar{\varphi} = (\varphi_1 + \varphi_2)/2$ oznaczają odpowiednio różnicę faz i średnią fazę fal. Z równania (5.3) wynika, że skutkiem nałożenia się fal o tej samej amplitudzie, częstotliwości i kierunku propagacji jest fala o niezmienionym kierunku, liczbie falowej k i częstości ω. Jej amplituda jest jednak określona przez różnicę faz obu fal $\Delta\varphi$, a faza wypadkowa stanowi średnią arytmetyczną faz składowych $\bar{\varphi}$.

Warto w tym miejscu poddać analizie fizyczny sens czynnika amplitudy $\cos(\Delta\varphi/2)$. Wynik ten dopuszcza ujemne wartości amplitudy pola elektrycznego, co w sensie fizycznym interpretuje się po prostu jako dodatkowe przesunięcie fazowe o π (odwrócenie fazy fali).

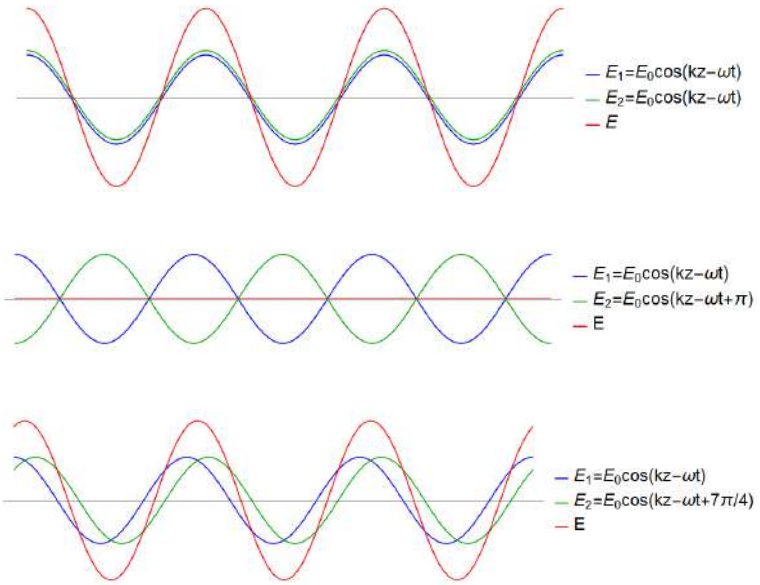
5.2.2 Dwie fale o identycznych częstotliwościach rozchodzące się w przeciwnych kierunkach

Rozważmy teraz fale o przeciwnych wektorach falowych $\mathbf{k}$ oraz $-\mathbf{k}$, ale o tych samych częstotliwościach:

$$\begin{aligned} E_1(\mathbf{r},t) &= E_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\varphi_1)}, \\ E_2(\mathbf{r},t) &= E_0^{(+)} e^{i(-\mathbf{k}\cdot\mathbf{r}-\omega t+\varphi_2)}. \end{aligned} \tag{5.4}$$

W wyniku superpozycji otrzymujemy:

$$\begin{aligned} E(\mathbf{r},t) &= E_0^{(+)} e^{-i\omega t+i\bar{\varphi}} \left(e^{i(\mathbf{k}\cdot\mathbf{r}+\Delta\varphi/2)} + e^{-i(\mathbf{k}\cdot\mathbf{r}+\Delta\varphi/2)}\right) \\ &= 2E_0^{(+)} \cos(\mathbf{k}\cdot\mathbf{r}+\Delta\varphi/2) e^{-i\omega t+i\bar{\varphi}}. \end{aligned} \tag{5.5}$$



Rysunek 5.1: Nałożenie fal o identycznych amplitudach, częstościach i kierunkach propagacji, lecz różnych fazach: $\Delta\varphi = 0$ (góra), $\Delta\varphi = \pi$ (środek), $\Delta\varphi = 7\pi/4$ (dół). Kolory niebieski i zielony reprezentują fale składowe, kolor czerwony — falę wypadkową.

Część rzeczywista tego wyrażenia opisuje **falę stojącą**:

$$\mathrm{Re}\{E(\mathbf{r},t)\} = 2E_0^{(+)}\cos(\mathbf{k}\cdot\mathbf{r}+\Delta\varphi/2)\cos(-\omega t+\bar{\varphi}). \tag{5.6}$$

Amplituda tej fali jest zmodulowana przestrzennie czynnikiem $\cos(\mathbf{k}\cdot\mathbf{r}+\Delta\varphi/2)$. Fala ta nie transportuje energii, co jest charakterystyczne dla układów w rezonansie.

Charakterystyczną cechą fali stojącej jest występowanie w przestrzeni punktów, w których amplituda drgań jest trwale równa zeru. Punkty te nazywane są **węzłami**. Zgodnie z otrzymanym równaniem, położenie węzłów wyznacza warunek $\cos(\mathbf{k}\cdot\mathbf{r}+\Delta\varphi/2)=0$. Z kolei miejsca, w których amplituda osiąga wartość maksymalną (równą $2E_0^{(+)}$), to **strzałki**. Odległość pomiędzy dwoma sąsiednimi węzłami lub strzałkami wzdłuż kierunku propagacji wynosi $\lambda/2$, co wynika bezpośrednio z relacji między wektorem falowym a długością fali ($k=2\pi/\lambda$). Dodatkowo, w przeciwieństwie do fali biegnącej, wszystkie punkty ośrodka zawarte między dwoma sąsiednimi węzłami drgają w tej samej fazie, a po przekroczeniu węzła faza drgań zmienia się skokowo o π. Na koniec zaznaczmy, że fala stojąca, jak sugeruje sama nazwa, nie niesie energii przestrzennej.

5.2.3 Dwie fale o jednakowych częstotliwościach rozchodzące się pod kątem

Dotychczasowe rozważania ograniczały się do fal współbieżnych lub przeciwbieżnych (zagadnienie jednowymiarowe). Niezwykle istotnym przypadkiem jest superpozycja fal płaskich, których wektory falowe przecinają się pod pewnym kątem w przestrzeni. Załóżmy, że dwie fale o jednakowej amplitudzie $E_0^{(+)}$, częstości ω i zerowej fazie początkowej propagują się w płaszczyźnie xz, a ich wektory falowe są symetrycznie odchylone od osi z o kąt θ. Wektory te można zapisać jako $\mathbf{k}_1=(k\sin\theta,0,k\cos\theta)$ oraz $\mathbf{k}_2=(-k\sin\theta,0,k\cos\theta)$. Sygnały analityczne obu fal przyjmują postać:

$$\begin{aligned} E_1(\mathbf{r},t) &= E_0^{(+)}e^{i(kx\sin\theta+kz\cos\theta-\omega t)},\\ E_2(\mathbf{r},t) &= E_0^{(+)}e^{i(-kx\sin\theta+kz\cos\theta-\omega t)}. \end{aligned} \tag{5.7}$$

Dodając do siebie te równania, można wyłączyć przed nawias czynnik odpowiadający za propagację wzdłuż osi z:

$$\begin{aligned} E(\mathbf{r},t) &= E_1(\mathbf{r},t)+E_2(\mathbf{r},t)\\ &= E_0^{(+)}e^{i(kz\cos\theta-\omega t)}\left(e^{ikx\sin\theta}+e^{-ikx\sin\theta}\right)\\ &= 2E_0^{(+)}\cos(kx\sin\theta)e^{i(kz\cos\theta-\omega t)}. \end{aligned} \tag{5.8}$$

Przejście do mierzalnego, fizycznego pola elektrycznego daje:

$$\mathrm{Re}\{E(\mathbf{r},t)\} = 2E_0^{(+)}\cos(kx\sin\theta)\cos(kz\cos\theta-\omega t). \tag{5.9}$$

Równanie (5.9) opisuje falę, która propaguje się wzdłuż osi z z efektywnym wektorem falowym zmodyfikowanym względem początkowego wektora falowego

o czynnik $\cos\theta$. Co jednak najważniejsze, jej amplituda jest zmodulowana w poprzek kierunku propagacji (wzdłuż osi x) przez czynnik $\cos(kx\sin\theta)$.

5.2.4 Dwie fale o nieco różnych częstotliwościach i liczbach falowych, ale tym samym kierunku propagacji

Doświadczenie Obserwacja sygnału dudnień z dwóch generatorów wytwarzających przebiegi o nieznacznie różniących się częstościach.

Warto przeanalizować przypadek dwóch fal rozchodzących się w tym samym kierunku, lecz charakteryzujących się różnymi częstotliwościami. Zakładając dla uproszczenia zerowe fazy początkowe ($\varphi_1 = \varphi_2 = 0$) oraz wykorzystując spójny z poprzednimi sekcjami zapis zespolony, fale składowe przyjmują postać:

$$\begin{aligned} E_1(\mathbf{r},t) &= E_0^{(+)} e^{i(\mathbf{k}_1\cdot\mathbf{r}-\omega_1 t)}, \\ E_2(\mathbf{r},t) &= E_0^{(+)} e^{i(\mathbf{k}_2\cdot\mathbf{r}-\omega_2 t)}. \end{aligned} \tag{5.10}$$

Wprowadzając średni wektor falowy $\bar{\mathbf{k}} = (\mathbf{k}_1+\mathbf{k}_2)/2$ i różnicę wektorów falowych $\Delta\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2$, a także średnią częstość $\bar{\omega} = (\omega_1+\omega_2)/2$ i różnicę częstości $\Delta\omega = \omega_1 - \omega_2$, superpozycję tych fal można zapisać jako:

$$\begin{aligned} E(\mathbf{r},t) &= E_1(\mathbf{r},t) + E_2(\mathbf{r},t) \\ &= E_0^{(+)} e^{i(\bar{\mathbf{k}}\cdot\mathbf{r}-\bar{\omega}t)} \left(e^{i(\Delta\mathbf{k}\cdot\mathbf{r}/2-\Delta\omega t/2)} + e^{-i(\Delta\mathbf{k}\cdot\mathbf{r}/2-\Delta\omega t/2)} \right) \\ &= 2E_0^{(+)} \cos\left(\frac{\Delta\mathbf{k}}{2}\cdot\mathbf{r} - \frac{\Delta\omega}{2}t \right) e^{i(\bar{\mathbf{k}}\cdot\mathbf{r}-\bar{\omega}t)}. \end{aligned} \tag{5.11}$$

Przejście do mierzalnego, fizycznego pola elektrycznego daje w wyniku:

$$\mathrm{Re}\{E(\mathbf{r},t)\} = 2E_0^{(+)} \cos\left(\frac{\Delta\mathbf{k}}{2}\cdot\mathbf{r} - \frac{\Delta\omega}{2}t \right) \cos(\bar{\mathbf{k}}\cdot\mathbf{r} - \bar{\omega}t). \tag{5.12}$$

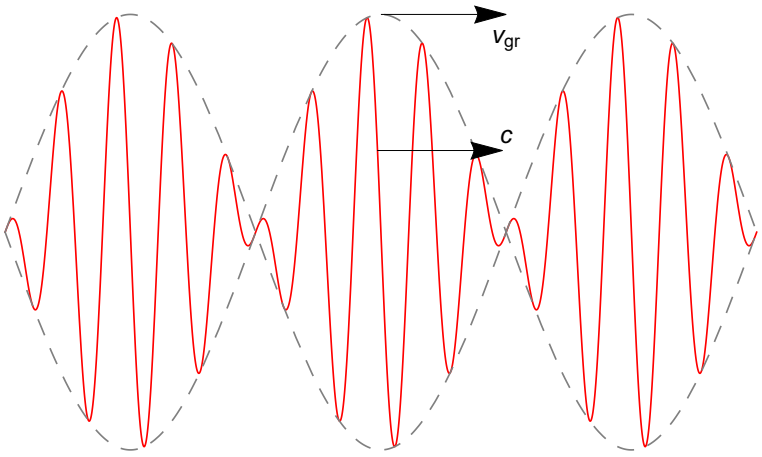
Przyjmując, że $\mathbf{k}_1 \approx \mathbf{k}_2 \approx \mathbf{k}$ oraz $\omega_1 \approx \omega_2 \approx \omega$, można utożsamić wartości średnie z parametrami pojedynczej fali ($\bar{\mathbf{k}} \approx \mathbf{k}$, $\bar{\omega} \approx \omega$). Wówczas równanie (5.12) przyjmuje postać:

$$\mathrm{Re}\{E(\mathbf{r},t)\} = 2E_0^{(+)} \cos\left(\frac{\Delta\mathbf{k}}{2}\cdot\mathbf{r} - \frac{\Delta\omega}{2}t \right) \cos(\mathbf{k}\cdot\mathbf{r} - \omega t). \tag{5.13}$$

W przypadku, gdy $\mathbf{k}_1 \parallel \mathbf{k}_2$, wynikiem superpozycji jest fala biegnąca, której amplituda jest wolnozmiennie modulowana w przestrzeni i w czasie. Takie nałożenie na siebie fal o zbliżonych częstotliwościach i liczbach falowych oraz identycznych kierunkach propagacji nosi nazwę **dudnień**.

Należy odróżnić częstość zmian samej obwiedni od częstości dudnień, którą faktycznie rejestruje detektor. Amplituda fali zmienia się z częstością $\Delta\omega/2$, jednakże natężenie fali (proporcjonalne do kwadratu amplitudy) osiąga maksimum dwa razy w ciągu jednego okresu obwiedni. Z tego powodu fizycznie mierzalna częstość dudnień ω_{dud} jest równa modułowi różnicy częstości fal składowych:

$$\omega_{\mathrm{dud}} = |\omega_1 - \omega_2| = |\Delta\omega|. \tag{5.14}$$



Rysunek 5.2: Dudnienia powstałe w wyniku superpozycji dwóch fal rozchodzących się w tym samym kierunku, różniących się nieznacznie częstością i liczbą falową.

Znosząc początkowe założenie o zerowych fazach ($\varphi_1 \neq \varphi_2$), w wyrażeniu na dudnienia pojawiają się dodatkowe przesunięcia fazowe. Różnica faz $\Delta\varphi = \varphi_1 - \varphi_2$ powoduje przestrzenno-czasowe przesunięcie wzoru obwiedni, natomiast średnia faza $\bar{\varphi} = (\varphi_1 + \varphi_2)/2$ modyfikuje fazę szybkookresowej fali nośnej propagującej się wewnątrz obwiedni:

$$\mathrm{Re}\{E(\mathbf{r},t)\} = 2E_0^{(+)} \cos\left(\frac{\Delta\mathbf{k}}{2}\cdot\mathbf{r} - \frac{\Delta\omega}{2}t + \frac{\Delta\varphi}{2}\right)\cos(\bar{\mathbf{k}}\cdot\mathbf{r} - \bar{\omega}t + \bar{\varphi}). \quad (5.15)$$

Dodatkowe fazy początkowe nie wpływają jednak na ogólny fizyczny charakter zjawiska dudnień, który pozostaje niezmieniony.

5.2.5 Dwie fale o jednakowych częstotliwościach, tym samym kierunku propagacji i ortogonalnych polaryzacjach liniowych

Dotychczasowe rozważania dotyczyły superpozycji fal skalarnych. W przypadku przejścia do fal wektorowych, kierunek drgań wektora pola elektrycznego w przestrzeni staje się kluczowym parametrem. Przeanalizujmy superpozycję dwóch współbieżnych fal (np. propagujących się wzdłuż osi z) o ortogonalnych polaryzacjach liniowych: jednej drgającej wzdłuż osi x, a drugiej wzdłuż osi y. Zakładając różnicę faz początkowych $\Delta\varphi$ oraz ewentualnie różne amplitudy $E_{0x}^{(+)}$ i $E_{0y}^{(+)}$, zespolone wektory pola można zapisać jako:

$$\begin{aligned}\mathbf{E}_1(\mathbf{r},t) &= E_{0x}^{(+)}\mathbf{x}e^{i(kz-\omega t)},\\ \mathbf{E}_2(\mathbf{r},t) &= E_{0y}^{(+)}\mathbf{y}e^{i(kz-\omega t+\Delta\varphi)}.\end{aligned} \quad (5.16)$$

Superpozycja i wyciągnięcie części rzeczywistej prowadzi do wektora wypadkowego pola elektrycznego o współrzędnych:

$$\mathrm{Re}\{\mathbf{E}(\mathbf{r},t)\} = \mathbf{x}E_{0x}^{(+)}\cos(kz-\omega t) + \mathbf{y}E_{0y}^{(+)}\cos(kz-\omega t+\Delta\varphi). \tag{5.17}$$

Krzywa zakreślana w ustalonej płaszczyźnie przestrzeni (np. $z = 0$) przez koniec wektora wypadkowego jest w ogólności elipsą. Wynika stąd fundamentalny wniosek: złożenie dwóch prostopadłych fal drgających z tą samą częstością prowadzi do powstania **polaryzacji eliptycznej**. To właśnie o takiej polaryzacji była po raz pierwszy mowa w rozdziale 1.5.1.

5.2.6 Dwie fale o jednakowych częstotliwościach, identycznych polaryzacjach kołowych i przeciwnych wektorach falowych

Dotychczasowe rozważania dotyczące superpozycji odnosiły się do fal skalarnych. Warto rozszerzyć ten opis na przypadek fal wektorowych. Jako pierwszy przykład należy przeanalizować nałożenie się dwóch fal o jednakowych częstościach, przeciwnych kierunkach propagacji i identycznych polaryzacjach kołowych. Przyjmując dla uproszczenia zerową fazę początkową oraz polaryzację kołową prawoskrętną (σ^+), celowe jest zastosowanie zapisu wykładniczego. Wektory pola elektrycznego obu fal, przy wykorzystaniu zespolonego wektora polaryzacji ($\mathbf{x} - i\mathbf{y}$), przyjmują wówczas postać:

$$\begin{aligned}\mathbf{E}_1(\mathbf{r},t) &= E_0^{(+)}(\mathbf{x}-i\mathbf{y})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)},\\ \mathbf{E}_2(\mathbf{r},t) &= E_0^{(+)}(\mathbf{x}-i\mathbf{y})e^{i(-\mathbf{k}\cdot\mathbf{r}-\omega t)}.\end{aligned} \tag{5.18}$$

Wynikiem superpozycji tych pól jest pole wypadkowe o postaci:

$$\begin{aligned}\mathbf{E}(\mathbf{r},t) &= \mathbf{E}_1(\mathbf{r},t) + \mathbf{E}_2(\mathbf{r},t)\\ &= E_0^{(+)}(\mathbf{x}-i\mathbf{y})e^{-i\omega t}\left(e^{i\mathbf{k}\cdot\mathbf{r}} + e^{-i\mathbf{k}\cdot\mathbf{r}}\right)\\ &= 2E_0^{(+)}(\mathbf{x}-i\mathbf{y})\cos(\mathbf{k}\cdot\mathbf{r})e^{-i\omega t}.\end{aligned} \tag{5.19}$$

Przejście do mierzalnego, fizycznego pola elektrycznego wymaga obliczenia części rzeczywistej z powyższego wyrażenia. Wykorzystując wzór Eulera, otrzymujemy:

$$\mathrm{Re}\{\mathbf{E}(\mathbf{r},t)\} = 2E_0^{(+)}\cos(\mathbf{k}\cdot\mathbf{r})[\mathbf{x}\cos(\omega t) - \mathbf{y}\sin(\omega t)]. \tag{5.20}$$

Otrzymane wyrażenie opisuje falę stojącą o polaryzacji kołowej. Jej amplituda jest zmodulowana przestrzennie czynnikiem $\cos(\mathbf{k}\cdot\mathbf{r})$, natomiast wektor pola elektrycznego w każdym punkcie przestrzeni rotuje w czasie, zachowując charakterystyczny dla polaryzacji kołowej stały moduł.

5.2.7 Dwie fale o jednakowych częstotliwościach, ortogonalnych polaryzacjach kołowych i przeciwnych wektorach falowych

Kolejnym istotnym przypadkiem jest superpozycja dwóch fal o ortogonalnych polaryzacjach kołowych, rozchodzących się w przeciwnych kierunkach. W zapisie zespolonym wektory pola przyjmują postać:

$$\begin{aligned}\mathbf{E}_1(\mathbf{r},t) &= E_0^{(+)}(\mathbf{x}-i\mathbf{y})e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)},\\ \mathbf{E}_2(\mathbf{r},t) &= E_0^{(+)}(\mathbf{x}+i\mathbf{y})e^{i(-\mathbf{k}\cdot\mathbf{r}-\omega t)}.\end{aligned} \tag{5.21}$$

Superpozycja tych fal prowadzi do wyrażenia:

$$\begin{aligned}\mathbf{E}(\mathbf{r},t) &= \mathbf{E}_1(\mathbf{r},t)+\mathbf{E}_2(\mathbf{r},t)\\ &= E_0^{(+)}e^{-i\omega t}\left[\mathbf{x}\left(e^{i\mathbf{k}\cdot\mathbf{r}}+e^{-i\mathbf{k}\cdot\mathbf{r}}\right)-i\mathbf{y}\left(e^{i\mathbf{k}\cdot\mathbf{r}}-e^{-i\mathbf{k}\cdot\mathbf{r}}\right)\right]\\ &= 2E_0^{(+)}e^{-i\omega t}\left[\mathbf{x}\cos(\mathbf{k}\cdot\mathbf{r})+\mathbf{y}\sin(\mathbf{k}\cdot\mathbf{r})\right].\end{aligned} \tag{5.22}$$

Wyznaczając część rzeczywistą pola, otrzymujemy:

$$\mathrm{Re}\{\mathbf{E}(\mathbf{r},t)\} = 2E_0^{(+)}\left[\mathbf{x}\cos(\mathbf{k}\cdot\mathbf{r})+\mathbf{y}\sin(\mathbf{k}\cdot\mathbf{r})\right]\cos(\omega t). \tag{5.23}$$

Równanie (5.23) opisuje falę stojącą o polaryzacji liniowej. Wspólny czynnik $\cos(\omega t)$ wskazuje, że wszystkie składowe pola drgają w fazie. Jednakże orientacja płaszczyzny polaryzacji, wyznaczona przez wektor $[\mathbf{x}\cos(\mathbf{k}\cdot\mathbf{r})+\mathbf{y}\sin(\mathbf{k}\cdot\mathbf{r})]$, zmienia się wzdłuż kierunku propagacji, tworząc strukturę helikalną (skręconą liniowo).

5.3 Prędkość fali

5.3.1 Prędkość fazowa

Analizę propagacji fal w przestrzeni należy uzupełnić o doprecyzowanie pojęcia prędkości. We wcześniejszych rozdziałach wprowadzono prędkość rozchodzenia się fali v, daną zależnością:

$$v = \frac{\omega}{k}. \tag{5.24}$$

Wielkość ta nosi nazwę **prędkości fazowej** i oznacza prędkość, z jaką w przestrzeni przemieszcza się ustalona faza fali.

5.3.2 Prędkość grupowa

Analiza równania (5.15) pozwala zauważyć, że w przypadku dudnień pojawia się jeszcze jedna prędkość. Opisuje ona ruch obwiedni fali i przez analogię do prędkości fazowej definiuje się ją jako:

$$v_{gr} = \frac{\Delta\omega}{\Delta k}. \tag{5.25}$$

W przypadku superpozycji wielu fal, definicję (5.25) uogólnia się do postaci różniczkowej:

$$v_{gr} = \frac{d\omega}{dk}. \tag{5.26}$$

Wielkość ta nosi nazwę **prędkości grupowej** i oznacza prędkość, z którą przemieszcza się obwiednia impulsu świetlnego w przestrzeni.

Prędkość grupowa i prędkość fazowa opisują odmienne własności propagacji, dlatego mogą przyjmować różne wartości. W ośrodkach dyspersyjnych prędkość grupowa może znacząco różnić się od prędkości fazowej, w szczególności przyjmując wartości ujemne (kierunek propagacji obwiedni jest wówczas przeciwny do kierunku propagacji fazy). Ponadto zarówno prędkość fazowa v, jak i grupowa v_{gr} mogą przekraczać prędkość światła w próżni c. Fakt ten jednak nie łamie postulatów szczególnej teorii względności. Ograniczenie relatywistyczne dotyczy prędkości przesyłu informacji, a w obszarach silnej dyspersji ani prędkość fazowa, ani grupowa nie są z nią tożsame. Kwestia ta wykracza poza ramy niniejszego rozdziału i została szerzej omówiona w rozdziale 9.6.1.

Wykorzystując relację dyspersji $\omega = vk$, równanie (5.26) można zapisać w postaci:

$$v_{gr} = v + k\frac{dv}{dk}. \tag{5.27}$$

Równanie to wprost wiąże ze sobą prędkość fazową i grupową. Prędkość grupową można uzależnić również bezpośrednio od współczynnika załamania n:

$$\frac{1}{v_{gr}} = \frac{dk}{d\omega} = \frac{d}{d\omega}\left(\frac{\omega n}{c}\right) = \frac{n}{c} + \frac{\omega}{c}\frac{dn}{d\omega}, \tag{5.28}$$

co po przekształceniach daje wyrażenie:

$$v_{gr} = \frac{c}{n + \omega\frac{dn}{d\omega}}. \tag{5.29}$$

W optyce często znacznie wygodniejsze jest wyrażenie prędkości grupowej w funkcji długości fali w próżni λ, co wynika ze sposobu pomiaru dyspersji materiałowej $n(\lambda)$. Wykorzystując relację między wektorem falowym a długością fali, można wyprowadzić tzw. wzór Rayleigha:

$$v_{gr} = v - \lambda\frac{dv}{d\lambda}, \tag{5.30}$$

co przy użyciu współczynnika załamania przyjmuje powszechnie stosowaną w fizyce doświadczalnej postać:

$$v_{gr} = \frac{c}{n - \lambda\frac{dn}{d\lambda}}. \tag{5.31}$$

Mianownik w równaniu (5.31) definiuje nową wielkość fizyczną, zwaną **grupowym współczynnikiem załamania** (oznaczanym jako n_g):

$$n_{gr} = n - \lambda \frac{dn}{d\lambda}. \tag{5.32}$$

Dzięki temu prędkość grupową można zapisać w postaci w pełni analogicznej do prędkości fazowej:

$$v_{gr} = \frac{c}{n_{gr}}. \tag{5.33}$$

Zależność ta pozwala na bardzo prostą klasyfikację dyspersji ośrodka. W obszarze **dyspersji normalnej** współczynnik załamania maleje wraz ze wzrostem długości fali ($dn/d\lambda < 0$). Oznacza to, że $n_{gr} > n$, a zatem impuls świetlny porusza się w ośrodku wolniej niż jego faza ($v_{gr} < v$). Odwrotna sytuacja występuje w obszarze **dyspersji anomalnej** ($dn/d\lambda > 0$), gdzie $v_{gr} > v$.

5.3.3 Prędkość fazowa w przewodnikach

W rozdziale 4.2 wykazano, że w przewodnikach liczba falowa k przyjmuje postać zespoloną (równanie (4.69)). Podstawiając część rzeczywistą wektora falowego [równanie (4.71b)] do definicji prędkości fazowej (5.24), otrzymujemy:

$$v = \frac{\omega}{\mathrm{Re}(k)} = \frac{1}{\sqrt{\varepsilon\mu}} \sqrt{\frac{2}{\sqrt{1 + \left(\frac{\sigma}{\omega\varepsilon}\right)^2} + 1}}. \tag{5.34}$$

Z analizy równania (5.34) wynika, że prędkość fazowa w dielektryku ($\sigma \approx 0$) jest większa niż w przewodniku.

5.4 Interferencja konstruktywna i destruktywna

Jak wielokrotnie wspominano, w zakresie widzialnym wielkością rejestrowaną przez detektory nie jest pole elektryczne światła, lecz jego energia lub natężenie. Natężenie fali powstałej przez nałożenie na siebie dwóch fal składowych (przy założeniu propagacji w próżni i przy wykorzystaniu notacji zespolonej) dane jest

zależnością:

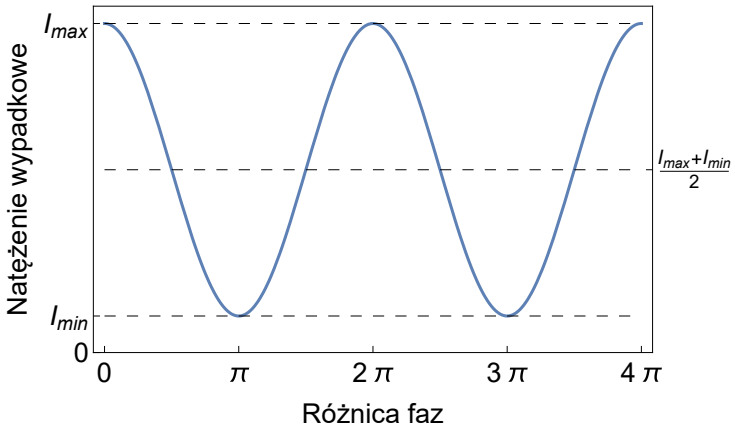
$$
\begin{aligned}
I &= \frac{1}{2}\varepsilon_0 c|E_1+E_2|^2 = \frac{1}{2}\varepsilon_0 c(E_1+E_2)(E_1^*+E_2^*)\\
&= \frac{1}{2}\varepsilon_0 c\left(|E_1|^2+|E_2|^2+E_1E_2^*+E_1^*E_2\right)\\
&= \frac{1}{2}\varepsilon_0 c\Big\{(E_{01}^{(+)})^2+(E_{02}^{(+)})^2+E_{01}^{(+)}E_{02}^{(+)}\times\\
&\quad\left[e^{i[(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}-(\omega_1-\omega_2)t+(\varphi_1-\varphi_2)]}+e^{-i[(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}-(\omega_1-\omega_2)t+(\varphi_1-\varphi_2)]}\right]\Big\} \qquad (5.35)\\
&= \frac{1}{2}\varepsilon_0 c\Big\{(E_{01}^{(+)})^2+(E_{02}^{(+)})^2\\
&\quad +2E_{01}^{(+)}E_{02}^{(+)}\cos[(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}-(\omega_1-\omega_2)t+(\varphi_1-\varphi_2)]\Big\}\\
&= I_1+I_2+2\sqrt{I_1I_2}\cos[(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{r}-(\omega_1-\omega_2)t+(\varphi_1-\varphi_2)].
\end{aligned}
$$

Analiza równania (5.35) wskazuje jednoznacznie, że oprócz prostej sumy dwóch natężeń, w wyrażeniu pojawia się człon mieszany, który moduluje wypadkowe natężenie w czasie i przestrzeni. Należy jednak podkreślić, że dla skutecznej obserwacji interferencji kluczowe jest, aby fazy interferujących wiązek nie zmieniały się w czasie w sposób przypadkowy. W przeciwnym razie wypadkowe natężenie ulegałoby stochastycznym fluktuacjom, a człon interferencyjny uległby uśrednieniu do zera.

Warto przeanalizować przypadek fal rozchodzących się w tę samą stronę i charakteryzujących się identycznymi częstościami ($\mathbf{k}_1 = \mathbf{k}_2$ oraz $\omega_1 = \omega_2$). Równanie (5.35) redukuje się wówczas do postaci:

$$
\boxed{I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(\Delta\varphi),} \qquad (5.36)
$$

gdzie $\Delta\varphi = \varphi_1 - \varphi_2$ oznacza różnicę faz obu fal. Zależność od różnicy faz sprawia, że wypadkowe natężenie oscyluje pomiędzy wartościami I_{min} oraz I_{max} (rys. 5.3). Gdy różnica faz przyjmuje wartości z przedziału od $-\pi/2$ do $\pi/2$ (z dokładnością do całkowitej wielokrotności 2π), wartość wypadkowego natężenia przekracza sumę natężeń fal składowych. Natomiast w przedziale od $\pi/2$ do $3\pi/2$ (również z dokładnością do wielokrotności 2π) natężenie wypadkowe jest mniejsze od tej sumy. W szczególnym przypadku, gdy $I_1 = I_2$, obserwowana modulacja rozpina się pomiędzy wartością 0 a $4I_1$. Maksimum natężenia określane jest mianem **konstruktywnej interferencji**, natomiast minimum — **destruktywnej interferencji**. Interferencja konstruktywna zachodzi, gdy różnica faz fal interferujących jest równa parzystej wielokrotności π, to znaczy $\Delta\varphi = 2m\pi$, gdzie m jest liczbą całkowitą. Interferencja destruktywna występuje natomiast dla nieparzystych wielokrotności, co wyraża się warunkiem $\Delta\varphi = (2m+1)\pi$.



Rysunek 5.3: Modulacja wypadkowego natężenia światła powstałego w wyniku superpozycji wiązek o natężeniach I_1 i I_2. W zaprezentowanym przykładzie natężenie pierwszej wiązki było czterokrotnie większe od natężenia drugiej wiązki ($I_1 = 4I_2$).

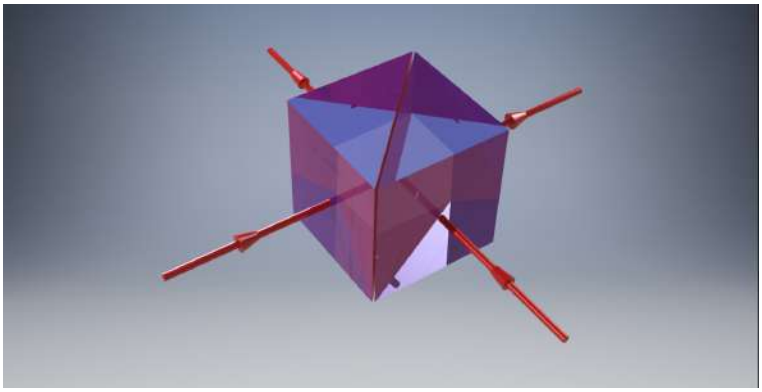
Dygresja

Ponieważ natężenie jest miarą uśrednionej w czasie energii, może powstać błędne wrażenie, że w procesie interferencji dochodzi do kreacji dodatkowej energii, co naruszałoby fundamentalną zasadę zachowania energii. Sprzeczność ta jest pozorna i rozwiązuje się ją poprzez całościową analizę zagadnienia.

Zajście zjawiska interferencji wymaga fizycznego nałożenia na siebie dwóch wiązek. Proces ten często realizowany jest za pomocą **dzielnika wiązki**. Element ten rozdziela padającą wiązkę na dwie ścieżki (wiązka odbita i transmitowana), których stosunek podziału zależy od parametrów technologicznych układu. Najbardziej typowym przypadkiem jest kostka światłodzieląca, realizująca podział w stosunku 50/50.

Warto rozważyć sytuację, w której na kostkę 50/50 padają z dwóch kierunków spójne wiązki światła. Fale transmitowane i odbite od obu wejść nakładają się na siebie parami w kanałach wyjściowych (rys. 5.4). Należy zauważyć, że przy odbiciu od granicy ośrodka optycznie gęstszego (np. na granicy powietrze-szkło wewnątrz warstwy światłodzielącej), jedna ze składowych doznaje przesunięcia fazowego o π. W konsekwencji, jeżeli w jednym z kanałów wyjściowych dochodzi do interferencji konstruktywnej (wzmocnienia), w drugim musi zachodzić interferencja destruktywna (wygaszenie). Natężenie w jednym z kanałów jest zatem większe niż suma natężeń wnoszonych do niego niezależnie, lecz w kanale drugim jest odpowiednio mniejsze. Całkowita moc wprowadzona do układu pozostaje rygo-

rystycznie zachowana – interferencja prowadzi wyłącznie do redystrybucji energii w przestrzeni, a nie do jej generacji czy strat.



Rysunek 5.4: Interferencja dwóch spójnych wiązek na kostce światłodzielącej. Jeżeli w jednym kanale dochodzi do konstruktywnej interferencji, w drugim interferencja jest destruktywna. Zjawisko to ilustruje zasadę zachowania energii w układach interferencyjnych: energia ulega jedynie redystrybucji w przestrzeni.

W optyce falowej klasyfikuje się dwa podstawowe mechanizmy uzyskiwania wiązek spójnych: **interferencję przez podział frontu falowego** oraz **interferencję przez podział amplitudy**. Oba te zjawiska zostały szczegółowo omówione w kolejnych podrozdziałach.

5.5 Interferencja przez podział amplitudy

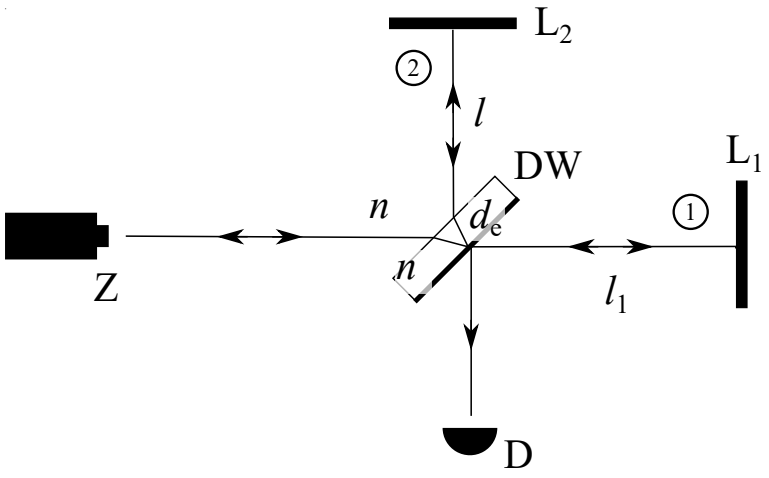
5.5.1 Interferometr Michelsona

Interferometr Michelsona jest sztandarowym przykładem urządzenia, w którym dochodzi do interferencji światła na skutek podziału amplitudy. Układ ten składa się z dzielnika wiązki, zwanego także **lustrem półprzepuszczalnym**, oraz dwóch luster całkowicie odbijających, które służą do kierowania promieni wewnątrz ramion interferometru (rys. 5.5).

Doświadczenie Prążki interferencyjne widoczne przy pomocy interferometru Michelsona w transmisji i odbiciu.

W układzie tym promień wychodzący ze źródła światła (Z) pada na dzielnik wiązki (DW), załamuje się na jego przedniej ściance i wnika do wnętrza szklanej płytki (odbicie Fresnela na tej powierzchni jest zazwyczaj pomijalne). Następnie promień dociera do specjalnie pokrytej, tylnej powierzchni dzielnika, gdzie następuje podział amplitudy wiązki (na część odbitą i załamaną) w stosunku 50/50.

Wiązka kierowana w ramię (1) pokonuje najpierw drogę optyczną nl_1, po czym odbija się od lustra L_1, doznając określonego skoku fazy φ_o. Następnie wraca po tej samej drodze do punktu podziału. Całkowita droga optyczna



Rysunek 5.5: Interferometr Michelsona. Oznaczenia: Z – źródło światła, L_1 i L_2 – lustra, DW – dzielnik wiązki, D – detektor.

pokonana przez tę wiązkę wynosi:

$$\Delta_1 = 2nl_1, \tag{5.37}$$

gdzie n jest współczynnikiem załamania ośrodka w ramieniu, a l_1 to jego długość geometryczna.

Z kolei druga wiązka, ulegając wewnętrznemu odbiciu na powłoce dielektrycznej dzielnika (od ośrodka gęstszego optycznie), doznaje skoku fazy o π. Przebywa ona w szkle drogę geometryczną d_e, po czym wychodzi do ośrodka o współczynniku załamania n, pokonując w nim odległość l_2. Następnie odbija się od lustra L_2 (doznając również skoku fazy φ_o) i wraca tą samą drogą. Łączna droga optyczna tej wiązki w ramieniu (2) wyraża się wzorem:

$$\Delta_2 = 2n_e d_e + 2nl_2, \tag{5.38}$$

gdzie n_e jest współczynnikiem załamania materiału, z którego wykonano dzielnik wiązki, natomiast l_2 i d_e to odpowiednio długość ramienia (2) oraz droga geometryczna pokonana przez promień w samym dzielniku.

Różnica faz $\Delta\varphi = \varphi_1 - \varphi_2$, jakiej doznają obie wiązki po powrocie do dzielnika (przy założeniu identycznego skoku fazy φ_o przy odbiciu na obu lustrach L), wynosi:

$$\Delta\varphi = (2k_0 nl_1 + \varphi_o) - (2k_0 n_e d_e + 2k_0 nl_2 + \varphi_o + \pi), \tag{5.39}$$

gdzie k_0 jest liczbą falową opisującą falę w próżni. Po uproszczeniu różnica faz dana jest zależnością:

$$\Delta\varphi = 2k_0 n(l_1 - l_2) - 2k_0 n_e d_e - \pi. \tag{5.40}$$

Jeżeli różnica faz $\Delta\varphi$ jest parzystą wielokrotnością π (włączając w to zero), wiązki interferują ze sobą konstruktywnie. Dla nieparzystych wielokrotności π następuje interferencja destruktywna.

Uwaga!

W klasycznych układach interferometru Michelsona w ramieniu (1) umieszcza się dodatkową płytkę płaskorównoległą (tzw. płytkę kompensacyjną), wykonaną z identycznego materiału co dzielnik wiązki DW i ustawioną pod tym samym kątem. Zadaniem tej płytki jest kompensacja dodatkowej fazy $2k_0 n_e d_e$ (wynikającej z przejścia przez szkło). Taka konfiguracja niweluje efekty dyspersji, co jest niezbędne do obserwacji prążków interferencyjnych przy użyciu niemonochromatycznego światła białego.

Należy pamiętać, że w układzie dochodzi również do interferencji wiązek kierowanych z powrotem w stronę źródła światła. Obraz interferencyjny powstający w tym kierunku jest komplementarny do obrazu na ekranie detektora – zachodzi tu zjawisko dopełniania (tam, gdzie w jednym kierunku pojawia się maksimum, w drugim występuje minimum natężenia). Wynika to z faktu asymetrii odbić i dodatkowego skoku fazy o π w jednym z kanałów.

5.5.2 Interferometr Michelsona a źródła rozciągłe

Z punktu widzenia obserwatora znajdującego się w miejscu detektora D, interferometr Michelsona nie jest układem dwóch luster i płytki płaskorównoległej (tej ostatniej obserwator nie widzi), ale układem luster L_1 i L_2 oddalonych od siebie o odległość $d = l_1 - l_2$. Rozważmy teraz promień wychodzący z punktu Z źródła rozciągłego. Efektywnie tworzy on dwa równoległe promienie, które są ogniskowane na ekranie przy użyciu soczewki S. Różnica dróg optycznych pomiędzy promieniem odbitym od lustra L_1 i od lustra L_2 wynosi:

$$\Delta_{12} = 2nd\cos\theta, \tag{5.41}$$

gdzie θ jest kątem padania na lustro. Różnica ta prowadzi do pojawienia się geometrycznej różnicy faz. Całkowita różnica faz ulega jednak dodatkowemu przesunięciu o π, co wynika z odbicia od ośrodka gęstszego optycznie. Wówczas warunek na konstruktywną interferencję przyjmuje postać:

$$2d\cos\theta = \left(m + \frac{1}{2}\right)\lambda, \tag{5.42}$$

gdzie λ jest długością fali światła w ośrodku wypełniającym interferometr, zaś m jest liczbą całkowitą.

Analogiczne rozważania można przeprowadzić dla drugiego promienia, w którym nie ma skoku fazy, zaś równanie na konstruktywną interferencję dane jest zależnością:

$$2d\cos\theta = m\lambda. \tag{5.43}$$

Oba równania potwierdzają sformułowaną już wcześniej hipotezę o komplementarności obrazów powstałych dla obu wiązek.

Warto na koniec wspomnieć, że podobne rozumowanie można sformułować w sytuacji, w której źródło jest punktowe, ale emituje światło w postaci fali sferycznej. Różne fale interferują wtedy, tworząc obraz w postaci koncentrycznych pierścieni. Kąt, pod którym powstają te promienie zależy od długości fali.

Dla interferometru Michelsona, podobnie jak dla innych interferometrów, można określić tzw. **widzialność prążków**. Jest ona równa unormowanej różnicy minimalnego I_{min} i maksymalnego I_{max} natężenia obserwowanego w interferometrze:

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \tag{5.44}$$

Ponieważ różne jest położenie prążków w zależności od długości fali, dla światła niemonochromatycznego prążki nakładają się na siebie. W konsekwencji dla tego światła spada widzialność prążków.

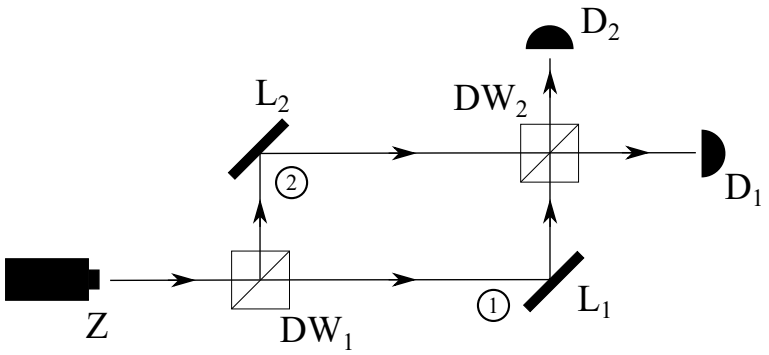
Dygresja

Jednym z najbardziej spektakularnych zastosowań interferometru Michelsona jest detektor fal grawitacyjnych LIGO. To monstrualne urządzenie, o ramionach długości 4 km, jest niczym innym jak urządzeniem wykrywającym zmiany długości ramion interferometru pod wpływem rozchodzącej się fali grawitacyjnej. Zmiany, które może wykryć LIGO sięgają 10^{-18} m, czyli ok. 0,001 rozmiarów protonu.

5.5.3 Interferometr Macha-Zehndera

W układzie interferometru Macha-Zehndera (rys. 5.6) pojedyncza wiązka światła dzielona jest na dwie wiązki przy pomocy dzielnika wiązki 50/50 (DW_1). Po przejściu przez każde z ramion interferometru wiązki są na siebie nakładane na wyjściowej kostce światłodzielącej DW_2, która również miesza je w stosunku 50/50. W konsekwencji powstają dwie wiązki wyjściowe, kierowane odpowiednio do detektora D_1 i D_2. Nałożenie na siebie wiązek propagujących się w ramieniu (1) oraz (2) sprawia, że dochodzi między nimi do interferencji, przy czym, podobnie jak w przypadku interferometru Michelsona, odbicie wewnątrz dzielnika wiązki sprawia, że fazy jednej ze składowych są w obu ramionach przesunięte o π. W ten sposób, kiedy w jednym z ramion dochodzi do interferencji konstruktywnej, w drugim ramieniu interferencja jest destruktywna (i na odwrót).

Wprowadzenie do jednego z ramion interferometru elementu optycznego kontrolującego opóźnienie fazy umożliwia stworzenie **modulatora natężenia światła**. Urządzenie takie można skonstruować, wykorzystując np. **kryształ elektrooptyczny** (rozdział 10.4), którego współczynnik załamania zależy od przyłożonego pola elektrycznego. Zmiana ta odbywa się bardzo szybko, co



Rysunek 5.6: Interferometr Macha-Zehndera. Oznaczenia: Z – źródło światła, L_1 i L_2 – lustra, DW_1 i DW_2 – dzielnik wiązki dzielący natężenie w stosunku pół na pół (50/50), D_1 i D_2 – detektory.

sprawia, że równie szybko modulowane może być światło. Dlatego też tego typu urządzenia wykorzystywane są w modulatorach natężenia światła w sieciach telekomunikacyjnych, a ich pasmo modulacji sięga nawet setek GHz.

Dygresja

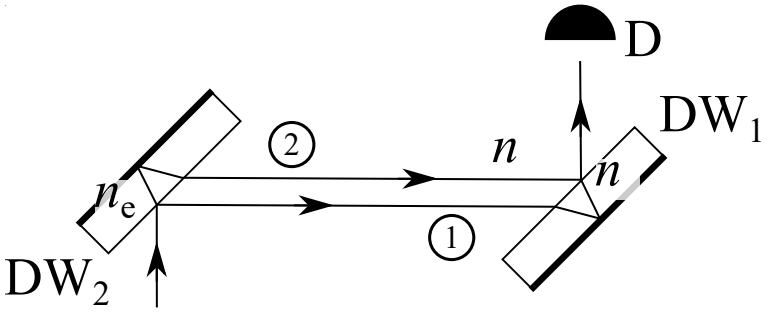
Bardzo ciekawym zastosowaniem interferometrów Macha-Zehndera jest interferencja jednofotonowa. W takim przypadku interferometr oświetlany jest światłem z urządzenia pozwalającego na emisję pojedynczych fotonów. Należy podkreślić, że źródło to musi emitować pojedyncze fotony, a nie bardzo słabe impulsy światła laserowego, ponieważ w tych dwóch przypadkach zupełnie inna jest statystyka i kwantowa charakterystyka pola świetlnego.

Wpuszczając pojedynczy foton do interferometru Macha-Zehndera jego funkcja falowa ulega podziałowi na dwa ramiona interferometru, które następnie interferują ze sobą na wyjściowym dzielniku wiązki. W przypadku gdy dwie drogi w interferometrze są kwantowo nierozróżnialne, na wyjściu dochodzi do interferencji fotonu samego ze sobą. Jej konsekwencją może być m.in. fakt zajścia konstruktywnej interferencji w jednym z ramion interferometru i kierowanie go zawsze w ten sam kanał. Gdy różnica faz w jednym z ramion zmieni się o π, wówczas dojdzie do destruktywnej interferencji w tym ramieniu i żadne fotony nie będą rejestrowane przez detektor umieszczony w tym kanale wyjściowym (wszystkie fotony rejestrowane będą przez detektor umieszczony na drugim wyjściu).

Warto na koniec zauważyć, że określenie ramienia, w którym rozchodził się rejestrowany foton, np. przez przypadkowe zasłonięcie jednego z ramion, sprawia, że interferencja znika, a fotony dzielą się równo pomiędzy oba kanały. Obserwacja ta dowodzi, że w układzie dochodzi do superpozycji stanów fotonu z samym sobą, a nie jakiegoś statystycznego uśredniania zachowania paczek wielu fotonów.

5.5.4 Interferometr Jamina

Swego rodzaju modyfikacją interferometru Macha-Zehndera jest **interferometr Jamina**, na który składają się dwie grube płytki płaskorównoległe (dzielniki wiązki), których jedna z powierzchni jest pokryta warstwą półprzepuszczającą, a druga warstwą całkowicie odbijającą (rys. 5.7). W ten sposób wiązka światła



Rysunek 5.7: Interferometr Jamina.

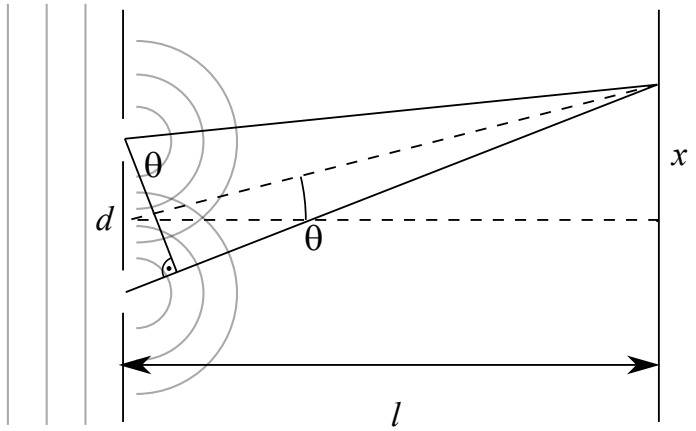
odbitego oraz wiązka światła załamanego i odbitego od wewnętrznej powierzchni mogą ze sobą interferować, a natężenie wypadkowego światła zależy od różnicy dróg optycznych każdej z wiązek.

5.6 Interferencja przez podział frontu falowego

5.6.1 Eksperyment Younga

Doświadczenie Obserwacja prążków interferencyjnych od światła laserowego przechodzącego przez dwie szczeliny

Rozważmy dwie szczeliny, które oświetlane są falą płaską. W oryginalnym **eksperymencie Thomasa Younga** (1773-1829) warunek ten zapewniony był przez umieszczenie przed szczelinami dodatkowej szczeliny, która oświetlana była światłem słonecznym. Ponieważ szczelina stawała się źródłem fal sferycznych, których wycinek, jak pokazano w rozdziale 3.3, może być traktowany jako fala płaska, światło docierające do dwóch szczelin było spójne, a sama fala była płaska. W dzisiejszych eksperymentach warunek oświetlenia szczelin spójną falą można stosunkowo łatwo osiągnąć wykorzystując lasery. Szczeliny, do których dociera światło z pierwszej szczeliny (kiedyś) lub z lasera (dziś) są następnie źródłem kolejnych fal kulistych, które oświetlają ekran umieszczony w dużej odległości w porównaniu z odległością między szczelinami (rys. 5.8). Jeżeli ekran ustawiony jest bardzo daleko, to można przyjąć, że promienie docierające do niego są równoległe. Wówczas różnica dróg optycznych tych promieni docierających do określonego punktu na ekranie zależy jedynie od kąta θ, pod którym ten punkt jest obserwowany ze szczelin (precyzyjniej jest to kąt promieni względem prostej prostopadłej do ekranu ze szczelinami). Prowadząc rozważania geometryczne, można pokazać, że dla konstruktywnej interferencji światła, przy której na ekranie obserwowany jest jasny prążek, spełniony musi



Rysunek 5.8: Eksperyment Younga. Dwie szczeliny oddalone od siebie o odległość d oświetlane są falą płaską. Szczeliny stają się źródłami fali kulistej, które docierając do ekranu, tworzą na nim obraz interferencyjny. Promienie docierające do ekranu obserwowane są pod kątem θ.

być warunek:

$$d \sin \theta_m = m\lambda, \tag{5.45}$$

gdzie d jest odległością między szczelinami, θ_m jest kątem pod którym obserwowany jest jasny prążek na ekranie, m jest liczbą całkowitą numerującą prążki, zaś λ jest długością fali światła.

Chcąc obliczyć odległość dwóch kolejnych jasnych prążków na ekranie, można zauważyć, że $\sin \theta_m \approx x_m/l$, skąd otrzymujemy, że:

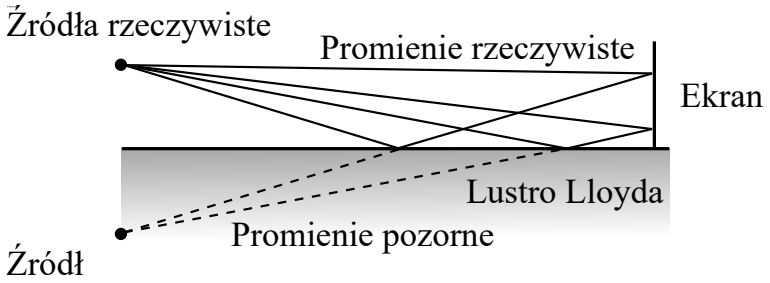
$$\Delta x = x_{m+1} - x_m = l \sin \theta_{m+1} - l \sin \theta_m = l \left(\frac{m+1}{d}\lambda - \frac{m}{d}\lambda \right) = \frac{l\lambda}{d}. \tag{5.46}$$

Jak widać zatem, odległość pomiędzy jasnymi prążkami obserwowanymi na ekranie zależy od długości fali λ. Warto na marginesie zauważyć, że użyte w eksperymencie fizyczne szczeliny w rzeczywistości wytwarzają fale cylindryczne, jednak w analizowanym przekroju 2D opis geometrii fal rygorystycznie pokrywa się z falami sferycznymi.

Eksperyment Younga jest sztandarowym przykładem interferencji przez podział frontu falowego.

5.6.2 Lustro Lloyda

Przykładem urządzenia realizującego interferencję przez podział frontu falowego jest **lustro Lloyda**. W układzie tym pojedyncze źródło światła oświetla ekran wprost lub po odbiciu od ustawionego poziomo lustra (rys. 5.9). W lustrze powstaje pozorny obraz źródła. W ten sposób na ekranie interferują dwa rodzaje



Rysunek 5.9: Lustro Lloyda. Światło ze źródła punktowego oświetla ekran bezpośrednio lub po odbiciu od lustra umieszczonego poziomo pod nim. Na skutek różnych dróg optycznych każdego z promieni na ekranie powstaje struktura dyfrakcyjna.

promieni, te pochodzące wprost ze źródła oraz te, które wydają się pochodzić ze źródła pozornego. Skutkiem interferencji „obu" wiązek jest powstanie obrazu interferencyjnego na ekranie.

5.6.3 Bipryzmat Fresnela

Bipryzmat Fresnela jest elementem szklanym złożonym z dwóch pryzmatów o podstawie w kształcie trójkąta prostokątnego. W układzie tym światło emitowane jest z pojedynczego źródła umieszczonego w osi z bipryzmatem Fresnela (rys. 5.10). Na skutek przejścia przez bipryzmat, wiązki wychodzące z pryzmatu uginane są w stronę osi optycznej i przecinają się w jednym miejscu. Przedłużenia tych promieni po stronie źródła tworzą dwa symetryczne względem osi pozorne źródła. W ten sposób wydaje się, że dochodzi do interferencji światła pochodzącego z dwóch niezależnych źródeł. Jej skutkiem jest powstanie obrazu interferencyjnego.

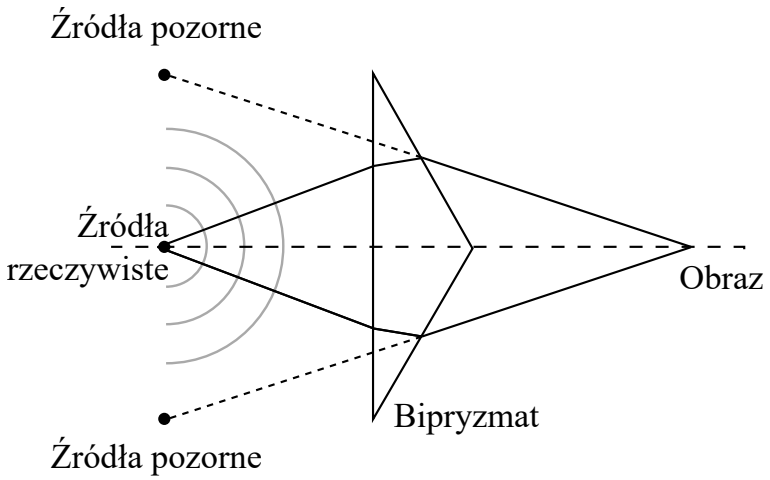
5.7 Długość i czas spójności

Wyobraźmy sobie, że interferometr Michelsona oświetlany jest światłem, które nie jest w pełni monochromatyczne, tzn. ma niezerową rozciągłość w skali długości fali $\Delta\lambda$. Skutkiem tej skończonej szerokości spektralnej źródła jest to, że w interferometrze przestaniemy móc rozróżniać poszczególne prążki wtedy, kiedy różnica faz dla fali centralnej λ i krańcowej $\lambda + \Delta\lambda$ przesunie się o 2π:

$$\Delta\varphi = \frac{2\pi}{\lambda} 2d\cos\theta - \frac{2\pi}{\lambda + \Delta\lambda} 2d\cos\theta = 2\pi. \tag{5.47}$$

Zakładając, że szerokość spektralna jest niewielka w stosunku do długości fali, tzn. $\Delta\lambda \ll \lambda$, możemy równanie (5.47) przepisać do postaci:

$$\frac{\Delta\lambda}{\lambda^2} 2d\cos\theta = 1. \tag{5.48}$$



Rysunek 5.10: Bipryzmat Fresnela. Promienie docierające ze źródła punktowego podlegają załamaniu w kierunku osi optycznej. Na skutek ich interferencji dochodzi do powstania obrazu interferencyjnego. Obraz ten jest taki sam jak w przypadku, gdyby interferowały ze sobą promienie wychodzące z dwóch źródeł pozornych.

Prawa strona otrzymanej granicznej różnicy dróg optycznych definiuje fundamentalną własność samego źródła światła, tzw. **długość spójności światła,** zwaną także **drogą koherencji światła**. Otrzymujemy zatem zależności:

$$l_{koh} = \frac{\lambda^2}{\Delta\lambda} = \lambda\frac{\nu}{\Delta\nu} = c\tau_{koh}, \tag{5.49}$$

gdzie τ_{koh} zwane jest **czasem koherencji światła**. Interpretacja drogi spójności jest następująca – jest to maksymalna różnica dróg optycznych jaką pokona fala odszczepiona od wiązki głównej (w naszym układzie równa granicznej wartości $2d\cos\theta$), dla której, po nałożeniu jej z wiązką pierwotną, wciąż obserwowana będzie wyraźna interferencja. Czas, w którym światło pokonuje tę drogę, jest czasem spójności.

W oparciu o równanie (5.49) można pokazać, że pomiędzy drogą i czasem spójności oraz szerokością spektralną, wyrażoną w dziedzinie częstotliwości, istnieje związek:

$$\tau_{koh}\Delta\nu = 1. \tag{5.50}$$

Związek ten wynika z tzw. **ograniczenia fourierowskiego**. Analizując równanie (5.50) łatwo można zauważyć, że im bardziej monochromatyczne jest światło, tzn. im mniejsza jest jego szerokość spektralna, tym dłuższy jest czas koherencji światła, a co za tym idzie, jego droga spójności.

Dygresja

Rozważmy termiczne źródło światła, które emituje promieniowanie o widmie ciągłym w zakresie od 450 nm do 950 nm. Wykorzystując tę informację, można zapisać, że

$$\Delta\nu = 3 \times 10^{14} \text{ Hz}. \tag{5.51}$$

Stąd otrzymujemy, że czas spójności dla tego światła to zaledwie 10 fs, zaś droga spójności wynosi 1 μm.

Prowadząc podobne rozważania dla światła laserowego emitującego promieniowanie o długości fali 800 nm i szerokości spektralnej 2 pm otrzymujemy, że czas spójności wynosi ok. 1 ns, zaś droga spójności 30 cm.

Droga spójności jest niezwykle ważnym parametrem wykorzystywanym w praktyce, ze szczególnym uwzględnieniem **holografii**. Klasyczna fotografia rejestruje jedynie natężenie światła (kwadrat amplitudy), bezpowrotnie tracąc informacje o fazie fali odbitej od przedmiotu, przez co tradycyjne zdjęcia są płaskie. Holografia, wynaleziona przez Dennisa Gabora, pozwala na zapisanie i odtworzenie pełnego, trójwymiarowego frontu falowego dzięki wykorzystaniu zjawiska interferencji.

W procesie zapisu hologramu spójna wiązka lasera dzielona jest na dwie części. Jedna z nich (wiązka przedmiotowa) oświetla obiekt i po rozproszeniu na nim pada na kliszę fotograficzną. Druga (wiązka odniesienia) pada bezpośrednio na tę samą kliszę. Fale te nakładają się na materiale światłoczułym, tworząc skomplikowany wzór prążków interferencyjnych. Niezbędnym warunkiem zajścia tego procesu jest użycie lasera o bardzo dużej drodze spójności, która musi przewyższać wymiary fizyczne fotografowanego obiektu (różnica dróg optycznych między różnymi punktami przedmiotu nie może przekroczyć l_{koh}).

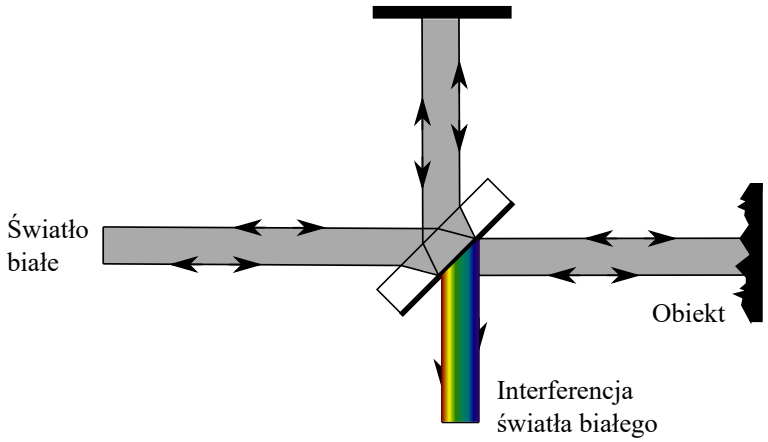
Ten mikroskopijny labirynt prążków to *de facto* niezwykle złożona siatka dyfrakcyjna (patrz rozdział 7.7), w której zamrożona została informacja o przesunięciach fazowych pochodzących od trójwymiarowej topografii przedmiotu. Gdy wywołaną kliszę oświetlimy ponownie samą wiązką odniesienia, zjawisko dyfrakcji zrekonstruuje oryginalny front falowy, tworząc przed oczami obserwatora idealnie trójwymiarowy, wirtualny obraz przedmiotu.

5.8 Interferencja światła białego

Interferencja światła może zachodzić nie tylko dla światła monochromatycznego, ale również dla światła białego. Przykładowo, oświetlenie obiektu wiązką światła białego i nałożenie odbitej od niego wiązki z referencyjną (spójną) wiązką światła białego może prowadzić do powstania wzoru interferencyjnego. W szczególności na ekranie obserwowane będą kolorowe struktury wynikające z konstruktywnej interferencji dla określonych długości fali. Z tego typu wzorem interferencyjnym możemy mieć do czynienia, gdy dwie szczeliny w eksperymencie Younga oświetlimy światłem białym. W takim przypadku na środku ekranu,

w równej odległości pomiędzy szczelinami, obserwowany będzie wyraźny prążek światła białego (dla którego różnica dróg wynosi zero niezależnie od fali), a wraz z oddalaniem się od środka pojawiać się będą tęczowe prążki, które bardzo szybko zaczną się rozmywać i na siebie nakładać.

Jednym z przykładów zastosowań interferencji światła białego są precyzyjne badania powierzchni profilometrycznych. Na tego typu badania pozwala interferometr Michelsona oświetlony światłem z szerokopasmowego źródła (rys. 5.11). Umieszczenie w jednym z ramion interferometru badanej powierzchni i później-



Rysunek 5.11: Interferencja światła białego w interferometrze Michelsona. Na skutek nałożenia na siebie światła odbitego od obiektu oraz spójnego światła referencyjnego dochodzi do interferencji na długościach fali determinowanych przez topografię powierzchni.

sze zinterferowanie światła odbitego z wiązką referencyjną pozwala na bardzo precyzyjne zbadanie topografii powierzchni, zwłaszcza dzięki temu, że światło białe "wygasza"interferencję wszędzie poza ścisłym obszarem równej drogi optycznej (z uwagi na mikroskopijną drogę spójności), co zapobiega błędnej identyfikacji rzędu prążków interferencyjnych.

5.9 Podsumowanie rozdziału

W niniejszym rozdziale zaprezentowano zjawisko interferencji, u którego podstaw leży addytywność równania falowego i wynikająca z niej zasada superpozycji w ośrodkach liniowych. Przeprowadzono szczegółową analizę nakładania się fal, ukazując bogactwo efektów fizycznych: od powstawania fal stojących i przestrzennej modulacji natężenia, po zjawisko dudnień wynikające z superpozycji fal o zbliżonych częstościach. Rozważania te stały się naturalnym punktem wyjścia do rygorystycznego rozróżnienia prędkości fazowej oraz prędkości grupowej, stanowiącej kluczowy parametr opisujący propagację obwiedni impulsu świetlnego w ośrodkach dyspersyjnych.

Kluczowym elementem analizy było matematyczne sformułowanie warunków interferencji konstruktywnej i destruktywnej dla wielkości mierzalnych makroskopowo, czyli natężenia światła. Wykazano przy tym, że powstawanie prążków nie łamie zasady zachowania energii, a stanowi jedynie wyraz jej przestrzennej redystrybucji. Przedstawiony aparat pojęciowy pozwolił na systematyczną klasyfikację technik interferometrycznych na te opierające się na podziale amplitudy (np. interferometry Michelsona i Macha-Zehndera) oraz na podziale frontu falowego (klasyczne doświadczenie Younga, lustro Lloyda, bipryzmat Fresnela).

Fundamentalnym ograniczeniem obserwowalności tych zjawisk okazała się rozciągłość spektralna rzeczywistych źródeł światła. Zdefiniowanie czasu i drogi koherencji uwypukliło rolę spójności, co znajduje bezpośrednie odzwierciedlenie w zaawansowanych technikach inżynierii optycznej, takich jak holografia czy profilometria światła białego. Opanowanie opisu superpozycji dyskretnej, skończonej liczby fal stanowi bezwzględnie niezbędny fundament dla kolejnego etapu naszej analizy. Jak zostanie wykazane w następnym rozdziale, uogólnienie tej zasady – polegające na nałożeniu na siebie ciągłego widma nieskończenie wielu fal cząstkowych – pozwoli nam na pełne zrozumienie i matematyczny opis zjawiska dyfrakcji.

Zadania

5.1. Określ parametry (rodzaj fali, kierunki wektorów pola elektrycznego i magnetycznego, amplitudy) fali powstałej w wyniku złożenia dwóch przeciwbieżnych fal elektromagnetycznych o tej samej częstotliwości i amplitudzie, spolaryzowanych liniowo:

(a) w tym samym kierunku,

(b) w kierunkach prostopadłych.

5.2. Określ parametry (rodzaj fali, kierunki wektorów pola elektrycznego i magnetycznego, amplitudy) fali powstałej w wyniku złożenia dwóch przeciwbieżnych fal elektromagnetycznych o tej samej częstotliwości i amplitudzie, które są spolaryzowane kołowo:

(a) z tą samą skrętnością,

(b) z przeciwnymi skrętnościami.

5.3. Oblicz długość spójności (koherencji) światła lasera He-Ne ($\lambda = 632{,}8$ nm), jeżeli szerokość spektralna emitowanej linii wynosi $\Delta\lambda = 1{,}33 \times 10^{-12}$ m.

5.4. Dwie fale monochromatyczne o długościach fali $\lambda_1 = 633$ nm i $\lambda_2 = 650$ nm nakładają się na detektorze. Określ częstotliwość dudnień powstających na skutek nałożenia się tych fal.

5.5. Zjonizowany gaz, tj. plazma, jest ośrodkiem dyspersyjnym dla fal elektromagnetycznych. Mając dane równanie dyspersji w postaci:

$$\omega^2 = \omega_{\mathrm{p}}^2 + c^2 k^2,$$

gdzie ω_{p} jest stałą częstością plazmową, wyznacz wyrażenia dla prędkości fazowej i grupowej oraz wykaż, że $v_f v_g = c^2$.

5.6. Fala elektromagnetyczna porusza się w ośrodku dyspersyjnym z prędkością fazową v daną zależnością:

$$v = a + b\lambda^2,$$

gdzie a i b są stałymi materiałowymi, a λ jest długością fali światła. Oblicz prędkość grupową v_g fali propagującej w tym ośrodku.

5.7. Określ warunki, dla których prędkość grupowa światła w ośrodku mogłaby przyjąć makroskopowo małą wartość, np. $v_g = 3$ km/h (tzw. zjawisko spowolnionego światła).

5.8. We włóknie światłowodowym wykonanym ze szkła kwarcowego propaguje się impuls świetlny o centralnej długości fali $\lambda = 500$ nm i początkowym czasie trwania równym:

(a) 1 ps,

(b) 0,1 ps.

Długość włókna wynosi 10 m, a współczynnik załamania kwarcu wynosi odpowiednio $n(400\text{ nm}) = 1{,}470$ oraz $n(600\text{ nm}) = 1{,}458$. Oblicz, jak długo będzie propagowało się światło z czerwonego i niebieskiego skrzydła spektralnego tego impulsu. Jakie ma to konsekwencje dla wydłużenia impulsu po przebyciu całego włókna?

5.9. Za jedną ze szczelin w eksperymencie Younga wstawiono cienką przezroczystą folię o współczynniku załamania $n = 1{,}48$, co spowodowało przesunięcie prążka centralnego o 4 prążki. Przyjmując długość fali $\lambda = 550$ nm oblicz o ile folia zwiększyła drogę optyczną w ramieniu i jaka jest jej grubość?

5.10. W eksperymencie Younga na ekranie obserwujemy pierwszy jasny prążek pod kątem $\theta = 2^\circ$. Oblicz odległość kątową między kolejnymi jasnymi prążkami obserwowanymi na ekranie, zakładając, że odległość między szczelinami wynosi $d = 0{,}2$ mm.

5.11. W doświadczeniu Younga wykorzystano oświetlenie światłem białym. Czy w tym przypadku możliwa jest obserwacja jasnego prążka w obrazie interferencyjnym? Uzasadnij, jak wygląda oświetlenie ekranu w miarę oddalania się od jego środka.

5.12. Znajdź położenie centralnego prążka interferencyjnego w eksperymencie Younga, w którym jedna ze szczelin została przesłonięta płytką płaskorównoległą o grubości a i współczynniku załamania n. Źródło światła umieszczone jest symetrycznie względem osi układu, odległość między szczelinami wynosi d, a odległość pomiędzy płaszczyzną szczelin i ekranem wynosi l.

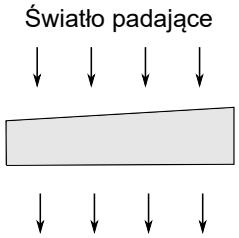
5.13. W układzie lustra Lloyda punktowe źródło światła o długości fali $\lambda = 600$ nm znajduje się w odległości $h = 2$ mm od płaszczyzny poziomego lustra. Ekran umieszczono w odległości $L = 2$ m od źródła.

(a) Oblicz odległość między sąsiednimi jasnymi prążkami na ekranie.

(b) Odpowiedz na pytanie, czy prążek brzegowy (stykający się bezpośrednio z płaszczyzną lustra, gdzie różnica dróg geometrycznych obu promieni wynosi zero) jest jasny, czy ciemny? Odpowiedź rygorystycznie uzasadnij, odwołując się do zasad fizyki.

5.14. Bipryzmat Fresnela użyty został do uzyskania obrazu interferencyjnego z punktowego źródła emitującego światło o długości fali $\lambda = 500$ nm. Zakładając, że pryzmat wykonany jest ze szkła o współczynniku załamania $n = 1{,}5$, znajduje się w odległości $a = 1$ m od źródła i $b = 2$ m od ekranu, oblicz kąty łamiące pryzmatu, dla których odległość między sąsiednimi prążkami interferencyjnymi wynosi 0,5 mm.

5.15. Szeroka wiązka światła o długości fali $\lambda = 630$ nm pada prostopadle na cienki klin z materiału o współczynniku załamania $n = 1{,}5$. Obserwator oglądający ten klin w świetle przechodzącym widzi 10 jasnych i 9 ciemnych prążków wzdłuż całej długości klina. O ile zmienia się grubość klina na tej długości?



5.16. W interferometrze Michelsona dzielnik wiązki idealnie dzieli padające światło w stosunku 50/50. Niestety, jedno z luster w ramionach uległo zabrudzeniu i odbija jedynie 64% padającego na nie natężenia światła (drugie lustro odbija 100%). Oblicz widzialność prążków V obserwowanych na ekranie detektora.

5.17. W jednym z ramion interferometru Michelsona umieszczono szklaną komórkę o długości 2 cm wypełnioną powietrzem. Zanim z komórki odpompowano powietrze, odległości luster zostały dobrane w taki sposób, że w środku ekranu widoczny był jasny prążek dla oświetlenia światłem laserowym o długości fali $\lambda = 550$ nm. Następnie otwarto zawór i powoli rozpoczęto odpompowywanie powietrza. O ile prążków przesunie się obraz na ekranie, jeżeli współczynnik załamania powietrza wynosi 1,00028, zaś docelową próżnię w komórce charakteryzuje $n = 1{,}00000$?

5.18. Różnica dróg geometrycznych w dwóch ramionach interferometru Michelsona jest równa $\Delta L = 0{,}5$ mm, a układ oświetlany jest światłem o długości fali $\lambda = 632{,}8$ nm. Zakładając, że ośrodkiem wypełniającym interferometr jest powietrze o współczynniku załamania $n = 1{,}02$, oblicz, ile maksimów interferencyjnych zaobserwuje detektor podczas przesuwania jednego z luster o dystans 1 mm.

5.19. W interferometrze Michelsona jedno z luster porusza się ze stałą prędkością $v = 2$ mm/s wzdłuż osi optycznej. Określ częstość ν modulacji natężenia światła $I(t)$ rejestrowanego na wyjściu z interferometru, jeżeli długość fali światła oświetlającego układ wynosi 633 nm.

5.20. Prążki interferencyjne uzyskiwane przy pomocy interferometru Jamina obserwowane są w płaszczyźnie ogniskowej soczewki o ogniskowej $f = 15$ cm umieszczonej za interferometrem. Płytki interferometru wykonane są z materiału o $n = 1{,}5$ i mają grubość $d = 2$ cm. Obrócono je względem siebie o kąt $1'$, a światło pada na pierwszą z nich pod kątem $45°$. Jeżeli prążki interferencyjne oddalone są od siebie o 5 mm, to jaka jest długość fali światła padającego na układ?

Rozdział 6

Interferencja wielowiązkowa

6.1 Wprowadzenie

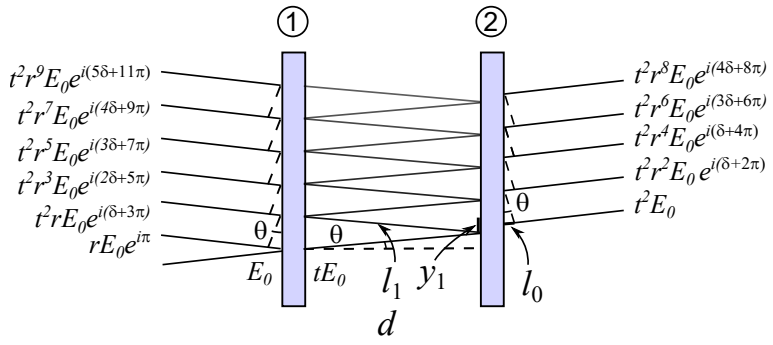
Rozważmy dwa lustra charakteryzujące się wysokim współczynnikiem odbicia (dla uproszczenia rozważań przyjmijmy, że ich reflektancja wynosi $R = 0{,}99$). W takim przypadku natężenie światła transmitowanego przez pierwsze lustro stanowi zaledwie 1% natężenia światła padającego, co jest w pełni zgodne z intuicją. Interesujące zjawiska zachodzą jednak w sytuacji, gdy za pierwszym lustrem umieścimy równolegle do niego kolejne lustro o takim samym współczynniku odbicia. Na pierwszy rzut oka można by przypuszczać, że skoro współczynnik odbicia drugiego lustra jest również równy 0,99, to całkowite natężenie światła transmitowanego przez ten układ wyniesie zaledwie 10^{-4} natężenia fali padającej na pierwsze lustro. Okazuje się jednak, że takie proste rozumowanie jest błędne.

Gdy fala ulega wielokrotnym odbiciom pomiędzy lustrami, który to układ od tego momentu będziemy nazywać **wnęką optyczną**, dochodzi do „uwięzienia" pomiędzy nimi promieniowania. Podczas każdego odbicia wewnątrz wnęki tylko niewielka część zgromadzonej energii opuszcza układ. Rozpatrując ten proces wyłącznie jako niekoherentne zjawisko sumowania natężeń (przy sumowaniu pomijamy fazę fali), można wyprowadzić matematyczny wniosek, że w ostatecznym bilansie w przybliżeniu połowa energii, która wniknęła do wnęki, zostanie przepuszczona zgodnie z początkowym kierunkiem propagacji, podczas gdy druga połowa ulegnie odbiciu z powrotem w stronę źródła. Jednak i to rozumowanie jest fałszywe. Pomija ono bowiem kluczową w tym przypadku rolę faz fal elektromagnetycznych opuszczających wnękę przy każdym odbiciu i związane z nimi zjawisko interferencji.

6.2 Interferometr Fabry'ego-Perota

Aby przyjrzeć się problemowi interferencji wielowiązkowej w sposób bardziej systematyczny, rozważmy układ dwóch płaskich luster umieszczonych równole-

gle do siebie w odległości d (rys. 6.1). Układ taki jest nazywany **interferometrem Fabry'ego-Perota** (Charles Fabry 1867–1945, Alfred Perot 1863–1925) i stanowi klasyczny przykład urządzenia realizującego **interferencję wielowiązkową**. Wyobraźmy sobie teraz, że interferometr ten oświetlany jest falą płaską, która pada na pierwsze lustro pod kątem θ (bez utraty ogólności możemy założyć, że lustra mają zaniedbywalną grubość, ośrodkiem między nimi jest próżnia, a odbicie na nich można traktować jako odbicie od pojedynczej powierzchni dielektrycznej, opisywanej efektywnymi współczynnikami odbicia r i transmisji t).



Rysunek 6.1: Układ interferometru Fabry'ego-Perota składającego się z dwóch luster umieszczonych równolegle do siebie w odległości d, oświetlanych falą płaską padającą pod kątem θ. Symbole r i t to amplitudowe współczynniki odpowiednio: odbicia i transmisji luster.

Na pierwszej granicy ośrodków pomiędzy lustrem (1) a ośrodkiem wypełniającym obszar wnęki dochodzi do podziału amplitudy wiązki padającej. Część fali odbija się od lustra, a część jest przez nie transmitowana. Po wniknięciu do wnętrza interferometru i propagacji w nim światło natrafia na drugie lustro (2). Tutaj zachodzi podobny proces, tj. część wiązki zostaje odbita i skierowana wstecz, ku lustru (1), podczas gdy pozostała część przechodzi przez lustro (2) i opuszcza interferometr. Proces ten powtarza się wielokrotnie, a przy każdym padaniu na lustro, niezależnie od tego, czy jest to lustro (1) czy (2), dochodzi do odbicia i transmisji światła. W konsekwencji powstaje nieskończony ciąg spójnych fal o malejących amplitudach i ściśle określonych przesunięciach fazowych, które wychodzą z obu stron interferometru.

6.2.1 Natężenie światła transmitowanego przez interferometr Fabry'ego-Perota

Rozważmy ten proces teraz ilościowo, opierając się na analitycznej reprezentacji fali elektromagnetycznej, co pozwoli nam w pełni uwzględnić koherentną superpozycję. Wykorzystując amplitudowe współczynniki Fresnela, możemy określić parametry każdej z wiązek opuszczających układ. Zakładając, że na wejście

interferometru pada fala, której zespolona amplituda wynosi $E_0^{(+)}$, amplituda pierwszej wiązki transmitowanej przez oba lustra jest równa:

$$E_1^{(+)} = E_0^{(+)} t^2, \tag{6.1}$$

gdzie t jest amplitudowym współczynnikiem transmisji (przyjmujemy lustra o identycznych parametrach). Kolejna wiązka, która opuszcza układ po doznaniu dwukrotnego odbicia wewnątrz wnęki (od lustra (2) i z powrotem od (1)), musi zostać pomnożona przez odpowiednie współczynniki odbicia wewnętrznego r' oraz musi uwzględniać narastające opóźnienie fazowe δ:

$$E_2^{(+)} = E_0^{(+)} t^2 (r')^2 e^{i\delta}. \tag{6.2}$$

Z uwagi na fakt, że dla typowych układów dielektrycznych przejście światła odbywa się z ośrodka rzadszego do gęstszego, współczynniki odbicia z zewnątrz (r) i od wewnątrz (r') wiąże relacja Stokesa $r' = -r$. Ponieważ jednak światło zamknięte we wnęce zawsze odbija się parzystą liczbę razy zanim ją opuści, zmiana znaku przy podniesieniu do kwadratu znika: $(r')^2 = (-r)^2 = r^2$. W ogólnym przypadku zespolona amplituda j-tej wiązki transmitowanej przez układ dana jest przez:

$$E_j^{(+)} = E_0^{(+)} t^2 r^{2(j-1)} e^{i(j-1)\delta}. \tag{6.3}$$

Jak pokazano powyżej, oprócz zmiany amplitudy każda z wiązek zyskuje przesunięcie fazowe podczas przejścia pomiędzy lustrami. Droga geometryczna, którą pokonuje fala między lustrami, jest równa

$$l_1 = \frac{d}{\cos\theta}, \tag{6.4}$$

gdzie d jest odległością między lustrami. Wynika stąd, że dodatkowa droga geometryczna, którą pokonuje fala odbijająca się jednokrotnie od obu luster, wynosi $2l_1$. Należy jednak zauważyć, że fala opuszczająca interferometr jest falą płaską, która rozchodzi się pod kątem θ do osi optycznej. Dlatego też jej front falowy nachylony jest pod kątem θ do powierzchni luster (na rys. 6.1 front falowy dla obu wiązek został zaznaczony linią przerywaną tuż za lustrem (2)). Wynika stąd, że rozważaną wiązkę trzeba porównywać z wiązką, która wyszła z interferometru jako pierwsza i pokonała w przestrzeni drogę l_0. Drogę tę można obliczyć, opierając się na analizie geometrycznej:

$$l_0 = 2y_2 \sin\theta = 2d \operatorname{tg}\theta \sin\theta = \frac{2\mathrm{d}\sin^2\theta}{\cos\theta}. \tag{6.5}$$

W konsekwencji opóźnienie fazowe δ pomiędzy dwiema kolejnymi wiązkami opuszczającymi interferometr wynosi:

$$\delta = k(2l_1 - l_0) = k\left(\frac{2d}{\cos\theta} - \frac{2d\sin^2\theta}{\cos\theta}\right) = 2kd\cos\theta, \tag{6.6}$$

gdzie założono, że współczynnik załamania wewnątrz i na zewnątrz wnęki jest taki sam.

Uwaga!

W przypadku, gdy do odbicia dochodzi na powierzchni dielektryka o wyższym współczynniku załamania, faza fali zmienia się skokowo o wartość π. Ponieważ kolejne wiązki transmitowane przez wnękę odbijają się od luster parzystą liczbę razy, zsumowana różnica faz od tych odbić wynosi zawsze całkowitą wielokrotność 2π. Z tego powodu w arytmetyce falowej skoki fazy przy transmisji znoszą się i można je w tym miejscu zaniedbać. Sytuacja wygląda jednak zgoła inaczej w przypadku analizy fali odbitej, gdzie pierwsze odbicie zachodzi od zewnątrz, co zostanie omówione w dalszej części rozdziału.

Wypadkowa fala transmitowana przez interferometr jest koherentną superpozycją wszystkich wyłaniających się fal składowych. Stanowi ona sumę nieskończonego szeregu geometrycznego o ilorazie $r^2e^{i\delta}$:

$$E_t^{(+)} = \sum_{j=1}^{\infty} E_j^{(+)} = E_0^{(+)} t^2 \sum_{j=0}^{\infty} \left(r^2 e^{i\delta}\right)^j = \frac{E_0^{(+)} t^2}{1 - r^2 e^{i\delta}}. \tag{6.7}$$

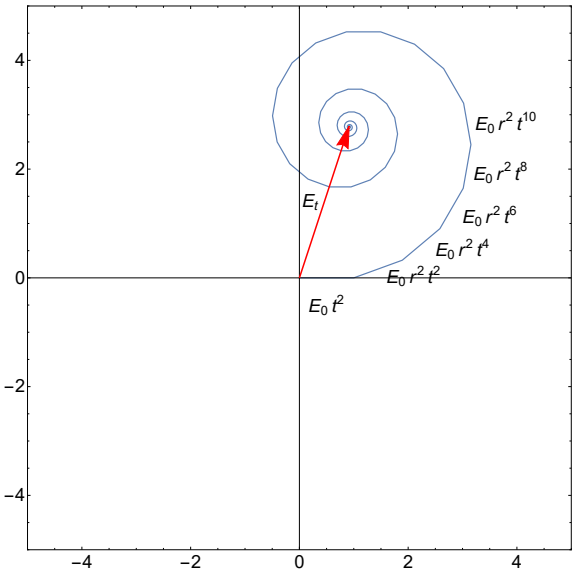
Dygresja

Proces powstawania fali transmitowanej można przedstawić w sposób niezwykle intuicyjny, posługując się tzw. **metodą wskazów**. W metodzie tej interferencję wielowiązkową obrazuje się jako dodawanie szeregu wektorów (wskazów) na płaszczyźnie zespolonej. Długość każdego wektora determinowana jest przez amplitudę rozpatrywanej wiązki częściowej, zaś jego kierunek określa faza fali względem poprzednika (kąt δ).

W warunku pełnego rezonansu ($\delta = 2m\pi$) wszystkie wskazy skierowane są zgodnie, tworząc pojedynczą linię prostą, co prowadzi do maksymalnej amplitudy wypadkowej. Gdy faza, choćby nieznacznie, odbiega od warunku rezonansu (rys. 6.2), ciąg wektorów zaczyna zwijać się w spiralę, co sprawia, że ich ostateczna suma (wektor łączący początek pierwszego i koniec ostatniego wskazu) gwałtownie maleje. Narzędzie to doskonale tłumaczy, dlaczego dla wysokich wartości współczynnika odbicia R (kiedy superpozycja uwzględnia bardzo długi ciąg istotnych amplitudowo wskazów) piki transmisyjne stają się tak niezwykle wąskie.

Przywołując zależność pomiędzy natężeniem i zespoloną amplitudą światła, $I = \frac{1}{2}\varepsilon_0 c|E^{(+)}|^2$, wyznaczamy w pierwszej kolejności kwadrat modułu fali transmitowanej, mnożąc wyrażenie (6.7) przez jego sprzężenie zespolone:

$$|E_t^{(+)}|^2 = E_t^{(+)} \left(E_t^{(+)}\right)^* = \frac{|E_0^{(+)}|^2 t^4}{(1 - r^2 e^{i\delta})(1 - r^2 e^{-i\delta})} = \frac{|E_0^{(+)}|^2 t^4}{1 - 2r^2 \cos\delta + r^4}. \tag{6.8}$$



Rysunek 6.2: Wykres wskazowy dla światła transmitowanego przez układ dwóch luster o współczynnikach odbicia 95% oraz fazie akumulowanej pomiędzy dwoma lustrami równej 20°.

W tym miejscu wygodnie jest powiązać te wielkości z makroskopowymi parametrami układu: reflektancją $R = |r|^2$ oraz transmitancją $T = |t|^2$ (należy podkreślić, że w rozważanym przypadku zarówno fala padająca, jak i transmitowana propagują się w tym samym ośrodku, dlatego współczynniki geometryczne występujące w ogólnej definicji transmitancji skracają się). Zakładając z zasady zachowania energii układ pozbawiony strat ($T+R=1$), możemy zastąpić czynnik t^4 w liczniku wyrażeniem $(1-R)^2$, a mianownik zapisać jako $1-2R\cos\delta+R^2$. Wykorzystując tożsamość $\cos\delta = 1 - 2\sin^2(\delta/2)$, otrzymujemy fundamentalną dla optyki rezonansowej zależność na natężenie światła przechodzącego:

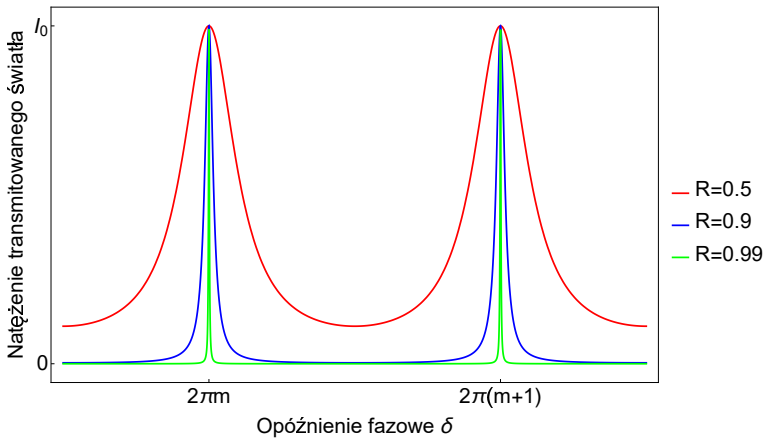
$$I_t = \frac{I_0}{1 + \frac{4R}{(1-R)^2}\sin^2(\delta/2)}. \tag{6.9}$$

Równanie (6.9) nosi nazwę **wzoru Airy'ego** (George Airy 1801–1892). Przewiduje ono, że przy warunku pełnego rezonansu fali z wnęką ($\delta = 2m\pi$, gdzie m to liczba całkowita), układ staje się idealnie „przezroczysty", tj. $I_{\max} = I_0$, niezależnie od tego, jak wysoki jest współczynnik odbicia samych luster R. Minimum transmisji zachodzi dla różnicy faz będącej nieparzystą wielokrotnością π:

$$\delta = 2\pi\left(m + \frac{1}{2}\right) \Rightarrow I_{\min} = \frac{I_0}{1 + \frac{4R}{(1-R)^2}}. \tag{6.10}$$

Wyjaśnieniem tego intrygującego zjawiska jest wzmocnienie wnękowe – spułapkowane natężenie rośnie do wartości tak znacznej, że nawet minimalny procent przepuszczany przez barierę drugiego lustra sumuje się do 100% natężenia zasilającego z zewnątrz.

Zależność rozkładu funkcji Airy'ego od współczynnika odbicia luster została zobrazowana na rys. 6.3. Im większa wartość R, tym smuklejsze i ostrzejsze stają się rezonanse, co bezpośrednio podnosi precyzję pomiarów widmowych.



Rysunek 6.3: Zależność natężenia światła transmitowanego przez interferometr Fabry'ego-Perota od narastającej różnicy faz δ (funkcja Airy'ego) dla trzech wartości współczynnika odbicia luster R.

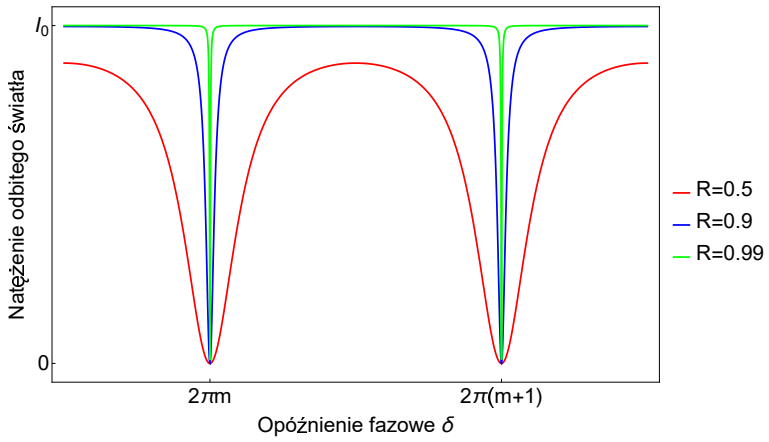
Ważnym parametrem określającym jakość uzyskiwanego obrazu jest **kontrast prążków interferencyjnych**, zwany też ich **widzialnością**. Zależy on bezpośrednio od współczynnika odbicia luster i definiuje się go jako:

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{2R}{1 + R^2}. \tag{6.11}$$

Zależność ta dowodzi, że nawet dla stosunkowo niskich wartości współczynnika odbicia (np. $R = 0{,}8$) widzialność pozostaje bardzo wysoka ($V \approx 0{,}975$). Należy jednak pamiętać, że w analizie widmowej i rezonatorach optycznych priorytetem rzadko jest sam kontrast — znacznie istotniejsze są wysoka rozdzielczość i ostre rezonanse, które w świetle wzoru Airy'ego można osiągnąć dopiero dla granicy $R \to 1$.

6.2.2 Natężenie światła odbitego od interferometru Fabry'ego-Perota

Z punktu widzenia zasady zachowania energii naturalne jest zapytanie o los fotonów, które nie uległy transmisji. Zespolona amplituda fali odbitej powstaje



Rysunek 6.4: Natężenie światła odbijanego od układu dwóch równoległych luster (interferometr Fabry-Perota) w funkcji różnicy faz pomiędzy dwoma kolejnymi promieniami propagującymi się w układzie.

przez nałożenie fali odbitej bezpośrednio od pierwszej zewnętrznej granicy interferometru (z amplitudą początkową $E_0^{(+)}r$) oraz szeregu fal przenikających z powrotem przez to lustro z wnętrza wnęki:

$$\begin{aligned} E_r^{(+)} &= E_0^{(+)}r + E_0^{(+)}tt'r'e^{i\delta} + E_0^{(+)}tt'(r')^3e^{2i\delta} + \dots \\ &= E_0^{(+)}r - E_0^{(+)}tt're^{i\delta}\sum_{j=0}^{\infty}\left(r^2e^{i\delta}\right)^j \\ &= E_0^{(+)}r\left(1 - \frac{tt'e^{i\delta}}{1-r^2e^{i\delta}}\right) = E_0^{(+)}r\frac{1-e^{i\delta}}{1-r^2e^{i\delta}}, \end{aligned} \tag{6.12}$$

gdzie ponownie odwołano się do relacji Stokesa ($r' = -r$ oraz $tt' = 1-r^2$), wiążących współczynniki wejścia i wyjścia z gęstszego ośrodka optycznego. Kwadrat modułu tego wyrażenia, przy założeniu braku strat absorpcyjnych, prowadzi wprost do równania na natężenie fali odbitej:

$$\boxed{I_r = I_0\frac{\frac{4R}{(1-R)^2}\sin^2(\delta/2)}{1+\frac{4R}{(1-R)^2}\sin^2(\delta/2)}.} \tag{6.13}$$

Natężenie światła odbitego od interferometru Fabry'ego-Perota przedstawiono na rys. 6.4.

Porównanie profilu I_r z rozkładem natężenia fali transmitowanej wprost dowodzi, że krzywe te są dla siebie idealnie komplementarne ($I_t + I_r = I_0$). Brak odbicia w stanie pełnego rezonansu wnękowego jest wynikiem całkowicie destruktywnej interferencji (o stałej różnicy faz wynoszącej dokładnie π) między

pierwszą falą odbitą z zewnątrz a wypadkową wszystkich fal opuszczających z powrotem rezonator.

6.2.3 Współczynnik finezji

Kształt („ostrość") prążków wyznaczony przez funkcję Airy'ego zależy od parametru $4R/(1-R)^2$. Definiuje on w zasadniczy sposób, z jaką stromością opadają zbocza krzywej w pobliżu rezonansu. Wielkość tę określa się mianem **współczynnika finezji** i tradycyjnie oznacza literą F:

$$F = \frac{4R}{(1-R)^2}. \tag{6.14}$$

Współczynnik ten determinuje głębokość minimów pomiędzy kolejnymi szczytami transmisyjnymi. Przepisując w oparciu o niego transmitancję wnęki, uzyskujemy najpopularniejszą w optyce instrumentalnej formę zapisu funkcji Airy'ego:

$$I_t = \frac{I_0}{1 + F\sin^2(\delta/2)}. \tag{6.15}$$

6.2.4 Finezja rezonatora ($\mathcal{F}$)

Pojęciem nierozłącznie związanym ze współczynnikiem finezji (i stanowiącym jego rozwinięcie o dużo głębszej interpretacji fizycznej) jest **finezja interferometru** ($\mathcal{F}$). Definiuje się ją jako stosunek odległości między dwoma sąsiednimi pikami transmisji interferometru, co nazywamy **przedziałem spektralnym** FSR, do całkowitej szerokości połówkowej pojedynczego rezonansu (FWHM), mierzonymi w domenie przesunięcia fazowego δ:

$$\mathcal{F} = \frac{\Delta\delta_{\mathrm{FSR}}}{\delta_{1/2}}. \tag{6.16}$$

Ponieważ układ rezonansowy powtarza swoje piki transmisyjne cyklicznie co 2π, licznik wynosi $\Delta\delta_{\mathrm{FSR}} = 2\pi$. Szerokość połówkową wyznacza się z warunku spadku natężenia we wzorze Airy'ego do połowy wielkości maksymalnej:

$$\frac{1}{1 + F\sin^2(\delta_{1/2}/4)} = \frac{1}{2} \quad \Rightarrow \quad F\sin^2\left(\frac{\delta_{1/2}}{4}\right) = 1. \tag{6.17}$$

Dla wysokich wartości współczynnika odbicia ($R \to 1$), rezonanse stają się niezwykle wąskie, a aproksymacja małych kątów $\sin(x) \approx x$ pozwala zapisać

$\delta_{1/2} \approx 4/\sqrt{F}$ (patrz Rozdział 6.3.1). Połączenie tych dwóch faktów daje jednoznaczny, bezwymiarowy wzór na finezję:

$$\mathcal{F} = \frac{\pi\sqrt{F}}{2} = \frac{\pi\sqrt{R}}{1-R}. \tag{6.18}$$

Finezja $\mathcal{F}$ ma bezpośrednie, namacalne odzwierciedlenie w fizyce urządzenia: stanowi przybliżoną miarę tego, ile średnio „cykli uwięzienia" (wewnętrznych odbić) pokonuje fala zmagazynowana pomiędzy zwierciadłami, zanim jej pierwotna energia wycieknie z rezonatora.

6.3 Interferometr Fabry'ego-Perota jako analizator widma

Wysoka czułość transmitancji na zmianę długości fali sprawia, że interferometr Fabry'ego-Perota stanowi niezwykle precyzyjne narzędzie do **analizy widmowej światła**. Ze względu na fakt, że fale o różnych długościach λ podlegają we wnęce innym opóźnieniom fazowym δ, układ ten pozwala na separację przestrzenną bardzo blisko leżących linii widmowych. Jakość i przydatność takiego analizatora optycznego określają jednoznacznie dwa fundamentalne parametry: zdolność rozdzielcza oraz wolny przedział spektralny.

6.3.1 Zdolność rozdzielcza

Rozważmy dwie wiązki światła o długościach fal λ_1 i λ_2, które wprowadzamy do interferometru Fabry'ego-Perota. Za interferometrem fale te tworzą podwójną strukturę prążków interferencyjnych. Załóżmy teraz, że dwa prążki m-tego rzędu znajdują się blisko siebie oraz że w ich maksimach zachodzi:

$$\sin\frac{\delta_1}{2} = \sin\frac{\delta_2}{2} = 0, \tag{6.19}$$

gdzie

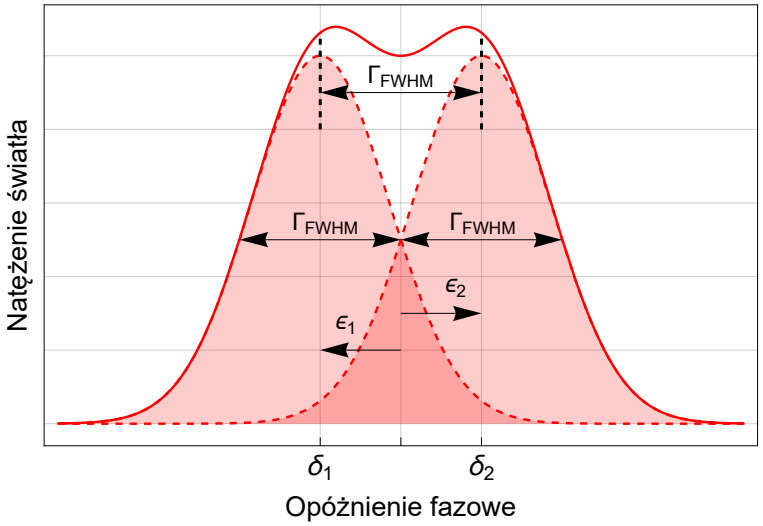
$$\begin{aligned} \delta_1 &= \frac{2\pi}{\lambda_1} 2d\cos\theta_1 = \frac{2\pi}{\lambda} 2d\cos\theta_1, \\ \delta_2 &= \frac{2\pi}{\lambda_2} 2d\cos\theta_2 = \frac{2\pi}{\lambda-\Delta\lambda} 2d\cos\theta_2, \end{aligned} \tag{6.20}$$

przy czym $\Delta\lambda$ jest różnicą długości fal obu wiązek, zaś θ_1 i θ_2 to kąty, pod którymi obserwowane są maksima m-tego rzędu. Pytanie, na które chcemy znaleźć odpowiedź, brzmi: jaka jest najmniejsza różnica długości fali, którą można rozróżnić przy użyciu interferometru?

Warunek rozróżnialności dwóch prążków interferencyjnych zależy od szeregu czynników optycznych. Do celów praktycznych przyjmuje się najczęściej **kryterium oparte na szerokości połówkowej** (zwane również kryterium Taylora),

w myśl którego dwa prążki są ledwo rozróżnialne wtedy, gdy odległość między ich maksimami jest równa szerokości połówkowej (rys. 6.5). W takim przypadku dla punktu leżącego w połowie dystansu między maksimami ułamek natężenia od każdego z prążków wynosi:

$$I = \frac{I_0}{2} \quad \Rightarrow \quad F \sin^2\left(\frac{\delta_{1/2}}{2}\right) = 1. \tag{6.21}$$



Rysunek 6.5: Kryterium szerokości połówkowej pozwalające na rozdzielenie dwóch prążków interferencyjnych. Symbole są zgodne z tymi wykorzystanymi w tekście.

Aby możliwe było rozróżnienie dwóch prążków, zakładamy, że natężenie każdego z nich w punkcie siodłowym spada do połowy swojej wartości maksymalnej. Wówczas spełnione muszą być następujące warunki dla małych przesunięć fazowych ε_1 oraz ε_2 wychylających się od maksimów obu fal:

$$\begin{aligned} F \sin^2 \frac{\delta_1 + \varepsilon_1}{2} &= 1, \\ F \sin^2 \frac{\delta_2 - \varepsilon_2}{2} &= 1. \end{aligned} \tag{6.22}$$

Korzystając ze wzorów na funkcje trygonometryczne sumy kątów oraz zakładając, że $\varepsilon_{1,2} \ll 1$, otrzymujemy:

$$\begin{aligned} \sin \frac{\delta_1 + \varepsilon_1}{2} &= \sin \frac{\delta_1}{2} \cos \frac{\varepsilon_1}{2} + \cos \frac{\delta_1}{2} \sin \frac{\varepsilon_1}{2} = \sin \frac{\varepsilon_1}{2} \approx \pm \frac{\varepsilon_1}{2}, \\ \sin \frac{\delta_2 - \varepsilon_2}{2} &= \sin \frac{\delta_2}{2} \cos \frac{\varepsilon_2}{2} - \cos \frac{\delta_2}{2} \sin \frac{\varepsilon_2}{2} = -\sin \frac{\varepsilon_2}{2} \approx \mp \frac{\varepsilon_2}{2}, \end{aligned} \tag{6.23}$$

gdzie skorzystaliśmy z definicji maksimów $\sin(\delta_{1,2}/2) = 0$ oraz wynikającego z tego faktu, że $\cos(\delta_{1,2}/2) = \pm 1$. Po podstawieniu równań (6.23) do równania (6.22) znak opada w wyniku podniesienia do kwadratu i ostatecznie otrzymujemy:

$$\begin{aligned} F\left(\frac{\varepsilon_1}{2}\right)^2 &= 1, \\ F\left(\frac{\varepsilon_2}{2}\right)^2 &= 1, \end{aligned} \tag{6.24}$$

skąd:

$$\varepsilon_1 = \varepsilon_2 = \frac{2}{\sqrt{F}} \quad \Rightarrow \quad \varepsilon_1 + \varepsilon_2 = \delta_2 - \delta_1 = \frac{4}{\sqrt{F}}. \tag{6.25}$$

Z drugiej strony:

$$\delta_2 - \delta_1 = \frac{2\pi}{\lambda - \Delta\lambda} 2d\cos\theta_2 - \frac{2\pi}{\lambda} 2d\cos\theta_1, \tag{6.26}$$

co przy założeniu, że analizujemy centralną część obrazu interferencyjnego, dla której $\theta_{1,2} \to 0$, pozwala zapisać:

$$\delta_2 - \delta_1 \approx 4\pi d\left(\frac{1}{\lambda - \Delta\lambda} - \frac{1}{\lambda}\right) \approx 4\pi d\frac{\Delta\lambda}{\lambda^2}. \tag{6.27}$$

Po połączeniu równania (6.27) z równaniem (6.25) otrzymujemy relację:

$$\boxed{\frac{\lambda}{\Delta\lambda} = \frac{\pi d}{\lambda}\sqrt{F} = \frac{2\pi d}{\lambda}\frac{\sqrt{R}}{1-R}.} \tag{6.28}$$

Powyższa wielkość nosi nazwę **zdolności rozdzielczej** i jest ona jedną z podstawowych wielkości charakteryzujących przyrządy spektralne, w tym interferometry Fabry'ego-Perota.

6.3.2 Przedział spektralny

Wysoka zdolność rozdzielcza nie rozwiązuje jednak wszystkich problemów spektroskopowych. Kolejnym ograniczeniem przyrządu jest zjawisko **przekrywania się rzędów interferencji**. Aby interpretacja widma była fizycznie jednoznaczna, pik rezonansowy rzędu m dla fali o długości λ nie może zbiec się z pikiem rzędu $(m+1)$ generowanym przez sąsiednią, krótszą długość fali $\lambda - \Delta\lambda$. Sytuacja krytyczna powstaje przy warunku:

$$(m+1)(\lambda - \Delta\lambda_{\mathrm{FSR}}) = m\lambda, \tag{6.29}$$

skąd bezpośrednio $\lambda = (m+1)\Delta\lambda_{\mathrm{FSR}}$.

Zakładając, że rząd interferencji we wnęce optycznej jest z definicji bardzo duży ($m \approx m+1 \approx 2nd/\lambda$), otrzymujemy wzór na **wolny przedział spektralny** (ang. *Free Spectral Range*, FSR). Wielkość ta, wyrażona w dziedzinie

długości fali, określa najszersze okno widmowe, w którym można prowadzić obserwację bez ryzyka aliasingu rzędów:

$$\Delta\lambda_{\mathrm{FSR}} = \frac{\lambda^2}{2nd}. \tag{6.30}$$

Znacznie bardziej naturalne w fizyce nowoczesnych wnęk i fizyce laserów jest wyrażenie FSR w dziedzinie częstotliwości. Opierając się na relacji $\nu = c/\lambda$, przyrost w domenie częstotliwości wiąże się ze zmianą długości fali poprzez różniczkę $|\Delta\nu| = \frac{c}{\lambda^2}\Delta\lambda$. Wykorzystując ten fakt w wyznaczonym przedziale, uzyskujemy wolny przedział spektralny w domenie częstotliwości:

$$\Delta\nu_{\mathrm{FSR}} = \frac{c}{2nd}. \tag{6.31}$$

Wielkość ta jest absolutnie kluczowa dla optyki instrumentalnej: stanowi ona fizyczny odstęp (wyrażony najczęściej w hercach) dzielący dwa sąsiednie, rezonujące we wnęce mody podłużne światła.

6.3.3 Dobroć wnęki i czas życia fotonu

Gdy interferometr traktujemy nie jako narzędzie do spektroskopii interferencyjnej fal, lecz jako fotonowy rezonator stratny, nadrzędnym modelem opisowym staje się dynamika czasowa ewakuacji energii świetlnej ze wnęki. Jeżeli do ustabilizowanego rezonatora dostarczymy światło o natężeniu stałym, a następnie raptownie wyłączymy to światło, to światło we wnęce zacznie zanikać wykładniczo zgodnie z rozkładem czasowym

$$I(t) = I_0 \exp(-t/\tau_p), \tag{6.32}$$

gdzie τ_p to średni czas życia fotonu we wnęce optycznej.

Parametr ten wiąże się wprost z charakterystykami stacjonarnymi opisanymi powyżej, poprzez zależność częstotliwościową szerokości piku z czasem relaksacji:

$$\tau_p = \frac{1}{2\pi\Delta\nu_{\mathrm{FWHM}}} = \frac{1}{2\pi}\frac{\mathcal{F}}{\Delta\nu_{\mathrm{FSR}}} = \frac{\mathcal{F}dn}{\pi c} \tag{6.33}$$

Powyższa relacja stanowi podstawę ultraprecyzyjnej techniki spektroskopowej zwanej **spektroskopią strat we wnęce**.

Dygresja

Spektroskopia strat we wnęce optycznej to niezwykle czuła technika absorpcyjna służąca do detekcji gazów śladowych. Jej działanie opiera się na wykorzystaniu wnęki rezonansowej (zbudowanej ze zwierciadeł o bardzo wysokim współczynniku odbicia, np. $R > 0{,}9999$) jako pułapki dla fotonów.

Pomiar rozpoczyna się od wpuszczenia do wnęki krótkiego impulsu świa-

tła laserowego. Gdy źródło światła zostaje wyłączone, uwięziony impuls odbija się wielokrotnie między lustrami, pokonując we wnęce efektywną drogę optyczną rzędu wielu kilometrów. Ponieważ przy każdym odbiciu wnękę opuszcza znikomy, ale stały ułamek energii, natężenie światła rejestrowanego przez detektor na zewnątrz powoli maleje, tworząc krzywą zaniku wykładniczego. Rejestrując ten spadek, wyznacza się tzw. czas zaniku, który jest wprost proporcjonalny do czasu życia fotonu w rezonatorze.

Obecność gazu absorbującego we wnęce stwarza dodatkowy kanał utraty energii dla propagującej fali, co zauważalnie przyspiesza jej zanik. Porównanie czasu zaniku w rezonatorze pustym z czasem zmierzonym dla wnęki wypełnionej badaną próbką pozwala bardzo precyzyjnie wyznaczyć stężenie absorbera. Fundamentalną przewagą tej techniki nad klasyczną spektroskopią jest to, że wynik opiera się wyłącznie na wyznaczeniu stałej czasowej zaniku, a nie na pomiarze bezwzględnej amplitudy osłabionego sygnału. Dzięki temu metoda ta staje się całkowicie niewrażliwa na szumy i chwilowe wahania mocy samego lasera pompującego.

W mechanice kwantowej i inżynierii mikrofalowej odpowiednikiem klasycznej finezji dla opisu układów rezonansowych i laserów jest **dobroć wnęki**. Opisuje ona stosunek energii zmagazynowanej w oscylatorze do energii rozproszonej w pojedynczym cyklu rezonansowym i powiązana jest wprost ze zjawiskami dyspersji wnęki:

$$Q = 2\pi\nu_0\tau_p = \frac{\nu_0}{\Delta\nu_{\mathrm{FWHM}}} \tag{6.34}$$

Dobroć z punktu widzenia wartości liczbowej stanowi de facto równoważnik absolutnej zdolności rozdzielczej interferometru. W wybitnie zbudowanych we współczesnych warunkach laboratoryjnych wnękach (o finezji rzędu stu tysięcy), czynnik dobroci Q potrafi sięgać gigantycznych wartości rzędu 10^{10}, a tym samym czynić wnęki Fabry'ego-Perota najbardziej czułymi urządzeniami znanymi fizyce optycznej.

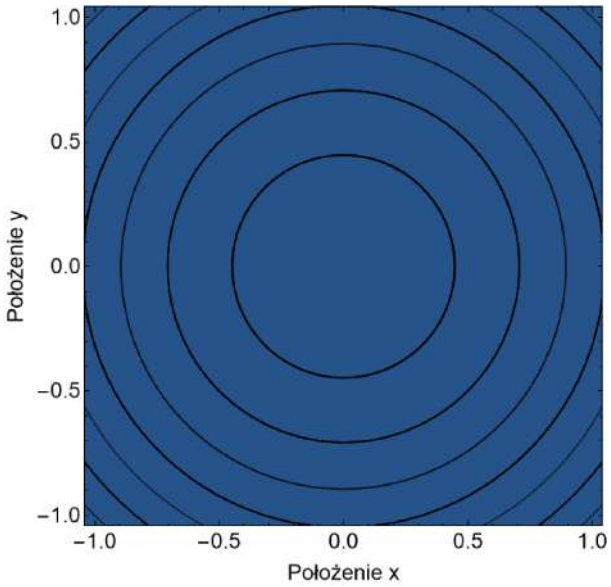
6.4 Interferencja wiązki rozbieżnej

Doświadczenie Obserwacja koncentrycznych pierścieni za interferometrem Fabry'ego-Perota oświetlanym rozbieżną wiązką światła.

Gdy na układ opisany powyżej pada wiązka światła rozbieżnego, np. wiązka emitowana ze źródła rozciągłego przestrzennie, to każda z fal cząstkowych zyskuje nieco inne przesunięcie fazowe. W konsekwencji na interferometr padają fale pod różnymi kątami, a opóźnienie fazowe δ zależy od kąta padania.

Aby na ekranie zaobserwować interferencję tych fal, obraz interferencyjny należałoby obserwować w nieskończoności. Naturalnie nie jest to możliwe, dlatego problem ten rozwiązuje się, wykorzystując soczewkę, która przenosi obraz z nieskończoności do płaszczyzny ogniskowej. Właśnie w tej płaszczyźnie obserwuje się zazwyczaj obrazy interferencyjne zwane **interferogramami**.

W przypadku rozbieżnej wiązki światła interferogram ma kształt współśrodkowych okręgów, których promienie zależą od długości fali, ale również od odległości między lustrami. Obraz taki, zwany **pierścieniami równego nachylenia**, pokazany został na rys. 6.6.



Rysunek 6.6: Koncentryczne pierścienie obserwowane za interferometrem Fabry'ego-Perota oświetlanym rozbieżną wiązką światła.

6.5 Ograniczenia interferometru Fabry'ego-Perota

Podczas wyprowadzania wzoru na interferencję wielowiązkową przyjęto, że w układzie nie występują straty energii, co pozwala zapisać bilans w postaci $R + T = 1$. W rzeczywistości jednak na powierzchniach zwierciadeł zawsze dochodzi do pewnego rozpraszania oraz pochłaniania światła, co stanowi źródło strat w rezonatorze. Co więcej, sam ośrodek wypełniający interferometr również może absorbować promieniowanie. Z tego powodu równanie oparte na zasadzie zachowania energii należy zmodyfikować, przyjmując postać:

$$R + T + A = 1, \tag{6.35}$$

gdzie A jest natężeniowym współczynnikiem strat we wnęce. Uwzględnienie tego czynnika prowadzi do bardziej złożonej zależności opisującej transmitancję interferometru Fabry'ego-Perota i skutkuje zauważalnym pogorszeniem parametrów obserwowanych prążków interferencyjnych.

Drugie założenie, przyjęte w początkowej fazie analizy, dotyczyło interferencji nieskończonej liczby fal opuszczających układ. W praktyce każda wiązka świetlna ma skończone rozmiary poprzeczne, podobnie jak fizycznie ograniczone są rozmiary zwierciadeł tworzących rezonator. Ponadto wiązki zawsze charakteryzują się pewną rozbieżnością, co wynika bezpośrednio z natury źródeł oświetlających układ oraz ze zjawiska dyfrakcji na krawędziach elementów optycznych.

Aby zniwelować ten problem, znacznie częściej niż interferometry z płaskimi lustrami wykorzystuje się urządzenia, w których lustra mają kształt sferyczny i pracują w tzw. **reżimie konfokalnym**.

6.6 Konfokalny interferometr Fabry'ego-Perota

Konfokalny interferometr Fabry'ego-Perota jest układem składającym się z dwóch wklęsłych luster, umieszczonych względem siebie w odległości równej ich promieniowi krzywizny. Takie ustawienie sprawia, że ogniska obu luster się ze sobą pokrywają.

W porównaniu z klasycznym interferometrem, w którym wykorzystywane są tylko dwa płaskie lustra, interferometr konfokalny wykazuje szereg istotnych różnic. Pierwszą z nich jest fakt, że wiązka uwięziona we wnęce odbija się od każdego lustra dwukrotnie, zanim powróci do punktu początkowego i zamknie swoją ścieżkę optyczną.

Z uwagi na taką geometrię propagacji, zmianie ulega wyrażenie na przedział spektralny interferometru, który w dziedzinie częstotliwości przyjmuje postać:

$$\Delta\nu_{\mathrm{FSR},k} = \frac{c}{4nd}. \tag{6.36}$$

Dodatkowo sferyczna geometria luster sprawia, że problemem przestają być skończone rozmiary luster, gdyż światło jest uwięzione („spułapkowane") w niewielkim, centralnym obszarze wnęki. Łatwiej jest też odpowiednio ustawić taki interferometr, gdyż nie trzeba zapewniać idealnej równoległości luster, a samego sprzęgnięcia wiązki z interferometrem dokonuje się, umieszczając przed nim soczewkę o odpowiednio dobranej ogniskowej.

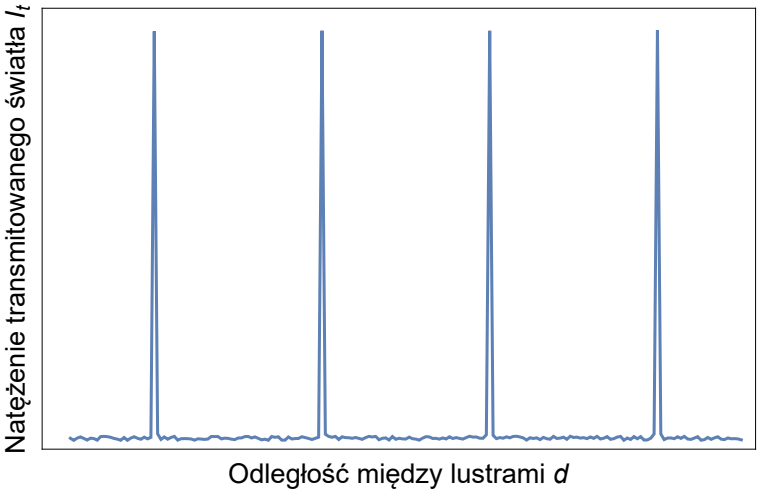
To właśnie te zalety sprawiają, że interferometry konfokalne, a nie te wykorzystujące płaskie lustra, są powszechnie stosowane w wielu zastosowaniach, w tym przede wszystkim w analizie widmowej światła.

6.7 Zastosowania interferencji wielowiązkowej

6.7.1 Analiza widmowa

W dotychczas rozważanych przypadkach interferometr oświetlano falą płaską lub wiązką rozbieżną, przy założeniu stałości parametrów samego układu. Jednak w wielu praktycznych zastosowaniach co najmniej jedno ze zwierciadeł interferometru jest zamontowane na elemencie ruchomym, co pozwala na precyzyjną regulację odległości między lustrami. Najczęściej wykorzystuje się do tego celu przetworniki piezoceramiczne (PZT), które zmieniają swoje wymiary liniowo pod wpływem przyłożonego napięcia. Możliwość „skanowania" transmisji interferometru pozwala na przeprowadzenie dokładnej analizy widmowej światła wprowadzanego do wnęki.

Rysunek 6.7 przedstawia przykładowe widmo promieniowania emitowanego przez laser półprzewodnikowy. Wraz ze zmianą długości wnęki rezonansowej



Rysunek 6.7: Zasada działania skanowanego interferometru Fabry'ego-Perota.

modyfikowane jest natężenie światła przenikającego przez układ. Każdorazowo, gdy dla danej długości fali i odległości między lustrami spełniony zostanie warunek rezonansu, tj.:

$$\frac{4\pi nd\cos\theta}{\lambda} = 2\pi m$$

przez interferometr transmitowana jest (przy zaniedbaniu strat) całkowita energia padającej fali. Dzięki analizie natężenia światła transmitowanego w funkcji długości wnęki możliwe jest precyzyjne odtworzenie widma promieniowania źródła.

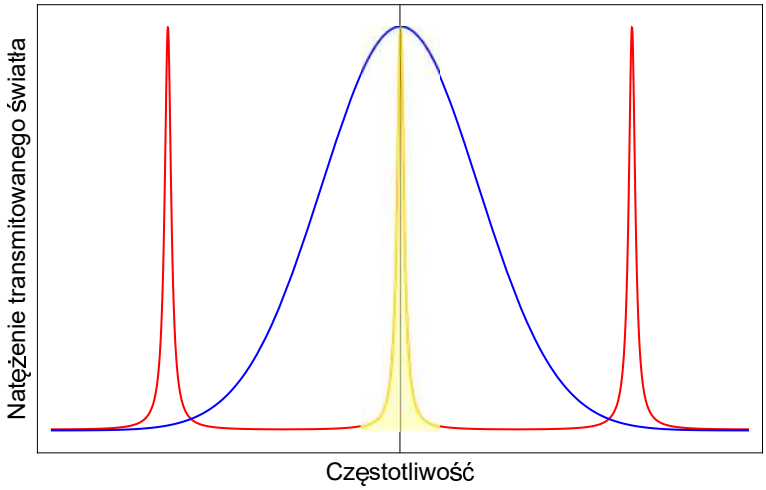
Uwaga!

Należy pamiętać, że profil natężenia światła transmitowanego przez interferometr w funkcji odległości d nie zależy wyłącznie od widma badanego światła. W rzeczywistości obserwowany kształt prążków jest splotem (konwolucją) rzeczywistego widma sygnału z funkcją aparaturową samego interferometru. Ta ostatnia jest determinowana przez współczynnik odbicia luster R. Dlatego kluczowe jest, aby lustra charakteryzowały się możliwie wysoką refleksyjnością, co pozwala na uzyskanie wąskich i ostrych prążków spektralnych, a tym samym wyższej zdolności rozdzielczej.

6.7.2 Filtry spektralne

Wąski zakres transmisyjny interferometru Fabry'ego-Perota pozwala na jego wykorzystanie w roli filtra selekcjonującego wąskie pasmo z szerokiego widma

źródła (rys. 6.8). Proces ten wiąże się z odbiciem części mocy padającej, jednak w zamian uzyskuje się promieniowanie o drastycznie zawężonym profilu spektralnym.



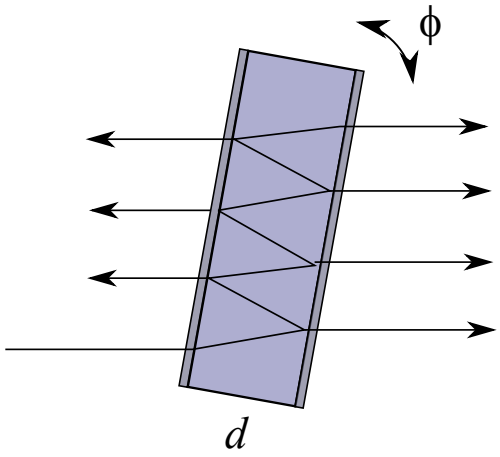
Rysunek 6.8: Mechanizm zawężenia widma promieniowania. Niebieska krzywa reprezentuje widmo źródła, czerwona — charakterystykę interferometru, a kolor żółty — widmo po filtracji.

Najprostszym przykładem takiego układu jest **etalon** (rys. 6.9), czyli przezroczysta płytka płaskorównoległa z napylonymi warstwami wysokorefleksyjnymi. W tym przypadku przestrajanie długości fali rezonansowej odbywa się poprzez zmianę kąta padania światła, co modyfikuje drogę optyczną wewnątrz materiału. Etalony są powszechnie stosowane wewnątrz rezonatorów laserowych do selekcji pojedynczego modu wzdłużnego, co wymusza stabilną pracę lasera na jednej, ściśle określonej częstotliwości.

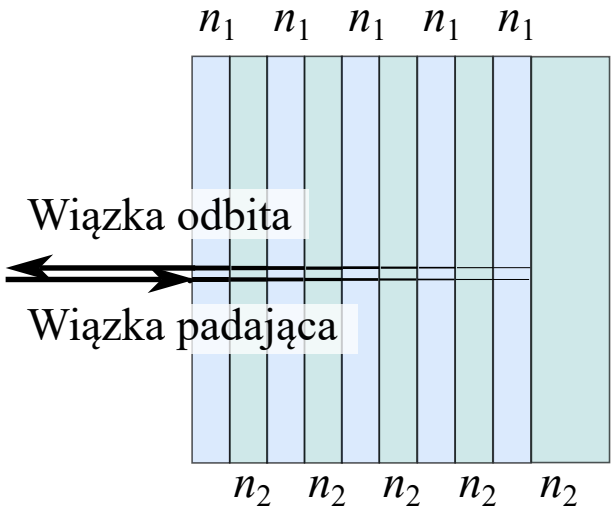
6.7.3 Lustra dielektryczne

W zaawansowanych układach optycznych niezbędne jest stosowanie luster o współczynniku odbicia R bliskim jedności. Tradycyjne lustra metaliczne, ze względu na straty omowe, osiągają refleksyjność rzędu $90 - 98\%$, co jest wartością niewystarczającą dla interferometrów o wysokiej finezji. Rozwiązaniem są **lustra dielektryczne** (lustra Bragga), zbudowane z wielu naprzemiennych warstw dielektryków o precyzyjnie dobranych współczynnikach załamania i grubościach (rys. 6.10).

Dzięki zjawisku konstruktywnej interferencji fal odbitych na wielu granicach ośrodków, możliwe jest uzyskanie współczynnika odbicia przekraczającego $99,99\%$. Przykładem ekstremalnego zastosowania tej technologii jest detektor fal grawitacyjnych LIGO, gdzie lustra o współczynniku $R = 0,999965$ umożliwiły pomiar zmian odległości rzędu 10^{-18} m.



Rysunek 6.9: Etalon jest przezroczystą płytką płaskorównoległą, na którą napylono powłoki o wysokim współczynniku odbicia R. Kąt ϕ jest kątem obrotu etalonu determinującym kąt padania θ wiązki na przednią powierzchnię etalonu.



Rysunek 6.10: Lustro Bragga (lustro dielektryczne).

Dygresja: Barwy strukturalne w przyrodzie
Wielowarstwowe struktury interferencyjne występują powszechnie w świecie ożywionym pod postacią **barw strukturalnych**. W przeciwieństwie do pigmentów, barwy te powstają w wyniku spójnego rozpraszania światła na periodycznych nanostrukturach (np. na skrzydłach motyli *Morpho* czy piórach pawi). Struktury te działają jak naturalne lustra Bragga, wzmacniając wybrane długości fal na drodze interferencji konstruktywnej. Charakterystyczną cechą barw strukturalnych jest **iryzacja**, czyli płynna zmiana koloru w zależności od kąta obserwacji.

6.7.4 Pokrycia antyrefleksyjne

W **pokryciach antyrefleksyjnych** wykorzystuje się zasadę interferencji destruktywnej. Grubości warstw są dobierane tak, aby fale odbite od poszczególnych granic wygaszały się wzajemnie. Pozwala to na niemal całkowitą redukcję niepożądanych odblasków, co jest kluczowe w obiektywach fotograficznych, okularach korekcyjnych oraz systemach laserowych o dużej mocy.

6.7.5 Lasery i stabilizacja częstotliwości

Szczególnym przypadkiem zastosowania wnęk rezonansowych są lasery. Wprowadzenie ośrodka aktywnego między dwa lustra zapewnia dodatnie sprzężenie zwrotne, niezbędne do podtrzymania **emisji wymuszonej**. Wnęka optyczna filtruje dostępne mody, co prowadzi do generacji wysoce spójnego i monochromatycznego światła.

W metrologii laserowej zewnętrzne interferometry są wykorzystywane do stabilizacji częstotliwości (tzw. „lockowania"). Standardem jest tu metoda Pounda-Drevera-Halla (PDH), polegająca na analizie fazy światła odbitego od ultrastabilnej wnęki. Pozwala to na budowę laserów o szerokości linii poniżej 1 Hz, co jest fundamentem działania współczesnych **zegarów optycznych**.

Dygresja
Zegary optyczne to najdokładniejsze urządzenia pomiarowe, jakie kiedykolwiek skonstruowano. Wykorzystują one przejścia elektronowe w atomach (np. strontu), które absorbują światło o częstotliwościach rzędu setek teraherców. Dzięki temu, że „tyknięcia" zegara są znacznie gęstsze niż w zegarach mikrofalowych, oferują one bezprecedensową precyzję. Stabilizacja lasera próbkującego do wnęki Fabry'ego-Perota o wysokiej dobroci przy użyciu techniki PDH jest niezbędna do odczytu tych częstotliwości. Osiągnięcia te torują drogę do redefinicji sekundy w układzie SI.

6.7.6 Macierze ABCD i stabilność wnęki optycznej

W celu pełnego opisu propagacji światła wewnątrz wnęk rezonansowych, szczególnie w kontekście wiązek gaussowskich, stosuje się formalizm macierzowy (tzw.

macierze ABCD). Pozwala on na śledzenie ewolucji parametrów wiązki po pełnym obiegu wnęki. Każdy element optyczny (lustro, soczewka, przestrzeń swobodna) reprezentowany jest przez macierz przejścia:

$$\mathbf{M} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{6.37}$$

Dla wnęki o długości L, ograniczonej dwoma lustrami o promieniach krzywizny R_1 i R_2, stabilność określa, czy wiązka po wielu odbiciach pozostanie wewnątrz układu. Wprowadza się parametry stabilności g:

$$g_1 = 1 - \frac{L}{R_1}, \quad g_2 = 1 - \frac{L}{R_2} \tag{6.38}$$

Warunek stabilności rezonatora przyjmuje wtedy postać:

$$0 \leq g_1 g_2 \leq 1 \tag{6.39}$$

Rezonatory spełniające ten warunek pozwalają na stabilną pracę lasera i efektywne magazynowanie energii wewnątrz wybranego modu.

6.8 Podsumowanie rozdziału

W niniejszym rozdziale rozszerzono opis zjawisk falowych na przypadek interferencji wielowiązkowej, zachodzącej w układach o wysokim współczynniku odbicia. Głównym obiektem analizy stał się interferometr Fabry'ego-Perota, w którym nieskończony ciąg uwięzionych we wnęce, wielokrotnych odbić prowadzi do powstania niezwykle ostrych rezonansów transmisyjnych. Rozkład tych rezonansów opisano analitycznie za pomocą funkcji Airy'ego. Wykazano, że kluczem do głębokiego zrozumienia tego zjawiska jest koherentna superpozycja fal – doskonale ilustrowana metodą wskazów na płaszczyźnie zespolonej – która w warunkach pełnego rezonansu zapewnia 100% transmisję energii, omijając pozorną barierę silnie odbijających luster.

W celu parametryzacji i oceny instrumentalnej wnęk rezonansowych wprowadzono rygorystyczny aparat pojęciowy, definiując finezję układu, przedział spektralny oraz zdolność rozdzielczą opartą na kryterium szerokości połówkowej. Parametry te, ściśle powiązane z makroskopowym pojęciem czasu życia fotonu i dobroci wnęki (Q), pozwoliły na fizyczne wyjaśnienie zasady działania wysoce precyzyjnych technik badawczych, takich jak spektroskopia strat we wnęce czy zaawansowana stabilizacja częstotliwości laserów. Przedstawiono również powszechne w inżynierii fotonicznej realizacje wielowarstwowych układów interferencyjnych: etalony, powłoki antyrefleksyjne oraz lustra dielektryczne (lustro Bragga). Rozważania nad wnękami zwieńczono analizą warunków stabilności rezonatorów sferycznych, wykazując po raz kolejny uniwersalność formalizmu macierzy ABCD.

Dotychczasowe analizy interferencji – zarówno dwu-, jak i wielowiązkowej – opierały się jednak w dużej mierze na założeniu propagacji nieograniczonych

przestrzennie fal płaskich i sferycznych. W rzeczywistości każda wiązka świetlna oraz każdy przyrząd optyczny posiadają skończone rozmiary poprzeczne (apertury). Narzucenie fizycznego ograniczenia na propagujący się front falowy prowadzi do załamania przybliżenia prostoliniowego rozchodzenia się promieni i skutkuje niemożliwym do pominięcia ugięciem fali na krawędziach. Obserwacja ta stanowi konieczny i naturalny punkt wyjścia do kolejnego rozdziału, w którym sformułowana zostanie pełna, rygorystyczna teoria dyfrakcji.

Zadania

6.1. Współczynnik odbicia luster próżniowego interferometru Fabry'ego-Perota wynosi $R = 0,99$. Oblicz:

(a) współczynnik finezji F występujący we wzorze Airy'ego,

(b) szerokość połówkową prążków interferencyjnych wyrażoną w dziedzinie częstotliwości ($\Delta\nu_{1/2}$), jeśli odległość między lustrami wynosi $d = 1,0$ cm,

(c) finezję $\mathcal{F}$ tego układu interferencyjnego.

6.2. W spektrometrze Fabry'ego-Perota odległość między płaskorównoległymi lustrami wynosi $d = 2,0$ cm, a ośrodek między nimi stanowi próżnia. Przyrząd ten oświetlono światłem o centralnej długości fali $\lambda = 500$ nm.

(a) Oblicz wolny zakres spektralny (FSR) tego interferometru w dziedzinie częstotliwości ($\Delta\nu_{\mathrm{FSR}}$) oraz długości fali ($\Delta\lambda_{\mathrm{FSR}}$).

(b) Wykaż na podstawie obliczonych wartości, czy przy tej konfiguracji przyrządu możliwe jest jednoznaczne rozdzielenie (bez zjawiska nakrywania się sąsiednich rzędów interferencji) dwóch bliskich linii widmowych oddalonych od siebie o $\Delta\lambda = 0,05$ nm.

6.3. Oblicz, jaka musi być minimalna odległość między lustrami d próżniowego interferometru Fabry'ego-Perota, aby za jego pomocą można było rozdzielić linie dubletu sodowego D_1 ($588,995$ nm) oraz D_2 ($589,592$ nm). Zdolność rozdzielczą przyrządu przyjmij zgodnie z kryterium Rayleigha, a finezję układu jako $\mathcal{F} = 30$.

6.4. Etalon zbudowano z płytki szklanej o grubości $d = 10$ μm i współczynniku załamania $n = 1,5$. Na jej równoległe powierzchnie naniesiono warstwy odbijające. Oblicz minimalną wymaganą wartość współczynnika odbicia R tych powłok, aby przy pomocy tego etalonu możliwe było rozdzielenie wspomnianych wyżej linii dubletu sodowego (D_1 i D_2).

6.5. Rezonator laserowy, działający na zasadzie interferometru wielowiązkowego, ma długość $L = 1,0$ m. Oblicz liczbę modów podłużnych mieszczących się w szerokości linii lasera, jeżeli szerokość ta wynosi $\Delta\lambda = 1,8$ pm przy centralnej długości fali $\lambda = 600$ nm.

6.6. Próżniowy interferometr Fabry'ego-Perota o odległości między lustrami $d = 5,0$ mm jest oświetlany rozbieżną wiązką światła z lasera He-Ne ($\lambda = 632,8$ nm). Przechodzące światło jest ogniskowane za pomocą soczewki o ogniskowej $f = 20,0$ cm na płaskim ekranie. Oblicz promienie trzech pierwszych jasnych pierścieni interferencyjnych obserwowanych na ekranie, zakładając, że w centrum obrazu znajduje się maksimum rzędu zerowego.

6.7. Na szklaną soczewkę o współczynniku załamania $n_s = 1,52$ nałożono pojedynczą powłokę przeciwodblaskową (AR) z fluorku magnezu (MgF_2) o współczynniku załamania $n_w = 1,38$. Powłoka sąsiaduje z powietrzem ($n_0 \approx 1,0$).

(a) Oblicz minimalną grubość powłoki d, dla której współczynnik odbicia dla światła o długości fali $\lambda = 550$ nm osiąga minimum przy padaniu normalnym.

(b) Oblicz wartość resztkowego współczynnika odbicia R dla tej długości fali i oceń, ile razy zmniejszyło się odbicie w stosunku do niepokrytej powierzchni szkła.

6.8. Na płaską powierzchnię szkła ($n_s = 1,50$) naniesiono cienką warstwę dielektryka o wysokim współczynniku załamania $n_w = 2,00$.

(a) Oblicz minimalną, niezerową grubość tej warstwy d, dla której natężenie światła odbitego o długości fali $\lambda = 700$ nm (padającego prostopadle z powietrza) będzie maksymalne.

(b) Oblicz efektywny współczynnik odbicia R dla tak zaprojektowanego układu.

6.9. Zaprojektuj klasyczne lustro Bragga pracujące przy długości fali $\lambda = 795$ nm, naniesione na podłoże szklane ($n_s = 1{,}50$). Lustro składa się z naprzemiennych warstw materiałów o współczynnikach załamania $n_H = 2,30$ (warstwa o wysokim współczynniku) oraz $n_L = 1,38$ (warstwa o niskim współczynniku). Warstwą zewnętrzną kontaktującą się z powietrzem jest n_H.

(a) Oblicz fizyczne grubości obu typów warstw (d_H oraz d_L).

(b) Oblicz minimalną liczbę par warstw N, która jest wymagana, aby całkowita reflektancja układu przekroczyła barierę $R = 0,99$.

Rozdział 7

Dyfrakcja

7.1 Strefy Fresnela

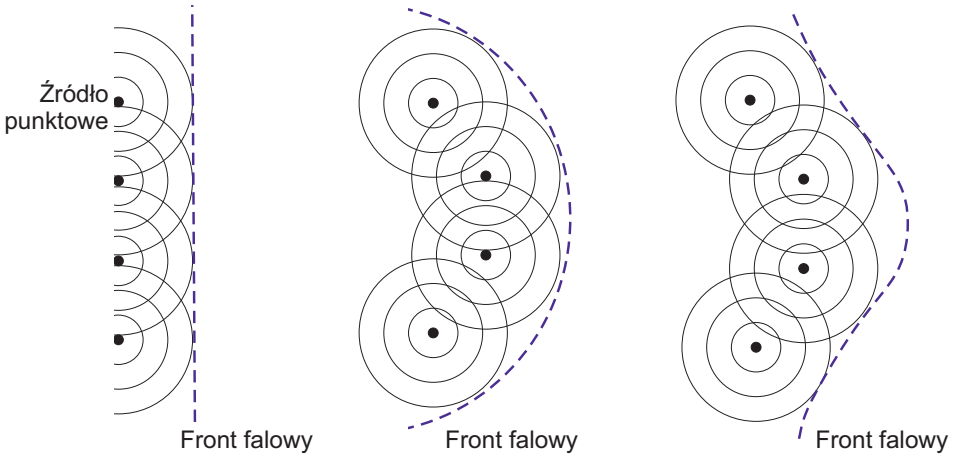
7.1.1 Zasada Huygensa-Fresnela

Jednym z argumentów w toczącej się w XVII w. debacie dotyczącej natury światła, mającym przemawiać za poglądem falowym, była możliwość wytłumaczenia praw rządzących światłem (np. prawa odbicia i załamania) **zasadą Huygensa**. Zgodnie z nią każdy punkt, do którego dociera fala, staje się źródłem **wtórnej fali sferycznej**. Powierzchnia styczna do tych fal wtórnych (ich obwiednia) wyznacza nowe położenie powierzchni o stałej fazie, zwanej **frontem falowym** (rys. 7.1).

Na początku XIX w. Augustin Fresnel zauważył, że emitowane ze źródeł wtórnych fale są wzajemnie spójne (**zasada Huygensa-Fresnela**), przez co mogą ulegać interferencji, co w rozważaniach wymaga uwzględnienia nie tylko ich amplitud, ale również faz. Fresnel był również przekonany, że amplituda fal wtórnych zależy od kierunku ich rozchodzenia się (cechuje się zależnością kątową), osiągając maksimum w kierunku propagacji fali pierwotnej i malejąc wraz ze wzrostem kąta. Dokładna zależność amplitudy od kąta (tzw. czynnik nachylenia) nie była mu jednak znana i dopiero Gustav Kirchhoff (1824–1887) wyprowadził analityczną postać tej relacji (zob. podrozdział 7.2).

7.1.2 Strefy Fresnela

Rozważmy falę sferyczną emitowaną z punktowego źródła światła Z. Wybierzmy punkt E, w którym chcemy obliczyć natężenie światła (rys. 7.2). W tym celu podzielmy front falowy Σ, znajdujący się w odległości R od źródła Z, na współosiowe pierścienie, których osią symetrii jest odcinek $\overline{\mathrm{ZE}}$. Promienie tych pierścieni są tak dobrane, że różnica dróg optycznych od krawędzi dwóch kolejnych pierścieni (tj. punktów P_m i P_{m-1}) do punktu obserwacji E jest równa $\lambda/2$. W



Rysunek 7.1: Różne kształty frontów falowych powstały w oparciu o zasadę Huygensa.

takim przypadku:

$$\begin{aligned} \overline{\mathrm{P_1E}} - \overline{\mathrm{P_0E}} &= \lambda/2, \\ \overline{\mathrm{P_2E}} - \overline{\mathrm{P_1E}} &= \lambda/2, \\ \overline{\mathrm{P_3E}} - \overline{\mathrm{P_2E}} &= \lambda/2, \\ &\dots \end{aligned} \tag{7.1}$$

Promienie r_m poszczególnych pierścieni mogą być obliczone w oparciu o geometryczną analizę problemu, odpowiednio rozważając trójkąty o wierzchołku w punkcie Z albo o wierzchołku w punkcie E. Zgodnie z tą analizą:

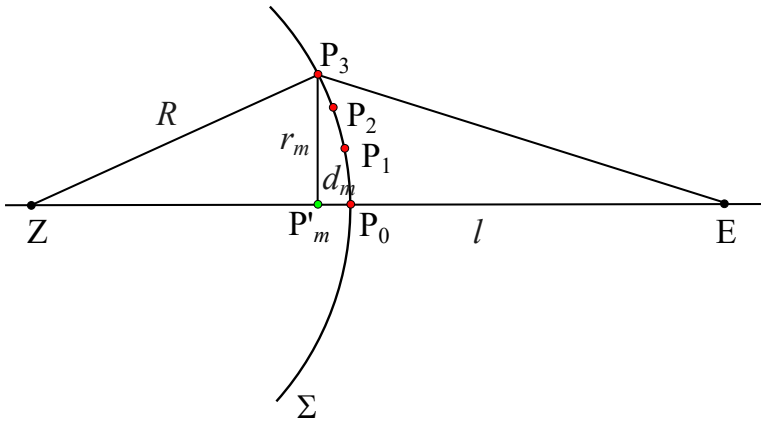
$$r_m^2 = R^2 - (R - d_m)^2 = \left(l + m\frac{\lambda}{2}\right)^2 - (l + d_m)^2\,, \tag{7.2}$$

gdzie l jest odległością $\overline{\mathrm{P_0E}}$, zaś d_m jest odległością pomiędzy punktem P_0 oraz punktem P'_m (będącym rzutem punktu P_m, znajdującego się na m-tym pierścieniu, na odcinek $\overline{\mathrm{ZE}}$). Z dokładnością do członu zależnego od λ^2 można pokazać, że:

$$\begin{aligned} d_m &\approx \frac{lm}{R+l}\frac{\lambda}{2}, \\ r_m^2 &\approx \frac{Rlm\lambda}{R+l}. \end{aligned} \tag{7.3}$$

Pozwala to obliczyć pole powierzchni wewnątrz poszczególnych okręgów. I tak pole wewnątrz pierwszego okręgu ($m = 1$), stanowiące **pierwszą strefę Fresnela**, jest równe:

$$S_1 = \pi r_1^2 = \frac{\pi Rl\lambda}{R+l}, \tag{7.4}$$



Rysunek 7.2: Schemat analizy dyfrakcji przy pomocy stref Fresnela dla punktowego źródła światła i fali sferycznej.

podczas gdy pole wewnątrz drugiego okręgu ($m = 2$) dane jest przez:

$$S_2 = \pi r_2^2 = \frac{2\pi Rl\lambda}{R+l}. \tag{7.5}$$

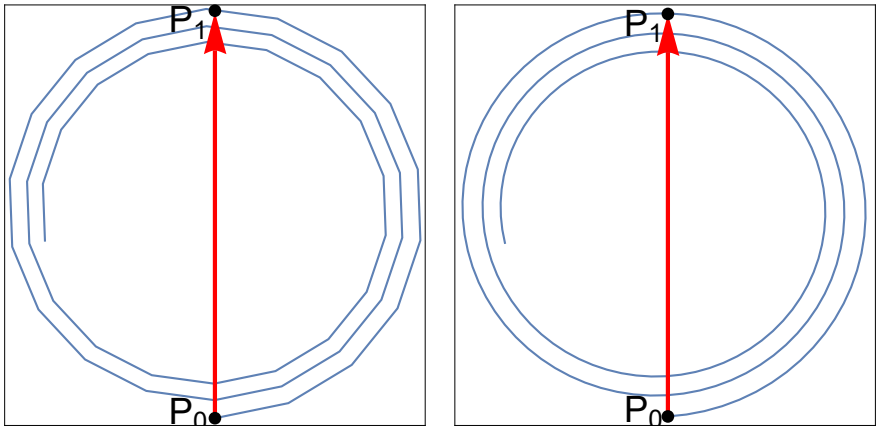
Odejmując od siebie pola kół zajętych oboma obszarami, można obliczyć pole pierwszego pierścienia, tj. drugiej strefy Fresnela:

$$S_{pierścień} = S_2 - S_1 = \frac{\pi Rl\lambda}{R+l}. \tag{7.6}$$

W oparciu o analogiczną analizę można pokazać, że pola powierzchni kolejnych stref Fresnela (czyli pierścieni pomiędzy okręgami r_m i r_{m-1}) są w przybliżeniu identyczne i równe S_1. Wynika stąd, że różnice długości kolejnych promieni r_m stają się coraz mniejsze ze wzrostem m.

W powyższych rozważaniach założono, że powierzchnie pierścieni zrzutowanych na płaszczyznę prostopadłą do odcinka $\overline{\mathrm{ZE}}$ są w przybliżeniu równe ich rzeczywistym powierzchniom na czaszy sferycznej. Przybliżenie to nie wprowadza poważnych błędów w sytuacji, w której promienie poszczególnych pierścieni są znacząco mniejsze od promienia krzywizny frontu falowego ($r_m \ll R$).

W celu uzyskania dokładniejszych wyników należy przeprowadzić bardziej skrupulatną analizę rozważanego problemu. W szczególności możemy podzielić pierścienie stref Fresnela na kolejne podstrefy tak, że faza pomiędzy kolejnymi podstrefami zmienia się o $\delta = \pi/m_p$, gdzie m_p jest liczbą całkowitą. Stosując tzw. **metodę wskazów**, możemy dodawać do siebie kolejne przyczynki pochodzące od fal wtórnych. Należy jednak pamiętać, że wraz z oddalaniem się od osi $\overline{\mathrm{ZE}}$ rośnie droga pomiędzy punktem leżącym na froncie falowym a punktem E. Spadek amplitudy kolejnych fal wtórnych docierających do punktu E wynika z dwóch faktów: po pierwsze, fale wtórne są sferyczne, więc ich amplituda jest odwrotnie proporcjonalna do pokonywanej odległości; po drugie, jak omówiono



Rysunek 7.3: Metoda wskazów ukazująca składanie amplitud poszczególnych fal wtórnych dla dyskretnej, tj. dla liczby całkowitej m_p przyjmującej stosunkowo niewielkie wartości (lewy wykres), oraz ciągłej, czyli dla $m_p \to \infty$ (prawy wykres). Zmiany fazy i amplitudy poszczególnych fal składowych składają się na falę wypadkową. Pierwszy wektor wskazuje amplitudę fali wypadkowej powstałej przez nałożenie na siebie kolejnych przyczynków w ramach pierwszej strefy Fresnela.

wcześniej, emisja maleje wraz ze wzrostem kąta z powodu czynnika nachylenia. Efekt ten został przedstawiony na rys. 7.3 dla przypadku dyskretnego (gdy wartość m_p nie jest zbyt duża) oraz dla przypadku ciągłego ($m_p \to \infty$).

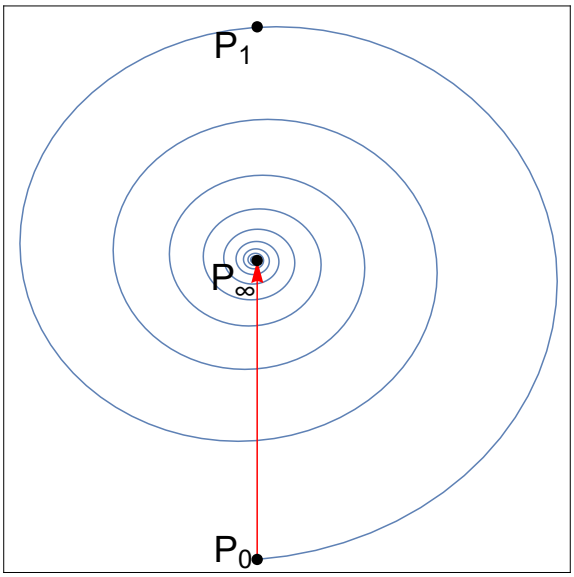
Doświadczenie Wykorzystanie źródła mikrofalowego do demonstracji działania soczewek Fresnela.

Posługując się metodą wskazów, można obliczyć wypadkową amplitudę światła powstałego przez nałożenie na siebie przyczynków pochodzących od kolejnych (pod)stref Fresnela. W szczególności pokazany na wykresie czerwony wektor odpowiada amplitudzie fali powstałej przez ograniczenie się wyłącznie do pierwszej strefy Fresnela. Można to np. osiągnąć poprzez użycie przesłony o odpowiednio małej aperturze. Choć na pierwszy rzut oka może wydawać się to zaskakujące, dalsze otwarcie takiej przesłony i uwzględnienie kolejnych stref Fresnela nie tylko nie powoduje wzrostu amplitudy światła, ale przeciwnie – powoduje jej spadek. W szczególności dla całkowicie otwartej przesłony amplituda wypadkowego światła jest mniej więcej o połowę mniejsza niż w przypadku ograniczenia się jedynie do pierwszej strefy (rys. 7.4). Fakt, że amplituda, a co za tym idzie również natężenie światła, maleje wraz z otwieraniem przesłony, jest dziś ważnym dowodem potwierdzającym falową naturę światła.

7.1.3 Plamka Poissona

Doświadczenie Obserwacja plamki Poissona na cieniu główki od szpilki.

Jeżeli na drodze wiązki światła umieści się niewielki, nieprzezroczysty przedmiot, na ekranie powstanie cień. Jednakże dokładna analiza pokazuje, że w jego samym środku występuje jasny punkt, który nazywany jest **plamką Poissona**



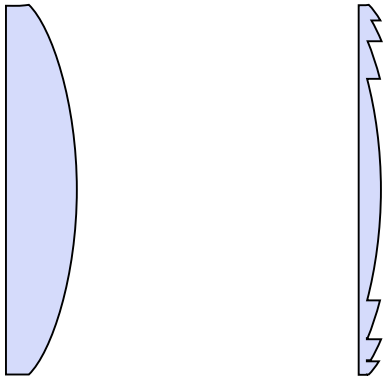
Rysunek 7.4: Spirala powstała przy pomocy metody wskazów, ukazująca wypadkową amplitudę światła dla całkowicie otwartej przestrzeni (punkt P_∞). W oparciu o ten wykres wyraźnie widać, że amplituda pola elektrycznego dla przesłony ograniczającej obszar do pierwszej strefy Fresnela, determinowana przez odległość pomiędzy punktami P_0 oraz P_1, jest zdecydowanie większa niż w sytuacji, w której odsłonięta jest cała przestrzeń (odległość pomiędzy punktami P_0 i P_∞).

(bądź **plamką Arago**). Istnienie tej plamki można dość prosto wytłumaczyć, posługując się teorią stref Fresnela, bowiem umieszczenie przedmiotu na drodze wiązki powoduje zasłonięcie szeregu stref Fresnela. Jednakże zewnętrzne strefy pozostają odsłonięte i mogą wnosić swój wkład do powstającego obrazu. Wielkość tego wkładu zależy od tego, jak dużo stref Fresnela zostało zasłoniętych. Im wyższy jest bowiem rząd strefy Fresnela, tym mniejsza jest różnica amplitud fal pochodzących z dwóch kolejnych stref, a co za tym idzie, tym lepsze jest ich wzajemne kompensowanie się. Widać to wyraźnie na spirali pokazanej na rys. 7.4. Dlatego też dla obiektów, które mają większe rozmiary, plamka Poissona jest prawie niewidoczna, za to dla małych obiektów jest ona wyraźnie widoczna na obszarze cienia.

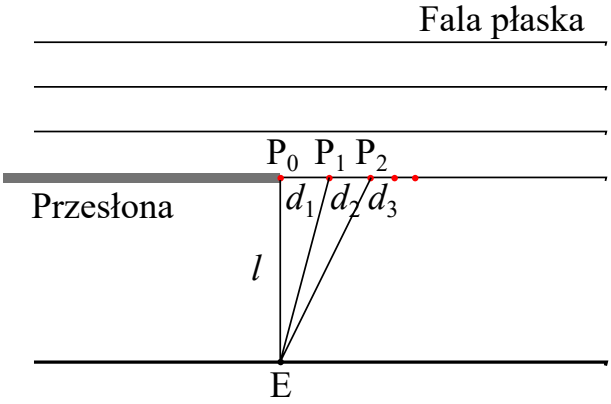
Dygresja

Opisana powyżej teoria stref Fresnela została zaproponowana w odpowiedzi na konkurs ogłoszony przez Francuską Akademię Nauk, którego celem było wyjaśnienie dyfrakcji światła. Wychodząc od założenia o falowej naturze światła, Augustin Fresnel (pracujący jako inżynier przy budowie latarni morskich) zaproponował wyjaśnienie tego procesu jako skutku nakładania się przyczynków od poszczególnych fal wtórnych. Propozycja ta wywołała jednak szereg kontrowersji i napotkała szczególnie silny opór ze strony Siméona Poissona (1781–1840), który uważał, że prowadzi ona do „absurdalnych" wniosków. Szczególnym ich przykładem miałoby być istnienie jasnego punktu w środku cienia rzucanego przez nieprzezroczysty obiekt. Poisson nie przeprowadził jednak eksperymentu, który mógłby zdyskredytować teorię Fresnela. Eksperyment taki zrealizował natomiast François Arago (1786–1853), dowodząc, że rzeczywiście w środku cienia niewielkich obiektów widoczny jest jasny punkt. Istnienie tego punktu jest dziś koronnym argumentem potwierdzającym poprawność teorii stref Fresnela, a sam jasny punkt widoczny w cieniu nazywany jest dość przewrotnie **plamką Poissona**.

Warto również wspomnieć o wybitnym zmyśle inżynieryjnym Fresnela, który swój geniusz naukowy przełożył na zastosowania praktyczne. Choć zasłynął on z matematycznego opisu falowej natury światła i zjawisk dyfrakcyjnych, jego najsłynniejszy wynalazek – zastosowana po raz pierwszy w 1822 roku soczewka do latarni morskich – opiera się na znacznie prostszej optyce geometrycznej. Fresnel zauważył, że za załamanie światła odpowiadają wyłącznie krzywizny powierzchni soczewki, a nie jej wewnętrzna masa czy nawet całkowita grubość (rys. 7.5). Pozwoliło to na stworzenie soczewki o budowie schodkowej, podzielonej na koncentryczne pierścienie, z których usunięto zbędną grubość materiału. Dzięki temu wyeliminowano konieczność budowania potężnych, niepraktycznych soczewek, otwierając możliwość wykonania cienkiej i lekkiej optyki wielkoformatowej. Taka konstrukcja sprawia, że znacznie łatwiej osiągnąć dobrą jednorodność szkła, a sama soczewka jest zdecydowanie lżejsza, co ułatwia jej montaż w latarni.



Rysunek 7.5: Porównanie klasycznej soczewki z soczewką Fresnela, która działa w identyczny sposób, ale wolna jest od szeregu problemów soczewki klasycznej.



Rysunek 7.6: Schemat analizy dyfrakcji przy pomocy stref Fresnela dla fali płaskiej padającej na półpłaszczyznę.

7.1.4 Ugięcie na krawędzi nieprzepuszczalnej półpłaszczyzny

Doświadczenie Dyfrakcja na żyletce.

Podobne rozważania do tych przeprowadzonych powyżej dla przesłony o aperturze kołowej można przeprowadzić dla nieprzezroczystej półpłaszczyzny o ostrej krawędzi, oświetlanej falą płaską (rys. 7.6). Rozważmy punkt obserwacji E znajdujący się dokładnie na granicy geometrycznego cienia. W takim przypadku wewnętrzną krawędź pierwszej strefy Fresnela wyznacza punkt P_0, który znajduje się na krawędzi półpłaszczyzny (w odległości l od ekranu), oraz punkt P_1, względem którego odległość do punktu E jest dłuższa o $\lambda/2$ (precyzyjniej nie są to punkty, a linie, gdyż układ posiada symetrię translacyjną w kierunku krawędzi półpłaszczyzny). Odległość zewnętrznej granicy każdej kolejnej strefy może być wyznaczona geometrycznie, przyjmując, że różnica dróg optycznych między

kolejnymi krawędziami a punktem E jest równa $\lambda/2$. Zakładając, że płaszczyzna jest prostopadła do wektora falowego fali padającej, można obliczyć położenie zewnętrznej granicy d_m kolejnej strefy Fresnela:

$$d_m^2 = \left(l + m\frac{\lambda}{2}\right)^2 - l^2, \tag{7.7}$$

co z dokładnością do wyrazu proporcjonalnego do λ^2 daje:

$$d_m \approx \sqrt{lm\lambda}. \tag{7.8}$$

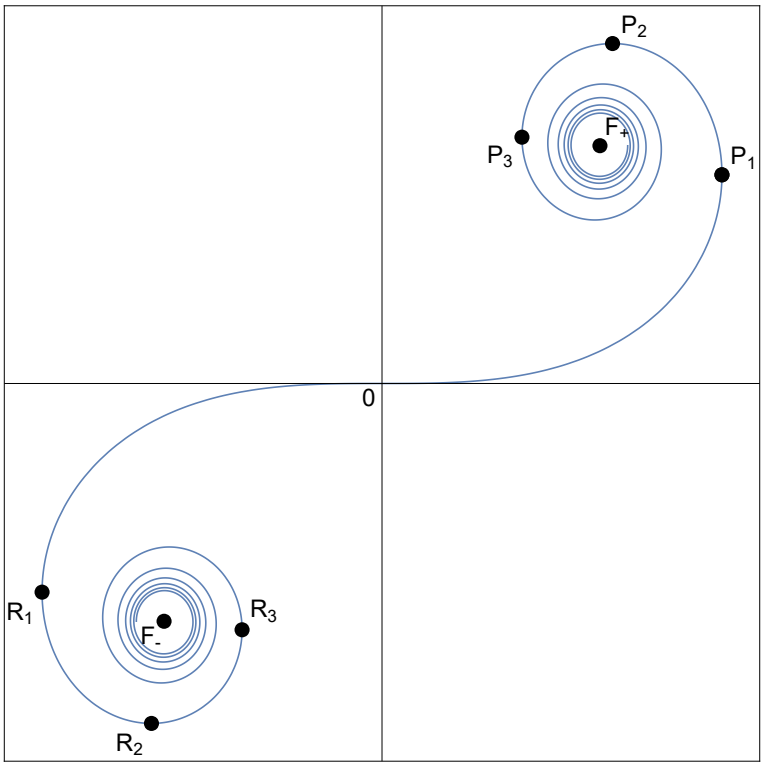
Ponieważ krawędzie stref Fresnela zorientowane są równolegle do krawędzi półpłaszczyzny, powierzchnie kolejnych stref determinowane są przez ich szerokości ($\Delta d_m = d_m - d_{m-1}$).

$$d_1 : (d_2 - d_1) : (d_3 - d_2) : \cdots = 1 : (\sqrt{2} - 1) : (\sqrt{3} - \sqrt{2}) : \ldots, \tag{7.9}$$

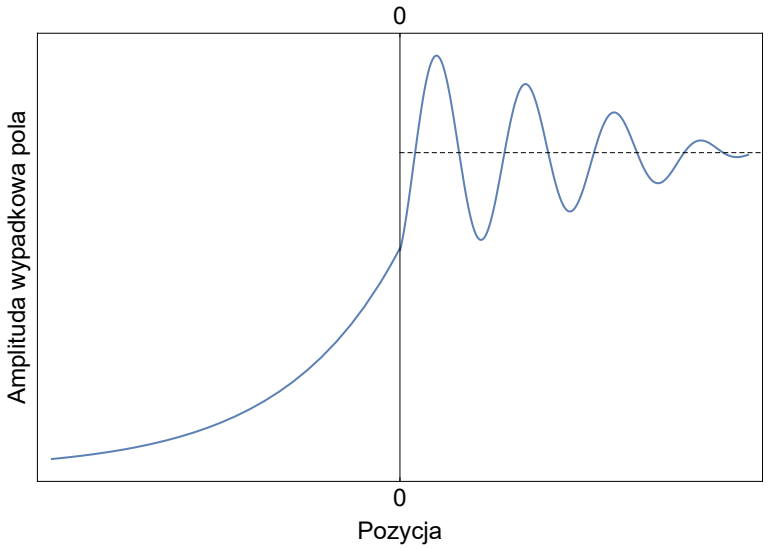
gdzie znak : oznacza stosunki odpowiednich długości. W oparciu o tę relację można zauważyć, że pola kolejnych stref Fresnela (pasów) szybko maleją wraz z oddalaniem się od krawędzi płaszczyzny. Oznacza to, że długość wskazów od kolejnych stref maleje dużo gwałtowniej niż dla apertury kołowej (gdzie pola stref były w przybliżeniu stałe). W przypadku podziału stref Fresnela na kolejne podstrefy różniące się drogą optyczną o $\lambda/(2m_p)$, gdy $m_p \to \infty$, otrzymujemy charakterystyczną strukturę wskazową zwaną **spiralą Cornu** (Marie Alfred Cornu 1841–1902) (rys. 7.7).

Analizując rys. 7.7, można zauważyć, że w przypadku gdy półpłaszczyzna „zamyka" lewą stronę przestrzeni, wektor poprowadzony od punktu 0 do punktów P_1, P_2, P_3 itd. odpowiada amplitudzie wypadkowej fali rejestrowanej w punkcie E, gdy odsłonięta jest odpowiednio pierwsza, druga, trzecia itd. strefa Fresnela. Gdy otwarta jest cała płaszczyzna (z prawej strony), amplitudę fali w punkcie E determinuje odległość $\overline{0F_+}$. Podobnie, jeżeli półpłaszczyzna na lewo od krawędzi jest otwarta, wówczas odległość pomiędzy punktem 0 oraz punktami R_1, R_2, R_3, ... odpowiada amplitudzie wypadkowej fali, gdy otwarte są kolejne strefy Fresnela po lewej stronie, zaś gdy otwarta jest cała lewa półpłaszczyzna, to amplitudę determinuje długość odcinka $\overline{0F_-}$. W końcu, gdy brak jest jakiejkolwiek przeszkody (otwarta jest cała przestrzeń), amplituda fali niezaburzonej determinowana jest przez długość odcinka $\overline{F_-F_+}$. Spirala Cornu pozwala także obliczyć wiele przypadków, w których wiązka światła podzielona jest szeregiem szczelin o określonych szerokościach.

Rysunek 7.8 przedstawia zależność wypadkowej amplitudy światła w funkcji położenia na ekranie przysłoniętym płaszczyzną o krawędzi w punkcie 0. Jak widać, wraz z oddalaniem się w stronę ekranu nieosłoniętego półpłaszczyzną (w obszar oświetlony) obserwowana jest oscylacyjna fluktuacja amplitudy. W przypadku gdy poruszamy się w stronę przeciwną (w obszar geometrycznego cienia), następuje monotoniczny spadek amplitudy do zera. Wynika z tego, że na krawędzi półpłaszczyzny pole różni się znacząco od zera – jego amplituda wynosi połowę amplitudy fali niezaburzonej.



Rysunek 7.7: Spirala Cornu, krzywa na płaszczyźnie zespolonej używana do graficznego wyznaczania wypadkowej amplitudy pola w problemach dyfrakcji na elementach o krawędziach prostoliniowych.



Rysunek 7.8: Wypadkowa amplituda pola elektrycznego światła padającego na ekran osłonięty półpłaszczyzną w funkcji położenia na ekranie. Punkt 0 oznacza miejsce znajdujące się dokładnie pod krawędzią półpłaszczyzny (granicę geometrycznego cienia).

7.1.5 Ograniczenia metody stref Fresnela

Metoda stref Fresnela, o ile daje dobrą zgodność dla wielu przypadków praktycznych, ma jednak szereg poważnych ograniczeń. Jednym z najpoważniejszych jest brak naturalnej możliwości wyznaczenia amplitudy fali w funkcji kierunku propagacji światła. Zależność czynnika nachylenia musiała zostać przez Fresnela zapostulowana całkowicie sztucznie, aby wyeliminować niefizyczną falę propagującą się wstecz, której powstawania pierwotna zasada Huygensa nie wykluczała.

Drugim fundamentalnym problemem jest fakt, że metoda Fresnela nie pozwala na obliczenie poprawnej fazy fali wypadkowej. Zsumowanie przyczynków od poszczególnych stref generuje falę przesuniętą w fazie dokładnie o $\pi/2$ (ćwierć okresu) względem fali pierwotnej. Aby wynik zgadzał się z rzeczywistością, Fresnel musiał poczynić kolejne założenie *ad hoc*: uznał, że wtórne źródła emitują fale z początkowym wyprzedzeniem fazowym $\pi/2$. Choć w wielu doświadczeniach (gdzie detektory rejestrują jedynie uśrednione w czasie natężenie) faza nie odgrywa kluczowej roli, to z punktu widzenia rygoru teoretycznego stanowiło to istotny mankament.

Obydwie te teoretyczne „łaty” znajdują swoje rygorystyczne, matematyczne uzasadnienie dopiero przy pomocy bardziej ścisłej teorii dyfrakcji bazującej na całkach Kirchhoffa.

7.2 Teoria dyfrakcji Kirchhoffa

7.2.1 Twierdzenie całkowe Helmholtza-Kirchhoffa

Teoria dyfrakcji Kirchhoffa pozwala na rygorystyczne sformułowanie zasady Huygensa-Fresnela, opierając się bezpośrednio na równaniu falowym. Zamiast heurystycznego sumowania fal wtórnych, metoda ta wykorzystuje twierdzenie całkowe do powiązania amplitudy pola świetlnego w dowolnym punkcie przestrzeni z wartościami tego pola na otaczającej ten punkt powierzchni zamkniętej.

Aby wyznaczyć zaburzenie optyczne w dowolnym punkcie obserwacji P, wolnym od źródeł światła, skorzystamy z **twierdzenia Greena** (George Green, 1793–1841):

$$\oint_S \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) \mathrm{d}a = \int_V \left(U \nabla^2 G - G \nabla^2 U \right) \mathrm{d}v, \tag{7.10}$$

gdzie różniczkowanie odbywa się wzdłuż kierunku normalnego do powierzchni zamkniętej S, zaś $U = U(\mathbf{r})$ oraz $G = G(\mathbf{r})$. Wybierzmy teraz funkcje U i G w taki sposób, aby:

$$\begin{aligned} G(\mathbf{r}) &= \frac{e^{ikr}}{r}, \\ U(\mathbf{r}) &= E(\mathbf{r}), \end{aligned} \tag{7.11}$$

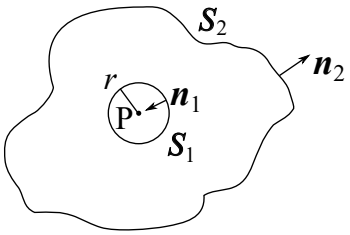
gdzie funkcja $E(\mathbf{r})$ została wybrana tak, aby spełniała równanie Helmholtza w przestrzeni pozbawionej źródeł (dla pewnej ustalonej liczby falowej k). Z kolei jako funkcję $G(\mathbf{r})$ przyjęto falę sferyczną, w której r oznacza odległość od wybranego punktu obserwacji P. Funkcja ta jest matematycznym narzędziem pomocniczym (tzw. funkcją Greena), które celowo konstruujemy tak, aby również spełniało równanie Helmholtza z dokładnie tą samą liczbą falową k wszędzie poza punktem P ($r = 0$), w którym posiada osobliwość. Podstawiając wyrażenia na $U(\mathbf{r})$ i $G(\mathbf{r})$ do równania (7.10), otrzymujemy:

$$\oint_S \left[E \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial n} \right] \mathrm{d}a = \int_V \left[E \nabla^2 \left(\frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \nabla^2 E \right] \mathrm{d}v, \tag{7.12}$$

gdzie na potrzeby bardziej kompaktowego zapisu opuściliśmy jawną zależność E od $\mathbf{r}$. Po skorzystaniu z faktu, że obie funkcje spełniają to samo jednorodne równanie Helmholtza wewnątrz objętości V ($\nabla^2 E = -k^2 E$ oraz $\nabla^2 G = -k^2 G$, skąd wynika, że $E(-k^2 G) - G(-k^2 E) \equiv 0$), prawa strona tożsamościowo się zeruje i otrzymujemy:

$$\oint_S \left[E \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial n} \right] \mathrm{d}a = 0. \tag{7.13}$$

Ponieważ równanie (7.13) musi być spełnione dla dowolnej powierzchni zamkniętej niezawierającej osobliwości, otoczmy punkt obserwacji P dwiema powierzchniami: dużą powierzchnią S_2 oraz znajdującą się w jej wnętrzu sferą S_1



Rysunek 7.9: Dwie powierzchnie S_1 i S_2 otaczające punkt obserwacji P. W dowolnym punkcie leżącym na powierzchni S_2 normalna skierowana jest na zewnątrz, podczas gdy na sferze S_1 skierowana jest do środka.

o promieniu r (i środku w punkcie P) (rys. 7.9). W takim przypadku całkowita powierzchnia to $S = S_1 + S_2$, więc możemy zapisać:

$$\oint_{S_2} \left[E \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial n} \right] \mathrm{d}a = - \oint_{S_1} \left[E \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial n} \right] \mathrm{d}a, \tag{7.14}$$

gdzie normalna do powierzchni sfery S_1 została wybrana do wewnątrz (zgodnie z konwencją, wektor normalny do powierzchni zamkniętej musi być skierowany na zewnątrz ograniczanej objętości V, co w przypadku wewnętrznej sfery S_1 oznacza zwrot do jej środka).

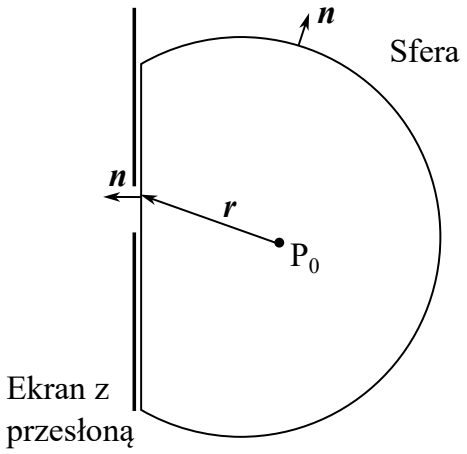
Celem naszych obliczeń jest wyznaczenie pola w punkcie P. Dlatego też chcemy, aby punkt ten był otoczony infinitezymalnie małą powierzchnią, tzn. aby $r \to 0$. Wprowadzając sferyczny układ współrzędnych i zauważając, że z powodu orientacji wektora normalnego do wewnątrz sfery zachodzi relacja $\frac{\partial}{\partial n} = -\frac{\partial}{\partial r}$, możemy zapisać:

$$\begin{aligned} &- \oint_{S_1} \left[E \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial n} \right] \mathrm{d}a \\ &\qquad = \lim_{r \to 0} \int_0^{2\pi} \int_0^{\pi} \left[E \left(\frac{\partial}{\partial r} \frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial r} \right] r^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi. \end{aligned} \tag{7.15}$$

Obliczając całkę w powyższym równaniu, otrzymujemy:

$$\begin{aligned} &\lim_{r \to 0} \int_0^{2\pi} \int_0^{\pi} \left[E \left(\frac{\partial}{\partial r} \frac{e^{ikr}}{r} \right) - \frac{e^{ikr}}{r} \frac{\partial E}{\partial r} \right] r^2 \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi \\ &\qquad = 4\pi \lim_{r \to 0} \left[r^2 \left(-\frac{e^{ikr}}{r^2} + ik \frac{e^{ikr}}{r} \right) E - r^2 \frac{e^{ikr}}{r} \frac{\partial E}{\partial r} \right] \\ &\qquad = 4\pi \lim_{r \to 0} \left[\left(-e^{ikr} + ikre^{ikr} \right) E - re^{ikr} \frac{\partial E}{\partial r} \right] \\ &\qquad = -4\pi E(P), \end{aligned} \tag{7.16}$$

gdzie drugi i trzeci wyraz w przedostatniej linii są równe zero z uwagi na mnożenie przez r, podczas gdy pierwszy wyraz redukuje się do $-E(P)$. Korzystając z



Rysunek 7.10: Przesłona z aperturą.

powyższego równania, jesteśmy w stanie wyznaczyć pole w punkcie obserwacji, które dane jest zależnością:

$$E(P) = \frac{1}{4\pi} \oint_{S_2} \left[\frac{e^{ikr}}{r} \frac{\partial E}{\partial n} - E \frac{\partial}{\partial n} \left(\frac{e^{ikr}}{r} \right) \right] \mathrm{d}a. \tag{7.17}$$

Z równania tego wynika, że znając wartość pola elektrycznego fali oraz jego pochodnej normalnej we wszystkich punktach na dowolnej powierzchni zamkniętej S_2, możemy obliczyć wartość pola w rozpatrywanym punkcie P. Relacja ta nosi nazwę **twierdzenia całkowego Helmholtza-Kirchhoffa** i stanowi ona podstawę **skalarnej teorii dyfrakcji**.

7.2.2 Amplituda fali w punkcie za przesłoną

Obliczmy teraz amplitudę fali w punkcie P_0 znajdującym się za nieprzezroczystym ekranem z aperturą (rys. 7.10). Wartość tego pola można wyznaczyć z równania (7.17), przy czym zamknięta powierzchnia S, którą otoczony jest ten punkt, składa się z dwóch części: płaszczyzny zlicowanej z ekranem wraz z przesłoną oraz części sfery o promieniu R odciętej tą płaszczyzną. W konsekwencji całka z równania (7.17) rozdziela się na sumę dwóch całek.

Zajmijmy się najpierw całką po części sferycznej. Na powierzchni sfery wektor normalny $\mathbf{n}$ pokrywa się z kierunkiem promienia r, co oznacza, że $\partial/\partial n = \partial/\partial r$. Obliczając pochodną radialną funkcji kulistej otrzymujemy:

$$\frac{\partial}{\partial r} \left(\frac{e^{ikr}}{r} \right) = \frac{ike^{ikr}r - e^{ikr}}{r^2} = \frac{e^{ikr}}{r} \left(ik - \frac{1}{r} \right). \tag{7.18}$$

Zakładając, że promień sfery R jest znacznie większy od długości fali, człon $1/R$

staje się zaniedbywalny. Zatem na powierzchni sfery ($r = R$) możemy przyjąć:

$$\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right) \approx ik\frac{e^{ikR}}{R}. \tag{7.19}$$

Ponieważ całkowanie po powierzchni sfery sprowadza się do sumowania po zmiennych kątowych, odległość od środka układu jest na niej z definicji stała i wynosi R. Czynnik $\exp(ikR)/R$ jest więc niezależny od zmiennych całkowania i może zostać wyłączony przed znak całki. Biorąc to pod uwagę, dla części sfery otrzymujemy:

$$\int_{\mathrm{sf}}\left[\frac{e^{ikr}}{r}\frac{\partial E}{\partial n} - E\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right)\right]\,\mathrm{d}a = \frac{e^{ikR}}{R}\int_{\mathrm{sf}}\left(\frac{\partial E}{\partial n} - ikE\right)\,\mathrm{d}a. \tag{7.20}$$

Zgodnie z tzw. warunkiem promieniowania Sommerfelda, który zakłada, że w nieskończoności nie ma źródeł emitujących fale z powrotem do układu, wyrażenie podcałkowe w równaniu (7.20) dąży do zera, gdy $R \to \infty$. Oznacza to, że wkład od części sferycznej znika, a pole w punkcie P_0 zależy wyłącznie od całki po płaszczyźnie ekranu i apertury. W celu rozwiązania tego problemu Kirchhoff przyjął dwa założenia (tzw. **warunki brzegowe Kirchhoffa**):

1. Pole E oraz jego pochodna $\partial E/\partial n$ w obszarze apertury przyjmują takie same wartości, jakie miałyby w przypadku braku przesłony.

2. Na nieoświetlonej (nieprzezroczystej) części przesłony zarówno pole, jak i jego pochodna tożsamościowo znikają, tzn. $E = 0$ oraz $\partial E/\partial n = 0$.

Przyjęcie tych dwóch założeń pozwala na rozwiązanie problemu dyfrakcji w ramach tzw. **przybliżenia Kirchhoffa**. Warto w tym miejscu zwrócić jednak uwagę, że powyższe założenia nigdy nie są całkowicie spełnione. W rzeczywistości bowiem na krawędzi apertury spełnione muszą być ścisłe warunki brzegowe wynikające z teorii elektromagnetyzmu. Dodatkowo, jak pokazaliśmy w przypadku metody stref Fresnela, cień geometryczny za przesłoną nigdy nie jest idealnie ostry na skutek ugięcia fal na jej krawędziach. Niemniej jednak, jeśli rozmiary apertury są znacznie większe od długości fali, to przybliżenie Kirchhoffa jest dobrze spełnione.

Wróćmy do analizy całki po powierzchni płaskiej. Zgodnie z przybliżeniem Kirchhoffa, pole na nieprzezroczystej części ekranu znika. Wynika stąd, że obszarem, w którym należy policzyć pole, jest obszar apertury. Zauważmy, że różniczkowanie wzdłuż kierunku normalnego odpowiada obliczeniu pochodnej kierunkowej. Ponieważ rozpatrywana funkcja zależy wyłącznie od odległości radialnej r, jej gradient jest równoległy do wektora wodzącego $\mathbf{r}$. Rzutowanie tego gradientu na wektor normalny $\mathbf{n}$ z definicji iloczynu skalarnego wprowadza czynnik w postaci cosinusa kąta pomiędzy tymi dwoma wektorami, wobec czego otrzymujemy:

$$\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right) = \frac{\partial r}{\partial n}\frac{\partial}{\partial r}\left(\frac{e^{ikr}}{r}\right) = \left(\frac{ik}{r} - \frac{1}{r^2}\right)e^{ikr}\cos[\sphericalangle(\mathbf{r},\mathbf{n})]. \tag{7.21}$$

Zakładając, że rozmiary obiektów, na których dochodzi do dyfrakcji, są znacząco większe od długości fali (przybliżenie optyczne), równanie (7.21) można przybliżyć przez:

$$\left(\frac{ik}{r} - \frac{1}{r^2}\right) e^{ikr} \cos[\sphericalangle(\mathbf{r}, \mathbf{n})] \approx \frac{ik}{r} e^{ikr} \cos[\sphericalangle(\mathbf{r}, \mathbf{n})]. \tag{7.22}$$

Wstawiając równania (7.20), (7.21) oraz (7.22) do równania (7.17), otrzymujemy wzór na amplitudę fali elektromagnetycznej w punkcie P_0 o współrzędnych (x, y, z):

$$E(x, y, z) = \frac{1}{4\pi} \int_{\text{ap}} \frac{e^{ikr}}{r} \left[\frac{\partial E}{\partial n} - ikE \cos[\sphericalangle(\mathbf{r}, \mathbf{n})]\right] \mathrm{d}a, \tag{7.23}$$

gdzie $\mathbf{r}$ jest wektorem łączącym element powierzchni przesłony da z punktem P_0.

7.2.3 Apertura oświetlana falą płaską

Ostatnim krokiem, który pozwala wyprowadzić jawną postać amplitudy pola w punkcie P_0, jest przyjęcie założenia co do rodzaju fali padającej na przesłonę. W najprostszym przypadku możemy założyć, że jest nią fala płaska w postaci $E(z) = E_0^{(+)} e^{ikz}$. W takim przypadku:

$$\frac{\partial E}{\partial n} = \frac{\partial z}{\partial n} \frac{\partial E}{\partial z} = ikE_0^{(+)} e^{ikz}(-1) = -ikE, \tag{7.24}$$

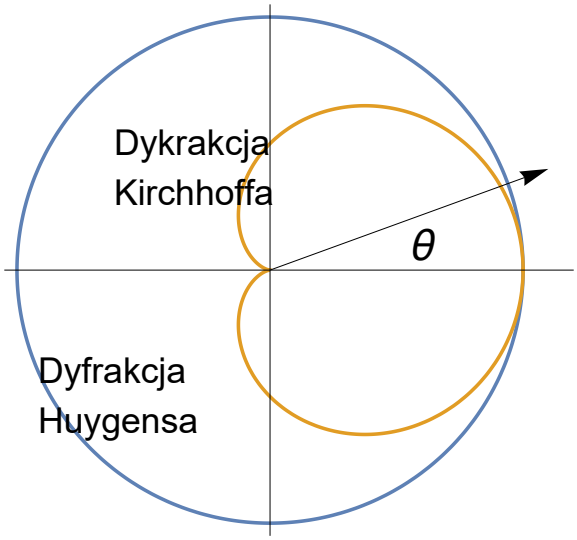
gdzie znak minus wynika z faktu, że kierunek propagacji fali i wektor wodzący wychodzący z punktu P_0 są przeciwnie zorientowane. Podstawiając równanie (7.24) do równania (7.23) oraz wykorzystując relację $k = 2\pi/\lambda$, otrzymujemy:

$$E(x, y, z) = -\frac{iE_0^{(+)}}{\lambda} \int_{\text{ap}} \frac{e^{ikr}}{r} \frac{1 + \cos[\sphericalangle(\mathbf{r}, \mathbf{n})]}{2} \mathrm{d}a. \tag{7.25}$$

Czynnik $-i$ pojawiający się przed całką informuje o przesunięciu fazowym (o $-\pi/2$), do jakiego dochodzi podczas dyfrakcji fali wtórnej względem fali pierwotnej. Wyrażenie zawierające cosinus jest przywoływanym przez nas już wcześniej czynnikiem nachylenia, który definiuje asymetrię emisji z wtórnych źródeł światła.

Równanie (7.25) może zostać również przepisane do postaci z wykorzystaniem zmiennych kartezjańskich. W takim przypadku pole elektryczne w dowolnym punkcie P_0 opisanym przez zestaw współrzędnych (x, y, z) dane jest zależnością:

$$E(x, y, z) = -\frac{iE_0^{(+)}}{\lambda} \iint_{\text{ap}} \frac{e^{ikr}}{r} \frac{1 + \cos[\sphericalangle(\mathbf{r}, \mathbf{n})]}{2} \mathrm{d}x' \, \mathrm{d}y', \tag{7.26}$$



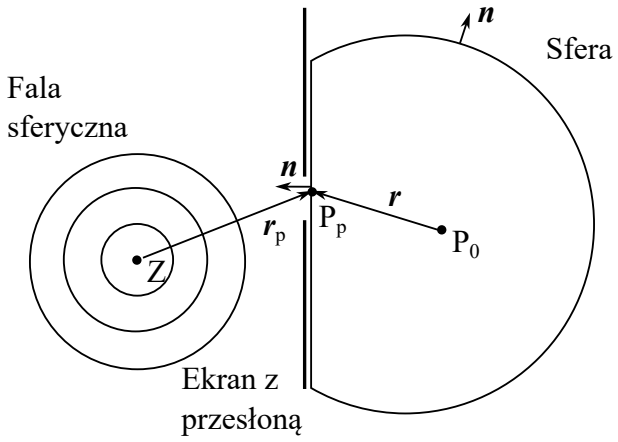
Rysunek 7.11: Porównanie amplitudy fali emitowanej w określonym kierunku, oznaczanej jako odległość odpowiedniej krzywej od środka układu współrzędnych, dla dyfrakcji Kirchhoffa (krzywa pomarańczowa) i Huygensa (krzywa niebieska).

gdzie

$$r = \sqrt{(x - x')^2 + (y - y')^2 + z^2}, \tag{7.27}$$

zaś sama apertura umieszczona jest w płaszczyźnie $z' = 0$.

Rysunek 7.11 porównuje dyfrakcję Kirchhoffa z klasyczną dyfrakcją Huygensa. Najważniejsza różnica polega na zmodulowaniu amplitudy emitowanej fali w funkcji kąta emisji. Należy w tym miejscu wyraźnie podkreślić, że analizowane krzywe reprezentują wyłącznie czynnik nachylenia określający zależność kierunkową amplitudy, a nie kształt frontu falowego. Same fale wtórne w obu modelach pozostają falami sferycznymi, co oznacza, że ich powierzchnie stałej fazy są kulami rozchodzącymi się od źródła. Jak widać, w klasycznym modelu Huygensa amplituda fali wtórnej jest izotropowa – taka sama w każdym kierunku na danym froncie sferycznym. W teorii Kirchhoffa rozkład natężenia na tym sferycznym froncie nie jest już jednorodny. Największą amplitudę ma fala propagująca się wzdłuż pierwotnego kierunku fali padającej. Amplituda fal wtórnych spada wraz ze zwiększaniem się kąta pomiędzy pierwotnym kierunkiem propagacji a kierunkiem dyfrakcji, spadając do zera dla fali emitowanej w tył. Wynik ten formalizuje wcześniejszą intuicję Fresnela, który postulował konieczność zmodulowania amplitudy fali wtórnej, aby wyeliminować niefizyczny problem powstawania fali wstecznej o natężeniu równym fali biegnącej w przód.



Rysunek 7.12: Dyfrakcja Kirchhoffa na aperturze oświetlanej falą sferyczną.

7.2.4 Apertura oświetlana falą sferyczną

Analogiczne rozważania można przeprowadzić dla sytuacji, w której przesłona oświetlona jest falą sferyczną ze źródła umieszczonego w punkcie Z znajdującym się przed przesłoną (rys. 7.12). W takim przypadku pole elektryczne w punkcie $\mathrm{P_{ap}}$ leżącym w płaszczyźnie przesłony dane jest zależnością:

$$E(\mathrm{P_{ap}}) = \mathrm{E_0^{(+)}} \frac{\mathrm{e^{ikr_{ap}}}}{\mathrm{r_{ap}}}, \tag{7.28}$$

gdzie r_{ap} jest odległością od źródła Z do punktu $\mathrm{P_{ap}}$ należącego do apertury. Zauważając, że dla wektora wodzącego $\mathbf{r}_{\mathrm{ap}}$ skierowanego od apertury do źródła zachodzi relacja

$$\frac{\partial r_{\mathrm{ap}}}{\partial n} = -\cos[\sphericalangle(\mathbf{r}_{\mathrm{ap}}, \mathbf{n})], \tag{7.29}$$

otrzymujemy, stosując ponownie przybliżenie optyczne (zaniedbując bardzo szybko malejące człony rzędu $1/r_{\mathrm{ap}}^2$):

$$\frac{\partial E}{\partial n} = -ikE_0^{(+)} \frac{e^{ikr_{\mathrm{ap}}}}{r_{\mathrm{ap}}} \cos[\sphericalangle(\mathbf{r}_{\mathrm{ap}}, \mathbf{n})]. \tag{7.30}$$

Wstawiając tę pochodną (oraz analogiczne przybliżenie pochodnej dla zmiennej r) do równania (7.23), otrzymujemy ostateczny wzór na amplitudę fali w przypadku dyfrakcji na przesłonie oświetlanej falą sferyczną:

$$E(P_0) = -\frac{iE_0^{(+)}}{\lambda} \int_{\mathrm{ap}} \frac{e^{ik(r+r_{\mathrm{ap}})}}{r r_{\mathrm{ap}}} \frac{\cos[\sphericalangle(\mathbf{r}, \mathbf{n})] + \cos[\sphericalangle(\mathbf{r}_{\mathrm{ap}}, \mathbf{n})]}{2} \,\mathrm{d}a. \tag{7.31}$$

Otrzymany wynik stanowi pełne sformułowanie całki dyfrakcyjnej Fresnela-Kirchhoffa dla punktowego źródła światła. Wyeksponowany pod całką czynnik

nachylenia jest w tym przypadku symetryczny względem wektorów kierunkowych padania ($\mathbf{r}_{\mathrm{ap}}$) oraz ugięcia ($\mathbf{r}$). Ta matematyczna symetria stanowi bezpośrednie odzwierciedlenie optycznego twierdzenia o wzajemności Helmholtza, które gwarantuje fizyczną odwracalność biegu promieni światła – zamiana miejscami punktu źródłowego Z i punktu obserwacji P_0 nie wpłynie na wartość rejestrowanej amplitudy pola.

7.2.5 Problem obliczeń dyfrakcji Kirchhoffa

Mimo że całka we wzorze dyfrakcyjnym Fresnela-Kirchhoffa przyjmuje stosunkowo zwartą postać, jej ścisłe analityczne rozwiązanie jest w wielu przypadkach praktycznie niemożliwe. Nawet w przypadku przybliżenia, w którym pomijamy czynnik nachylenia warunkujący kątową zależność amplitudy, tzn. gdy zakładamy, że $\frac{\cos[\sphericalangle(\mathbf{r},\mathbf{n})]+\cos[\sphericalangle(\mathbf{r}_{\mathrm{ap}},\mathbf{n})]}{2} \approx 1$, obliczenie całki często bywa niezwykle złożone matematycznie. Dlatego też konieczne staje się zastosowanie dalszych przybliżeń, które — choć ograniczają ogólność opisu — pozwolą na znaczne uproszczenie wyrażenia podcałkowego i analityczne wyznaczenie rozkładu pola w określonych reżimach propagacji (co prowadzi bezpośrednio do dyfrakcji Fresnela oraz Fraunhofera).

7.3 Reżimy dyfrakcyjne

Punktem wyjścia dla opisu zjawisk dyfrakcyjnych jest teoria Kirchhoffa, która w przybliżeniu optycznym definiuje funkcję pola elektrycznego fali wtórnej zależnością:

$$E(x,y,z) \propto \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{r}\cos[\sphericalangle(\mathbf{r},\mathbf{n})]. \tag{7.32}$$

Analityczne rozwiązanie tej zależności dla dowolnych odległości jest niezwykle złożone. Jednak wprowadzając rygor geometryczny wynikający z odległości ekranu z od apertury, możemy wyodrębnić dwa kluczowe reżimy przybliżone: **pole bliskie** (dyfrakcję Fresnela) oraz **pole dalekie** (dyfrakcję Fraunhofera).

7.3.1 Przybliżenie Fresnela (Pole bliskie)

Dla obrazów dyfrakcyjnych odległych od apertury o dystans znacznie większy od jej charakterystycznych rozmiarów, tj. gdy $z^2 \gg (x-x')^2+(y-y')^2$, odległość r pomiędzy punktem na aperturze a punktem na ekranie możemy rozwinąć w szereg Taylora. Przyjmując przybliżenie przyosiowe (paraksjalne), z dokładnością do wyrazów kwadratowych, otrzymujemy:

$$\begin{aligned} r &= \sqrt{(x-x')^2+(y-y')^2+z^2} \\ &= z\sqrt{1+\frac{(x-x')^2+(y-y')^2}{z^2}} \\ &\approx z+\frac{(x-x')^2+(y-y')^2}{2z}. \end{aligned} \tag{7.33}$$

W tym reżimie przyjmujemy również, że kąt ugięcia jest na tyle mały, iż $\cos[\sphericalangle(\mathbf{r},\mathbf{n})] \approx 1$. Pozwala to przekształcić ogólne równanie Kirchhoffa do postaci opisującej **dyfrakcję Fresnela**:

$$E(x,y,z) = -\frac{ie^{ikz}e^{i\frac{k}{2z}(x^2+y^2)}}{\lambda z} \times \int\int_{ap} E(x',y',0)e^{i\frac{k}{2z}(x'^2+y'^2)}e^{-i\frac{k}{z}(xx'+yy')}dx'dy'. \tag{7.34}$$

Obraz dyfrakcyjny w tym obszarze nazywamy obrazem pola bliskiego. Charakteryzuje się on tym, że niesie ze sobą bezpośrednie informacje o kształcie przedmiotu. Obraz ten odtwarza geometrię apertury, choć wraz z oddalaniem się od niej ulega stopniowej ewolucji i przestrzennym oscylacjom natężenia.

7.3.2 Przybliżenie Fraunhofera (Pole dalekie)

Mimo poczynionych uproszczeń, równanie (7.34) wciąż często wymaga metod numerycznych ze względu na obecność członu kwadratowego w wykładniku funkcji podcałkowej. W sytuacjach, gdy odległość z jest na tyle duża w stosunku do maksymalnych rozmiarów apertury, że spełniony jest warunek $z \gg \frac{k}{2}(x'^2+y'^2)_{\max}$, możemy przyjąć dalsze uproszczenie fazy:

$$e^{i\frac{k(x'^2+y'^2)}{2z}} \approx 1. \tag{7.35}$$

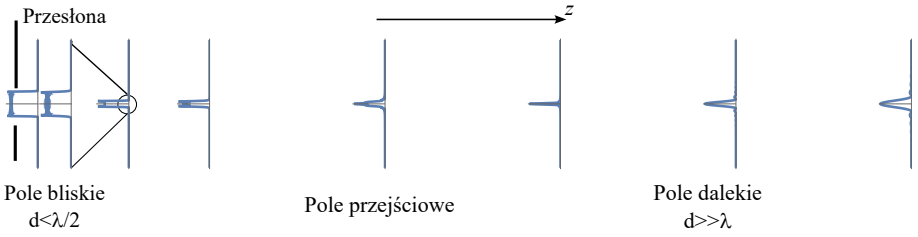
Wówczas wzór (7.34) przyjmuje znacznie prostszą postać opisującą **dyfrakcję Fraunhofera** (Joseph von Fraunhofer 1787–1826):

$$E(x,y,z) = -\frac{ie^{ikz}e^{i\frac{k}{2z}(x^2+y^2)}}{\lambda z}\int\int_{ap} E(x',y',0)e^{-i\frac{k}{z}(xx'+yy')}dx'dy'. \tag{7.36}$$

Jest to reżim pola dalekiego. Usunięcie członu kwadratowego ze współrzędnymi apertury (x',y') pod całką zdecydowanie zwiększa szansę na znalezienie rozwiązania analitycznego. Zauważmy, że całka ta jest w istocie dwuwymiarową odwrotną transformatą Fouriera rozkładu amplitudy pola na aperturze $E(x',y',0)$, gdzie zmienne kx/z i ky/z pełnią rolę dwóch częstotliwości przestrzennych. W tym reżimie rejestrowany obraz nie przypomina już geometrycznego zarysu przedmiotu (jak ma to miejsce blisko apertury), lecz jest bezpośrednią wizualizacją jego widma częstotliwości przestrzennych. Struktura obrazu dyfrakcyjnego jest zatem w pełni zdeterminowana przez przedmiot, podczas gdy właściwości ośrodka (poprzez długość fali λ) oraz odległość od ekranu z decydują wyłącznie o jego przestrzennej skali.

7.3.3 Ewolucja obrazu dyfrakcyjnego

Przejście między polem bliskim a dalekim ma fundamentalną naturę fizyczną. Jak przedstawiono na rys. 7.13, blisko przesłony obraz odtwarza jej kształt z widocznymi, drobnymi zafalowaniami. W miarę oddalania się od apertury, wpływ dyfrakcji na krawędziach narasta, aż do osiągnięcia obszaru pola dalekiego. W tym obszarze obraz przestaje być definiowany bezpośrednio przez początkowy geometryczny kształt przedmiotu, a zaczyna zależeć od jego widma częstotliwości przestrzennych. Wraz z dalszym oddalaniem się, struktura obrazu nie ulega już zmianie, a następuje jedynie jego proporcjonalne przeskalowanie.



Rysunek 7.13: Zmiana obrazu dyfrakcyjnego wraz z oddalaniem się od szczeliny jednowymiarowej. Blisko przesłony obraz dyfrakcyjny odtwarza kształt przesłony z drobnymi zafalowaniami. Obraz ulega ewolucji wraz z rosnącym dystansem i w polu dalekim osiąga kształt, który strukturalnie nie ulega już zmianie, a modyfikacji ulega jedynie jego skala.

7.4 Dyfrakcja światła na różnych obiektach

Doświadczenie Obserwacja obrazów dyfrakcyjnych na falownicy za szczeliną o różnej szerokości.

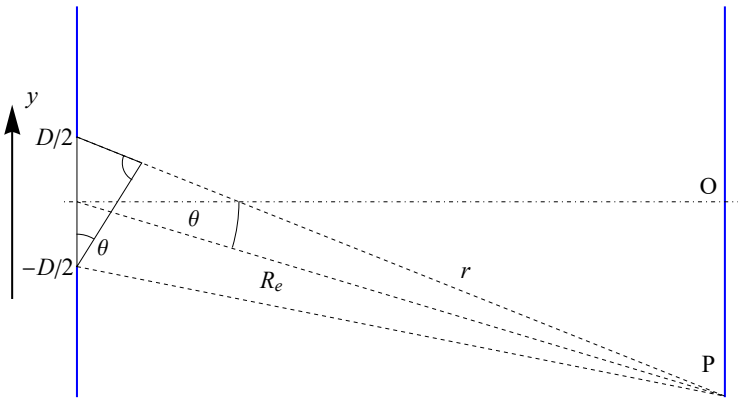
Skoro ustaliliśmy już, że obraz w reżimie Fraunhofera stanowi wizualizację widma częstotliwości przestrzennych apertury, prześledźmy teraz to zjawisko na przykładzie kilku podstawowych geometrii. Zanim przejdziemy do szczegółowych wyprowadzeń, należy podkreślić, że o ostatecznym charakterze obserwowanego obrazu decyduje przede wszystkim stosunek charakterystycznego wymiaru obiektu (np. szerokości apertury D) do długości fali uginanego światła λ. To ten parametr określa, w jakim stopniu zjawiska falowe determinują rozkład przestrzenny natężenia. W skrajnym przypadku, gdy obiekt jest makroskopowy ($D \gg \lambda$), dyfrakcja na jego krawędziach staje się pomijalnie mała. Przechodzimy wówczas do reżimu optyki geometrycznej, w którym za przesłoną obserwujemy po prostu ostry cień bez zauważalnej struktury falowej.

Dla uproszczenia, we wszystkich poniższych rozważaniach będziemy się poruszać w ramach omówionego wcześniej przybliżenia pola dalekiego.

7.4.1 Ugięcie światła na szczelinie jednowymiarowej

Rozważmy światło padające na nieprzezroczystą, nieskończoną przesłonę, w której wycięto (quasi)jednowymiarową szczelinę o szerokości D umieszczoną syme-

trycznie względem punktu $y' = 0$ (rys. 7.14). W przypadku dyfrakcji Fraun-



Rysunek 7.14: Ugięcie światła na jednowymiarowej aperturze o szerokości D.

hofera, amplituda pola elektrycznego, emitowanego z infinitezymalnie małego fragmentu przesłony dy', obserwowana na ekranie w punkcie P, dana jest zależnością

$$dE = \frac{dy'}{r} E_0^{(+)} e^{ikr}, \tag{7.37}$$

gdzie $E_0^{(+)}$ jest amplitudą fali padającej, zaś r jest odległością pomiędzy źródłem fali wtórnej w punkcie y' oraz punktem P na ekranie. Odległość r można wyrazić przez odległość R_e od środka szczeliny do punktu P oraz kąt θ, pod którym punkt P widoczny jest ze środka szczeliny. W pierwszym przybliżeniu pola dalekiego wyrażenie to przyjmuje postać

$$r \approx R_e - y' \sin\theta, \tag{7.38}$$

co pozwala zapisać równanie (7.37) jako

$$dE \approx \frac{dy'}{R_e} E_0^{(+)} e^{ik(R_e - y' \sin\theta)}. \tag{7.39}$$

Warto zauważyć, że w równaniu tym przybliżyliśmy odległość r w mianowniku wyrażenia przez R_e, pomijając nieistotny z punktu widzenia zmiany amplitudy człon $-y' \sin\theta$. Tego samego przybliżenia nie możemy jednak zastosować dla funkcji eksponencjalnej, z uwagi na mnożenie odległości r przez wektor falowy k, co prowadziłoby do utraty kluczowych informacji o zmianie fazy. Całkowitą amplitudę pola świetlnego w punkcie P obliczamy jako całkę po całej szerokości

szczeliny

$$\begin{aligned} E &\approx \frac{E_0^{(+)}}{R_e}\int_{-D/2}^{D/2} e^{ik(R_e - y'\sin\theta)}dy' \\ &= \frac{E_0^{(+)}}{R_e}e^{ikR_e}\left[\frac{e^{-iky'\sin\theta}}{-ik\sin\theta}\right]_{-D/2}^{D/2} \\ &= \frac{E_0^{(+)}}{R_e}e^{ikR_e}\frac{e^{-ik(D/2)\sin\theta} - e^{ik(D/2)\sin\theta}}{-ik\sin\theta} \\ &= \frac{E_0^{(+)}D}{R_e}e^{ikR_e}\frac{\sin\left(k\frac{D}{2}\sin\theta\right)}{k\frac{D}{2}\sin\theta}. \end{aligned} \tag{7.40}$$

Korzystając z ogólnej definicji natężenia światła w ośrodku, $I = v\varepsilon|E|^2/2$, możemy zapisać wyrażenie na natężenie światła docierającego do ekranu pod określonym kątem θ:

$$I = \frac{v\varepsilon}{2}\left|\frac{E_0^{(+)}D}{R_e}\right|^2\left(\frac{\sin\beta}{\beta}\right)^2 = I_{\max}\left(\frac{\sin\beta}{\beta}\right)^2, \tag{7.41}$$

gdzie $I_{\max}$ określa maksymalne natężenie światła w prążku centralnym (obserwowane na osi optycznej układu, tj. dla kąta $\theta = 0$), natomiast β zdefiniowana jest wzorem:

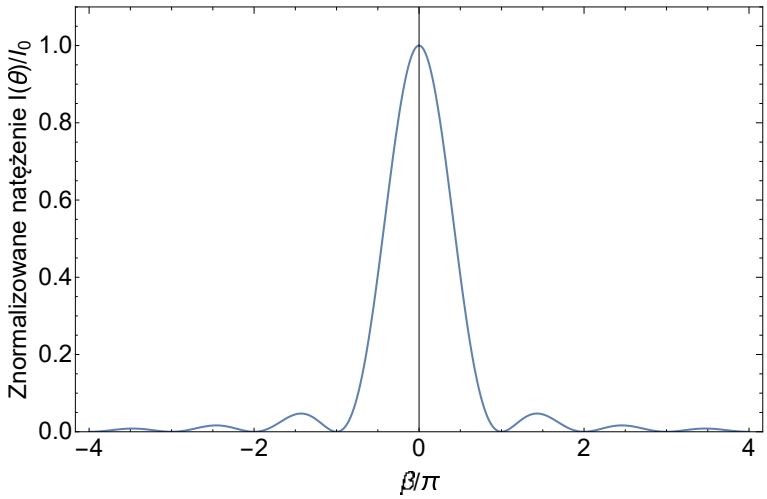
$$\beta = k\frac{D}{2}\sin\theta. \tag{7.42}$$

Rysunek 7.15 przedstawia wykres zależności natężenia światła w funkcji kąta θ dla ugięcia na szczelinie jednowymiarowej.

Posługując się równaniem (7.41), możemy teraz rygorystycznie określić rozkład natężenia światła w funkcji kąta dla dwóch rozważanych wcześniej skrajnych przypadków:

- **dyfrakcja na wąskiej szczelinie**
  W tym przypadku argument funkcji $\beta = k\frac{D}{2}\sin\theta = \frac{\pi D}{\lambda}\sin\theta$ jest małą wielkością ($D < \lambda/2$). Wówczas $\sin\beta \approx \beta$ i iloraz $\sin\beta/\beta \approx 1$, co oznacza, że natężenie fali w każdym kierunku ugięcia jest niemal identyczne, a fala wtórna za otworem staje się falą cylindryczną obserwowaną daleko od źródła (stąd identyczne oświetlenie na ekranie).

- **dyfrakcja na szerokiej szczelinie**
  Dla $D \gg \lambda$ argument β szybko zyskuje duże wartości wraz ze zmianą kąta. Funkcja $\sin\beta/\beta$ gwałtownie spada i ulega oscylacjom, co sprawia, że emisja jest silnie zmodulowana i wysoce kierunkowa, tworząc układ maksimów i minimów, tak jak to pokazano na rys. 7.15.



Rysunek 7.15: Dyfrakcja na szczelinie.

7.4.2 Ugięcie światła na otworze prostokątnym

Powyżej rozważaliśmy przypadek ugięcia fali na nieskończenie długiej szczelinie jednowymiarowej. Zajmijmy się teraz uogólnieniem tego problemu dla dwuwymiarowego otworu prostokątnego o wymiarach D_x i D_y. W takim przypadku elementarne pole dE od fragmentu apertury $dx'dy'$ obserwowanego na ekranie w odległości R_e przyjmuje postać:

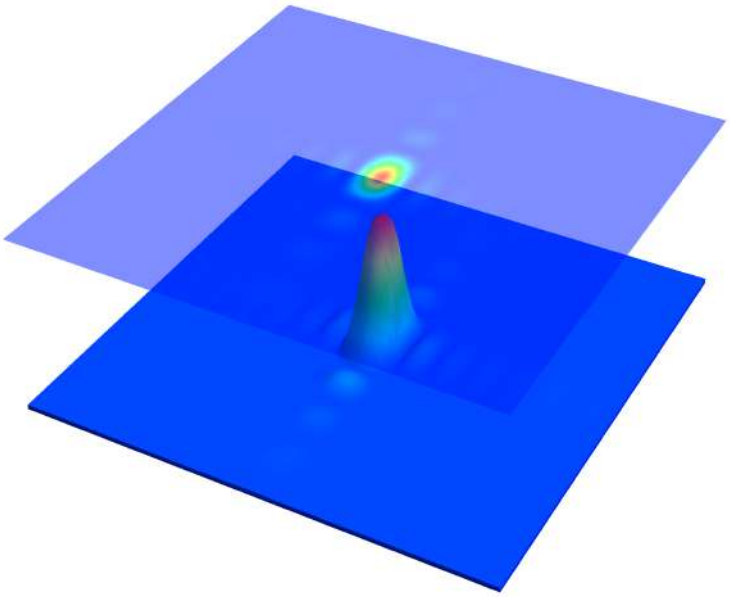
$$dE \approx \frac{dx'dy'}{R_e} E_0^{(+)} e^{ikr}, \tag{7.43}$$

gdzie przyjęliśmy, że amplituda maleje proporcjonalnie do $1/R_e$. W przybliżeniu pola dalekiego odległość w wykładniku fazowym wynosi $r \approx R_e - x' \sin\theta_x - y' \sin\theta_y$. Ponieważ granice całkowania dla prostokątnej apertury są niezależne, podwójna całka z równania (7.43) rozdziela się na iloczyn dwóch niezależnych całek jednowymiarowych po kierunkach x' i y'. Prowadzi to ostatecznie do wyrażenia na natężenie światła na ekranie, które jest iloczynem dwóch funkcji typu sinc^2:

$$\boxed{I(\theta_x, \theta_y) = I_{\max} \left(\frac{\sin\beta_x}{\beta_x}\right)^2 \left(\frac{\sin\beta_y}{\beta_y}\right)^2,} \tag{7.44}$$

gdzie $I_{\max}$ to maksymalne natężenie w centrum obrazu dyfrakcyjnego, a parametry β_x i β_y zdefiniowane są jako:

$$\begin{aligned} \beta_x &= k\frac{D_x}{2}\sin\theta_x, \\ \beta_y &= k\frac{D_y}{2}\sin\theta_y, \end{aligned} \tag{7.45}$$



Rysunek 7.16: Dyfrakcja na otworze prostokątnym.

zaś θ_x i θ_y są kątami widzenia punktu obserwacji rzutowanymi odpowiednio na płaszczyzny xz i yz.

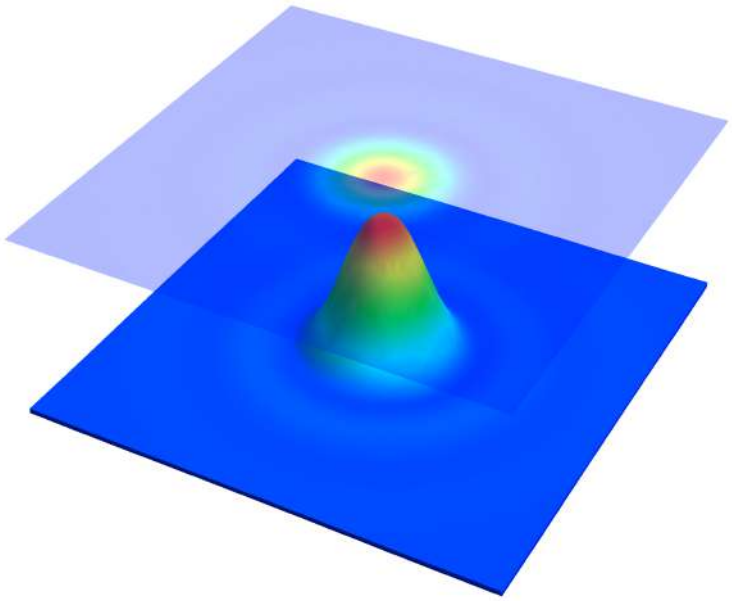
Rysunek 7.16 przedstawia natężenie światła obserwowanego na ekranie, gdy wymiar pionowy szczeliny jest dwukrotnie większy od poziomego. W obrazie dwuwymiarowym wyraźnie widoczna jest struktura krzyżowa powstała przez ortogonalne nałożenie na siebie niezależnych modulacji w kierunku x oraz y.

7.4.3 Ugięcie światła na otworze okrągłym

Obliczmy teraz natężenie fali świetlnej ugiętej na otworze kołowym o promieniu r_a. Całkowitą amplitudę pola w punkcie P znajdującym się na ekranie wyznaczamy, całkując przyczynki dE pochodzące od poszczególnych elementów powierzchni przesłony. Z uwagi na obrotową symetrię problemu, całkę najwygodniej jest zapisać przy użyciu współrzędnych biegunowych (r', ϕ') na aperturze. Element pola przyjmuje wówczas postać:

$$dE \approx \frac{r' dr' d\phi'}{R_e} E_0^{(+)} e^{ikr}. \tag{7.46}$$

W przybliżeniu Fraunhofera odległość r w wykładniku fazowym przyjmuje postać $r \approx R_e - r' \sin\theta \cos\phi'$, gdzie θ to obserwowany na ekranie kąt ugięcia fali. Podstawiając to rozwinięcie do równania (7.46), wykonując dwuwymiarowe całkowanie po całej powierzchni otworu, a następnie przechodząc z amplitudy do



Rysunek 7.17: Dyfrakcja na otworze kołowym.

natężenia (biorąc kwadrat modułu pola), otrzymujemy rozkład dany równaniem:

$$I(\theta) = I_{\max}\left[\frac{2J_1(kr_a\sin\theta)}{kr_a\sin\theta}\right]^2, \tag{7.47}$$

gdzie $J_1(u)$ to funkcja Bessela pierwszego rodzaju i rzędu pierwszego, pojawiająca się jako matematyczny wynik całkowania czynnika fazowego po zmiennej ϕ'.

Rysunek 7.17 obrazuje strukturę powstającą na ekranie w wyniku ugięcia fali na otworze okrągłym. Ze względu na symetrię kołową obraz dyfrakcyjny składa się ze współśrodkowych pierścieni (naprzemiennie jasnych i ciemnych) otaczających jasny obszar centralny. Ten centralny pik, wewnątrz którego zogniskowane jest około 84% całkowitej mocy światła, nazywamy **dyskiem Airy'ego**. Otaczające go pierścienie mają znacznie mniejsze natężenia. Przykładowo, maksimum pierwszego pierścienia osiąga 1,75%, zaś drugiego już tylko 0,42% wartości natężenia z centrum obrazu.

Aby obliczyć promień dysku Airy'ego, a co za tym idzie położenie **pierwszego minimum dyfrakcyjnego**, należy zauważyć, że funkcja Bessela $J_1(u)$ oscyluje, wielokrotnie powracając do wartości zerowych dla $u \approx 3{,}83;\ 7{,}02;\ \ldots$. Odpowiadający pierwszemu miejscu zerowemu warunek wygaszania przyjmuje postać:

$$kr_a\sin\theta \approx 3{,}83. \tag{7.48}$$

Przekształcając to wyrażenie, otrzymujemy fundamentalną relację określającą kątową wielkość dysku:

$$\theta \approx \sin\theta \approx 0,61\frac{\lambda}{r_a} = 1{,}22\frac{\lambda}{D_a}, \tag{7.49}$$

gdzie $D_a = 2r_a$ jest całkowitą średnicą apertury.

Rozpraszanie światła powodowane skończoną aperturą czynną układu (np. oprawą soczewki) sprawia, że nawet światło z pojedynczego punktowego źródła będzie widoczne na matrycy jako rozkład dany funkcją (7.47). Oznacza to, że każdy obrazowany punkt ma skończone wymiary przestrzenne, co determinuje absolutną rozdzielczość instrumentów optycznych.

7.5 Rozdzielczość układów optycznych

Zastanówmy się teraz, jaka będzie najmniejsza odległość dwóch obiektów punktowych, które będziemy mogli wyraźnie zobrazować na ekranie przy pomocy układu optycznego. Zgodnie z **kryterium Rayleigha**, dwa obrazy pochodzące od źródeł punktowych będą rozróżnialne wtedy, gdy środek obrazu dyfrakcyjnego (maksimum centralne) pochodzącego od pierwszego punktu pokrywa się z pierwszym minimum dyfrakcyjnym obrazu pochodzącego z drugiego punktu. Przykład obrazów dyfrakcyjnych pochodzących od dwóch źródeł punktowych pokazano na rys. 7.18.

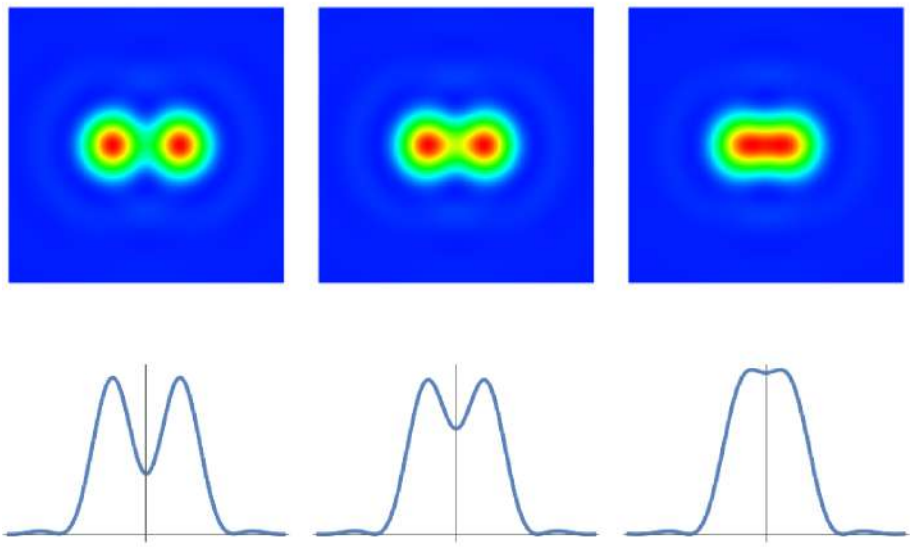
Warto w tym miejscu zauważyć, że kryterium Rayleigha, choć umowne (w praktyce można dokonać jeszcze dokładniejszego pomiaru, wykorzystując chociażby techniki superrozdzielcze), dostarcza prostej i intuicyjnej interpretacji **granicy rozdzielczości**, którą można osiągnąć w klasycznym obrazowaniu.

Na podstawie równania (7.49) można zdefiniować tzw. **rozdzielczość liniową**, zwaną także **zdolnością rozdzielczą**, czyli najmniejszą odległość dwóch przedmiotów (np. źródeł punktowych), które można rozróżnić przy pomocy układu optycznego:

$$d_u = \frac{1,22\lambda f}{D_a}. \tag{7.50}$$

Jak widać, rozdzielczość liniowa jest tym większa, tzn. odległość d_u tym bardziej maleje, im krótsza jest długość fali wykorzystywanej do obrazowania. To właśnie ta zależność tłumaczy, dlaczego mikroskop elektronowy – w którym obrazowania dokonuje się przy pomocy fali materii związanej z rozpędzoną wiązką elektronów, o dużo krótszej długości fali niż światło widzialne – pozwala na obserwację mniejszych obiektów od tych obserwowanych w tradycyjnej mikroskopii optycznej. Co więcej, poprawia się ona wraz ze skróceniem ogniskowej f i zwiększeniem średnicy apertury D_a układu obrazującego.

Przeprowadzone powyżej rozważania pozwalają na sformułowanie jednego z fundamentalnych z punktu widzenia mikroskopii kryteriów, tzw. **kryterium**



Rysunek 7.18: Dyfrakcja na dwóch otworach kołowych: obrazy rozróżnialne (lewa kolumna), obrazy dokładnie spełniające kryterium Rayleigha (środkowa kolumna) oraz niespełniające kryterium rozróżnialności (prawa kolumna).

Abbego (Ernst Abbe 1840–1905). Zgodnie z nim rozdzielczość mikroskopu ograniczona jest dyfrakcyjnie i dana jest zależnością:

$$d_u = \frac{1,22\lambda}{2\mathrm{NA}} \approx \frac{0,61\lambda}{\mathrm{NA}}, \tag{7.51}$$

gdzie NA jest **aperturą numeryczną** układu optycznego daną zależnością $\mathrm{NA} = n \sin \alpha_a$, przy czym α_a jest tzw. **kątem akceptacji układu optycznego**, który jest połową kąta wierzchołkowego stożka światła wpadającego do soczewki, zaś n jest współczynnikiem załamania ośrodka, w którym ten układ jest umieszczony.

Warto na koniec zauważyć, że wyprowadzone wcześniej wyrażenie na wielkość dysku Airy'ego może zostać bezpośrednio wykorzystane do zdefiniowania tzw. **rozdzielczości kątowej**, czyli najmniejszego kąta $\delta\theta_{\min}$, dla którego dwa obserwowane oddalone obiekty (np. gwiazdy) mogą zostać rozdzielone:

$$\delta\theta_{\min} = \frac{1,22\lambda}{D_a}. \tag{7.52}$$

Dygresja

Mikroskopia superzdolności rozdzielczej to dziedzina współczesnej optyki (uhonorowana Nagrodą Nobla z chemii w 2014 roku), która pozwala na obrazowanie struktur w skali nano, z precyzją wielokrotnie przekraczającą klasyczną granicę dyfrakcji Abbego. Przez ponad sto lat wydawało się, że fizyczna natura światła bezwzględnie uniemożliwia rozróżnienie obiektów leżących od siebie w odległości mniejszej niż około połowa długości fali (czyli ok. 200 nm dla światła widzialnego).

Eleganckie obejście tego problemu polega na przejściu od analizy przestrzennej do analizy w dziedzinie czasu. Znakomitym tego przykładem są metody oparte na lokalizacji pojedynczych cząsteczek (takie jak PALM czy STORM). Badany preparat biologiczny znakuje się fluoroforami (cząsteczkami świecącymi pod wpływem lasera), które posiadają zdolność przełączania się z ciemnego w jasny stan optyczny. Zamiast oświetlać wszystkie molekuły naraz – co doprowadziłoby do zlania się ich obrazów dyfrakcyjnych w jedną, nieczytelną plamę – badacze używają słabego światła, które w danym momencie wzbudza jedynie losowy, znikomy odsetek fluoroforów.

W efekcie świecące punkty są od siebie na tyle oddalone, że ich dyski Airy'ego nie nakładają się na siebie. Choć każda taka plamka światła z osobna pozostaje dyfrakcyjnie rozmyta, analiza matematyczna (dopasowanie) pozwala z ogromną precyzją (ograniczoną w zasadzie jedynie liczbą zebranych fotonów) wyznaczyć środek tej plamki, odpowiadający dokładnemu położeniu cząsteczki. Powtarzając ten proces błyskawicznie tysiące razy dla różnych podzbiorów molekuł, komputer punkt po punkcie rekonstruuje ostateczny obraz o niewyobrażalnej wręcz rozdzielczości rzędu 20 nm. Techniki te udowadniają, że klasyczny limit dyfrakcji dotyczy jednoczesnego oświetlenia ciągłego, a nie możliwości detekcji pojedynczych źródeł światła.

7.6 Dyfrakcja na dwóch szczelinach

Wyobraźmy sobie, że fala płaska pada normalnie na nieprzezroczysty ekran, w którym wycięto dwie równoległe szczeliny o szerokościach D, oddalone od siebie o odległość d (mierzoną między ich środkami). Na każdej ze szczelin dochodzi do ugięcia światła. Przyjmując początek układu współrzędnych w połowie odległości między szczelinami, możemy zapisać przestrzenny profil pola elektrycznego pochodzący od szczeliny górnej ($y_1 = d/2$) i dolnej ($y_2 = -d/2$) w punkcie obserwacji na odległym ekranie [porównaj z równaniem (7.40)]:

$$\begin{aligned} E_1 &= E_{\max} \frac{\sin\beta}{\beta} e^{ik(R_e - \frac{d}{2}\sin\theta)}, \\ E_2 &= E_{\max} \frac{\sin\beta}{\beta} e^{ik(R_e + \frac{d}{2}\sin\theta)}, \end{aligned} \tag{7.53}$$

gdzie $E_{\max}$ to amplituda pola emitowana w kierunku osi optycznej. W powyższych równaniach założyliśmy pole dalekie, przybliżając odległość w mianowniku jako R_e, ale w wykładniku fazowym jawnie uwzględniliśmy przesunięcia geometryczne obu źródeł względem środka układu.

Zgodnie z **zasadą superpozycji**, wypadkowe pole rejestrowane w danym punkcie na ekranie jest sumą wskazów pól składowych:

$$E = E_1 + E_2 = E_{\max}\frac{\sin\beta}{\beta}e^{ikR_e}\left(e^{-ik\frac{d}{2}\sin\theta} + e^{ik\frac{d}{2}\sin\theta}\right). \tag{7.54}$$

Wykorzystując tożsamość Eulera ($e^{ix} + e^{-ix} = 2\cos x$), możemy uprościć to wyrażenie do postaci:

$$E = 2E_{\max}\frac{\sin\beta}{\beta}e^{ikR_e}\cos\gamma, \tag{7.55}$$

gdzie wprowadziliśmy nowy parametr $\gamma = kd\sin\theta/2 = \pi d\sin\theta/\lambda$, opisujący przesunięcie fazowe wynikające z odległości między szczelinami.

Przechodząc do natężenia światła ($I \propto |E|^2$), otrzymujemy fundamentalny wzór dla doświadczenia z dwiema szczelinami:

$$I(\theta) = 4I_{\max}\left(\frac{\sin\beta}{\beta}\right)^2\cos^2\gamma, \tag{7.56}$$

gdzie $I_{\max}$ to maksymalne natężenie światła w centrum układu pochodzące od pojedynczej szczeliny (współczynnik 4 wynika z faktu, że pola dodają się konstruktywnie w fazie).

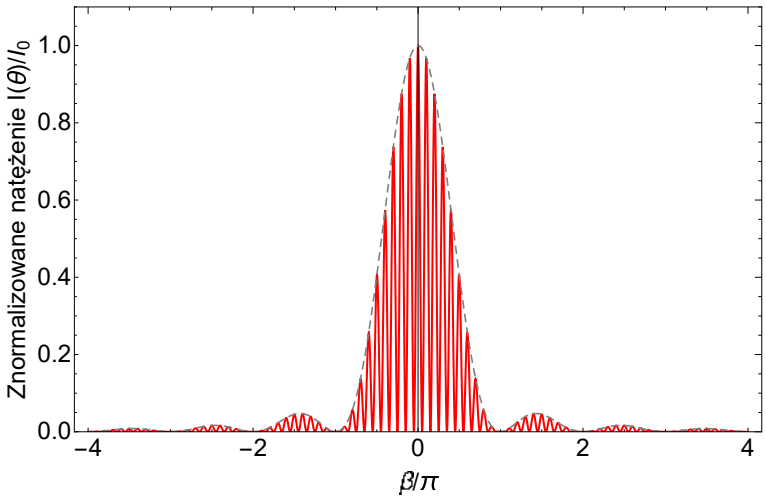
Analiza równania (7.56) wykazuje, że obok występującego wcześniej sygnału dyfrakcyjnego (funkcja $\sin^2\beta/\beta^2$), w rozkładzie pojawia się dodatkowy, szybkozmienny człon cosinusoidalny, będący bezpośrednim skutkiem **interferencji** fal ugiętych na obu szczelinach. Z uwagi na to, że odległość między szczelinami jest z reguły znacznie większa od ich szerokości ($d > D$), częstość modulacji wnoszona przez parametr γ jest odpowiednio większa niż wolniejsza modulacja dyfrakcyjna wnoszona przez β.

W konsekwencji, w polu dalekim otrzymujemy charakterystyczny obraz (Rys. 7.19), którego szeroka obwiednia determinowana jest przez dyfrakcję na pojedynczej szczelinie, podczas gdy drobna, regularna struktura prążków wewnątrz tej obwiedni tworzona jest przez interferencję typu Younga.

7.7 Układ N_s szczelin czyli siatka dyfrakcyjna

7.7.1 Rozkład natężenia światła

Rozszerzmy analizowany wcześniej przypadek na układ N_s identycznych szczelin o szerokości D, oddalonych od siebie o stałą odległość d. Układ taki stanowi naturalne uogólnienie doświadczenia z dwiema szczelinami, a jego praktyczna



Rysunek 7.19: Obraz powstawający na ekranie w wyniku oświetlenia dwóch szczelin falą płaską. Wolnozmienna obwiednia dyfrakcyjna moduluje szybkozmienne prążki interferencyjne.

realizacja – element optyczny zbudowany z bardzo wielu regularnie rozmieszczonych nacięć – nosi nazwę **siatki dyfrakcyjnej**. Warto zauważyć, że współczesne siatki przeznaczone dla światła widzialnego posiadają zazwyczaj od 1200 do 2400 rys na milimetr, co odpowiada odległościom d rzędu ułamków mikrometra.

Aby wyznaczyć rozkład natężenia na ekranie, musimy zsumować pola elektrycznego pochodzące już nie z dwóch, lecz ze wszystkich N_s szczelin. Przyjmując pierwszą z nich jako punkt odniesienia dla fazy, całkowitą amplitudę pola możemy zapisać w postaci sumy skończonej:

$$E = \sum_{j=0}^{N_s-1} E_j = E_{\max} \frac{\sin\beta}{\beta} e^{ikR_e} \sum_{j=0}^{N_s-1} e^{ijk\Delta r}, \tag{7.57}$$

gdzie $\Delta r = d\sin\theta$ jest różnicą dróg optycznych pomiędzy promieniami wychodzącymi z dwóch sąsiednich szczelin. Warto w tym miejscu podkreślić, że tak prosta postać geometryczna Δr wynika bezpośrednio z przybliżenia pola dalekiego (Fraunhofera) – ze względu na ogromną odległość do ekranu w porównaniu z rozmiarami siatki, promienie ugięte na poszczególnych szczelinach i biegnące do jednego punktu obserwacji traktujemy jako idealnie równoległe. Wykorzystując znany z matematyki wzór na sumę skończonego ciągu geometrycznego i wyciągając przed nawias odpowiednie czynniki fazowe, można zapisać amplitudę w postaci:

$$E = E_{\max} \frac{\sin\beta}{\beta} e^{ikR'_e} \frac{\sin(N_s\gamma)}{\sin\gamma}, \tag{7.58}$$

gdzie R'_e to zmodyfikowana odległość rzutująca na całkowitą fazę, a $\gamma = (kd\sin\theta)/2 =$

$(\pi d \sin\theta)/\lambda$ to zdefiniowany już wcześniej parametr przesunięcia fazowego między sąsiednimi szczelinami (zauważmy, że $k\Delta r = 2\gamma$).

Po wzięciu kwadratu modułu z powyższego wyrażenia, otrzymujemy ostateczne równanie na rozkład kątowy natężenia światła:

$$I(\theta) = I_{\max}\left(\frac{\sin\beta}{\beta}\right)^2\left[\frac{\sin(N_s\gamma)}{\sin\gamma}\right]^2. \tag{7.59}$$

Pierwszy czynnik, zgodnie ze wzorem (7.41), uwzględnia ugięcie na każdej ze szczelin z osobna (obwiednia dyfrakcyjna). Drugi człon związany jest z interferencją fal. Warto sprawdzić fenomenalną spójność tego zapisu: dla $N_s = 2$ (dwie szczeliny), ułamek interferencyjny wynosi $[\sin(2\gamma)/\sin\gamma]^2 = (2\cos\gamma)^2 = 4\cos^2\gamma$, co idealnie odtwarza wzór (7.56) z poprzedniej sekcji.

Zależność interferencyjna dla wielu szczelin nie jest już łagodną funkcją cosinus, ale kwadratem stosunku dwóch sinusów, z których ten w liczniku oscyluje N_s-razy szybciej. Prowadzi to do rozkładu, w którym natężenie pomiędzy maksimami bardzo szybko spada blisko zera. Z drugiej jednak strony, gdy argument w mianowniku zmierza do zera (tzn. gdy $\gamma = m\pi$, gdzie m jest liczbą całkowitą), z reguły de L'Hospitala wynika, że cały człon ulega wzmocnieniu do wartości N_s^2.

Doświadczenie Demonstracja dyfrakcji na wielu szczelinach.

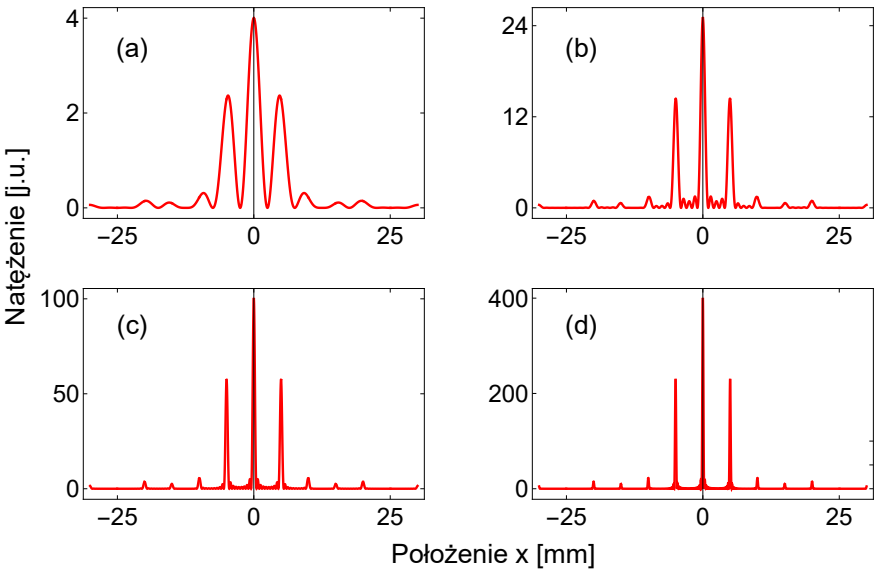
Rysunek 7.20 pokazuje obrazy powstałe przy dyfrakcji na układach o rosnącej liczbie szczelin. Jak wyraźnie widać, wzrost liczby szczelin powoduje drastyczne zawężenie obserwowanych maksimów interferencyjnych. Oprócz tych niezwykle jasnych pików głównych, w sygnale można również zaobserwować $N_s - 2$ słabszych maksimów wtórnych, których pojawienie się jest związane z lokalnymi ekstremami funkcji $\sin(N_s\gamma)$ z licznika. Natężenie całego tego złożonego układu prążków jest zawsze bezwzględnie ograniczane przez powolną obwiednię dyfrakcyjną daną wyrażeniem $(\sin\beta/\beta)^2$.

7.7.2 Równanie siatki i dyspersja

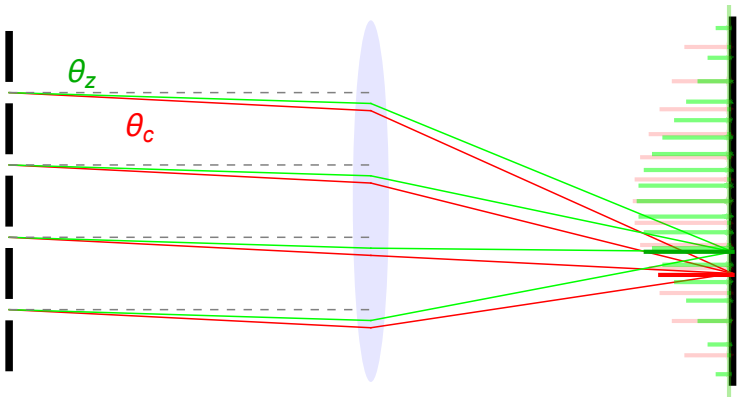
Z wyprowadzonych powyżej zależności wynika, że główne maksima dyfrakcyjne pojawiają się, gdy mianownik w członie interferencyjnym dąży do zera ($\gamma = m\pi$). Po podstawieniu definicji γ, otrzymujemy klasyczne **równanie siatki dyfrakcyjnej**:

$$d\sin\theta_m = m\lambda, \tag{7.60}$$

które oznacza, że konstruktywna interferencja zachodzi, gdy różnica dróg optycznych pochodzących z dwóch kolejnych szczelin równa jest całkowitej wielokrotności długości fali ($m = 0, \pm1, \pm2, \dots$). Najsilniejsze maksimum dla światła padającego prostopadle na siatkę osiągane jest dla kąta ugięcia zerowego ($\theta_0 = 0°$). Jest to tzw. prążek zerowego rzędu ($m = 0$), którego położenie jest niezależne od długości fali.



Rysunek 7.20: Dyfrakcja na układzie o różnej liczbie szczelin, ale tej samej szerokości D i odległości między szczelinami d: (a) dwie szczeliny, (b) pięć szczelin, (c) dziesięć szczelin oraz (d) dwadzieścia szczelin. Wraz ze zwiększaniem liczby szczelin maleje szerokość maksimów głównych, podczas gdy ich natężenie szczytowe rośnie proporcjonalnie do N_s^2. W konsekwencji całkowita moc optyczna (całka pod wykresem) jest rygorystycznie zachowana.



Rysunek 7.21: Schemat dyfrakcji światła czerwonego (700 nm) i zielonego (500 nm) oświetlającego tę samą siatkę dyfrakcyjną. Dłuższe fale uginają się pod większym kątem.

Położenie pozostałych wyższych rzędów zależy już silnie od długości fali λ. Fakt, że różne kolory uginane są pod różnym kątem, czyni z siatki dyfrakcyjnej niezwykle wydajny **element dyspersyjny**. Może być on wykorzystany do przestrzennego rozszczepienia światła i analizy widmowej źródła promieniowania. Dyfrakcja światła czerwonego (700 nm) i zielonego (500 nm) na tej samej siatce dyfrakcyjnej przedstawiona została schematycznie na rys. 7.21. Parametrem opisującym, jak silnie siatka rozsuwa kątowo fale o różnych długościach, jest **dyspersja kątowa**:

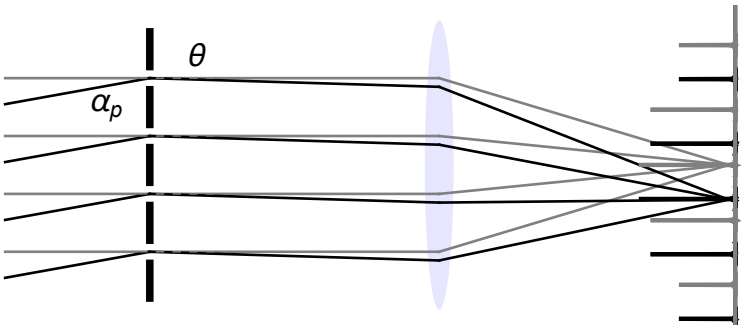
$$D_\theta = \frac{d\theta}{d\lambda} = \frac{m}{d\cos\theta}. \tag{7.61}$$

Z powyższej zależności wynika fundamentalny wniosek praktyczny: aby uzyskać jak największe rozszczepienie przestrzenne badanego widma, należy stosować siatki o dużej gęstości rys (małej stałej siatki d) oraz analizować promieniowanie ugięte w wyższych rzędach widma (m).

7.7.3 Padanie pod kątem i siatki profilowane

Równanie (7.60) jest prawdziwe dla sytuacji, w której siatka oświetlana jest falą płaską padającą prostopadle do przesłony. Gdy jednak fala pada pod niezerowym kątem α_p, faza fal docierających do kolejnych szczelin na starcie nie jest taka sama. Dla dwóch sąsiednich szczelin początkowa różnica dróg optycznych wynosi $\Delta r_{in} = d\sin\alpha_p$. W takim przypadku warunek na konstruktywną interferencję promieni musi bilansować różnicę dróg przed i za siatką, przyjmując uogólnioną postać:

$$d(\sin\theta - \sin\alpha_p) = m\lambda. \tag{7.62}$$



Rysunek 7.22: Dyfrakcja na siatce dyfrakcyjnej oświetlanej pod kątem do normalnej. Zbliżone kolory promieni (szary i czarny) mają sugerować, że różnica w ugięciu nie wynika z różnicy długości fali, a z kątów padania na siatkę.

Analiza równania (7.62) wykazuje, że skutkiem oświetlenia pod kątem nie jest zmiana samej struktury prążków, lecz przesunięcie całego obrazu dyfrakcyjnego (w tym przesunięcie maksimum zerowego rzędu do kąta równego kątowi padania). Sytuacja ta została zilustrowana na rys. 7.22.

W klasycznych siatkach szczelinowych najwięcej energii skupia się zawsze w „bezużytecznym" (niedyspersyjnym) prążku zerowego rzędu. Niezwykle ważnym rozwiązaniem technologicznym są tzw. **siatki profilowane**, w których większość energii kierowana jest w wybrany, wyższy rząd ugięcia (najczęściej pierwszy). Osiąga się to poprzez geometryczne nachylenie elementów dyfrakcyjnych (np. rzeźbienie w kształcie piłokształtnym) w stosunku do płaszczyzny przesłony, tak aby prawo odbicia faworyzowało konkretny kąt ugięcia. Innym nowoczesnym sposobem realizacji wysoko wydajnych siatek jest przestrzenna modulacja współczynnika załamania materiału, realizowana m.in. za pomocą oświetlenia laserowego, co prowadzi do powstania **siatek holograficznych**.

7.7.4 Zdolność rozdzielcza siatki

Skoro siatka dyfrakcyjna jest elementem dyspersyjnym analizującym widmo, kluczowym pytaniem jest: jak małą różnicę długości fal $\Delta\lambda_{\min}$ jesteśmy w stanie za jej pomocą zaobserwować? Miarą tej precyzji jest **chromatyczna zdolność rozdzielcza** $\mathcal{R}$, definiowana jako:

$$\mathcal{R} = \frac{\lambda}{\Delta\lambda_{\min}}. \tag{7.63}$$

Aby wyznaczyć tę granicę, posłużymy się **kryterium Rayleigha**: dwie bliskie linie widmowe (dla fali λ oraz $\lambda+\Delta\lambda_{\min}$) są ledwie rozróżnialne, gdy maksimum główne dla jednej z nich pokrywa się dokładnie z pierwszym minimum leżącym tuż obok maksimum drugiej fali.

Zastanówmy się, gdzie znajduje się pierwsze minimum obok m-tego maksimum. Warunek na maksimum główne dla fali λ to $d\sin\theta = m\lambda$. Jeśli pomno-

żymy go przez całkowitą liczbę oświetlonych szczelin N_s, otrzymamy relację dla całej siatki: $N_s d \sin\theta = N_s m\lambda$. Pierwsze minimum interferencyjne obok tego piku pojawi się, gdy sumaryczna różnica dróg optycznych na szerokości całej siatki wzrośnie o dokładnie jedną długość fali λ. W takiej sytuacji promienie pochodzące od kolejnych szczelin utworzą pełny cykl przesunięć fazowych, co doprowadzi do ich całkowitego destruktywnego wygaszenia. Zatem dla minimum kąt ugięcia spełnia warunek:

$$N_s d \sin\theta = N_s m\lambda + \lambda. \tag{7.64}$$

Z drugiej strony, maksimum główne m-tego rzędu dla sąsiedniej, nieco dłuższej fali $(\lambda+\Delta\lambda_{\min})$ musi wypaść dokładnie pod tym samym kątem θ (jest to wymóg nałożenia z kryterium Rayleigha):

$$N_s d \sin\theta = N_s m(\lambda + \Delta\lambda_{\min}). \tag{7.65}$$

Porównując ze sobą prawe strony równań (7.64) i (7.65), otrzymujemy natychmiast:

$$N_s m\lambda + \lambda = N_s m\lambda + N_s m\Delta\lambda_{\min}, \tag{7.66}$$

co po skróceniu wyrazów podobnych daje $\lambda = N_s m\Delta\lambda_{\min}$. Korzystając z definicji (7.63), otrzymujemy ostateczny, niezwykle prosty i elegancki wzór na **zdolność rozdzielczą siatki dyfrakcyjnej**:

$$\boxed{\mathcal{R} = mN_s.} \tag{7.67}$$

Ze wzoru (7.67) płynie fundamentalny wniosek techniczny: zdolność rozdzielcza zależy wyłącznie od rzędu ugięcia m oraz całkowitej liczby oświetlonych szczelin N_s. Z tego powodu, aby zbudować spektrometr o wysokiej rozdzielczości, konieczne jest poszerzenie wiązki optycznej (np. za pomocą kolimatora optycznego) tak, aby oświetlała ona możliwie największą powierzchnię siatki dyfrakcyjnej (jak największą liczbę czynnych rys).

Dygresja

Zrozumienie zjawiska dyfrakcji było kluczem do jednego z największych odkryć w historii biologii – rozwikłania struktury przestrzennej kwasu deoksyrybonukleinowego (DNA). Skoro granica dyfrakcji Abbego nie pozwala na obrazowanie pojedynczych cząsteczek za pomocą światła widzialnego, naukowcy musieli użyć promieniowania o znacznie krótszej długości fali. Wybór padł na promieniowanie rentgenowskie, którego długość fali (rzędu dziesiątych części nanometra) jest porównywalna z odległościami między atomami w cząsteczkach.

Gdy wiązka promieni X przechodzi przez skrystalizowaną próbkę, atomy tworzące cząsteczkę działają jak skomplikowana, trójwymiarowa siatka dyfrakcyjna. Fale ugięte na poszczególnych atomach interferują ze sobą, tworząc na kliszy fotograficznej wzór maksimów dyfrakcyjnych. Słynne „Zdję-

cie 51", wykonane w 1952 roku przez Rosalindę Franklin, ukazało charakterystyczny układ plamek w kształcie litery „X". Dla fizyków (takich jak James Watson i Francis Crick, którzy zinterpretowali te wyniki), ten specyficzny obraz dyfrakcyjny w polu dalekim był jednoznacznym, matematycznym dowodem na to, że cząsteczka DNA ma strukturę podwójnej helisy.

7.8 Zasada Babineta

W naszych dotychczasowych rozważaniach przyjmowaliśmy, że obiektem rozpraszającym jest nieprzezroczysta przesłona ze szczeliną. Można jednak się zastanowić, jaka będzie dyfrakcja, gdy odwrócimy sytuację, tzn. wtedy, gdy obiektem rozpraszającym nie jest otwór, a np. pojedynczy drut lub pasek o wymiarach dokładnie odpowiadających wymiarom tej szczeliny. Takie przesłony nazywamy **przesłonami komplementarnymi**.

Jasnym jest, że w polu bliskim, gdzie dyfrakcja determinowana jest bezpośrednio przez geometrię obiektu, obserwowany obraz będzie „negatywem" obrazu powstałego w przypadku szczeliny. Co jednak stanie się w polu dalekim, gdzie obraz dyfrakcyjny determinowany jest przez właściwości falowe światła?

Oznaczmy rozkład pola elektrycznego fali ugiętej na szczelinie (otworze w nieprzezroczystym ekranie) jako E_1, zaś pole wytworzone przez nieprzezroczysty pasek (umieszczony w otwartej przestrzeni) jako E_2. Zgodnie z **zasadą superpozycji**, nałożenie na siebie tych dwóch komplementarnych przesłon wirtualnie odtwarza sytuację, w której na drodze fali nie znajduje się żadna przeszkoda. Wynika stąd, że suma obu pól musi być równa polu fali całkowicie niezaburzonej E_0:

$$E_1 + E_2 = E_0. \tag{7.68}$$

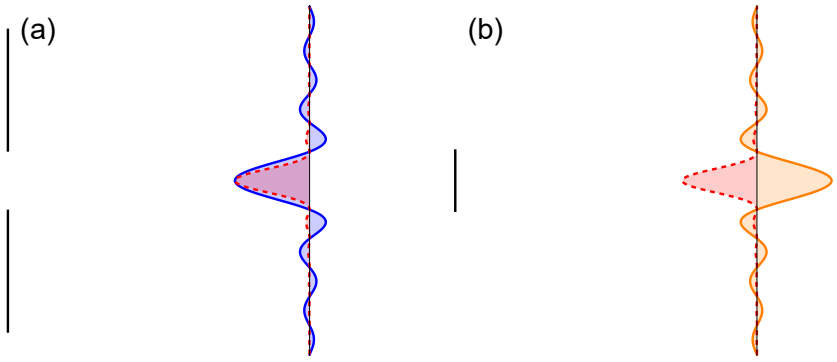
W przybliżeniu pola dalekiego (Fraunhofera) fala niezaburzona propaguje się wyłącznie wzdłuż pierwotnego kierunku padania, tworząc jasny punkt w samym centrum ekranu (dla kąta $\theta = 0$). Oznacza to, że dla wszystkich pozostałych kierunków ugięcia ($\theta \neq 0$) pole fali niezaburzonej jest równe zeru ($E_0 = 0$). Z relacji tej wynika zatem, że poza ścisłym środkiem obrazu pola ugięte na przesłonach komplementarnych muszą być sobie przeciwne:

$$E_1 = -E_2. \tag{7.69}$$

Ponieważ rejestrowane przez detektory natężenie światła jest proporcjonalne do kwadratu modułu amplitudy pola ($I \propto |E|^2$), oba układy dadzą w polu dalekim identyczny rozkład natężenia:

$$I_1(\theta) = I_2(\theta) \quad \text{dla} \quad \theta \neq 0. \tag{7.70}$$

Relacja ta stanowi treść **zasady Babineta** (Jacques Babinet 1794–1872). Mówi ona, że obrazy dyfrakcyjne komplementarnych przesłon są identyczne wszędzie poza bezpośrednim sąsiedztwem prążka zerowego rzędu.



Rysunek 7.23: Zasada Babineta. Dyfrakcja na przedmiocie i jego „negatywie" prowadzi do powstania przeciwnych fazowo rozkładów pola elektrycznego, jednak mierzony obraz dyfrakcyjny (natężenie) pozostaje w obu przypadkach taki sam.

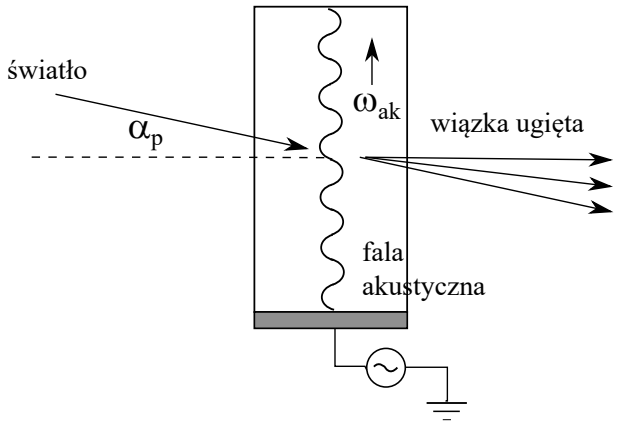
Zasada Babineta ma ogromne znaczenie praktyczne. Pozwala ona na badanie dyfrakcji na obiektach trudnych do matematycznego opisania (jak np. drobne cząstki, włókna czy krwinki), zastępując je w obliczeniach komplementarnymi otworami o tym samym kształcie. Przykładowo, mierząc obraz dyfrakcyjny pojedynczego włosa, możemy precyzyjnie wyznaczyć jego grubość, korzystając bezpośrednio ze wzoru na dyfrakcję na szczelinie o tej samej szerokości.

7.9 Dyfrakcja na fali akustycznej

Bardzo ciekawym przykładem dyfrakcji jest ugięcie światła na fali akustycznej. W krysztale (lub cieczy) fala akustyczna wywołuje okresową, przestrzenną modulację gęstości ośrodka. Prowadzi to do powstania periodycznego rozkładu współczynnika załamania światła (zjawisko elastooptyczne), dzięki czemu ośrodek zaczyna działać jak objętościowa siatka dyfrakcyjna. Parametry tej siatki, takie jak stała siatki czy amplituda modulacji, mogą być precyzyjnie kontrolowane poprzez zmianę charakterystyki fali akustycznej (częstotliwości i mocy).

Urządzeniem wykorzystującym to zjawisko jest **modulator akustooptyczny** (ang. *Acousto-Optic Modulator*, AOM).

AOM pozwala nie tylko na zmianę kierunku propagacji wiązki, ale również na błyskawiczną kontrolę jej natężenia (poprzez zmianę mocy fali akustycznej) oraz częstotliwości. Dzięki tym właściwościom urządzenia te są nieodzownym elementem nowoczesnych eksperymentów w optyce kwantowej i fizyce atomowej, np. w procesach chłodzenia laserowego atomów.



Rysunek 7.24: Schemat działania modulatora akustooptycznego. Fala akustyczna wytwarzana przez przetwornik piezoelektryczny tworzy w krysztale siatkę dyfrakcyjną.

7.9.1 Ugięcie jako oddziaływanie fotonu z fononem

Działanie modulatora można intuicyjnie wytłumaczyć w języku fizyki kwantowej jako nieliniowe rozpraszanie, w którym dochodzi do oddziaływania fotonu z **fononem**. Fonon jest kwantem wzbudzenia sieci krystalicznej (pola akustycznego), analogicznym do fotonu w polu elektromagnetycznym.

Proces dyfrakcji światła w j-ty rząd polega na pochłonięciu (lub emisji wymuszonej) j fononów przez foton. Taka interpretacja pozwala natychmiast wyznaczyć parametry fali ugiętej. Częstość fali ω_j ugiętej w j-ty rząd wynika z zasady zachowania energii:

$$\omega_j = \omega + j\omega_{ak}, \tag{7.71}$$

gdzie ω jest częstością światła padającego, a ω_{ak} częstością fali akustycznej.

Zasada zachowania pędu determinuje natomiast kierunek fali ugiętej:

$$\hbar\mathbf{k}_j = \hbar\mathbf{k} + j\hbar\mathbf{k}_{ak} \implies \mathbf{k}_j = \mathbf{k} + j\mathbf{k}_{ak}, \tag{7.72}$$

gdzie $\mathbf{k}$ i $\mathbf{k}_{ak}$ są odpowiednio wektorami falowymi światła i dźwięku wewnątrz kryształu.

Powyższe równania prowadzą do dwóch istotnych wniosków. Po pierwsze, zmiana częstości światła o wielokrotność ω_{ak} pozwala na niezwykle precyzyjne przesuwanie linii widmowych lasera (częstość akustyczna $\sim 10^8$ Hz jest znikoma wobec optycznej $\sim 10^{14}$ Hz, ale istotna w spektroskopii). Po drugie, ponieważ wektor $\mathbf{k}_{ak}$ zależy od częstości dźwięku, każda zmiana modulacji zmienia kąt ugięcia wiązki. Aby wyeliminować ten niepożądany efekt w układach wymagających stabilnego kierunku, stosuje się **układy z podwójnym przejściem** (ang. *double-pass*). Wiązka po przejściu przez AOM jest zawracana przez lustro i przechodzi przez kryształ ponownie. Powoduje to sumaryczne podwojenie

przesunięcia częstości ($2\omega_{ak}$), ale kompensuje zmianę kąta, przywracając wiązce pierwotny kierunek.

7.9.2 Reżimy ugięcia: Raman-Nath oraz Bragg

W zależności od geometrii oddziaływania i grubości kryształu Δz, wyróżniamy dwa reżimy pracy:

- **Reżim Ramana-Natha**: Zachodzi, gdy fala świetlna przechodzi przez cienką warstwę fali akustycznej. Zjawisko to przypomina dyfrakcję na cienkiej siatce transmisyjnej – światło ugina się jednocześnie w wiele rzędów dyfrakcyjnych ($j = 0, \pm 1, \pm 2, \dots$).
- **Reżim Bragga**: Zachodzi w grubych kryształach, gdzie fala akustyczna tworzy strukturę objętościową. W tym przypadku, ze względu na konieczność spełnienia warunku konstruktywnej interferencji w całej objętości kryształu, światło ugina się wydajnie tylko w jeden, konkretny rząd dyfrakcyjny.

W reżimie Bragga optymalne ugięcie następuje, gdy światło pada pod charakterystycznym **kątem Bragga** θ_B. Ponieważ $\omega \gg \omega_{ak}$, możemy przyjąć, że długość wektora falowego światła prawie się nie zmienia ($k_j \approx k$). Z relacji wektorowej wynika wtedy:

$$\sin\theta_B = \frac{|j|\mathbf{k}_{ak}}{2k} = \frac{|j|\lambda}{2\lambda_{ak}}, \tag{7.73}$$

gdzie λ i λ_{ak} to długości fal w ośrodku. Całkowity kąt ugięcia wiązki względem wiązki padającej wynosi $2\theta_B$.

Do klasyfikacji reżimu pracy służy bezwymiarowy **parametr Kleina-Cooka** Q:

$$Q = \frac{k_{ak}^2 \Delta z}{k} = \frac{2\pi\lambda\Delta z}{\lambda_{ak}^2}. \tag{7.74}$$

Przyjmuje się, że dla $Q \ll 1$ mamy do czynienia z reżimem Ramana-Natha, natomiast dla $Q \gg 1$ (zazwyczaj $Q > 10$) modulator pracuje w reżimie Bragga.

7.10 Dyfrakcja światła białego

Choć nasze dotychczasowe rozważania dotyczyły głównie dyfrakcji światła monochromatycznego, warto na koniec zauważyć, że w przyrodzie powszechnie obserwujemy dyfrakcję światła białego. Ponieważ kąt ugięcia (lub położenie minimów i maksimów) jest wprost proporcjonalny do długości fali λ, każda składowa widmowa światła białego tworzy własny obraz dyfrakcyjny o innej skali przestrzennej.

Gdy światło białe przechodzi przez wąską szczelinę lub otwór, dochodzi do nakładania się tych obrazów, co skutkuje powstaniem barwnego widma dyfrakcyjnego. W samym centrum układu ($\theta = 0$) wszystkie barwy nakładają się konstruktywnie, tworząc jasny, biały prążek centralny. Jednak wraz z oddalaniem

się od osi optycznej, poszczególne maksima zaczynają się rozsuwać: składowa fioletowa i niebieska (krótsze fale) ulegają najmniejszemu ugięciu i pojawiają się najbliżej centrum, natomiast składowa czerwona (najdłuższa fala widzialna) ugięta jest najsilniej.

Zjawisko to odpowiada za szereg efektów optycznych obserwowanych w życiu codziennym:

- **Wieńce i iryzacja:** Kolorowe pierścienie wokół tarczy Słońca lub Księżyca powstające na skutek dyfrakcji na kropelkach wody lub kryształkach lodu w chmurach.
- **Aureole entoptyczne:** Efekt „kolorowych obwódek" wokół źródeł światła, wynikający z dyfrakcji na krawędziach pomocy optycznych (np. oprawek okularów) lub na strukturach wewnątrz gałki ocznej.
- **Diagnostyka:** Analiza obrazów dyfrakcyjnych światła białego jest wykorzystywana w okulistyce do badania przejrzystości ośrodków optycznych oka oraz w inżynierii materiałowej do badania mikrostruktur powierzchniowych.

7.11 Optyka fourierowska

Jak wykazano w sekcji 7.3.3, w reżimie pola dalekiego rozkład natężenia światła przestaje odwzorowywać bezpośrednio fizyczną geometrię przesłony, a staje się odzwierciedleniem jej widma częstotliwości przestrzennych. Związek ten stanowi fundament dziedziny nazywanej **optyką fourierowską**, która pozwala opisać zjawiska ugięcia światła za pomocą formalizmu przetwarzania sygnałów.

Aby ująć to zjawisko analitycznie, należy wprowadzić **zespoloną funkcję apertury** $\mathcal{A}(y, z)$. Funkcja ta opisuje całkowity rozkład amplitudy oraz opóźnienia fazowe fali świetlnej w płaszczyźnie przesłony. W przybliżeniu Fraunhofera kąty ugięcia fal płaskich mapują się na składowe wektora częstotliwości przestrzennych:

$$\begin{aligned} k_Y &= k \sin\theta_y, \\ k_Z &= k \sin\theta_z, \end{aligned} \tag{7.75}$$

gdzie k to liczba falowa światła padającego. Podstawienie tych zmiennych do całki dyfrakcyjnej pozwala wyrazić zespoloną amplitudę fali na odległym ekranie w postaci:

$$E(k_Y, k_Z) = C \iint_{-\infty}^{\infty} \mathcal{A}(y, z) \mathrm{e}^{-\mathrm{i}(k_Y y + k_Z z)} \mathrm{d}y \mathrm{d}z. \tag{7.76}$$

Zespolona stała C zawiera czynniki fazowe związane z propagacją, a granice całkowania można rozszerzyć do nieskończoności, ponieważ funkcja $\mathcal{A}(y, z)$ znika tożsamościowo poza obszarem szczeliny.

Powyższe równanie przyjmuje postać klasycznej, dwuwymiarowej transformaty Fouriera funkcji apertury. Rozkład pola elektrycznego w obrazie dyfrakcyjnym Fraunhofera jest zatem wprost dokładnym widmem częstotliwości przestrzennych analizowanego obiektu, tj. $E \propto \mathcal{F}\{\mathcal{A}\}$.

Bezpośrednia rejestracja obrazu w idealnym polu dalekim jest w laboratoriach zazwyczaj niepraktyczna ze względu na wymagane ogromne dystanse propagacji. Właściwość transformująca dyfrakcji Fraunhofera jest jednak w pełni zachowana, gdy zastosujemy soczewkę skupiającą. Jeśli oświetlimy obiekt spójną falą płaską, a następnie umieścimy go dokładnie w **przedniej płaszczyźnie ogniskowej** soczewki (w odległości f przed nią), to w jej tylnej płaszczyźnie ogniskowej (tzw. płaszczyźnie transformaty) powstanie fizyczny, ostry i pozbawiony zniekształceń fazowych rozkład dyfrakcyjny Fraunhofera.

Soczewka pełni w tym układzie rolę analogowego komputera, który natychmiastowo oblicza dwuwymiarową transformatę Fouriera rozkładu światła z płaszczyzny obiektu. Zjawisko to otwiera drogę do tzw. filtracji przestrzennej (realizowanej np. w układzie 4f). Wprowadzając mechaniczne przysłony w płaszczyźnie transformaty (na przykład blokując wysokie częstotliwości przestrzenne odpowiadające za ostry szum optyczny), można w kontrolowany sposób modyfikować widmo sygnału optycznego, a po ponownej transformacji uzyskać przefiltrowany obraz końcowy.

Dygresja

Jednym z najnowocześniejszych i najbardziej fascynujących zastosowań optyki fourierowskiej są fotoniczne sieci neuronowe. Architektury te stanowią sprzętową realizację obliczeń neuromorficznych, naśladując działanie ludzkiego mózgu za pomocą zjawisk falowych.

W klasycznych, elektronicznych sieciach neuronowych, wykorzystywanych powszechnie do analizy i rozpoznawania obrazów, operacja splotu dwóch wielowymiarowych macierzy jest niezwykle kosztowna obliczeniowo i stanowi wąskie gardło pochłaniające ogromne ilości energii. Tymczasem fizyka falowa pozwala na wykonanie tej samej operacji całkowicie pasywnie i niemal natychmiastowo – z prędkością propagacji światła.

Podstawą tego procesu jest optyczna realizacja twierdzenia o splocie, wykorzystująca omawiany wcześniej układ $4f$. Pierwsza soczewka wykonuje dwuwymiarową transformatę Fouriera wejściowego obrazu (lub rozkładu danych). W płaszczyźnie transformaty umieszcza się przestrzenny modulator światła, będący programowalną maskę fazową dla światła. Element ten pełni rolę fizycznego filtru, który sprzętowo realizuje mnożenie widma przestrzennego sygnału przez odpowiednie wagi sieci (jądro splotu). Następnie druga soczewka dokonuje odwrotnej transformaty Fouriera tak zmodyfikowanego pola optycznego, a wynik rejestrowany jest na matrycy detektorów.

Dzięki wykorzystaniu zjawisk dyfrakcyjnych oraz wrodzonego, potężnego zrównoleglenia propagacji fal świetlnych, układ optyczny ogromnie

redukuje złożoność obliczeniową konwolucji. Taka fourierowska akceleracja otwiera drogę do budowy hybrydowych procesorów sztucznej inteligencji, które osiągają drastycznie wyższą energooszczędność i szybkość niż tradycyjne układy krzemowe.

7.12 Podsumowanie rozdziału

Niniejszy rozdział przedstawił ewolucję opisu dyfrakcji, począwszy od zasady Huygensa-Fresnela, traktującej front falowy jako zbiór źródeł fal wtórnych, aż po rygorystyczne sformułowanie całkowe Kirchhoffa. Zależnie od odległości od przeszkody, zjawisko to analizujemy w dwóch reżimach. W polu bliskim (dyfrakcja Fresnela) obraz zachowuje widoczny związek z geometrycznym kształtem obiektu, co dogodnie opisuje się za pomocą metody stref Fresnela. W polu dalekim (dyfrakcja Fraunhofera) obraz staje się przestrzenną transformatą Fouriera apertury, ukazując rozkład jej częstotliwości przestrzennych. W sytuacjach, w których analiza matematyczna skomplikowanej przeszkody jest zbyt złożona, z pomocą przychodzi zasada Babineta, pozwalająca wyznaczyć rozkład pola w oparciu o łatwiejsze do opisania przesłony komplementarne.

Zjawisko dyfrakcji determinuje fundamentalne granice wszystkich układów optycznych. Zgodnie z kryterium Rayleigha, ugięcie światła na skończonej aperturze obiektywu prowadzi do przestrzennego rozmycia obrazu punktu (powstania dysku Airy'ego), narzucając nieprzekraczalną granicę rozdzielczości przestrzennej. Te same zasady interferencji i ugięcia fal można jednak wykorzystać konstruktywnie – tak jak w przypadku siatek dyfrakcyjnych, gdzie nałożenie fal z wielu szczelin pozwala na niezwykle precyzyjną analizę widmową promieniowania. Zrozumienie tych mechanizmów jest w związku z tym niezbędne zarówno do określania fizycznych limitów obrazowania mikroskopowego i teleskopowego, jak i do projektowania zaawansowanych instrumentów spektroskopowych.

Zadania

7.1. Równoległa wiązka światła monochromatycznego o długości fali $\lambda = 550$ nm pada na przesłonę, w której wycięto nieskończenie długą szczelinę. Prążki dyfrakcyjne, odległe od siebie o 0,5 mm, są obserwowane na ekranie umieszczonym prostopadle do wiązki w odległości 1 m za szczeliną. Oblicz szerokość szczeliny.

7.2. Równoległa wiązka skolimowanego światła o długości fali $\lambda = 632$ nm pada na nieprzezroczysty ekran, w którym wycięto nieskończenie długą szczelinę. Za ekranem powstaje obraz dyfrakcyjny.

(a) Zakładając, że 10. ciemny prążek obserwowany jest pod kątem 3,1° od osi optycznej, oblicz szerokość szczeliny.

(b) Pod jakim kątem będzie obserwowany ten sam prążek, jeśli cały układ zanurzymy w wodzie ($n_w = 1{,}33$)?

7.3. Wykorzystując zasadę Babineta, udowodnij, że rozkład natężenia światła ugiętego w polu dalekim na nieprzezroczystym cienkim drucie o szerokości D jest identyczny z rozkładem natężenia od szczeliny o tej samej szerokości D, z wyjątkiem nieskończenie małego obszaru w okolicach kąta $\theta = 0$. Następnie oblicz grubość ludzkiego włosa, na którym dyfrakcja światła lasera ($\lambda = 532$ nm) tworzy na ekranie odległym o 2 m minima dyfrakcyjne oddalone od siebie o 1,5 cm.

7.4. Dwie długie szczeliny o szerokościach 0,10 mm w nieprzezroczystym ekranie, oddalone od siebie o 0,20 mm, są oświetlone prostopadle światłem o długości fali $\lambda = 500$ nm. Ile prążków interferencyjnych Younga będzie dostrzegalnych w obrębie centralnego jasnego pasma dyfrakcyjnego (dysku zerowego rzędu)?

7.5. Przypuśćmy, że w doświadczeniu z dwiema szczelinami pod obwiednią obrazu dyfrakcyjnego (wewnątrz centralnego maksimum) znajduje się 11 jasnych prążków interferencyjnych. Wiadomo również, że pierwsze minima dyfrakcyjne dokładnie tłumią kolejne jasne prążki interferencyjne, gdyż wypadają w tym samym miejscu. Ile jasnych prążków interferencyjnych znajduje się pod obwiednią dyfrakcyjną między pierwszym i drugim minimum dyfrakcyjnym?

7.6. Wiązka laserowa ($\lambda = 632{,}8$ nm) pada normalnie na prostokątną aperturę o wymiarach $D_x = 0{,}1$ mm oraz $D_y = 0{,}5$ mm. Ekran znajduje się w odległości $L = 4$ m od apertury.

(a) Oblicz wymiary geometryczne (szerokość i wysokość) centralnego jasnego maksimum dyfrakcyjnego na ekranie.

(b) W którym kierunku wiązka rozbiega się silniej?

7.7. Proszę pokazać, że dyfrakcja Fraunhofera w polu dalekim dla apertury kołowej o promieniu a, oświetlonej jednorodnie (fala płaska padająca normalnie), prowadzi do rozkładu amplitudy pola:

$$E(r,z) = E_0^{(+)} 2\pi a^2 \frac{J_1(2\pi ar/\lambda z)}{2\pi ar/\lambda z},$$

gdzie $E_0^{(+)}$ jest amplitudą fali padającej, r jest odległością od środka ekranu, a $J_1(x)$ jest funkcją Bessela pierwszego rodzaju i rzędu pierwszego, zdefiniowaną całkową relacją:

$$J_n(x) = \frac{1}{\pi}\int_0^{\pi} \cos(n\phi - x\sin\phi)\,\mathrm{d}\phi.$$

7.8. Laser emituje promieniowanie o długości fali $\lambda = 632,8$ nm. Wiązka światła ma początkową średnicę 2 mm. Zakładając, że rozbieżność wiązki jest ograniczona wyłącznie dyfrakcyjnie, wyznacz liniowy rozmiar plamki laserowej na powierzchni Księżyca. Dla uproszczenia zaniedbaj wpływ atmosfery.

7.9. Georges Seurat należał do grupy malarzy pointylistów, których obrazy składały się z wielkiej liczby malutkich punktów (o średnicy ok. 6 mm) czystej farby ułożonych blisko siebie. Wrażenie odcienia barwy powstaje w oku obserwatora na skutek zjawisk dyfrakcyjnych. Zakładając, że średnica źrenicy oka wynosi 4 mm, a efektywna długość fali dla światła widzialnego to $\lambda = 550$ nm, oblicz, z jakiej minimalnej odległości należy podziwiać obraz, aby osiągnąć efekt całkowitego zlewania się sąsiadujących punktów.

7.10. Na podstawie kryterium Rayleigha oceń, czy orzeł latający na wysokości $h = 1$ km może wyraźnie rozpoznać kształt myszy o długości 2 cm, czy też jest w stanie jedynie zauważyć jej obecność jako nierozdzielonego punktu? Przyjmij średnicę źrenicy ptaka równą 6 mm oraz długość fali $\lambda = 550$ nm.

7.11. Samochód z zapalonymi światłami przednimi zbliża się po prostej drodze. Światła są umieszczone w odległości 120 cm od siebie. Jak daleko od obserwatora znajdował się samochód w momencie, gdy obserwator zorientował się, że widzi dwa światła, a nie jedno? Przyjmij, że źrenica oka ma średnicę 5 mm, a efektywna długość fali wynosi 550 nm.

7.12. Znajdź najmniejszą odległość dwóch punktów na powierzchni Księżyca, które mogą być ledwie rozdzielone przez $5,1$-metrowy teleskop zwierciadlany. Przyjmij długość fali $\lambda = 550$ nm i załóż, że rozdzielczość jest ograniczona wyłącznie przez efekty dyfrakcyjne. Odległość Ziemi od Księżyca wynosi 384 000 km.

7.13. System radarowy krążownika wysyła falę o długości $1,6$ cm z kolistej anteny o średnicy $2,3$ m. Jaka może być najmniejsza odległość między dwiema łodziami motorowymi, żeby system radarowy był w stanie z odległości $6,2$ km rozróżnić je jako pojedyncze obiekty?

7.14. Dwie szczeliny zostały jednocześnie oświetlone dwiema wiązkami światła o różnej barwie. Siódmy prążek dyfrakcyjny dla światła o długości fali $\lambda = 555$ nm widoczny jest na ekranie odległym od szczelin o 5 m w odległości 9,5 cm od jego środka. Oblicz długość fali drugiej wiązki, jeśli w tym samym miejscu widoczny jest jej piąty prążek, oraz wyznacz odległość między szczelinami.

7.15. Siatka dyfrakcyjna posiada 400 rys/mm. Światło białe pokrywa zakres widzialny $\lambda \in [400, 700]$ nm.

(a) Ile pełnych rzędów widma, oprócz rzędu zerowego, można teoretycznie zaobserwować przy pomocy tej siatki?

(b) Czy widma poszczególnych rzędów będą się ze sobą przekrywać? Jeśli tak, wskaż, które rzędy i dla jakich długości fal ulegają nałożeniu.

7.16. Źródło światła zawierające mieszaninę wodoru i deuteru emituje dwie bliskie sobie długości fali światła czerwonego. Centralna długość fali to

656,3 nm, a różnica między nimi wynosi $\Delta\lambda = 0,180$ nm. Oblicz minimalną wymaganą liczbę rys w siatce dyfrakcyjnej, niezbędną do rozdzielenia tych linii w pierwszym rzędzie widma. Jaką minimalną szerokość musiałaby mieć ta siatka, gdyby gęstość nacięć wynosiła 600 rys/mm?

7.17. Zgodnie z teorią optyki fourierowskiej, pole dalekie jest transformatą Fouriera funkcji transmitancji apertury. Oblicz i naszkicuj rozkład natężenia światła na ekranie w reżimie Fraunhofera po przejściu płaskiej fali monochromatycznej przez cienką siatkę amplitudową, której transmitancja zmienia się w jednym kierunku harmonicznie, zgodnie ze wzorem:

$$t(y) = \frac{1}{2}\left[1 + \cos\left(\frac{2\pi}{d}y\right)\right].$$

7.18. Na siatkę dyfrakcyjną o gęstości 100 rys/mm pada prostopadle równoległa wiązka światła białego (od 400 nm do 700 nm). Za pomocą soczewki skupiającej o ogniskowej $f = 30$ cm umieszczonej tuż za siatką, wytwarzane jest widmo w płaszczyźnie ogniskowej. Oblicz odległość liniową na ekranie między:

(a) czerwonym i fioletowym prążkiem w widmie drugiego rzędu,

(b) końcem widma pierwszego rzędu a początkiem widma drugiego rzędu.

7.19. Na ekran z dwiema równoległymi szczelinami padają bezpośrednio promienie słoneczne. Średnica kątowa Słońca widocznego z Ziemi wynosi $\alpha \approx 0,01$ rad. Zakładając, że tarczę Słońca można zastąpić jednorodnym kwadratem o stałej jasności powierzchniowej, oblicz, dla jakiej maksymalnej odległości D między szczelinami można jeszcze zaobserwować wyraźne prążki interferencyjne.

Rozdział 8

Polaryzacja światła

Doświadczenie
Dipol Hertza

Jak wiemy z rozdziału 1, jednym z kluczowych parametrów opisujących światło jako poprzeczną falę elektromagnetyczną jest jego polaryzacja, określająca przestrzenno-czasową ewolucję drgań wektora pola elektrycznego. Choć trzy podstawowe klasy — **polaryzację liniową**, **kołową** i **eliptyczną** — omówiliśmy już jakościowo w rozdziale 1.5.1, to rygorystyczne opisywanie zjawisk takich jak dwójłomność czy elementów optycznych takich jak modulatory optyczne wymaga przejścia na język ścisłego opisu ilościowego. W tym celu wprowadzimy poniżej użyteczne formalizmy matematyczne, stanowiące fundament współczesnej optyki wektorowej.

8.1 Formalizm opisu stanu polaryzacji

8.1.1 Formalizm Jonesa

Podstawowym aparatem matematycznym służącym do opisu światła o dobrze określonym stanie polaryzacji jest **formalizm wektorów Jonesa**, wprowadzony przez Roberta Clarka Jonesa (1916–2004). Punktem wyjścia dla tego formalizmu jest możliwość zapisu dowolnej spolaryzowanej płaskiej fali elektromagnetycznej w postaci superpozycji dwóch ortogonalnych składowych liniowych [równanie (1.27)]. Ponieważ drgania pola elektrycznego wzdłuż obu kierunków osi poprzecznych można rozważać niezależnie, a obie składowe oscylują z tą samą częstością kołową, pełna informacja o stanie polaryzacji fali zostaje jednoznacznie zakodowana w dwuelementowym zespolonym wektorze kolumnowym, zwanym **wektorem Jonesa**:

$$\mathbf{E}_0^{(+)} = \begin{pmatrix} E_{0x} e^{i\varphi_x} \\ E_{0y} e^{i\varphi_y} \end{pmatrix}, \tag{8.1}$$

gdzie E_{0x} i E_{0y} są rzeczywistymi (dlatego bez $^{(+)}$) amplitudami składowej światła wzdłuż kierunku x i y, zaś φ_x i φ_y są fazami początkowymi każdej z tych składowych. Wygodną praktyką jest stosowanie wektorów znormalizowanych,

w których całkowite natężenie światła zostaje wyciągnięte przed nawias jako czynnik skalujący. Poniżej przedstawiono znormalizowane wektory Jonesa dla kilku przykładowych stanów polaryzacji:

- polaryzacja liniowa pozioma (wzdłuż osi x): $\begin{pmatrix}1\\0\end{pmatrix}$,
- polaryzacja liniowa obrócona o kąt $\pi/4$ względem osi x: $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$,
- polaryzacja kołowa prawoskrętna: $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-i\end{pmatrix}$,
- polaryzacja kołowa lewoskrętna: $\frac{1}{\sqrt{2}}\begin{pmatrix}1\\i\end{pmatrix}$.

Oprócz opisu samych stanów polaryzacji, formalizm Jonesa pozwala na eleganckie i rygorystyczne modelowanie transformacji fali podczas jej przejścia przez ośrodki optycznie anizotropowe (fizyczny opis zjawiska dwójłomności znajduje się w rozdziale 8.2.4). Przekształcenie to opisuje się za pomocą zespolonych macierzy kwadratowych wymiaru 2×2, zwanych **macierzami Jonesa**. Ich ogólna postać reprezentowana jest jako:

$$\mathsf{J} = \begin{pmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{pmatrix}, \tag{8.2}$$

gdzie współczynniki j_{ij} przyjmują określone wartości w zależności od konkretnego typu elementu optycznego i jego orientacji przestrzennej. Wiele elementów optycznych może być opisywanych przy pomocy macierzy Jonesa. Przykładami mogą być:

- idealny polaryzator liniowy zorientowany wzdłuż kierunku x: $\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$,
- idealny polaryzator liniowy obrócony o kąt $\pi/4$ względem osi x: $\frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$,
- płytka ćwierćfalowa o osi szybkiej zorientowanej wzdłuż osi x: $e^{i\pi/4}\begin{pmatrix}1 & 0\\0 & -i\end{pmatrix}$,

gdzie zasada działania samych elementów optycznych opisana jest w dalszej części tego rozdziału. Działając macierzą Jonesa J odpowiadającą określonemu elementowi optycznemu z lewej strony na początkowy wektor Jonesa fali padającej, otrzymujemy wektor Jonesa reprezentujący stan polaryzacji światła po przejściu przez ten element:

$$\begin{pmatrix} E'_{0x}e^{i\varphi'_x} \\ E'_{0y}e^{i\varphi'_y} \end{pmatrix} = \mathsf{J}\begin{pmatrix} E_{0x}e^{i\varphi_x} \\ E_{0y}e^{i\varphi_y} \end{pmatrix}. \tag{8.3}$$

Macierze Jonesa mogą być również zastosowane do opisu zmian polaryzacji zachodzących przy odbiciu i transmisji na granicy dwóch ośrodków dielektrycznych. W modelu tym zakłada się, że fali poruszają się w przybliżeniu wzdłuż osi z niezależnie od kątów padania i załamania (efekty geometryczne są w pełni uwzględnione w wartościach współczynników Fresnela). Przyjmując, że współrzędna x reprezentuje składową pola prostopadłą do płaszczyzny padania (polaryzacja s), a współrzędna y składową równoległą (polaryzacja p), wektory składowe odpowiadają bezpośrednio współrzędnym kartezjańskim. Zgodnie z tą bazą macierze Jonesa dla odbicia i transmisji przyjmują spójną postać diagonalną:

- dla wiązki odbitej: $\mathsf{J}_r = \begin{pmatrix} r_\perp & 0 \\ 0 & -r_\parallel \end{pmatrix}$,
- dla wiązki załamanej: $\mathsf{J}_t = \begin{pmatrix} t_\perp & 0 \\ 0 & t_\parallel \end{pmatrix}$,

gdzie $r_\parallel$ i $r_\perp$ są amplitudowymi współczynnikami odbicia Fresnela, natomiast $t_\parallel$ i $t_\perp$ są współczynnikami transmisji. Znak minus przy składowej równoległej dla wiązki odbitej wynika z przyjętej konwencji zwrotów wektorów pola elektrycznego przy zmianie kierunku propagacji fali.

Kolejną istotną zaletą formalizmu Jonesa jest możliwość kaskadowego łączenia elementów optycznych. Stan fali na wyjściu układu składającego się z wielu elementów wyznacza się poprzez sukcesywne mnożenie początkowego wektora Jonesa przez macierze poszczególnych komponentów. Operację tę wykonuje się od prawej do lewej, zgodnie z kierunkiem propagacji fali:

$$\begin{pmatrix} E'_{0x} e^{i\varphi'_x} \\ E'_{0y} e^{i\varphi'_y} \end{pmatrix} = \mathsf{J}_N \mathsf{J}_{N-1} \dots \mathsf{J}_2 \mathsf{J}_1 \begin{pmatrix} E_{0x} e^{i\varphi_x} \\ E_{0y} e^{i\varphi_y} \end{pmatrix}, \tag{8.4}$$

gdzie J_i jest macierzą Jonesa reprezentującą i-ty element optyczny w układzie.

Warto pamiętać, że formalizm Jonesa opiera się na zasadzie superpozycji pól elektrycznych. W tym kontekście formalizm Jonesa zakłada, że fale są w pełni spójne i całkowicie spolaryzowane. W rzeczywistych układach fizycznych mamy jednak często do czynienia ze **światłem częściowo spolaryzowanym** lub **niespolaryzowanym**, w którym faza i kierunek drgań pola elektrycznego ulegają szybkim, stochastycznym fluktuacjom. Ponieważ nakładające się fale niekoherentne nie zachowują stałej relacji fazowej, ich superpozycja nie polega na dodawaniu amplitud, lecz na sumowaniu średnich natężeń. W takich sytuacjach formalizm Jonesa traci swoją użyteczność i konieczne jest zastosowanie formalizmu Stokesa.

8.1.2 Formalizm Stokesa

W **formalizmie Stokesa** stan polaryzacji światła opisujemy za pomocą czteroelementowego wektora Stokesa:

$$\mathbf{S}_S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_0 \\ I_x - I_y \\ I_{45^\circ} - I_{-45^\circ} \\ I_l - I_p \end{pmatrix}, \tag{8.5}$$

gdzie I_0 oznacza całkowite natężenie światła, I_x oraz I_y to natężenia składowych o polaryzacji liniowej wzdłuż osi x i y, I_{45° i I_{-45° odpowiadają natężeniom światła zmierzonym pod kątem 45° oraz -45° względem osi x, zaś I_p i I_l to natężenia składowej kołowej prawoskrętnej i lewoskrętnej. Główną zaletą formalizmu Stokesa jest jego uniwersalność — pozwala on na równoprawny opis zarówno światła całkowicie spolaryzowanego, częściowo spolaryzowanego, jak i całkowicie niespolaryzowanego. W tym drugim przypadku wektor pola elektrycznego wykazuje stochastyczne, nieuporządkowane fluktuacje, przez co faza i kierunek drgań ulegają losowym zmianom w czasie.

Do ilościowego określenia stopnia uporządkowania fali służy wielkość zwana **stopniem polaryzacji**, zdefiniowana jako:

$$V_p = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \tag{8.6}$$

gdzie $I_{\max}$ i $I_{\min}$ oznaczają odpowiednio maksymalne i minimalne natężenie światła rejestrowane po przejściu fali przez idealny polaryzator liniowy podczas pełnego obrotu jego osi anizotropii wokół kierunku propagacji światła.

Przeanalizujmy, przez analogię do formalizmu Jonesa, wektory Stokesa reprezentujące podstawowe stany polaryzacji fali (dla jednostkowego natężenia całkowitego):

- światło całkowicie spolaryzowane wzdłuż osi x: $\mathbf{S}_S = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$,
- światło całkowicie spolaryzowane pod kątem $\pi/4$: $\mathbf{S}_S = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$,
- światło całkowicie spolaryzowane kołowo prawoskrętnie: $\mathbf{S}_S = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$,

- światło niespolaryzowane: $\mathbf{S}_S = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Analogicznie do formalizmu Jonesa, również w formalizmie Stokesa można wprowadzić macierze, które pozwalają określić zmiany stanu polaryzacji światła przy przejściu przez określony ośrodek dwójłomny. Macierze te, zwane **macierzami Müllera**, można zapisać w ogólnej formie jako:

$$\mathsf{M}_M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix}. \tag{8.7}$$

Przejście światła przez element optyczny opisane jest wówczas równaniem:

$$\mathbf{S}'_S = \mathsf{M}_M \mathbf{S}_S. \tag{8.8}$$

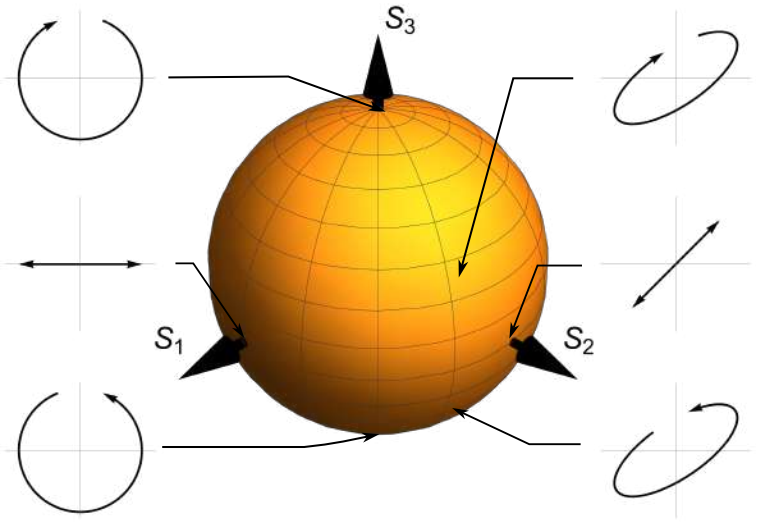
Poniżej przedstawiono przykłady kilku macierzy Müllera M_M:

- idealny polaryzator liniowy o osi ustawionej wzdłuż x: $\frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,
- idealny polaryzator liniowy o osi ustawionej pod kątem $\pi/4$: $\frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,
- idealna (bezstratna) płytka ćwierćfalowa o osi szybkiej zorientowanej wzdłuż osi x: $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$.

Jak widać, formalizm Müllera jest ogólniejszy od formalizmu Jonesa i może być z powodzeniem wykorzystywany nie tylko do opisu przejścia światła przez ośrodki dwójłomne, ale również przez takie, w których zachodzi rozpraszanie i stopień polaryzacji światła ulega pogorszeniu.

8.1.3 Sfera Poincarégo

Ciekawym sposobem przedstawiania stanu polaryzacji światła jest wykorzystanie tzw. **sfery Poincarégo** (Henri Poincaré 1854–1912). Taka geometryczna reprezentacja jest możliwa, ponieważ po zaniedbaniu parametru S_0 (całkowitego natężenia), trzy pozostałe parametry Stokesa (S_1, S_2, S_3) mogą pełnić rolę współrzędnych kartezjańskich determinujących punkt na sferze w przestrzeni



Rysunek 8.1: Sfera Poincarégo.

$\mathbb{R}^3$. W przypadku sfery Poincarégo bieguny odpowiadają ortogonalnym polaryzacjom kołowym, podczas gdy punkty znajdujące się na równiku odpowiadają polaryzacjom liniowym. Co więcej, każdy punkt znajdujący się między równikiem a biegunem odpowiada określonej polaryzacji eliptycznej (rys. 8.1). W konsekwencji każdemu stanowi polaryzacji światła przyporządkowane jest tylko jedno miejsce na sferze Poincarégo.

Sfera Poincarégo jest wygodnym formalizmem również dlatego, że przejście przez określone elementy optyczne wiąże się z konkretną transformacją punktu na sferze. W szczególności przejście światła przez tzw. płytki fazowe, tj. elementy optyczne opóźniające składową pola elektrycznego wzdłuż danego kierunku względem składowej prostopadłej o zadaną fazę, tożsame jest z obrotem na sferze o zadany kąt wokół określonego kierunku (kąt i kierunek determinowane są przez typ płytki oraz orientację przestrzenną jej osi). W ten sposób transformacja polaryzacji przy przejściu przez określone elementy optyczne zyskuje prostą i intuicyjną interpretację geometryczną.

Dygresja

Geometryczne właściwości obrotów grupy $\mathsf{SO}(3)$ prowadzą do fascynującego wniosku: przejście między dwoma dowolnymi punktami na sferze wymaga zaledwie trzech obrotów o ściśle określone, niezmienne kąty, w których kontrolujemy jedynie kierunki osi obrotu. Ponieważ każdy obrót na sferze Poincarégo odpowiada fizycznemu przejściu fali przez odpowiednią płytkę

fazową, oznacza to, że dowolny stan polaryzacji możemy przekształcić w każdy inny przy użyciu układu złożonego z zaledwie trzech takich elementów. W praktyce laboratoryjnej tę uniwersalną zdolność do manipulacji polaryzacją uzyskujemy poprzez kaskadowe zestawienie płytki ćwierćfalowej, płytki półfalowej i kolejnej płytki ćwierćfalowej. Ten niezwykle elegancki układ znajduje szerokie zastosowanie zarówno w precyzyjnych eksperymentach optycznych, jak i w telekomunikacji światłowodowej.

8.2 Sposoby polaryzacji światła

Światło możemy spolaryzować, wykorzystując różne zjawiska fizyczne towarzyszące jego oddziaływaniu z materią. Do najważniejszych z nich należą: odbicie od granic dielektryków, rozproszenie, selektywna absorpcja (dichroizm) oraz dwójłomność w ośrodkach anizotropowych. Poniżej szczegółowo analizujemy fizyczne mechanizmy leżące u podstaw każdego z tych procesów.

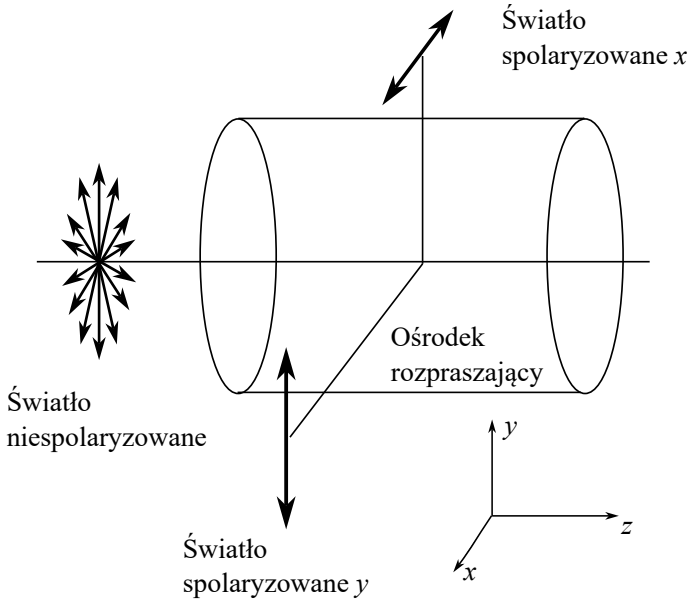
8.2.1 Polaryzacja przez odbicie

Doświadczenie
Stos Arago

Jak wykazaliśmy w rozdziale 4.2, na granicy dwóch ośrodków dielektrycznych dochodzi do odbicia i załamania fali świetlnej. Amplitudowe współczynniki odbicia określają odpowiednio równania (4.27) i (4.31). Dla ściśle określonego kąta padania, zwanego kątem Brewstera, współczynnik odbicia składowej równoległej (p) maleje do absolutnego zera. W tym reżimie promień odbity staje się całkowicie spolaryzowany prostopadle do płaszczyzny padania (składowa s), co pozwala na wykorzystanie zjawiska odbicia jako prostego polaryzatora liniowego.

Analizując możliwość polaryzacji światła przez odbicie, warto zwrócić uwagę na dwa istotne ograniczenia. Po pierwsze, natężenie tak spolaryzowanego światła odbitego stanowi stosunkowo niewielką część całkowitego światła padającego. Przykładowo, dla światła padającego pod kątem Brewstera na granicę szkło-powietrze, odbiciu ulega zaledwie około 15% natężenia składowej s, przy jednoczesnej całkowitej eliminacji składowej p. Po drugie, wiązka załamana (przechodząca) ulega jedynie częściowej polaryzacji. Ponieważ zawiera ona w 100% składową p oraz wciąż znaczną część składowej s (tej, która nie uległa odbiciu), nigdy nie osiąga polaryzacji całkowitej przy przejściu przez pojedynczą granicę ośrodków. Z tych względów technika pojedynczego odbicia rzadko służy jako samodzielne, wydajne źródło światła spolaryzowanego w zaawansowanych układach optycznych.

W celu drastycznego zwiększenia stopnia polaryzacji wiązki przechodzącej możemy zastosować układ kaskadowy złożony z wielu równoległych powierzchni granicznych, np. stos cienkich szkiełek, zwany **stosem Arago**. Idea jego działania opiera się na sukcesywnym „odfiltrowywaniu” niepożądanej polaryzacji. Ponieważ każda kolejna powierzchnia przepuszcza 100% składowej równoległej (p) i jednocześnie odbija (usuwa z głównej wiązki) część składowej prostopadłej (s), wiązka przechodząca przez cały stos staje się w coraz większym stopniu spo-



Rysunek 8.2: Emisja światła spolaryzowanego podczas rozpraszania światła niespolaryzowanego w ośrodku składającym się z dipoli elektrycznych.

laryzowana liniowo w płaszczyźnie padania (dominacja składowej p). Z uwagi na to, że dla odpowiednio grubych płytek kolejne wiązki nie nakładają się na siebie spójnie, możemy całkowicie pominąć zjawisko interferencji. Metoda ta jest bardzo tania i prosta w konstrukcji, przez co, mimo strat natężenia, znajduje szerokie zastosowanie w polaryzacyjnych dzielnikach wiązki.

8.2.2 Polaryzacja przez rozproszenie

Doświadczenie Rozpraszanie światła w zawiesinie w akwarium

W rozdziale 9.2.4 omawiamy emisję fali przez drgający dipol elektryczny. Wyprzedzając jednak nieco te rozważania, możemy zauważyć, że emisja promieniowania z dipola jest symetryczna wokół kierunku jego drgań oraz że dipol nie emituje promieniowania wzdłuż swojej osi. Zakładając, że pole elektryczne fali drga jedynie w kierunku prostopadłym do kierunku propagacji, światło rozproszone będzie spolaryzowane liniowo pod kątem prostym względem kierunku padania, a kierunek drgań jego pola elektrycznego będzie prostopadły do płaszczyzny rozpraszania (rys. 8.2).

Dygresja

Przykładem obiektu, który rozprasza światło jest niebo (do rozproszenia dochodzi na skutek fluktuacji gęstości atmosfery). Obserwując niebo przy pomocy polaryzatora, można dostrzec, że fragmenty nieba obserwowane

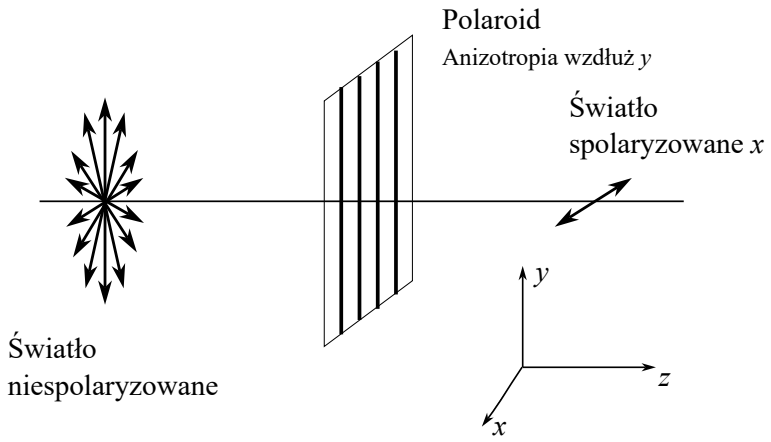
w kierunku prostopadłym do kierunku padania promieni słonecznych są silnie spolaryzowane. Najlepiej efekt ten widać, obserwując niebo tuż nad horyzontem. W takim przypadku obrót polaryzatora powoduje wyraźną zmianę intensywności światła transmitowanego przez polaryzator.

8.2.3 Polaryzacja przez selektywną absorpcję

Doświadczenie
Polaroid

Niektóre materiały charakteryzują się silną anizotropią budowy mikroskopowej, wynikającą na przykład z obecności długich i uporządkowanych przestrzennie cząsteczek. Prowadzi to do zjawiska selektywnej, kierunkowej absorpcji fal świetlnych, zwanego **dichroizmem**. Fizyczny mechanizm tego procesu opiera się na zróżnicowanym oddziaływaniu wektora pola elektrycznego z elektronami zdolnymi do ruchu wzdłuż różnych kierunków w materiale.

Jeśli pole elektryczne fali oscyluje równolegle do wydłużonej osi cząsteczek, wprawia ono elektrony w drgania o dużej amplitudzie. Te drgające ładunki intensywnie oddziałują z siecią materiału, co skutkuje konwersją energii fali na ciepło i jej silnym tłumieniem (absorpcją). Z kolei składowa pola elektrycznego drgająca w kierunku prostopadłym do osi molekuł w bardzo małym stopniu wzbudza ruch ładunków i dzięki temu przechodzi przez ośrodek bez znaczących strat.



Rysunek 8.3: Polaryzacja światła przy pomocy zjawiska dichroizmu. Składowa pola elektrycznego równoległa do osi molekuł jest absorbowana, a składowa prostopadła — transmitowana.

Opisany powyżej mechanizm stanowi podstawę działania popularnych polimerowych filtrów polaryzacyjnych, zwanych **polaroidami**. Cząsteczki polimeru (np. alkoholu poliwinylowego) są w nich mechanicznie rozciągane, aby uzyskać ich równoległe ułożenie, a następnie impregnowane jodem. Atomy jodu przyczepiają się do długich łańcuchów polimeru, tworząc mikroskopijne ścieżki przewodzące. Jak przedstawiono na rys. 8.3, oś przepuszczania takiego polaryzatora

jest zorientowana prostopadle do kierunku ułożenia łańcuchów, co pozwala na transmisję składowej pola elektrycznego drgającej „pomiędzy" nimi.

8.2.4 Polaryzacja przez dwójłomność

Doświadczenie
Rzutnik pisma i obiekty między polaryzatorami (kalcyt, folia)

Podobnie jak współczynnik absorpcji, w **ośrodkach anizotropowych** również współczynnik załamania nie jest wielkością izotropową. Oznacza to, że prędkość fali świetlnej wewnątrz takiego materiału zależy bezpośrednio od kierunku drgań jej pola elektrycznego. Zjawisko to nosi nazwę **dwójłomności** i jest powszechnie wykorzystywane do generowania światła o bardzo wysokim stopniu polaryzacji liniowej.

Rozważmy propagację monochromatycznej fali płaskiej w krysztale anizotropowym. Wektory pól $\mathbf{E}$ i $\mathbf{B}$ dane są standardowymi zależnościami harmonicznymi:

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \\ \mathbf{B} &= \mathbf{B}_0^{(+)} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \end{aligned} \tag{8.9}$$

gdzie $\mathbf{E}_0^{(+)}$ oraz $\mathbf{B}_0^{(+)}$ są zespolonymi amplitudami wektorowymi. Aby zrozumieć strukturę geometryczną fali wewnątrz kryształu, musimy odwołać się do prawa Gaussa dla dielektryków. Przyjmując dla wektora indukcji elektrycznej analogiczną postać harmoniczną, otrzymujemy:

$$\begin{aligned} \nabla\cdot\mathbf{D} &= i\mathbf{k}\cdot\mathbf{D} = 0, \\ \nabla\cdot\mathbf{B} &= i\mathbf{k}\cdot\mathbf{B} = 0. \end{aligned} \tag{8.10}$$

Powyższe zależności dowodzą, że wektory indukcji elektrycznej $\mathbf{D}$ i indukcji magnetycznej $\mathbf{B}$ są zawsze prostopadłe do wektora falowego $\mathbf{k}$. Oznacza to, że niezależnie od własności ośrodka, wektor indukcji $\mathbf{D}$ pozostaje zawsze ściśle poprzeczny względem kierunku propagacji czoła fali.

Kluczowa różnica względem ośrodków izotropowych tkwi jednak w sposobie, w jaki materia odpowiada na wymuszenie polem elektrycznym. W tym miejscu musimy wprowadzić pojęcie **wektora polaryzacji ośrodka $\mathbf{P}$**, który reprezentuje gęstość indukowanych w materiale momentów dipolowych. Ogólna relacja łącząca pole przyłożone $\mathbf{E}$ z indukcją $\mathbf{D}$ ma postać:

$$\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P}. \tag{8.11}$$

O ile w ośrodkach izotropowych odpowiedź materii jest zawsze równoległa do pola $\mathbf{E}$, o tyle w ośrodku anizotropowym (np. w krysztale) tak być nie musi. Wynika to z faktu, że elektrony w sieci krystalicznej są wiązane z różną siłą w różnych kierunkach, co sprawia, że podatność elektryczna staje się wielkością kierunkową, opisywaną za pomocą **tensora podatności elektrycznej** χ:

$$\mathbf{P} = \varepsilon_0\chi\mathbf{E}. \tag{8.12}$$

Podstawiając tę zależność do ogólnego związku między polami, otrzymujemy **równanie materiałowe** w postaci tensorowej:

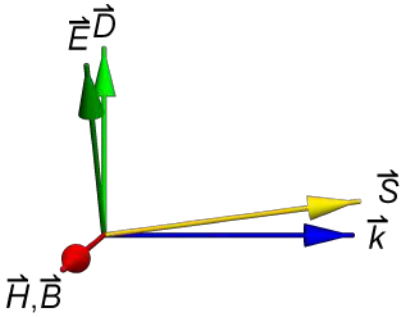
$$\mathbf{D} = \varepsilon_0(\mathsf{I} + \chi)\mathbf{E} = \varepsilon\mathbf{E}, \tag{8.13}$$

gdzie ε jest **tensorem przenikalności elektrycznej**. Zawsze można wyznaczyć taki ortogonalny układ współrzędnych, w którym tensor ten staje się diagonalny. Kierunki te nazywamy **osiami głównymi ośrodka**, a w układzie tym tensor przyjmuje postać:

$$\varepsilon = \begin{pmatrix} \varepsilon_x & 0 & 0 \\ 0 & \varepsilon_y & 0 \\ 0 & 0 & \varepsilon_z \end{pmatrix}, \tag{8.14}$$

gdzie ε_i są przenikalnościami magnetycznymi wzdłuż osi i.

Fakt, że wektory $\mathbf{E}$ i $\mathbf{D}$ nie są do siebie równoległe, prowadzi do niezwykle ważnej konsekwencji geometrycznej: wektor natężenia pola elektrycznego $\mathbf{E}$ przestaje być prostopadły do wektora falowego $\mathbf{k}$ (czyli $\mathbf{k}\cdot\mathbf{E} \neq 0$). Z kolei wektor Poyntinga $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, który wyznacza kierunek przepływu energii, jest z definicji prostopadły do $\mathbf{E}$, a nie do $\mathbf{D}$. W efekcie w ośrodku anizotropowym kierunek promienia świetlnego przestaje pokrywać się z kierunkiem wektora falowego $\mathbf{k}$. Geometrię tę przedstawiono na rys. 8.4. Warto w tym miejscu podkreślić, że



Rysunek 8.4: Przestrzenna orientacja wektorów natężenia i indukcji pól oraz wektora falowego i Poyntinga w ośrodku anizotropowym.

nawet jeśli czoło fali pada prostopadle na powierzchnię kryształu (wektor $\mathbf{k}$ nie zmienia kierunku), energia promienia nadzwyczajnego „skręca" wewnątrz materiału, co prowadzi do separacji przestrzennej obu polaryzacji.

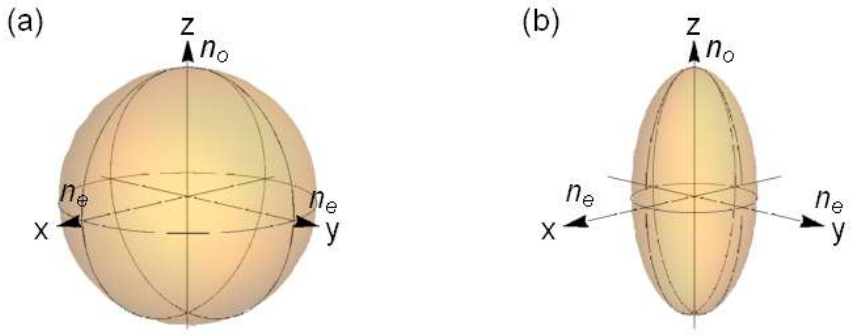
Zachowanie, które zostało opisane powyżej pozwala na wyróżnienie dwóch klas promieni (fal) rozchodzących się w materiale anizotropowym. Nazywane są one:

- **promieniem zwyczajnym**, dla którego ośrodek zachowuje się jak izotropowy ($\mathbf{k} \parallel \mathbf{S}$ oraz $\mathbf{D} \parallel \mathbf{E}$),

- **promieniem nadzwyczajnym**, dla którego kierunek energii ucieka od wektora falowego ($\mathbf{k} \nparallel \mathbf{S}$ oraz $\mathbf{D} \nparallel \mathbf{E}$).

Dla każdego z tych promieni definiuje się osobny współczynnik załamania: n_o (zwyczajny) oraz n_e (nadzwyczajny).

Podstawowym narzędziem wizualizacji tych własności jest **elipsoida współczynnika załamania** (rys. 8.5). Pozwala ona określić współczynnik załamania dla dowolnego kierunku propagacji.



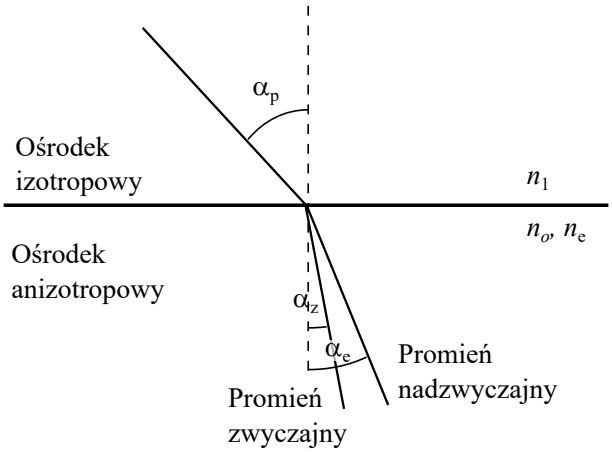
Rysunek 8.5: Elipsoida współczynnika załamania dla ośrodka jednorodnego (a) i dwójłomnego (b).

Na tej podstawie klasyfikujemy kryształy jako:

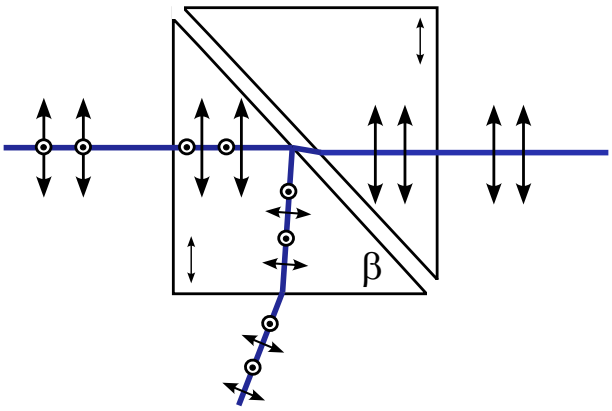
- **izotropowe** ($n_x = n_y = n_z$), gdzie elipsoida jest sferą,
- **jednoosiowe** ($n_x = n_y \neq n_z$), posiadające jedną oś optyczną,
- **dwuosiowe** ($n_x \neq n_y \neq n_z$), posiadające dwie osie optyczne.

Gdy promień pada na granicę z kryształem jednoosiowym, ulega rozszczepieniu na dwie wiązki o prostopadłych polaryzacjach liniowych, ponieważ każda ze składowych porusza się z inną prędkością i załamuje pod innym kątem (rys. 8.6). Zjawisko to wykorzystuje się do budowy precyzyjnych polaryzatorów krystalicznych, takich jak **polaryzator Glana**. Urządzenie to składa się z dwóch pryzmatów dwójłomnych, oddzielonych wąską szczeliną powietrzną. Wykorzystuje ono różnicę współczynników załamania dla promienia zwyczajnego i nadzwyczajnego: przy odpowiednim kącie padania jedna ze składowych ulega całkowitemu wewnętrznemu odbiciu na granicy ze szczeliną, podczas gdy druga zostaje transmitowana.

Jakość polaryzatora opisuje **współczynnik ekstynkcji** (I_{min}/I_{max}), który dla kryształów wynosi ok. 10^{-6}, a dla polaroidów ok. 10^{-2}.



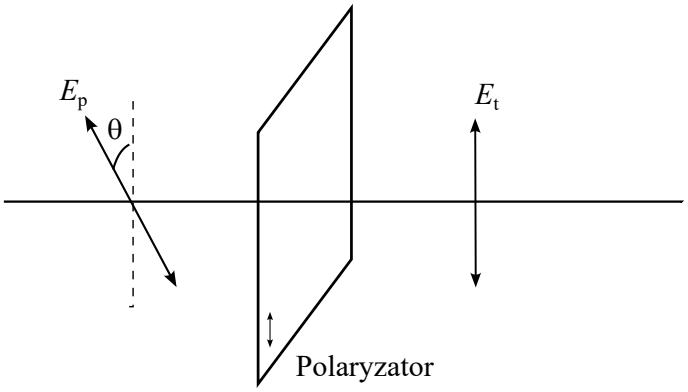
Rysunek 8.6: Załamanie światła na granicy ośrodka izotropowego i kryształu dwójłomnego.



Rysunek 8.7: Zasada działania polaryzatora Glana: wykorzystanie zjawiska całkowitego wewnętrznego odbicia jednej ze składowych polaryzacyjnych na granicy pryzmatów.

8.3 Prawo Malusa

Zastanówmy się teraz, w jaki sposób orientacja polaryzatora wpływa na natężenie transmitowanej fali. Wyobraźmy sobie, że światło spolaryzowane liniowo o natężeniu $I_{0,p}$ pada na polaryzator, którego oś przepuszczania tworzy kąt θ z płaszczyzną drgań pola fali padającej (rys. 8.8).



Rysunek 8.8: Ilustracja do prawa Malusa.

W wyniku oddziaływania światła z materiałem polaryzatora, pole elektryczne fali padającej zostaje zrzutowane na oś przepuszczania. Wyrażenie na rzeczywistą amplitudę fali (bo fala jest liniowo spolaryzowana) transmitowanej przez polaryzator $E_{0,t}$ dane jest przez:

$$E_{0,t}(\theta) = E_{0,p} \cos\theta, \tag{8.15}$$

gdzie $E_{0,p}$ jest amplitudą fali padającej (również założono polaryzację liniową). Ponieważ natężenie światła I jest proporcjonalne do kwadratu modułu amplitudy pola ($I \propto |E_0^{(+)}|^2$), to dla rozważanego przypadku otrzymujemy zależność:

$$\boxed{I_t(\theta) = I_0 \cos^2\theta.} \tag{8.16}$$

To fundamentalne równanie nosi nazwę **prawa Malusa** (Étienne Louis Malus 1785–1812). Pozwala ono na precyzyjne sterowanie natężeniem światła poprzez obrót polaryzatora.

8.4 Płytki fazowe

Własności dwójłomne materiałów wykorzystywane są w tzw. **płytkach fazowych**. Są to elementy optyczne działające na zasadzie wprowadzania kontrolowanego opóźnienia jfazy ednej składowej polaryzacyjnej światła względem

drugiej. Z fizycznego punktu widzenia płytka fazowa rozkłada padający wektor pola **E** na dwie ortogonalne składowe, które ze względu na anizotropię ośrodka propagują się z różnymi prędkościami fazowymi. Różnica faz $\Delta\varphi$ nabywana podczas przejścia przez płytkę dana jest zależnością:

$$\Delta\varphi = k_0 \Delta z (n_o - n_e), \tag{8.17}$$

gdzie k_0 jest wektorem falowym światła w próżni, zaś Δz oznacza grubość materiału. Kierunki drgań odpowiadające tym dwóm współczynnikom wyznaczają tzw. **oś szybką** oraz **oś wolną** płytki. Oś szybka odpowiada temu kierunkowi drgań pola elektrycznego, dla którego współczynnik załamania jest mniejszy, a co za tym idzie, prędkość fazowa fali jest większa.

Wyróżniamy trzy podstawowe typy płytek fazowych, klasyfikowane ze względu na wartość wprowadzanego opóźnienia:

- **płytka jednofalowa** (λ)

  Przesuwa fazę jednej składowej polaryzacyjnej względem drugiej o całkowitą wielokrotność 2π. Przy normalnym (prostopadłym) padaniu wiązki nie powoduje to zmiany stanu polaryzacji. Jednakże obrócenie płytki o mały kąt zmienia drogę optyczną w materiale, co sprawia, że faza ulega niewielkiej zmianie. Dzięki temu możliwa jest precyzyjna kompensacja niepożądanych eliptyczności polaryzacji, wynikających na przykład z odbicia wiązki od powierzchni metalicznych.

- **płytka półfalowa** ($\lambda/2$)

  Opóźnia fazę jednej ze składowych o π. Oznacza to, że orientacja osi polaryzacji liniowej (dla światła spolaryzowanego liniowo) lub wielkiej półosi (dla polaryzacji eliptycznej) światła padającego odbija się symetrycznie względem osi anizotropii kryształu, tj. płaszczyzna polaryzacji obraca się o kąt 2θ, gdzie θ jest kątem między polaryzacją wejściową a osią płytki. Dodatkowo, w przypadku fal spolaryzowanych eliptycznie, kierunek obrotu wektora pola (skrętność) zmienia się na przeciwny (prawoskrętna $\leftrightarrow$ lewoskrętna).

- **płytka ćwierćfalowa** ($\lambda/4$)

  Opóźnia fazę jednej składowej względem drugiej o $\pi/2$. Sprawia to, że fala spolaryzowana liniowo przechodzi w falę spolaryzowaną eliptycznie. W szczególnym przypadku, gdy oś szybka płytki tworzy kąt $\pi/4$ z kierunkiem drgań fali padającej, światło staje się spolaryzowane kołowo. Analogicznie, płytka ćwierćfalowa pozwala na zamianę polaryzacji kołowej na liniową.

Dla celów obliczeniowych działanie płytki fazowej o opóźnieniu $\Delta\varphi$, której oś szybka jest zorientowana pod kątem θ względem osi x, opisuje się za pomocą odpowiednich macierzy. W **formalizmie Jonesa** macierz retardera przyjmuje postać:

$$\mathsf{J}(\theta, \Delta\varphi) = \begin{pmatrix} \cos^2\theta + e^{-i\Delta\varphi}\sin^2\theta & (1 - e^{-i\Delta\varphi})\sin\theta\cos\theta \\ (1 - e^{-i\Delta\varphi})\sin\theta\cos\theta & \sin^2\theta + e^{-i\Delta\varphi}\cos^2\theta \end{pmatrix}. \tag{8.18}$$

Z kolei w **formalizmie Stokesa** działanie płytki o osi szybkiej zorientowanej wzdłuż osi x ($\theta = 0$) opisuje macierz Müllera:

$$\mathsf{M}(0, \Delta\varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\Delta\varphi & \sin\Delta\varphi \\ 0 & 0 & -\sin\Delta\varphi & \cos\Delta\varphi \end{pmatrix}. \tag{8.19}$$

Dla celów obliczeniowych, macierz Müllera dla płytki fazowej o dowolnym opóźnieniu $\Delta\varphi$, której oś szybka jest zorientowana pod kątem θ względem osi x, wyznacza się poprzez transformację:

$$\mathsf{M}(\theta, \Delta\varphi) = \mathsf{R}(-2\theta)\mathsf{M}(0, \Delta\varphi)\mathsf{R}(2\theta), \tag{8.20}$$

gdzie **macierz obrotu wektora Stokesa** $\mathsf{R}(\alpha)$ przyjmuje jawną postać:

$$\mathsf{R}(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{8.21}$$

Oprócz klasyfikacji ze względu na opóźnienie, płytki fazowe dzieli się pod kątem sposobu ich realizacji i odporności na zmiany długości fali:

- **płytki zerowego rzędu**

  Wprowadzają one dokładnie zadaną różnicę faz (np. $\pi/2$). Są to elementy bardzo cienkie i kruche, ale charakteryzują się najmniejszą czułością na odchyłki długości fali od wartości projektowej.

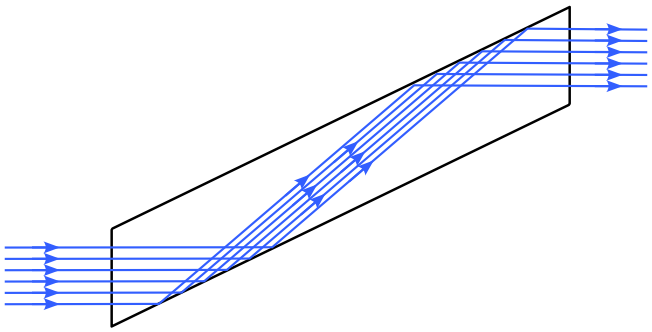
- **płytki wielorzędowe**

  Są to elementy znacznie grubsze, w których przesunięcie fazy odbywa się modulo 2π (np. $1000\pi + \pi/2$). Są prostsze w produkcji i wytrzymalsze, ale cechują się dużą chromatycznością — ich poprawne działanie ograniczone jest do bardzo wąskiego zakresu spektralnego.

- **płytki złożone zerowego rzędu**

  Składają się z dwóch płytek wielorzędowych sklejonych tak, aby ich **osie szybkie** były względem siebie prostopadłe. Dzięki temu duże opóźnienia w obu elementach niemal całkowicie się znoszą, a wynikowa różnica faz odpowiada rzędowi zerowemu. Łączy to zalety mechaniczne płytek grubych z dobrą charakterystyką spektralną płytek cienkich.

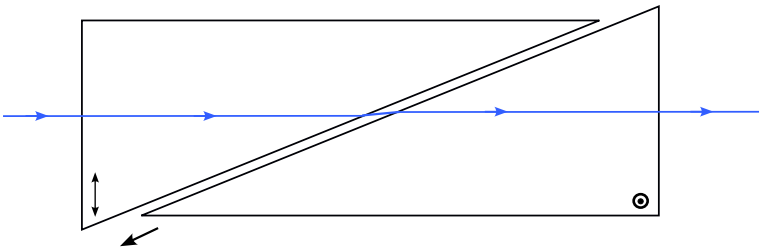
W układach wymagających poprawnej pracy w bardzo szerokim zakresie długości fal (np. światło białe) tradycyjne płytki krystaliczne okazują się niewystarczające ze względu na dyspersję. W takich przypadkach stosuje się **romb Fresnela** (rys. 8.9). Jego działanie opiera się na dwukrotnym całkowitym wewnętrznym odbiciu światła od granicy kryształ-powietrze. Odpowiednie dobranie kąta padania sprawia, że każde odbicie wprowadza fazę $\pi/4$, co łącznie daje



Rysunek 8.9: Romb Fresnela. Przesunięcie fazowe zachodzi tu podczas całkowitego wewnętrznego odbicia, co zapewnia niemal całkowitą achromatyczność elementu.

opóźnienie ćwierćfalowe ($\pi/2$), które w szerokim zakresie długości fali zależy niemal wyłącznie od geometrii elementu.

Przyrządem o jeszcze większej elastyczności jest **kompensator Babineta** (rys. 8.10). Składa się on z dwóch klinów wykonanych z kryształu dwójłomnego



Rysunek 8.10: Kompensator Babineta. Wzajemny przesuw dwóch klinów krystalicznych o prostopadłych osiach optycznych umożliwia uzyskanie dowolnej różnicy dróg optycznych dla obu składowych polaryzacyjnych.

o wzajemnie prostopadłych osiach optycznych. Przesuwając jeden klin względem drugiego, możemy w sposób ciągły zmieniać wypadkową różnicę dróg optycznych dla obu składowych polaryzacyjnych. Dzięki temu kompensator Babineta pozwala na precyzyjne wprowadzenie (lub zmierzenie) dowolnego przesunięcia fazy w wybranym punkcie wiązki.

8.5 Aktywność optyczna

Aktywność optyczna jest zjawiskiem polegającym na stopniowym skręcaniu płaszczyzny polaryzacji światła podczas jego przejścia przez ośrodek materialny. Zjawisko to, zaobserwowane po raz pierwszy w 1811 roku przez François Arago, związane jest z brakiem symetrii lustrzanej budowy określonych związków, czyli

Doświadczenie Skręcenie płaszczyzny polaryzacji w cukrze

tzw. **chiralnością**. W przypadku takich molekuł naturalną bazą do rozważania propagacji światła nie jest polaryzacja liniowa, lecz polaryzacja kołowa. Wynika to z faktu, że to właśnie fale spolaryzowane kołowo stanowią **mody własne** ośrodka chiralnego. Chiralna struktura materii (przypominająca mikroskopijne sprężyny) oddziałuje inaczej z polem elektrycznym wirującym zgodnie z jej skrętem, a inaczej z polem wirującym przeciwnie. W efekcie każda ze składowych kołowych odczuwa inny współczynnik załamania ($n_+ \neq n_-$), co prowadzi do narastania różnicy faz między nimi w miarę postępu fali. Sam materiał, w którym dochodzi do tego typu oddziaływania, cechuje się **dwójłomnością kołową**.

Aby przedstawić to zagadnienie w sposób ilościowy, rozważmy światło spolaryzowane początkowo wzdłuż kierunku $\mathbf{x}$, rozchodzące się wzdłuż osi $\mathbf{z}$:

$$\begin{aligned}\mathbf{E}(z,t) &= \mathbf{x}E_0^{(+)}e^{i(kz-\omega t)} \\ &= \frac{\sqrt{2}}{2}\left(\frac{\mathbf{x}+i\mathbf{y}}{\sqrt{2}}+\frac{\mathbf{x}-i\mathbf{y}}{\sqrt{2}}\right)E_0^{(+)}e^{i(kz-\omega t)} \\ &= {E'}_0^{(+)}\boldsymbol{\sigma}_+e^{i(kz-\omega t)}+{E'}_0^{(+)}\boldsymbol{\sigma}_-e^{i(kz-\omega t)},\end{aligned} \tag{8.22}$$

gdzie $\boldsymbol{\sigma}_+ = (\mathbf{x}+i\mathbf{y})/\sqrt{2}$ oraz $\boldsymbol{\sigma}_- = (\mathbf{x}-i\mathbf{y})/\sqrt{2}$ są wersorami odpowiadającymi polaryzacji kołowej odpowiednio lewo- i prawoskrętnej, zaś ${E'}_0^{(+)} = E_0^{(+)}/\sqrt{2}$. Równanie (8.22) pokazuje, że polaryzację liniową można przedstawić jako superpozycję dwóch ortogonalnych polaryzacji kołowych o identycznych amplitudach i fazach.

Załóżmy teraz, że światło porusza się przez ośrodek cechujący się dwójłomnością kołową, w którym współczynniki załamania dla każdej ze składowych kołowych wynoszą $n_\pm = \bar{n} \pm \Delta n$. Podstawiając te zależności do równania (8.22), otrzymujemy postać pola po przebyciu drogi z:

$$\begin{aligned}\mathbf{E}(z,t) &= E_0^{(+)}\boldsymbol{\sigma}_+e^{i(n_+k_0z-\omega t)}+E_0^{(-)}\boldsymbol{\sigma}_-e^{i(n_-k_0z-\omega t)} \\ &= \frac{E_0}{\sqrt{2}}\boldsymbol{\sigma}_+e^{i\Delta nk_0z}e^{i(\bar{n}k_0z-\omega t)}+\frac{E_0}{\sqrt{2}}\boldsymbol{\sigma}_-e^{-i\Delta nk_0z}e^{i(\bar{n}k_0z-\omega t)} \\ &= E_0\left[\mathbf{x}\cos(\Delta nk_0z)-\mathbf{y}\sin(\Delta nk_0z)\right]e^{i(\bar{n}k_0z-\omega t)}.\end{aligned} \tag{8.23}$$

Z powyższych rozważań wynika, że pole elektryczne pozostaje spolaryzowane liniowo, jednak kierunek jego drgań ulega obrotowi. Skręcenie to jest tym większe, im dłuższą drogę pokonuje światło w ośrodku i im silniejsza jest jego dwójłomność kołowa. W ogólności kąt skręcenia płaszczyzny polaryzacji θ dany jest zależnością:

$$\theta = k_0\Delta z\Delta n = \frac{\pi\Delta z}{\lambda_0}(n_+-n_-). \tag{8.24}$$

W zależności od tego, czy $n_+ > n_-$ czy $n_+ < n_-$, mówimy o substancjach odpowiednio prawoskrętnych lub lewoskrętnych. Warto zaznaczyć, że kąt skręcenia zależy od długości fali światła (zjawisko dyspersji rotacyjnej), co sprawia, że przy użyciu światła białego roztwory aktywne optycznie wykazują mieniące się barwy.

Dygresja

Jednym z najważniejszych zastosowań praktycznych dwójłomności jest **mikroskopia polaryzacyjna**. Technika ta polega na umieszczeniu badanej próbki pomiędzy dwoma skrzyżowanymi polaryzatorami (polaryzatorem i analizatorem). W układzie bez próbki lub przy próbce całkowicie izotropowej, pole fali po przejściu przez pierwszy polaryzator jest prostopadłe do osi drugiego, co zgodnie z prawem Malusa skutkuje całkowitym wygaszeniem światła.

Jeśli jednak próbka wykazuje lokalną dwójłomność (liniową lub kołową), stan polaryzacji fali ulega zmianie — powstaje składowa pola równoległa do osi analizatora, która zostaje transmitowana do detektora. Dzięki temu mikroskopia polaryzacyjna pozwala na niezwykle precyzyjną wizualizację wewnętrznych struktur próbek anizotropowych, takich jak kryształy, minerały, polimery czy włókna biologiczne, ujawniając ich orientację krystalograficzną oraz lokalne naprężenia.

8.6 Wymuszone efekty optyczne

Doświadczenie
Elastometria

Oprócz dwójłomności naturalnej, wynikającej z anizotropii struktury krystalicznej, istnieje szereg zjawisk fizycznych, w których dwójłomność jest indukowana przez czynniki zewnętrzne. Z punktu widzenia optyki kryształów, zjawiska te opisuje się jako modyfikację kształtu lub orientacji elipsoidy współczynników załamania pod wpływem zewnętrznych pól mechanicznych lub elektrycznych. W ośrodku elipsoida współczynnika załamania jest sferą; oddziaływanie zewnętrzne deformuje ją, nadając jej cechy elipsoidy o różnych półosiach.

8.6.1 Fotoelastyczność

Sztucznie wywołana dwójłomność, powstająca pod wpływem naprężeń mechanicznych w ośrodku pierwotnie izotropowym, nosi nazwę **fotoelastyczności**. Zjawisko to pojawia się, gdy do materiału (np. szkła lub polimeru) przykładane są siły zewnętrzne, powodujące lokalne odkształcenia sieci lub łańcuchów cząsteczkowych. Prowadzi to do zmiany rozkładu gęstości elektronowej, co objawia się zmianą wartości głównych współczynników załamania.

Różnica współczynników załamania dla składowych polaryzacyjnych drgających równolegle i prostopadle do kierunku naprężenia opisana jest przez **prawo Brewstera**:

$$\Delta n = C\sigma, \tag{8.25}$$

gdzie σ jest wartością przyłożonego naprężenia (lub różnicą naprężeń głównych), a C jest **stałą elastooptyczną** (stałą Brewstera), zależną od modułu sprężystości i podatności polaryzacyjnej materiału.

Fotoelastyczność znajduje szerokie zastosowanie w wizualizacji rozkładu naprężeń w inżynierii materiałowej. Dzięki tej technice możliwe jest precyzyjne wskazanie miejsc koncentracji naprężeń poprzez obserwację prążków interferencyjnych w świetle spolaryzowanym, co jest kluczowe dla projektowania bezpiecznych komponentów mechanicznych oraz w medycynie (np. analiza obciążeń w protetyce).

8.6.2 Efekty elektrooptyczne: Kerra i Pockelsa

Zjawiskami, w których dochodzi do indukowania dwójłomności liniowej przez zewnętrzne pole elektryczne E_s są **zjawiska elektrooptyczne**. Ich fizyczna natura wiąże się z polaryzacją elektronową lub, w przypadku cieczy, z reorientacją molekuł posiadających trwały moment dipolowy.

Jeżeli indukowana anizotropia zależy od kwadratu natężenia przyłożonego pola elektrycznego, zjawisko nazywamy **efektem Kerra** (John Kerr 1824–1907). Zmiana współczynnika załamania dana jest wzorem:

$$\Delta n = \lambda_0 K E_s^2, \tag{8.26}$$

gdzie K oznacza **stałą Kerra**, a λ_0 jest długością fali w próżni. Ponieważ jest to efekt trzeciego rzędu ($\chi^{(3)}$, patrz rodział 10), występuje on w każdym ośrodku, również w gazach i cieczach izotropowych (np. nitrobenzenie).

W kryształach dielektrycznych **pozbawionych środka symetrii** dominuje zazwyczaj liniowy **efekt Pockelsa** (Friedrich Pockels 1865–1913). Jest to efekt drugiego rzędu ($\chi^{(2)}$), w którym indukowana dwójłomność zależy liniowo od przyłożonego pola E_s:

$$\Delta n = n_o^3 r_{ij} E_s, \tag{8.27}$$

gdzie n_o jest współczynnikiem załamania promienia zwyczajnego, a r_{ij} oznacza odpowiedni element tensora elektrooptycznego (np. r_{63} dla kryształów grupy $\bar{4}2m$, takich jak KDP). Brak symetrii inwersji jest warunkiem koniecznym wystąpienia tego zjawiska.

Wywołana przez pole elektryczne różnica faz $\Delta\varphi$ po przebyciu przez światło drogi L w krysztale wynosi:

$$\Delta\varphi = \frac{2\pi L}{\lambda_0}\Delta n. \tag{8.28}$$

Przy różnicy faz wynoszącej π (warunek płytki półfalowej), układ umieszczony między skrzyżowanymi polaryzatorami staje się transparentny, co pozwala na konstrukcję ultraszybkich modulatorów natężenia światła.

Dygresja: Modulatory elektrooptyczne

Kluczowym zastosowaniem zjawisk elektrooptycznych są **modulatory elek-**

trooptyczne. Pozwalają one na zmianę stanu polaryzacji światła z częstotliwościami rzędu dziesiątek gigaherców (GHz). Umożliwia to generację impulsów o czasie trwania rzędu pikosekund oraz precyzyjne kodowanie informacji w fazie lub amplitudzie fali świetlnej.

Współczesna telekomunikacja światłowodowa opiera się na modulatorach zintegrowanych (np. interferometrach Macha-Zehndera opartych na efekcie Pockelsa). Pozwalają one na ekstremalnie gęste upakowanie danych przy jednoczesnym wyeliminowaniu zakłóceń elektromagnetycznych, które są problemem w tradycyjnej elektronice. Dzięki technikom multipleksowania z podziałem długości fali , każde włókno światłowodowe może transmitować setki niezależnych kanałów informacyjnych.

8.6.3 Efekt Faradaya

Doświadczenie
Efekt Faradaya

Efekt Faradaya (Michael Faraday 1791-1867) jest zjawiskiem polegającym na indukcji dwójłomności kołowej w ośrodku pod wpływem zewnętrznego, podłużnego pola magnetycznego. Powstała różnica faz, jaką nabywają składowe kołowe fali (lewo- i prawoskrętna), prowadzi do wypadkowego skręcenia płaszczyzny polaryzacji światła. Z mikroskopowego punktu widzenia pole magnetyczne oddziałuje z elektronami materiału poprzez siłę Lorentza, co modyfikuje ich drgania w sposób zależny od skrętności fali. W rezultacie każda ze składowych kołowych „odczuwa" nieco inny współczynnik załamania ($n_+ \neq n_-$).

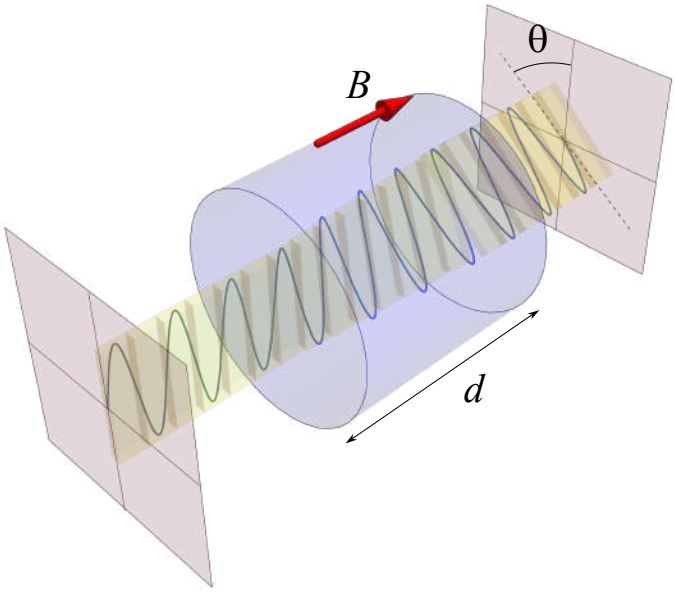
Fundamentalną cechą odróżniającą to zjawisko od naturalnej aktywności optycznej jest jego nieodwracalność. Oznacza to, że zwrot obrotu płaszczyzny polaryzacji zależy wyłącznie od kierunku linii pola magnetycznego, a nie od kierunku propagacji fali. Jeśli światło po przejściu przez ośrodek zostanie odbite i wróci przez ten sam materiał, kąt skręcenia nie wyzeruje się (jak w przypadku chiralności naturalnej), lecz ulegnie podwojeniu.

W najbardziej ogólnym przypadku kąt skręcenia płaszczyzny polaryzacji θ dany jest zależnością:

$$\theta = V B_0 \Delta z, \tag{8.29}$$

gdzie B_0 indukcją pola magnetycznego przykładanego do ośrodka, a wielkość V jest stałą materiałową zwaną **stałą Verdeta**, charakteryzującą odpowiedź ośrodka na pole magnetyczne. Stała Verdeta ta zależy od długości fali światła (dyspersja stałej Verdeta) oraz temperatury. Zazwyczaj V jest proporcjonalna do dyspersji współczynnika załamania $dn/d\lambda$, co sprawia, że efekt jest szczególnie silny w pobliżu krawędzi pasma absorpcji materiału.

Jednym z najważniejszych zastosowań tego zjawiska są **izolatory optyczne**. Są to układy składające się z kryształu (np. granatu terbowo-galowego) umieszczonego w silnym polu magnetycznym pomiędzy dwoma polaryzatorami, których osie są względem siebie skręcone o kąt 45° ($\pi/4$). Długość kryształu jest tak dobrana, aby w danym polu płaszczyzna polaryzacji skręciła się dokładnie o 45°. Światło biegnące w kierunku uprzywilejowanym przechodzi przez pierwszy polaryzator, zostaje skręcone o 45° w krysztale i dzięki temu trafia



Rysunek 8.11: Zasada efektu Faradaya. Skręcenie płaszczyzny polaryzacji fali liniowo spolaryzowanej podczas propagacji wzdłuż linii zewnętrznego pola magnetycznego.

idealnie w oś przepuszczania drugiego polaryzatora. Sytuacja zmienia się radykalnie dla fali biegnącej w stronę przeciwną (np. odbitej od dalszych części układu). Ze względu na nieodwracalność efektu Faradaya, światło wracające zostaje skręcone o kolejne 45° w tym samym zwrocie. W efekcie pada ono na polaryzator wejściowy pod kątem 90° względem jego osi przepuszczania i zgodnie z prawem Malusa [równanie (8.15)] zostaje całkowicie wygaszone. Izolatory te są niezbędne w systemach laserowych o dużej mocy, chroniąc źródło przed szkodliwym wpływem odbić wstecznych.

Dygresja

Szczególną formą opisywanego zjawiska jest tzw. **nieliniowy efekt Faradaya**. Zachodzi on, gdy skręcenie płaszczyzny polaryzacji zaczyna zależeć od natężenia samego światła (optyczne zjawiska nieliniowe zostaną szczegółowo omówione w rozdziale 10). W określonych gazach, takich jak pary metali alkalicznych, efekt ten może być o wiele rzędów wielkości silniejszy niż w ciałach stałych.

Tak czuła zależność kąta skręcenia od pola magnetycznego stanowi fundament **magnetometrii optycznej**. Jest to obecnie najczulsza metoda pomiaru stałych i wolnozmiennych pól magnetycznych, pozwalająca osiągać czułość rzędu pojedynczych femtotesli (10^{-15} T). Znajduje to zastosowanie w biomagnetyzmie (np. nieinwazyjne badanie aktywności magnetycznej serca i mózgu), defektoskopii przemysłowej, a nawet w fundamentalnych badaniach z zakresu fizyki cząstek elementarnych i poszukiwań ciemnej materii.

8.7 Podsumowanie rozdziału

Niniejszy rozdział stanowił rygorystyczne wprowadzenie do opisu polaryzacji światła jako fali poprzecznej, łącząc matematyczne formalizmy Jonesa i Stokesa z fizyczną naturą oddziaływania fali z materią. Przeanalizowano procesy prowadzące do polaryzacji światła, takie jak odbicie pod kątem Brewstera, rozpraszanie oraz dwójłomność krystaliczną opisaną za pomocą elipsoidy współczynnika załamania, a także metody manipulacji fazą składowych przy użyciu płytek fazowych i kompensatorów. Całość zamknęła analiza efektów wymuszonych (fotoelastyczności oraz zjawisk elektro- i magnetooptycznych), które ukazują możliwości dynamicznego sterowania światłem, znajdujące zastosowanie w nowoczesnej telekomunukacji światłowodowej oraz precyzyjnej metrologii.

Zadania

8.1. Opisz stan polaryzacji (rodzaj polaryzacji, skrętność, orientację) dla każdej z następujących fal:

(a) $\mathbf{E} = \mathbf{x}E_0 \cos(kz - \omega t) - \mathbf{y}E_0 \cos(kz - \omega t)$,

(b) $\mathbf{E} = \mathbf{x}E_0 \sin[2\pi(z/\lambda - \nu t)] - \mathbf{y}E_0 \sin[2\pi(z/\lambda - \nu t)]$,

(c) $\mathbf{E} = \mathbf{x}E_0 \sin(kz - \omega t) - \mathbf{y}E_0 \sin(kz - \omega t - \pi/4)$,

(d) $\mathbf{E} = \mathbf{x}E_0 \cos(kz - \omega t) - \mathbf{y}E_0 \cos(kz - \omega t + \pi/2)$.

8.2. Wyznacz orientację osi głównej polaryzacji eliptycznej oraz stopień eliptyczności światła dla następujących wektorów Jonesa:

(a) $A = B = \frac{1}{\sqrt{2}}$ oraz $\varphi = 0$,

(b) $A = B = \frac{1}{\sqrt{2}}$ oraz $\varphi = \pi/2$,

(c) $A = \frac{1}{2}$, $B = \frac{\sqrt{3}}{2}$ oraz $\varphi = \pi/4$.

8.3. Znajdź wektor Jonesa $\mathbf{E}_2$ przedstawiający stan polaryzacji ortogonalny do polaryzacji:

$$\mathbf{E}_1 = \begin{pmatrix} 1 \\ -2i \end{pmatrix}.$$

Naszkicuj oba stany polaryzacji na płaszczyźnie poprzecznej.

8.4. Światło spolaryzowane wzdłuż kierunku x przechodzi przez dwa polaryzatory, których osie zorientowane są względem polaryzacji początkowej pod kątem $\theta_1 = 45°$ oraz $\theta_2 = 90°$. Oblicz amplitudę pola elektrycznego światła za tym układem oraz porównaj wynik z sytuacją, w której w układzie znajduje się tylko drugi polaryzator ustawiony pod kątem $90°$ do polaryzacji wejściowej.

8.5. Światło niespolaryzowane pada na układ trzech polaryzatorów, których kierunki optyczne tworzą kąty $\theta_1 = 40°$, $\theta_2 = -20°$ oraz $\theta_3 = 40°$ z kierunkiem osi y. Jaki procent początkowego natężenia światła zostanie przepuszczony przez ten układ?

8.6. Chcemy obrócić kierunek polaryzacji światła spolaryzowanego liniowo o $90°$ za pomocą polaryzatorów liniowych.

(a) Ile co najmniej polaryzatorów należy użyć, aby osiągnąć ten cel?

(b) Ile polaryzatorów należy użyć, aby natężenie wiązki wyjściowej było większe niż 60% natężenia wiązki padającej?

8.7. Światło spolaryzowane liniowo pod kątem $45°$ pada na pionową płaszczyznę szklaną pod kątem $\theta_i = 80°$. Korzystając z formalizmu Jonesa i współczynników Fresnela, określ stan polaryzacji światła po odbiciu. Przyjmij współczynnik załamania szkła $n = 1{,}5$.

8.8. Płytka ćwierćfalowa dla niebieskiego światła o długości fali $\lambda = 434{,}0$ nm ma być wykonana z kwarcu. Współczynniki załamania dla promienia zwyczajnego i nadzwyczajnego wynoszą odpowiednio 1,55396 oraz 1,56340. Oblicz minimalną grubość płytki niezbędną do uzyskania założonej różnicy faz.

8.9. Materiał optyczny wykazuje współczynniki załamania $n_\mathrm{o} = 1{,}55335$ oraz $n_\mathrm{e} = 1{,}54424$.

(a) Oblicz grubości płytek półfalowej i ćwierćfalowej dla światła o długości fali $\lambda_0 = 780$ nm.

(b) Oblicz grubość płytki, która będzie pełniła rolę płytki półfalowej jednocześnie dla światła czerwonego ($\lambda_1 = 670$ nm) i niebieskiego ($\lambda_2 = 450$ nm).

8.10. Pryzmat Wollastona wykonano z kryształu o współczynnikach $n_\mathrm{o} = 1{,}658$ oraz $n_\mathrm{e} = 1{,}486$. Kąt cięcia obu pryzmatów wynosi $\alpha = 15°$. Narysuj bieg promieni w pryzmacie i po wyjściu z niego oraz oblicz kąt separacji pomiędzy promieniem zwyczajnym i nadzwyczajnym po opuszczeniu układu.

8.11. Oblicz kąt θ ostrej krawędzi rombu Fresnela wykonanego ze szkła o $n = 1{,}51$, tak aby po dwóch całkowitych wewnętrznych odbiciach dwie prostopadłe składowe polaryzacyjne nabyły różnicy faz $\pi/2$.

8.12. Aktywność optyczna sacharozy w wodzie przy $\lambda = 589{,}3$ nm wynosi 66,45° dla 10 cm kolumny roztworu o stężeniu 1 g/cm^3. Wiązka światła pada na komorę o długości 1 m zawierającą 1000 cm^3 roztworu, w którym rozpuszczono 10 g sacharozy. Wyznacz kąt skręcenia płaszczyzny polaryzacji na wyjściu z kolumny.

8.13. Wiązka światła charakteryzuje się natężeniem całkowitym I_0. Po przejściu przez idealny polaryzator liniowy o osi poziomej zmierzone natężenie wynosi $0{,}7I_0$, a po przejściu przez polaryzator o osi pod kątem 45° wynosi $0{,}6I_0$. Wyznacz parametry Stokesa tej wiązki, wiedząc, że nie posiada ona składowej kołowej.

8.14. Światło niespolaryzowane przechodzi kolejno przez polaryzator liniowy o osi pionowej oraz płytkę ćwierćfalową o osi szybkiej zorientowanej pod kątem 45° do pionu. Korzystając z macierzy Müllera, wyznacz wektor Stokesa wiązki wyjściowej i opisz jej stan polaryzacji.

8.15. Przedstaw na sferze Poincarégo ewolucję stanu polaryzacji światła początkowo spolaryzowanego liniowo poziomo podczas przejścia przez płytkę ćwierćfalową, której oś szybka jest obrócona o kąt 45° względem poziomu.

8.16. Oblicz natężenie pola magnetycznego wymagane do obrotu płaszczyzny polaryzacji o 45° w pręcie szklanym o długości 5 cm, jeśli stała Verdeta materiału wynosi $V = 160$ rad/(T·m).

8.17. Kryształ niobianu litu o grubości 10 mm wykazuje stałą elektrooptyczną $r = 30 \cdot 10^{-12}$ m/V. Przyjmując $n_o \approx 2{,}3$, oblicz napięcie półfalowe V_π, przy którym indukowana przez pole elektryczne różnica faz między składowymi wyniesie dokładnie π.

8.18. Wiązka światła jest mieszaniną światła niespolaryzowanego i spolaryzowanego kołowo. Po przejściu przez idealną płytkę ćwierćfalową i obracający się polaryzator liniowy, zmierzone natężenie maksymalne jest trzykrotnie większe od minimalnego. Oblicz stopień polaryzacji V_p pierwotnej wiązki.

Rozdział 9

Model Lorentza

9.1 Wprowadzenie

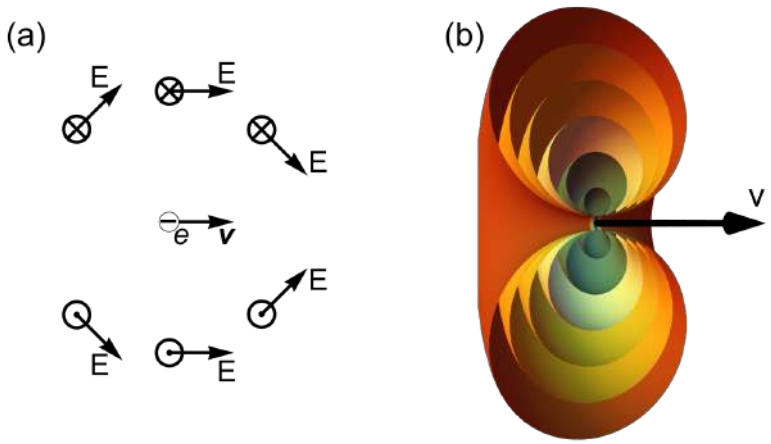
Zanim przystąpimy do szczegółowej analizy mechanizmów emisji światła, musimy poczynić istotne zastrzeżenie metodologiczne. Całość poniższych rozważań opiera się na gruncie **elektrodynamiki klasycznej**, w której promieniowanie elektromagnetyczne ma charakter ciągłej fali generowanej przez przyspieszające ładunki. Choć opis ten doskonale wyjaśnia zjawiska makroskopowe (takie jak działanie anten czy klasyczną dyspersję), fizyka w skali mikroskopowej rządzi się prawami mechaniki kwantowej. Zgodnie z teorią kwantową, emisja światła nie zachodzi w sposób ciągły, lecz dyskretny, jako emisja pojedynczych fotonów, towarzysząca przeskokom (tzw. przejściom kwantowym) elektronów pomiędzy określonymi stanami energetycznymi w atomie. Pełne i spójne połączenie tych dwóch światów zapewnia **elektrodynamika kwantowa**. Na tym etapie pozostaniemy jednak przy formalizmie klasycznym, który stanowi fundament dla zrozumienia oddziaływania światła z materią.

9.2 Emisja fali przez ładunek elektryczny

9.2.1 Ładunek w spoczynku lub poruszający się ruchem jednostajnym prostoliniowym

Zgodnie z elektrodynamiką klasyczną fundamentalnym źródłem pola elektromagnetycznego są ładunki elektryczne. Rozważmy w pierwszej kolejności spoczywający ładunek punktowy i przeanalizujmy, czy może on być źródłem fali elektromagnetycznej. Ładunek taki wytwarza wokół siebie statyczne pole elektrostatyczne (kulombowskie), nie generuje natomiast pola magnetycznego. Ponieważ rozkład pola elektrostatycznego nie wykazuje zmienności czasowej, układ ten nie emituje fali elektromagnetycznej.

Analogiczne wnioski płyną z analizy ładunku poruszającego się ruchem jednostajnym prostoliniowym. Chociaż ładunek taki jest źródłem pola elektrycz-



Rysunek 9.1: Orientacje przestrzenne pola elektrycznego i magnetycznego wytwarzanego przez ujemnie naładowaną cząstkę poruszającą się ruchem prostoliniowym z przyspieszeniem **a** (a) oraz przekrój przez profil przestrzenny emisji promieniowania (b).

nego oraz stałego pola magnetycznego (ściśle związanego z jego ruchem), to pola te nie wykazują zmienności czasowej w układzie współporuszającym się z ładunkiem. Oznacza to, że ruch jednostajny również nie prowadzi do emisji fali elektromagnetycznej.

Powyższe rozważania prowadzą do fundamentalnego wniosku: źródłem fali elektromagnetycznej mogą być wyłącznie ładunki poruszające się ruchem zmiennym, przy czym przyspieszenie to może wynikać ze zmiany zarówno wartości, jak i kierunku wektora prędkości. W dalszej części rozdziału przeanalizujemy klasyczne scenariusze ruchu ładunku elektrycznego skutkujące emisją promieniowania.

9.2.2 Emisja fali przez ładunek rozpędzany liniowo

Jeżeli cząstka naładowana porusza się ruchem prostoliniowym z przyspieszeniem **a**, to generowane przez nią pola elektryczne i magnetyczne wykazują zmienność czasową.

Aby intuicyjnie zrozumieć fizyczny mechanizm tej emisji, wyobraźmy sobie linie pola elektrycznego spoczywającej cząstki jako radialne linie rozchodzące się równomiernie we wszystkich kierunkach. Gdy cząstka zostaje nagle przyspieszona w krótkim przedziale czasu δt, informacja o zmianie jej położenia nie rozchodzi się w przestrzeni natychmiastowo, lecz propaguje się na zewnątrz z prędkością światła c. W rezultacie na liniach pola powstaje poprzeczne zaburzenie (załamanie w kształcie linii), które rozchodzi się w formie rozszerzającej się sferycznej powłoki o grubości $c\delta t$. To właśnie ta poprzeczna, czyli styczna do

czoła fali składowa pola elektrycznego, oznaczana na rys. 9.1(a) jako $\mathbf{E}_t$, stanowi pole promieniowania. Analiza geometryczna tego załamu pokazuje, że stosunek natężenia pola poprzecznego $\mathbf{E}_t$ do pola radialnego (normalnego do czoła fali, oznaczanego jako $\mathbf{E}_r$) rośnie proporcjonalnie do odległości R od źródła. W konsekwencji, o ile statyczne pole bliskie cząstki $\mathbf{E}_r$ zanika szybko jak $1/R^2$, o tyle pole promieniowania $\mathbf{E}_t$ maleje znacznie wolniej, tj. proporcjonalnie do $1/R$, co umożliwia transport energii na nieskończone odległości.

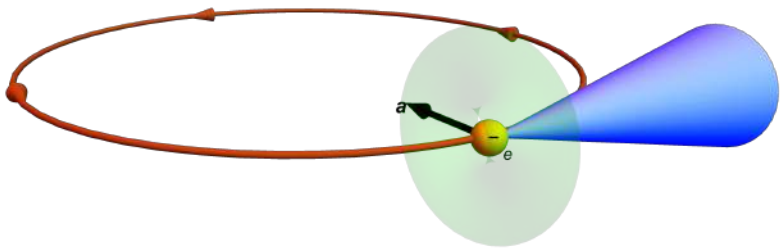
Interpretacja ta bezpośrednio tłumaczy rozmieszczenie przestrzenne generowanych pól. Dla kierunku zgodnego z osią przyspieszenia cząstki linie pola nie ulegają poprzecznemu przesunięciu i załamanie nie powstaje. Oznacza to całkowity brak poprzecznej składowej pola elektrycznego ($\mathbf{E}_t = 0$) wzdłuż osi ruchu. Najsilniejszy załam linii pola pojawia się pod kątem prostym do osi przyspieszenia, gdzie poprzeczne „szarpnięcie" linii jest najgłębsze, dając maksymalną amplitudę pola $\mathbf{E}_t$.

Z kolei ruch cząstki naładowanej ze zmienną prędkością stanowi prąd elektryczny o natężeniu zmiennym w czasie. Zgodnie z prawem Ampère'a-Maxwella, ten zmienny prąd oraz towarzyszące mu zmienne pole elektryczne indukują poprzeczne pole magnetyczne fali promieniowania $\mathbf{B}_t$. Linie sił tego pola układają się w zamknięte okręgi współosiowe z kierunkiem ruchu cząstki (mają charakter azymutalny), a ich amplituda w próżni jest ściśle związana z polem elektrycznym relacją $E_t = cB_t$.

Wzajemne sprzężenie tych czasowo zmiennych, ortogonalnych pól prowadzi do formowania się samodzielnej fali elektromagnetycznej, która odrywa się od ładunku i rozchodzi w przestrzeni. Poprzeczny charakter obu pól sprawia, że dla prędkości nierelatywistycznych strumień energii fali (opisywany wektorem Poyntinga $\mathbf{S} = \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}$) jest najsilniej emitowany w kierunkach prostopadłych do osi ruchu cząstki. Tworzy to przestrzenny rozkład gęstości emitowanej mocy promieniowania o charakterystycznym kształcie torusa (obwarzanka), co zostało graficznie zilustrowane na przekroju bryły promieniowania na rys. 9.1(b). Zgodnie z tym modelem, w kierunku wyznaczonym przez oś przyspieszenia nie jest emitowana żadna energia.

9.2.3 Promieniowanie synchrotronowe

Promieniowanie elektromagnetyczne emituje również ładunek poruszający się z niezmienną szybkością v, którego wektor prędkości $\mathbf{v}$ zmienia jednak kierunek. Przykładem takiego procesu jest ruch po okręgu. W układzie własnym ładunek emituje promieniowanie prostopadle do wektora przyspieszenia dośrodkowego $\mathbf{a}$ (rys. 9.1(a)). Jednakże po przejściu do laboratoryjnego układu odniesienia taka geometria emisji promieniowania musi zostać przetransformowana do nowego układu. W przypadku nierelatywistycznym kształt emisji pozostaje w zasadzie niezmieniony. W reżimie relatywistycznym rozkład ten należy przekształcić za pomocą transformacji Lorentza. Zgodnie z nią początkowa symetria względem wektora przyspieszenia ulega zmianie i emisja zaczyna wykazywać bardziej kierunkowy charakter (rys. 9.2). W szczególności wskutek efektów relatywistycznych obszar emisji przesuwa się w stronę wektora prędkości cząstki,



Rysunek 9.2: Emisja promieniowania elektromagnetycznego przez relatywistyczny elektron rozpędzony w akceleratorze pierścieniowym (synchrotronie). Zielony torus pokazuje kierunek emisji w układzie własnym elektronu, zaś niebieski stożek emisję kierunkową po przetransformowaniu do laboratoryjnego układu odniesienia.

a rozbieżność emitowanego promieniowania drastycznie maleje wraz ze wzrostem prędkości ładunku. Kąt rozbieżności tej emisji jest determinowany przez bezwymiarowy parametr $1/\gamma_L$ (określający połowę kąta rozwarcia stożka promieniowania), gdzie $\gamma_L = 1/\sqrt{1-(v/c)^2}$ jest czynnikiem Lorentza.

Szczególnym przypadkiem opisanego procesu jest **promieniowanie synchrotronowe**. Promieniowanie takie emitowane jest przez cząstki naładowane podczas zakrzywiania ich torów w akceleratorach kołowych lub w kosmosie, w obecności naturalnych, bardzo silnych pól magnetycznych. Jest to promieniowanie wysokoenergetyczne, którego częstość i natężenie mogą być kontrolowane zarówno przez dobór parametrów rozpędzonych cząstek, jak i przez zmianę kąta obserwacji. Jednym z zastosowań tego typu promieniowania jest krystalografia białek, gdyż promieniowanie synchrotronowe cechuje się wystarczająco krótką długością fali, by możliwe było badanie dyfraktogramów takich molekuł i odzyskiwanie na ich podstawie informacji o rozmieszczeniu atomów w molekule. W astrofizyce powszechnie przyjmuje się, że to właśnie ten mechanizm jest odpowiedzialny za tak silną emisję promieniowania z kwazarów.

Dygresja

Interesującym przykładem doświadczenia, w którym dochodzi do oddziaływania relatywistycznych cząstek z promieniowaniem elektromagnetycznym, jest przygotowywany w CERN projekt Gamma Factory. W ramach tego projektu silnie zjonizowane (mające zaledwie kilka elektronów na najbardziej wewnętrznych powłokach), ciężkie jony są rozpędzane do prędkości bliskich prędkości światła, a następnie oświetlane spójną wiązką światła laserowego. Ponieważ energia wiązania elektronów na najbardziej wewnętrznych powłokach jest ogromna, jonów tych nie można wzbudzić tradycyjnymi metodami optycznymi. Relatywistyczny ruch jonów sprawia jednak, że we własnym układzie odniesienia energia oświetlającego promieniowania zostaje znacząco zwiększona przez relatywistyczny efekt Dopplera. Transformacja Lorentza pomiędzy układem laboratoryjnym a układem własnym jonu zwiększa energię fotonu docierającego do jonu o czynnik $(1+\beta)\gamma_L \approx 2\gamma_L$ (gdzie $\beta = v/c \approx 1$). W konsekwencji możliwe

staje się wzbudzenie bardzo energetycznych przejść w jonach nawet za pomocą standardowego światła laserowego. Co więcej, ponieważ wzbudzone jony powracają do stanu podstawowego w skończonym czasie, emitują wówczas promieniowanie. W układzie własnym jonu emisja ta będzie zgodna z charakterystyką promieniowania dipolowego, jednakże transformacja do układu laboratoryjnego sprawi, że emitowane promieniowanie będzie ukierunkowane w podobny sposób jak w przypadku promieniowania synchrotronowego (rys. 9.2), tj. w stronę wektora prędkości jonu. Dodatkowo transformacja do laboratoryjnego układu odniesienia ponownie przeskaluje tę energię o czynnik $\approx 2\gamma_L$ dla fotonów emitowanych w kierunku zgodnym z wektorem prędkości jonu. W ten sposób energia fali rozproszonej wzdłuż kierunku ruchu jonu będzie $4\gamma_L^2$ razy większa od energii fali padającej. Oznacza to, że w akceleratorze LHC, gdzie czynnik Lorentza może dochodzić nawet do 4000, można osiągnąć energię kilkudziesięciu megaelektronowoltów, rozpraszając wstecznie fotony o energii rzędu 2 eV.

W porównaniu z innymi sztucznymi źródłami promieniowania gamma, Gamma Factory umożliwia generowanie wysokoenergetycznego, bardzo intensywnego, a jednocześnie przestrajanego (przez zmianę prędkości jonów i kontrolę czynnika Lorentza γ_L) promieniowania. Otwiera to pole dla wielu ciekawych badań podstawowych, ale także intrygujących zastosowań.

9.2.4 Dipol elektryczny

Jeszcze innym przykładem obiektu emitującego promieniowanie elektromagnetyczne jest drgający **dipol elektryczny**. Jak można pokazać poprzez analogię do wcześniejszych rozważań, jeśli elektryczny moment dipolowy zmienia się w czasie, tzn. jeśli tworzące go ładunki rozsuwają się, to ruch ładunków jest źródłem fali elektromagnetycznej. Fala ta jest emitowana prostopadle do wektora przyspieszenia ładunków stanowiących dipol, a sama emisja odbywa się zgodnie z profilem amplitudowym pokazanym na rys. 9.1(b). Średnia moc $\langle P_d \rangle$ emitowana przez taki układ w pełnym kącie bryłowym wynosi:

$$\langle P_d \rangle = \frac{\mu_0 p_0^2 \omega_d^4}{12\pi c} \tag{9.1}$$

gdzie p_0 jest amplitudą drgań momentu dipolowego, a ω_d oznacza częstość kołową tych drgań. Zależność mocy od czwartej potęgi częstości (ω_d^4) wyjaśnia, dlaczego dipole w skali mikroskopowej niezwykle wydajnie emitują światło, podczas gdy makroskopowe układy radiowe wymagają doprowadzania ogromnych mocy i budowy bardzo dużych struktur nadawczych.

Dygresja

Drgające dipole elektryczne są m.in. wykorzystywane do wytwarzania fal radiowych. W takim przypadku źródłem fali są swobodne elektrony w przewodniku. Elektrony te, pod wpływem przyłożonego do przewodnika

napięcia, przesuwają się, tworząc dipol elektryczny (jonowa sieć krystaliczna pozostaje nieruchoma). Zmienność dipola, wywoływana modulacją napięcia, powoduje powstanie zmiennego pola elektrycznego i magnetycznego, a w konsekwencji emisję fali.

Optymalna długość anten dipolowych jest równa połowie długości emitowanej fali ($\lambda/2$). Oznacza to, że w zakresie używanych powszechnie w przeszłości fal o częstotliwości setek kiloherców, konieczne było budowanie anten o kilkusetmetrowej długości. Co ciekawe jednak, w wielu przypadkach połowa takiej anteny była zakopywana pod ziemią. Dzięki temu wysokość anteny mogła być o połowę niższa.

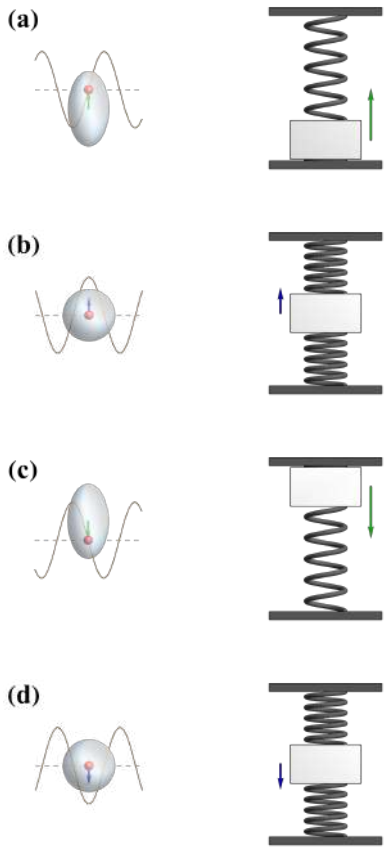
Przez długie lata najwyższą budowlą na świecie był wybudowany w 1974 roku maszt radiowy w Konstantynowie. Miał on wysokość 646 m, a emitowany z niego sygnał radiowy o częstotliwości 225 kHz odbierany był na całym świecie. Maszt ten jednak zawalił się podczas remontu in 1991 roku. Jego odbudowa okazała się nieopłacalna z uwagi na rozwój technologii i rozpoczęcie nadawania sygnału radiowego w zakresie fal o częstotliwości ok. 90 MHz, które nie wymagały budowy tak wysokich anten.

9.3 Model Lorentza

Szczególnym przykładem zjawiska, w którym dochodzi do powstania drgających dipoli elektrycznych, jest oddziaływanie fali elektromagnetycznej z atomami. Klasycznie zjawisko to modeluje się przy pomocy tzw, **modelu Lorentza**. W myśl tego modelu każdy atom opisywany jest jako układ złożony z masywnego, nieruchomego jądra o dodatnim ładunku elektrycznym oraz lekkich elektronów, które ulegają wychyleniu ze swoich położeń równowagi wskutek oddziaływania z oscylującym polem elektrycznym fali (rys. 9.3). To sprzężenie prowadzi do względnego przesunięcia różnoimiennych ładunków, czyli do powstania indukowanych dipoli, których **momenty dipolowe** $\mathbf{d}_q$ zmieniają się w czasie zgodnie z periodycznymi zmianami pola fali. Model ten, opracowany pod koniec XIX wieku przez holenderskiego fizyka Hendrika Lorentza (Hendrik Lorentz 1853–1928), stanowi podstawę klasycznego opisu oddziaływania światła z materią i pozwala na wyprowadzenie wielu wielkości opisujących propagację światła przez ośrodek od strony makroskopowej.

Aby precyzyjnie opisać to przejście od mikroskopowych drgań do mierzalnych wielkości fizycznych, musimy najpierw zestawić ze sobą dwa fundamentalne sposoby opisywania materii. W klasycznej elektrodynamice Maxwella ośrodki materialne są traktowane w sposób ciągły. Ich odpowiedź na fale elektromagnetyczne charakteryzuje się za pomocą parametrów makroskopowych, takich jak podatność elektryczna χ, przenikalność dielektryczna ε oraz współczynnik załamania n. Z kolei w myśl założeń modelu Lorentza, ośrodki te stanowią zbiór mikroskopijnych dipoli elektrycznych, które drgają w wyniku oddziaływania atomów z oscylującym polem elektrycznym E fali świetlnej. Skutkiem tego oddziaływania jest siła działająca na elektrony:

$$F_e = qE, \tag{9.2}$$



Rysunek 9.3: Schematyczna wizualizacja powstawania indukowanego dipola elektrycznego w ramach klasycznego modelu Lorentza. W stanie równowagi, przy braku zewnętrznego pola elektrycznego ($\mathbf{E} = 0$), „środek ciężkości" (efektywna pozycja ładunku) ujemnej chmury elektronowej pokrywa się z dodatnio naładowanym jądrem atomowym. Przyłożenie zewnętrznego oscylującego pola elektrycznego powoduje deformację chmury elektronowej i pojawienie się momentu dipolowego atomu $\mathbf{d}_q$ oraz siły, która próbuje przywrócić układ do stanu równowagi F_h. Rysunki (a)–(d) pokazują kolejne etapy ewolucji takiego układu wraz z siłami które działają w takim układzie. Prawa kolumna pokazuje jak taki układ modelowany jest jako mechaniczny układ harmoniczny.

gdzie q jest ładunkiem elektrycznym. W naszych rozważaniach, dla uproszczenia ograniczymy się jedynie do przypadku jednowymiarowego, w którym fali świetlna jest spolaryzowana liniowo w kierunku x. Rozsunięcie różnoimiennych ładunków elektrycznych powoduje pojawienie się między nimi siły elektrostatycznej, która w pierwszym rzędzie rozwinięcia przyjmuje postać siły harmonicznej:

$$F_h = -k_q x, \tag{9.3}$$

gdzie k_q jest stałą sprężystości charakteryzującą dany układ.

W rzeczywistym ośrodku oscylujący elektron nieustannie traci swoją energię. Proces tłumieniem, wynika z dwóch głównych mechanizmów fizycznych: ciągłego wypromieniowywania fal wtórnych przez przyspieszany ładunek (**tłumienia radiacyjnego**) oraz zderzeń elektronu z sąsiednimi atomami, co prowadzi do zamiany uporządkowanej energii drgań elektronów w nieuporządkowane drania sieci krystalicznej, czyli na ciepło (**tłumienie termiczne**). W klasycznym modelu Lorentza oba te zjawiska opisuje się łącznie w sposób fenomenologiczny. Przyjmuje się, że siła tłumienia, analogicznie do siły oporu lepkościowego, jest skierowana przeciwnie do zwrotu wektora prędkości elektronu i pozostaje wprost proporcjonalna do jego szybkości:

$$F_t = -\beta_q \frac{dx}{dt}, \tag{9.4}$$

gdzie β_q jest współczynnikiem tłumienia.

Korzystając z II zasady dynamiki Newtona, można zapisać równanie ruchu tego ładunku jako:

$$m\frac{d^2x}{dt^2} = F_e + F_h + F_t, \tag{9.5}$$

co po podstawieniu równań (9.2), (9.3) oraz (9.4) i po przekształceniu można przedstawić w postaci:

$$m\frac{d^2x}{dt^2} + \beta_q \frac{dx}{dt} + k_q x = qE. \tag{9.6}$$

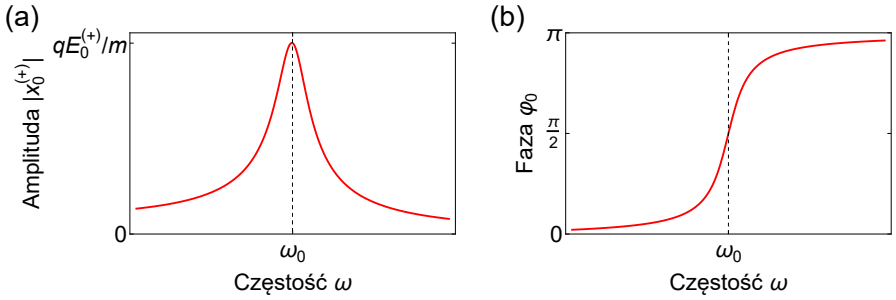
Równanie to jest równaniem różniczkowym niejednorodnym rzędu drugiego. Wyznaczenie położenia x w funkcji czasu polega zatem na znalezieniu rozwiązań tego równania.

Zespolona postać harmonicznego pola elektrycznego fali świetlnej, zapisana w postaci zeaspolonej, wyraża się zależnością:

$$E(t) = E_0^{(+)} e^{-i\omega t}, \tag{9.7}$$

gdzie, dla uproszczenia, pominięto zależność przestrzenną pola E oraz założono, że faza początkowa fali jest zerowa ($\varphi = 0$). Postać siły wymuszającej naturalnie implikuje rozwiązanie równania w postaci:

$$x(t) = x_0^{(+)} e^{-i\omega t}, \tag{9.8}$$



Rysunek 9.4: Amplituda (a) i faza (b) drgań dipola elektrycznego w funkcji częstości ω fali wymuszającej.

gdzie $x_0^{(+)}$ jest zespoloną amplitudą wychylenia ładunku pod wpływem siły $qE(t)$. Ze względu na zespolony charakter amplituda $x_0^{(+)}$ zawiera w sobie dwie kluczowe informacje fizyczne. Pierwsza dotyczy amplitudy wychylenia i zawarta jest w module $\left|x_0^{(+)}\right|$. Druga określa przesunięcie fazowe pomiędzy polem elektrycznym fali wymuszającej a wychyleniem elektronu i jest zawarta w argumencie $x_0^{(+)}$, czyli $\varphi_0 = \arg\left(x_0^{(+)}\right)$.

Podstawiając postać pola elektrycznego fali (9.7), postulowaną postać rozwiązania (9.7) oraz zależność pomiędzy stałą sprężystości k_q a częstością własną ω_0 drgającego dipola ($k_q = m\omega_0^2$) do równania (9.6) można obliczyć wyrażanie na zespoloną amplitudę drgań elektronów:

$$x_0^{(+)} = \frac{qE_0^{(+)}/m}{\omega_0^2 - \omega^2 - i\gamma_q\omega}, \tag{9.9}$$

gdzie $\gamma_q = \beta_q/m$ jest stałą relaksacji tego układu, która mówi o tym, jak szybko układ powraca do stanu równowagi po ustaniu wymuszenia. Obliczając amplitudę i fazę tych drgań, otrzymuje się:

$$\begin{aligned} \left|x_0^{(+)}\right| &= \frac{qE_0^{(+)}/m}{\sqrt{(\omega^2 - \omega_0^2)^2 + \gamma_q^2\omega^2}}, \\ \varphi_0 &= \operatorname{arctg}\left(-\frac{\gamma_q\omega}{\omega^2 - \omega_0^2}\right). \end{aligned} \tag{9.10}$$

Zależności te przedstawiono na rys. 9.4. Jak widać, układ wykazuje wyraźny rezonans, tzw. **rezonans optyczny**, dla częstości $\omega = \omega_0$, dla której amplituda osiąga maksimum, a przesunięcie fazowe między siłą a wychyleniem wynosi $\pi/2$.

Na podstawie przeprowadzonych analiz możemy wyznaczyć mikroskopowy indukowany moment dipolowy $d_q(t)$ rozważanego pojedynczego atomu:

$$d_q(t) = qx(t) = qx_0^{(+)}e^{-i\omega t}. \tag{9.11}$$

Aby przejść do opisu makroskopowego, musimy powiązać ten indywidualny moment dipolowy z polaryzacją ośrodka $P(t)$, która z definicji reprezentuje gęstość wyindukowanych momentów dipolowych (czyli całkowity moment dipolowy przypadający na jednostkę objętości). Zakładając, że w jednostce objętości znajduje się N_d jednorodnych, niezależnych dipoli, makroskopowa polaryzacja (w ujęciu skalarnym dla naszego przypadku jednowymiarowego) wyraża się jako:

$$P(t) = N_d d_q(t) = N_d qx(t) = N_d qx_0^{(+)}e^{-i\omega t}. \tag{9.12}$$

Ponieważ pole wymuszające fali oscyluje z częstością ω, wyindukowana polaryzacja ośrodka również będzie zmieniać się periodycznie z tą samą częstością. Pozwala nam to wprowadzić **zespoloną amplitudę polaryzacji ośrodka** $P_0^{(+)}$, definiując pełną zależność czasową jako:

$$P(t) = P_0^{(+)}e^{-i\omega t}. \tag{9.13}$$

Bezpośrednie porównanie równań (9.12) i (9.13) prowadzi do fundamentalnego powiązania amplitudy makroskopowej z parametrami mikroskopowymi:

$$P_0^{(+)} = N_d qx_0^{(+)}. \tag{9.14}$$

Z drugiej strony, w klasycznej elektrodynamice makroskopowej odpowiedź dielektryczna ośrodka na przyłożone pole fali opisywana jest fenomenologicznie za pomocą podatności elektrycznej χ. Zależność ta ma postać:

$$P(t) = \varepsilon_0 \chi E(t). \tag{9.15}$$

Zapisując obie wielkości za pomocą ich amplitudes zespolonych, możemy wyeliminować zależność od czasu. Pozwala nam to wyrazić podatność elektryczną χ jako stosunek rzeczonych amplitud:

$$\chi = \frac{P_0^{(+)}}{\varepsilon_0 E_0^{(+)}}. \tag{9.16}$$

Zauważmy teraz, że pomiędzy podatnością a względną przenikalnością elektryczną ośrodka ε_r istnieje bezpośredni związek:

$$\varepsilon_r = 1 + \chi, \tag{9.17}$$

co po podstawieniu relacji (9.16) pozwala zapisać ogólne makroskopowe wyrażenie na przenikalność:

$$\varepsilon_r = 1 + \frac{P_0^{(+)}}{\varepsilon_0 E_0^{(+)}}. \tag{9.18}$$

Podstawienie mikroskopowego wyrażenia na zespoloną amplitudę polaryzacji ośrodka $P_0^{(+)}$ (poprzez podstawienie jawnej postaci $x_0^{(+)}$ z równania (9.9) do relacji (9.14)) do równania (9.18) pozwala ostatecznie na pełne i eleganckie powiązanie mierzalnych parametrów makroskopowych z atomowymi parametrami mikroskopowymi:

$$\varepsilon_r = 1 + \frac{N_d q^2}{m\varepsilon_0} \frac{1}{\omega_0^2 - \omega^2 - i\gamma_q \omega}. \tag{9.19}$$

Zgodnie z definicją wprowadzoną w rozdziale 1, współczynnik załamania światła n jest określony jako stosunek prędkości światła w próżni c do prędkości światła w tym ośrodku v. Ponieważ w ośrodkach dielektrycznych przenikalność magnetyczna jest w przybliżeniu równa jedności, można łatwo wykazać, że współczynnik załamania wyraża się przez pierwiastek ze względnej przenikalności elektrycznej ε_r. Ponieważ w naszym mikroskopowym modelu względna przenikalność elektryczna jest wielkością zespoloną, współczynnik załamania staje się wielkością zespoloną $\eta = n + i\kappa$, o czym było wcześniej wspomniane w rozdziale 4. Korzystając z równania (9.19) oraz zależności $\eta = \sqrt{\varepsilon_r}$, możemy zapisać jawne wyrażenie na zespolony współczynnik załamania:

$$\eta = \sqrt{1 + \frac{N_d q^2}{m\varepsilon_0} \frac{1}{\omega_0^2 - \omega^2 - i\gamma_q \omega}}. \tag{9.20}$$

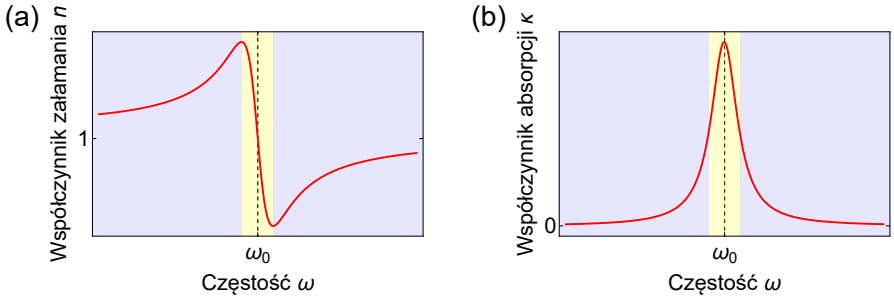
Gdy drugi człon pod pierwiastkiem w równaniu (9.20) jest znacznie mniejszy od jedności – co zachodzi w wielu rzeczywistych układach fizycznych – wyrażenie to można rozwinąć w szereg Taylora, ograniczając się do wyrazów pierwszego rzędu:

$$\eta \approx 1 + \frac{N_d q^2}{2m\varepsilon_0} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma_q^2 \omega^2} + i \frac{N_d q^2}{2m\varepsilon_0} \frac{\gamma_q \omega}{(\omega_0^2 - \omega^2)^2 + \gamma_q^2 \omega^2}. \tag{9.21}$$

Z postaci równania (9.21) wynika, że zespolony współczynnik załamania składa się z części rzeczywistej n (będącej rzeczywistym współczynnikiem załamania) oraz części urojonej κ (odpowiadającej współczynnikowi absorpcji nazywanemu też współczynnikiem ekstynkcji):

$$\begin{aligned} n &= 1 + \frac{N_d q^2}{2m\varepsilon_0} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma_q^2 \omega^2}, \\ \kappa &= \frac{N_d q^2}{2m\varepsilon_0} \frac{\gamma_q \omega}{(\omega_0^2 - \omega^2)^2 + \gamma_q^2 \omega^2}. \end{aligned} \tag{9.22}$$

Zależności współczynnika załamania n oraz współczynnika absorpcji κ od częstości przedstawione zostały na rys. 9.5. Jak widać, współczynnik załamania



Rysunek 9.5: Współczynnik załamania (a) oraz współczynnik absorpcji (b) w funkcji częstości. Obszar zaznaczony na niebiesko jest obszarem dyspersji normalnej, w której absorpcja jest słaba. Obszar zaznaczony na żółto jest obszarem dyspersji anomalnej, w której dochodzi do silnej absorpcji światła.

n cechuje się zależnością dyspersyjną przechodzącą przez wartość jednostkową dla częstości rezonansowej ω_0. Co ciekawe, dla częstości większych od częstości rezonansowej, wartość współczynnika załamania przyjmuje wartość mniejszą od 1, co oznacza, że prędkość fazowa fali w tym zakresie spektralnym jest większa od prędkości światła c w próżni. Na wykresach zaznaczono obszary dyspersji normalnej, w ramach których współczynnik załamania rośnie wraz z częstością, oraz anomalnej, w których maleje. Więcej na ten temat piszemy w rozdziale 9.6.

Warto teraz rozważyć odstrojenia $\omega - \omega_0$ dużo większe od szerokości rezonansu γ_q, $\omega - \omega_0 \gg \gamma_q$, ale wciąż małe w porównaniu z wartościami obu częstości. Na początek zauważmy, że $\omega^2 - \omega_0^2 = (\omega + \omega_0)(\omega - \omega_0) \approx 2\omega(\omega - \omega_0)$, co pozwala przepisać zależności na współczynnik załamania n oraz współczynnik absorpcji κ do postaci:

$$\begin{aligned} n &= 1 + \frac{N_d q^2}{4m\varepsilon_0\omega} \frac{\omega - \omega_0}{(\omega - \omega_0)^2 + \gamma_q^2/4}, \\ \kappa &= \frac{N_d q^2}{8m\varepsilon_0\omega} \frac{\gamma_q}{(\omega - \omega_0)^2 + \gamma_q^2/4}. \end{aligned} \tag{9.23}$$

Korzystając z tych wzorów, można pokazać, że dla dużych odstrojeń współczynnik załamania n zależy odwrotnie proporcjonalnie od odstrojenia, tzn. $n \propto 1/(\omega - \omega_0)$. Jednocześnie współczynnik absorpcji κ zależy odwrotnie proporcjonalnie od kwadratu odstrojenia, tzn. $\kappa \propto 1/(\omega - \omega_0)^2$. Z matematycznego punktu widzenia oznacza to, że skrzydła zależności pokazanych na rys. 9.5 zanikają wolniej dla współczynnika załamania niż dla współczynnika absorpcji. Z fizycznego punktu widzenia oznacza to, że dla odpowiednio dużego odstrojenia współczynnik absorpcji może być mały, podczas gdy współczynnik załamania może się znacząco różnić od jedności. To właśnie z taką sytuacją mamy do czynienia w przypadku wielu ośrodków przezroczystych w świetle widzialnym, w tym m.in. w przypadku szkła oraz wody.

Analizując zapisane powyżej równania, można zauważyć, że w wielu z nich pojawia się wyrażenie $N_d q^2/\varepsilon_0 m$, które ma wymiar kwadratu częstości. Wyrażenie to jest wielkością charakterystyczną dla określonego ośrodka, gdyż zależy od koncentracji indukowanych momentów dipolowych N_d. Ponieważ pierwiastek z tego wyrażenia ma wymiar częstości, wprowadza się nową wielkość ω_p:

$$\omega_p = \sqrt{\frac{N_d q^2}{\varepsilon_0 m}}, \tag{9.24}$$

zwaną **częstością plazmową**.

9.4 Propagacja fali w ośrodku materialnym

Wróćmy teraz do równania opisującego pole elektryczne fali świetlnej, uwzględniającego zależność pola fali od położenia (bez utraty ogólności rozważamy propagację w kierunku z):

$$E = E_0^{(+)} e^{i(kz-\omega t)} = E_0^{(+)} e^{i\left(\frac{2\pi}{\lambda_0}\eta z - \omega t\right)} = E_0^{(+)} e^{-\frac{2\pi}{\lambda_0}\kappa z} e^{i\left(\frac{2\pi}{\lambda_0} n z - \omega t\right)}, \tag{9.25}$$

gdzie skorzystaliśmy z definicji wiążącej zespolony współczynnik załamania η oraz współczynnik załamania n i współczynnik absorpcji κ:

$$\eta = n + i\kappa. \tag{9.26}$$

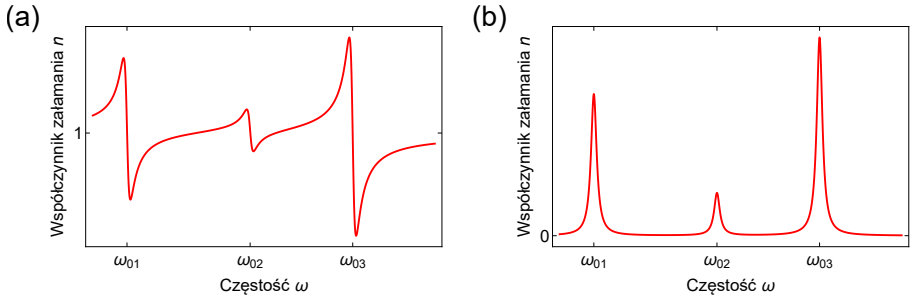
Równanie (9.25) potwierdza przyjęte przez nas powyżej założenie, że część rzeczywista zespolonego współczynnika załamania η odpowiada za opóźnienie fazowe fali podczas jej przechodzenia przez ośrodek, zaś jego część urojona – za jej absorpcję.

Uwzględniając, że natężenie światła zależy od kwadratu natężenia pola elektrycznego, możemy zapisać wyrażenie na absorpcję światła w ośrodku:

$$I = \frac{n\varepsilon_0 c}{2}|E|^2 = \frac{n\varepsilon_0 c}{2}\left|E_0^{(+)}\right|^2 e^{-\frac{4\pi}{\lambda_0}\kappa z} = \frac{n\varepsilon_0 c}{2}\left|E_0^{(+)}\right|^2 e^{-\alpha z} = I_0 e^{-\alpha z}. \tag{9.27}$$

Przekształcając wyrażenie na natężenie światła, sformułowaliśmy **prawo Lamberta-Beera**, w myśl którego natężenie światła rozchodzącego się w ośrodku materialnym maleje e-krotnie po pokonaniu w nim drogi równej $1/\alpha$. Otrzymaliśmy również związek pomiędzy współczynnikiem absorpcji κ oraz współczynnikiem absorpcji światła α:

$$\alpha = \frac{4\pi\kappa}{\lambda_0}. \tag{9.28}$$



Rysunek 9.6: Widmo współczynnika załamania i współczynnika absorpcji układu posiadającego wiele częstości rezonansowych.

9.5 Wiele linii rezonansowych

W ośrodkach materialnych zazwyczaj nie istnieje pojedyncza częstość rezonansowa, lecz wiele częstości rezonansowych. W rozrzedzonych gazach fakt ten można uwzględnić, przepisując współczynnik załamania do postaci:

$$\eta^2 = 1 + \frac{N_d q^2}{\varepsilon_0 m} \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 - i\gamma_{qj}\omega}, \tag{9.29}$$

gdzie sumowanie przebiega po wszystkich liniach rezonansowych, zaś f_j jest bezwymiarową wielkością zwaną **siłą oscylatora** spełniającą zależność:

$$\sum_j f_j = 1. \tag{9.30}$$

Przykładowa zależność współczynnika załamania n oraz współczynnika absorpcji κ ośrodka składającego się z atomów charakteryzujących się kilkoma liniami rezonansowymi o różnych siłach oscylatorów została przedstawiona na rys. 9.6.

Uwaga!

Powyższe równanie jest poprawne dla ośrodków zbudowanych z niezależnych, nieoddziałujących ze sobą oscylatorów (jedyna forma oddziaływania manifestuje się przez stałą relaksacji). Gdy ośrodek jest gęsty, tj. gdy dipole doświadczają lokalnych pól elektrycznych generowanych przez sąsiednie atomy, równanie (9.29), znane jako wzór Lorentza-Lorenza, przyjmuje postać:

$$\frac{\eta^2 - 1}{\eta^2 + 2} = \frac{N_d q^2}{3\varepsilon_0 m} \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2 - i\gamma_{qj}\omega}. \tag{9.31}$$

9.6 Dyspersja normalna i anomalna

Doświadczenie. Obserwacja transmisji światła białego przez szkło neodymowe za pomocą pryzmatu

Z analizy wzorów na współczynnik załamania n (równanie fali tłumionej) wynika, że jego wartość zależy od częstości kołowej fali, co określa się mianem **dyspersji**. W obszarach spektralnych oddalonych od rezonansu, tj. gdy $|\omega - \omega_0| > \gamma_q$, współczynnik załamania rośnie wraz ze wzrostem częstości kołowej, co nazywamy **dyspersją normalną**. Jednakże w pobliżu rezonansu, gdy $|\omega - \omega_0| \lesssim \gamma_q$, charakter dyspersji ulega odwróceniu i współczynnik załamania maleje wraz ze wzrostem częstości kołowej. Zjawisko to nosi nazwę **dyspersji anomalnej**. Warto podkreślić, że zakres dyspersji anomalnej pokrywa się z pasmem silnej absorpcji fali. Oznacza to, że ośrodki skondensowane (takie jak ciała stałe czy ciecze), które są przezroczyste w zakresie widzialnym, mają swoje pasma absorpcyjne w innych zakresach spektralnych. Klasycznym przykładem jest szkło krzemionkowe, którego główne pasmo absorpcyjne leży w zakresie ultrafioletu (dlatego w całym obszarze widzialnym wykazuje ono dyspersję normalną).

9.6.1 Nadświetlność i wolne światło

Propagacja fali w różnych obszarach dyspersyjnych niesie ze sobą istotne konsekwencje dla prędkości rozchodzenia się impulsów światłą (prędkość grupowa). Dzieje się tak dlatego, że zgodnie z definicją prędkości grupowej (5.26) jej wartość jest określona m.in. przez pochodną współczynnika załamania względem częstości, $dn/d\omega$. W rezultacie, gdy pochodna ta jest dodatnia (co odpowiada obszarowi dyspersji normalnej), prędkość grupowa jest mniejsza od prędkości fazowej fali w tym ośrodku ($v_\mathrm{g} < v$). Równocześnie w obszarze dyspersji anomalnej, gdzie pochodna $dn/d\omega$ jest ujemna, prędkość grupowa może przyjmować wartości większe od prędkości fazowej (a nawet od prędkości fali w próżni c, czyli $v_\mathrm{g} > c$).

Niezwykle interesujący fizycznie przypadek pojawia się wówczas, gdy iloczyn częstości kołowej fali oraz pochodnej współczynnika załamania, $\omega\, dn/d\omega$, przyjmuje wartości porównywalne z wartością samego współczynnika załamania n lub od niego większe. W tej sytuacji różnica między prędkością fazową a prędkością grupową fali staje się bardzo duża. Zjawisko to występuje m.in. dla wąskich linii rezonansowych, gdzie pochodna $dn/d\omega$ osiąga bardzo duże wartości, mimo że całkowita amplituda zmian współczynnika załamania pozostaje niewielka. W szczególności z taką sytuacją mamy do czynienia w zjawisku elektromagnetycznie indukowanej przezroczystości, które jest optycznym zjawiskiem nieliniowycm (rozdział 10), gdzie w pasmie absorpcji wytwarzane jest wąskie okno przezroczystości, a zależność współczynnika załamania od częstości ma nachylenie dodatnie. Gwałtowne zmiany współczynnika załamania sprawiają, że wspomniany iloczyn osiąga olbrzymie wartości dodatnie. Dzięki temu w ośrodkach materialnych możliwe jest spowolnienie prędkości grupowej do wartości rzędu zaledwie kilku metrów na sekundę. Zjawisko tak drastycznego spowolnienia rozchodzenia się impulsów światła określa się mianem **wolnego światła**.

Z odwrotnym przypadkiem mamy do czynienia, gdy obserwujemy bardzo wąski rezonans absorpcyjny. Wówczas strome, ujemne nachylenie krzywej współ-

czynnika załamania sprawia, że prędkość grupowa impulsu może przewyższyć prędkość fali w próżni c. Zjawisko to nosi nazwę **szybkiego światła** lub **propagacji nadświetlnej**. Dalsze zwiększanie wartości tego iloczynu może sprawć, że prędkość grupowa spadnie do zera (co odpowiada koherentnemu zatrzymaniu i zmagazynowaniu impulsu w ośrodku), a następnie przyjmie wartości ujemne. Ujemna wartość prędkości grupowej prowadzi do niezwykle intrygującej konsekwencji: szczyt impulsu wyjściowego opuszcza ośrodek, zanim jego maksimum zdąży dotrzeć do powierzchni wejściowej. Zjawisko to nie narusza jednak przyczynowości – wynika ze specyficznego przegrupowania (rekonstrukcji) składowych harmonicznych fali, w którym czoło impulsu ulega selektywnemu wzmocnieniu.

Zarówno propagacja nadświetlna, jak i ujemna prędkość grupowa silnie działają na wyobraźnię. Nasuwają one bowiem pytania o możliwość nadświetlnego transmisji informacji czy podróży kosmicznych. Okazuje się jednak, że przekroczenie prędkości światła w próżni przez prędkość grupową ($v_g > c$) nie narusza fundamentalnej zasady przyczynowości (kauzalności). Wynika to z faktu, że szczególna teoria względności ogranicza prędkość propagacji sygnałów (informacji), a nie prędkość geometrycznych punktów stałej fazy czy obwiedni. Sygnał (informacja) jest przenoszony wyłącznie przez czoło impulsu (nieciągłość fali), które w każdym ośrodku rozchodzi się ściśle z prędkością c_0.

Dygresja

Możliwość manipulacji prędkością światła, a zwłaszcza dynamicznej kontroli jej wartości za pomocą nieliniowych procesów optycznych, otwiera nowe horyzonty w dziedzinie przetwarzania informacji kwantowej. Fotony są doskonałymi nośnikami stanów kwantowych (informacja może być kodowana np. w ich polaryzacji), ponieważ wykazują niezwykle słabe oddziaływanie z otoczeniem, co chroni je przed dekoherencją. Z tego samego powodu fotony słabo oddziałują między sobą, co utrudnia realizację logicznych operacji dwukubitowych (manipulację informacją). Z drugiej strony atomy o niezerowym spinie mogą długo przechowywać stany kwantowe i łatwo podlegać zewnętrznej kontroli (np. za pomocą pól magnetycznych). Możliwość zatrzymania impulsu światła w ośrodku ($v_g = 0$) i przeniesienia stanu kwantowego pola na stany atomowe (tworząc tzw. ciemne polarytony) pozwala na połączenie zalet obu tych światów. Co kluczowe, proces ten pozwala na bezstratny i w pełni odwracalny transfer całego stanu kwantowego (amplitudy i fazy), co stanowi unikalną cechę pamięci kwantowych, nieosiągalną w klasycznych systemach.

9.7 Zależność pomiędzy absorpcją i dyspersją

Głębokim wnioskiem płynącym z klasycznego modelu Lorentza jest stwierdzenie, że dyspersja i absorpcja nie są niezależnymi zjawiskami, lecz stanowią odpowiednio część rzeczywistą i urojoną tej samej fizycznej wielkości – zespolonego

współczynnika załamania.

Aby zrozumieć to głębokie powiązanie, musimy odwołać się do fundamentalnej **zasady przyczynowości**. W kontekście optyki oznacza ona po prostu, że ośrodek materialny nie może zareagować (np. poprzez wyindukowanie makroskopowej polaryzacji) zanim dotrze do niego czoło fali elektromagnetycznej. Skutek nigdy nie może wyprzedzać przyczyny. Innymi słowy, funkcja opisująca odpowiedź ośrodka w dziedzinie czasu musi być tożsamościowo równa zeru dla $t < 0$. Jak udowadnia zaawansowany aparat analizy matematycznej (twierdzenie Titchmarsha), transformata Fouriera każdej takiej „przyczynowej" (ściętej w czasie) funkcji posiada unikalną właściwość: jej część rzeczywista i urojona w dziedzinie częstotliwości nie są niezależne, lecz powiązane ze sobą całkową transformacją Hilberta. W fizyce to sztywne matematyczne powiązanie nosi nazwę **relacji Kramersa-Kroniga** (Hendrik Kramers 1894–1952, Ralph Kronig 1904–1995).

Zważywszy, że dla nieskończenie wielkich częstości ośrodek przestaje reagować na pole fali i zespolony współczynnik załamania dąży do jedności próżni ($\eta(\infty) = 1$), omawiane relacje stosuje się dla funkcji $\eta(\omega) - 1$. Całkując po pełnym spektrum częstości, otrzymujemy sprzężony układ równań:

$$\begin{aligned} n(\omega) - 1 &= \frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \frac{\kappa(\omega')}{\omega' - \omega}\, d\omega', \\ \kappa(\omega) &= -\frac{1}{\pi}\mathcal{P}\int_{-\infty}^{\infty} \frac{n(\omega') - 1}{\omega' - \omega}\, d\omega', \end{aligned} \tag{9.32}$$

gdzie $\mathcal{P}$ oznacza wartość główną całki w sensie Cauchy'ego.

W praktyce fizycznej bezpośredni pomiar ujemnych częstości jest niemożliwy i pozbawiony sensu fizycznego. Wykorzystując jednak właściwości symetrii rzeczywistych pól elektromagnetycznych (parzystość części rzeczywistej $n(-\omega) = n(\omega)$ oraz nieparzystość części urojonej $\kappa(-\omega) = -\kappa(\omega)$), relacje te można przekształcić do znacznie bardziej użytecznej postaci, zdefiniowanej wyłącznie dla widma częstości dodatnich:

$$\begin{aligned} n(\omega) - 1 &= \frac{2}{\pi}\mathcal{P}\int_{0}^{\infty} \frac{\omega'\kappa(\omega')}{\omega'^2 - \omega^2}\, d\omega', \\ \kappa(\omega) &= -\frac{2\omega}{\pi}\mathcal{P}\int_{0}^{\infty} \frac{n(\omega') - 1}{\omega'^2 - \omega^2}\, d\omega'. \end{aligned} \tag{9.33}$$

Powyższe relacje mają niebagatelne znaczenie doświadczalne i inżynieryjne. Wynika z nich, że pełna wiedza o właściwościach optycznych materiału jest zawarta już w jednej z tych wielkości. Pozwala to na przykład na precyzyjne wyznaczenie pełnej charakterystyki dyspersyjnej materiału [czyli współczynnika załamania $n(\omega)$ w szerokim paśmie] wyłącznie na podstawie łatwiejszego w realizacji pomiaru jego widma absorpcyjnego $\kappa(\omega)$.

9.8 Podsumowanie rozdziału

Niniejszy rozdział dokonał fundamentalnej syntezy pomiędzy mikroskopową strukturą materii a jej makroskopowymi właściwościami optycznymi, opierając się na klasycznym modelu oscylatora Lorentza. Wykazano, że oddziaływanie monochromatycznej fali elektromagnetycznej z atomem indukuje zmienny w czasie elektryczny moment dipolowy, którego dynamiczną odpowiedź determinują siła wiązania elektronu oraz mechanizmy dyssypacji energii (tłumienie radiacyjne i zderzeniowe). Poprzez uśrednienie przestrzenne tych mikroskopijnych dipoli wyprowadzono zespolony współczynnik załamania $\eta(\omega)$, którego część rzeczywista $n(\omega)$ odpowiada za zjawisko dyspersji (zmiany prędkości fazowej i uwarunkowanie dyspersji normalnej oraz anomalnej), natomiast część urojona $\kappa(\omega)$ opisuje procesy tłumienia (absorpcję) światła, ujęte makroskopowo prawem Lamberta-Beera. Dowiedziono, że te dwa pozornie niezależne procesy są nierozerwalnie sprzężone relacjami całkowymi Kramersa-Kroniga, będącymi bezpośrednią konsekwencją relatywistycznej zasady przyczynowości. Analiza ta pozwoliła na głęboke zrozumienie natury prędkości grupowej oraz zjawisk wolnego i szybkiego światła, stanowiąc klasyczny fundament pod nowoczesne układy optyki nieliniowej i kwantowego przetwarzania informacji.

Zadania

9.1. Dla drgającego harmonicznie ładunku q wyznacz, korzystając ze wzorów na promieniowanie dipolowe, wyrażenie na uśredniony w czasie wektor Poyntinga $\langle \mathbf{S} \rangle$, a następnie pokaż, że przekrój czynny na rozpraszanie σ jest proporcjonalny do ω^4 (rozpraszanie Rayleigha). Na podstawie tej zależności oblicz teoretyczny stosunek mocy rozproszonego w atmosferze światła fioletowego ($\lambda_1 = 400$ nm) do światła czerwonego ($\lambda_2 = 700$ nm).

9.2. Oblicz częstość własną oscylatora ω_0 dla elektronu w modelu Lorentza, wiedząc, że stała sprężystości wynosi $k_q = 120$ N/m, a masa elektronu $m_e = 9{,}109 \times 10^{-31}$ kg.

9.3. Oblicz częstość plazmową ω_p dla sodu, wiedząc, że koncentracja elektronów przewodnictwa wynosi $N_d = 2{,}6 \times 10^{28}$ m^{-3}, ładunek elementarny $e = 1{,}602 \times 10^{-19}$ C, a masa elektronu $m_e = 9{,}109 \times 10^{-31}$ kg.

9.4. W warstwie F jonosfery koncentracja elektronów wynosi $N_d = 1{,}00 \times 10^{12}$ m^{-3}. Oblicz częstotliwość plazmową f_p w MHz. Czy faza radiowa o częstotliwości $f = 8$ MHz przeniknie przez tę warstwę?

9.5. Pokaż, że równanie:

$$n^2(\omega) = 1 + \frac{N_d e^2}{\varepsilon_0 m_e} \frac{1}{\omega_0^2 - \omega^2},$$

gdzie N_d to koncentracja elektronów, e ładunek elementarny, m_e masa elektronu, ω_0 częstość rezonansowa, a ε_0 przenikalność elektryczna próżni,

można zapisać jako:

$$\frac{1}{n^2-1} = \frac{C}{\lambda_0^2} - \frac{C}{\lambda^2},$$

gdzie $C = 4\pi^2 c^2 \varepsilon_0 m_e / N_d e^2$, przy czym $\lambda_0 = 2\pi c/\omega_0$.

9.6. Wykorzystując ogólne równanie dyspersyjne bez tłumienia:

$$n^2(\omega) = 1 + \frac{N_d e^2}{\varepsilon_0 m_e} \sum_j \frac{f_j}{\omega_{0j}^2 - \omega^2}, \tag{9.34}$$

rozwiń prawą stronę w szereg Taylora dla pasma widzialnego ($\omega \ll \omega_{0j}$), wyprowadź równanie Cauchy'ego:

$$n(\lambda) = C_1 + \frac{C_2}{\lambda^2}$$

oraz wyznacz analitycznie stałe C_1 i C_2 w funkcji parametrów atomowych f_j.

9.7. Kryształ kwarcu ma współczynniki załamania $n_1 = 1{,}557$ dla $\lambda_1 = 400$ nm oraz $n_2 = 1{,}547$ dla $\lambda_2 = 500$ nm. Wykorzystując równanie Cauchy'ego, wyznacz stałe C_1 oraz C_2 i oblicz współczynnik załamania dla $\lambda_3 = 610$ nm.

9.8. Dla szkła BK7 stałe Cauchy'ego wynoszą $C_1 = 1{,}5046$ oraz $C_2 = 4{,}20 \times 10^{-15}$ m^2. Oblicz współczynnik załamania dla linii D sodu ($\lambda = 589{,}3$ nm) oraz wyznacz dyspersję $dn/d\lambda$ w tym punkcie.

9.9. Wyprowadź analitycznie rozmycie czasowe impulsu Δt_{pulse} w światłowodzie o długości d:

$$\Delta t_{\text{pulse}} \approx \frac{d}{c_0} \lambda_0^2 \left| \frac{d^2 n}{d\lambda_0^2} \right| \Delta\lambda_0. \tag{9.35}$$

Dla światłowodu kwarcowego przy $\lambda_0 = 1550$ nm oraz $d^2n/d\lambda_0^2 = -0{,}00472\ \mu\text{m}^{-2}$, wyznacz poszerzenie impulsu na jednostkę długości.

9.10. Udowodnij matematycznie, korzystając ze związku na prędkość grupową $v_g = c_0/n_g$, gdzie wskaźnik grupowy zdefiniowany jest jako:

$$n_g = n(\omega) + \omega \frac{dn}{d\omega}, \tag{9.36}$$

że wewnątrz pasma o silnej absorpcji (obszar dyspersji anomalnej) pochodna $\frac{dn}{d\omega}$ przyjmuje wartości ujemne o dużej wartości bezwzględnej.

9.11. W eksperymencie wykorzystującym światło lasera i komórkę wypełnioną oparami cezu uzyskano ekstremalną dyspersję anomalną, osiągając ujemną

wartość wskaźnika grupowego $n_{\mathrm{g}} = -310$. Wiedząc, że komórka ma fizyczną długość $L = 6{,}0$ cm, oblicz tzw. relatywistyczny czas wyprzedzenia impulsu, zdefiniowany relacją:

$$\Delta t = \frac{(n_{\mathrm{g}} - 1)L}{c_0}.$$

O ile nanosekund maksimum obwiedni impulsu świetlnego pojawi się na płaszczyźnie wyjściowej komórki wcześniej, niż gdyby poruszało się ono na tym samym dystansie w próżni? Wyjaśnij na podstawie analizy przyczynowości Sommerfelda (Arnold Sommerfeld 1868–1951), dlaczego ten efekt nie umożliwia nadświetlnej transmisji informacji.

Rozdział 10

Optyka nieliniowa

10.1 Wprowadzenie

W poprzednich rozdziałach opisywaliśmy propagację światła przy założeniu, że pola elektryczne fal świetlnych są relatywnie słabe w porównaniu do wewnętrznych pól atomowych. Wraz z rozwojem techniki laserowej możliwe stało się uzyskanie natężeń światła, dla których odpowiedź ośrodka przestaje być liniową funkcją przyłożonego pola. W niniejszym rozdziale wprowadzimy formalizm pozwalający na opis takich silnych oddziaływań, gdzie klasyczna zasada superpozycji przestaje obowiązywać, a własności optyczne materiału zaczynają zależeć od samego natężenia światła.

10.2 Alternatywna postać równania falowego

Dla rozważanego w poprzednim rozdziale oddziaływania światła z materią w tzw. **reżimie liniowym** spełnione jest równanie falowe (1.8) oraz obowiązuje zasada superpozycji. W konsekwencji możliwe jest proste dodawanie do siebie fal, a wielkości opisujące ośrodek pod kątem optyki, takie jak współczynnik załamania oraz współczynnik absorpcji, nie zależą od natężenia światła. Podczas propagacji światła w ośrodkach liniowych nie zmienia się także długość rozchodzącej się fali.

Przyjrzyjmy się jednak raz jeszcze propagacji fali w ośrodkach liniowych. Zgodnie ze wzorem (1.8), równanie falowe można zapisać w postaci

$$\nabla^2\mathbf{E} - \frac{1}{v^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = \nabla^2\mathbf{E} - \frac{\eta^2}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = 0. \tag{10.1}$$

W tym miejscu należy jednak podkreślić, że w przeciwieństwie do wcześniej wyprowadzonej zależności, w równaniu (10.1) wykorzystany został zespolony współczynnik załamania η, a nie współczynnik załamania n, który rozważaliśmy wcześniej.

Skorzystajmy teraz ze związku pomiędzy zespolonym współczynnikiem załamania oraz podatnością elektryczną, $\eta^2 = \varepsilon_r = 1 + \chi$ i przepiszmy równanie (10.1) do postaci

$$\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\chi}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{10.2}$$

Wiedząc, że polaryzacja ośrodka $\mathbf{P}$ wiąże się z wektorem pola elektrycznego fali świetlnej $\mathbf{E}$ relacją $\mathbf{P} = \varepsilon_0 \chi \mathbf{E}$, otrzymujemy równanie:

$$\boxed{\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \mathbf{P}}{\partial t^2}.} \tag{10.3}$$

Ze względu na jawną zależność pola od polaryzacji ośrodka, zależność (10.3) będzie wykorzystywana w dalszych rozważaniach.

10.3 Zjawiska nieliniowe

10.3.1 Fizyczne źródła nieliniowej polaryzacji

Gdy w poprzednim rozdziale rozważaliśmy model Lorentza, założyliśmy, że siłę oddziaływania elektrostatycznego wewnątrz atomu można przybliżyć zależnością liniową. Skutkiem tego podejścia było pojawienie się w równaniu ruchu elastycznej siły harmonicznej $F_h = -k_q x$. Jednakże warto zauważyć, że dla wyższych natężeń światła odchylenia elektronu od punktu równowagi stają się na tyle duże, że konieczne jest uwzględnienie wyższych rzędów rozwinięcia w szereg Taylora, czyli rozważenie **siły anharmonicznej**:

$$F_{\text{ah}} = -k_q x - k_q^{(2)} x^2 - k_q^{(3)} x^3 + \ldots, \tag{10.4}$$

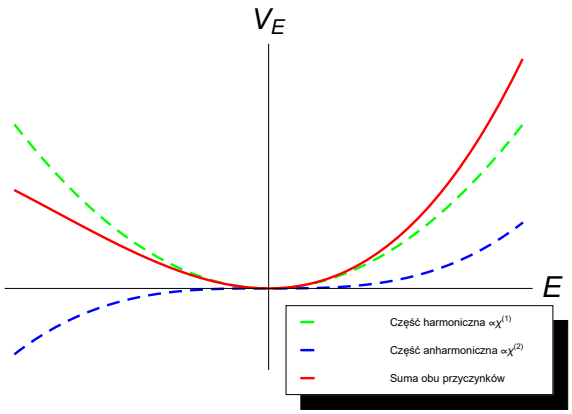
gdzie k_q jest liniową stałą sprężystości, zaś $k_q^{(j)}$ określa nieliniową stałą sprężystości j-tego rzędu.

Korzystając ze związku pomiędzy siłą a potencjałem ($V = -\int F\,dx$), można pokazać, że potencjał elektronu V_E przyjmuje postać:

$$V_E(x) = \frac{k_q x^2}{2} + \frac{k_q^{(2)} x^3}{3} + \frac{k_q^{(3)} x^4}{4} + \ldots \tag{10.5}$$

Przykładowy kształt takiego potencjału przedstawiono na rys. 10.1. Jak widać, dla siły anharmonicznej z dominującym wyrazem drugiego rzędu (niebieska krzywa) potencjał traci symetrię środkową, tzn. $V_E(x) \neq V_E(-x)$, w przeciwieństwie do parabolicznego potencjału oscylatora harmonicznego (krzywa zielona).

Istotnym pytaniem w kontekście powyższego profilu jest to, w jakich materiałach pojawiają się człony nieliniowe łamiące symetrię (np. $k_q^{(2)}$). Otóż najprościej efekt ten wytłumaczyć budową sieci krystalicznej i brakiem środka inwersji. Z taką sytuacją mamy do czynienia w strukturach asymetrycznych (np. w kryształach takich jak niobian litu). Z kolei w gazach atomowych lub



Rysunek 10.1: Różne przyczynki do potencjału elektronu w anharmonicznym polu elektrycznym jądra.

szkłach (czyli w ośrodkach posiadających środek symetrii) z racji wymogów symetrii przestrzennej, wszystkie parzyste człony nieliniowe (np. $k_q^{(2)}$) muszą tożsamościowo się zerować. W konsekwencji w ośrodkach centrosymetrycznych dominującym wkładem do nieliniowości są zjawiska opisywane przez pierwsze nieparzyste rzędy rozwinięcia wyższe niż liniowe, takie jak wkład od $k_q^{(3)}$.

10.3.2 Równanie Lorentza w ośrodku nieliniowym

Rozpatrzmy ruch ładunku wymuszany polem elektrycznym fali świetlnej, w sytuacji, w której siła przywracająca układ do stanu równowagi dana jest w postaci anharmonicznej. W takim przypadku wyprowadzone wcześniej równanie ruchu (9.5) musi zostać zmodyfikowane o dodatkowe wyrazy:

$$m\frac{d^2x}{dt^2} + \beta\frac{dx}{dt} + k_q x + k_q^{(2)}x^2 + k_q^{(3)}x^3 = qE, \tag{10.6}$$

gdzie założyliśmy obecność tylko dwóch najniższych wyrazów rozwinięcia anharmonicznego.

Równanie (10.6) nie ma ogólnego rozwiązania analitycznego. Możemy jednak rozważyć przypadek, w którym pole świetlne jest na tyle słabe, że zakładamy, iż człony $k_q^{(2)}x^2$ oraz $k_q^{(3)}x^3$ są o wiele mniejsze od członu liniowego $k_q x$. W takim przypadku równanie to może być rozwiązane przy pomocy rachunku zaburzeń. Na potrzeby tego rachunku, wprowadzamy podstawienie $E(t) \to \zeta E(t)$, gdzie ζ jest parametrem w przedziale od zera do jedności. Równanie (10.6) przyjmuje wówczas postać

$$\frac{d^2x}{dt^2} + \gamma\frac{dx}{dt} + \omega_0^2 x + a^{(2)}x^2 + a^{(3)}x^3 = \zeta\frac{qE}{m}, \tag{10.7}$$

gdzie $a^{(i)} = k_q^{(i)}/m$. Z uwagi na perturbacyjny charakter zaburzenia, rozwiązania na położenie x będziemy poszukiwać w postaci

$$x = \zeta x^{(1)} + \zeta^2 x^{(2)} + \zeta^3 x^{(3)} + \dots . \tag{10.8}$$

Wstawiając powyższą postać rozwiązania do równania (10.7), podnosząc do odpowiednich potęg x, a następnie grupując wyrazy w odpowiednich potęgach ζ, otrzymujemy układ równań:

$$\begin{aligned} \frac{d^2x^{(1)}}{dt^2} + \gamma\frac{dx^{(1)}}{dt} + \omega_0^2 x^{(1)} &= \frac{qE}{m}, \\ \frac{d^2x^{(2)}}{dt^2} + \gamma\frac{dx^{(2)}}{dt} + \omega_0^2 x^{(2)} + a^{(2)}\left(x^{(1)}\right)^2 &= 0, \\ \frac{d^2x^{(3)}}{dt^2} + \gamma\frac{dx^{(3)}}{dt} + \omega_0^2 x^{(3)} + 2a^{(2)}x^{(1)}x^{(2)} + a^{(3)}\left(x^{(1)}\right)^3 &= 0, \end{aligned} \tag{10.9}$$

gdzie ograniczyliśmy się do pierwszych trzech potęg ζ.

Rozważmy teraz odpowiedź ośrodka opisanego układem równań (10.9) na pole elektryczne fali świetlnej. W przeciwieństwie jednak do przypadku rozważanego w poprzednim rozdziale zakładamy, że światło składa się z dwóch pól o amplitudach E_1 i E_2 oraz częstościach ω_1 i ω_2. Dodatkowo nie możemy już używać zapisu zespolonego, a musimy podać pełną postać pola elektrycznego, co wynika z faktu, że $(\mathrm{Re}[A])^2 \neq \mathrm{Re}[A^2]$, gdzie A jest dowolną liczbą zespoloną. W takim przypadku:

$$E(t) = E_{10}^{(+)}e^{-i\omega_1 t} + E_{20}^{(+)}e^{-i\omega_2 t} + (E_{10}^{(+)})^* e^{i\omega_1 t} + (E_{20}^{(+)})^* e^{i\omega_2 t}. \tag{10.10}$$

W najniższej potędze ζ [pierwsza linia w równaniach (10.9)], dynamika układu rządzona jest przez (liniowy) model Lorentza. Oznacza to, że w najniższym rzędzie zaburzenia rozwiązanie $x^{(1)}$ można zapisać w postaci

$$x^{(1)}(t) = x_0^{(1)}(\omega_1)e^{-i\omega_1 t} + x_0^{(1)}(\omega_2)e^{-i\omega_2 t} + (x_0^{(1)})^*(\omega_1)e^{-i\omega_1 t} + (x_0^{(1)})^*(\omega_2)e^{-i\omega_2 t}, \tag{10.11}$$

gdzie każda z amplitud ma postać typu

$$x_0^{(1)}(\omega_j) = \frac{qE_j}{m}\frac{1}{\omega_0^2 - \omega_j^2 - i\gamma\omega_j}, \tag{10.12}$$

czego można łatwo dowieść przez niezależne rozpatrzenie wyrazów w odpowiedniej potędze $e^{\pm i\omega_)t}$.

Aby poprawnie obliczyć wymuszenie w członie nieliniowym ζ^2, do wyrażenia $\left(x^{(1)}\right)^2$ do kwadratu należy podnieść pełną postać $x^{(1)}$ [równanie (10.11)]. To właśnie przez to w wyrażeniu $\left(x^{(1)}\right)^2$ pojawi się aż cały szereg wyrazów na różnych częstotliwościowych: $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$, $\omega_1 - \omega_2$ plus ich sprzężenia zespolone (ujemne częstotliwości) oraz wyrazay niezależne od czasu. Każdy z nich odpowiada innemu zjawisku nieliniowemu (rozdział 10.4).

Wyliczając odpowiedź układu możemy rozseparować drugie równanie w układzie (10.9) na różne zależności częstotliwościowe. Przykładowo dla częstości $2\omega_1$ otrzymujemy:

$$\frac{d^2x^{(2)}}{dt^2} + \gamma\frac{dx^{(2)}}{dt} + \omega_0^2 x^{(2)} = -\frac{a^{(2)}q^2E_1^2}{m^2D^2(\omega_1)}e^{-2i\omega_1 t}, \tag{10.13}$$

gdzie wprowadziliśmy oznaczenie $D(\omega_j) = \omega_0^2 - \omega_j^2 - i\gamma\omega_j$. Poszukując rozwiązania w postaci

$$x^{(2)}(t) = x_0^{(2)}(2\omega_1)e^{-2i\omega_1 t}, \tag{10.14}$$

otrzymujemy

$$x_0^{(2)}(2\omega_1) = -\frac{a^{(2)}q^2E_1^2}{m^2D(2\omega_1)D^2(\omega_1)}. \tag{10.15}$$

W pełni analogiczne rozwiązania można przeprowadzić dla pozostałych częstości.

Zastanówmy się teraz nad opisem tego problemu z punktu widzenia makroskopowej polaryzacji ośrodka. Przez analogię do przeprowadzonych powyżej rozważań mikroskopowych możemy założyć, że związek pomiędzy indukowaną polaryzacją a polem elektrycznym można zapisać w postaci szeregu potęgowego:

$$\begin{aligned} P &= \varepsilon_0\left(\chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \dots\right) \\ &= P^{(1)} + P^{(2)} + P^{(3)} + \dots \\ &= P_\mathrm{l} + P_\mathrm{nl}, \end{aligned} \tag{10.16}$$

gdzie $\chi^{(j)}$ jest tensorem podatności elektrycznej j-tego rzędu, $P^{(j)}$ polaryzacją ośrodka j-tego rzędu, zaś P_l oraz P_nl oznaczają odpowiednio liniową i nieliniową część polaryzacji. W zapisie wielowymiarowym (uwzględniającym polaryzację fal i anizotropię) $\chi^{(j)}$ staje się tensorem o 3^{j+1} elementach, stąd poszczególne składowe przestrzenne polaryzacji wyrażają się sumami:

$$P_i = \varepsilon_0\left(\sum_j \chi_{ij}^{(1)}E_j + \sum_{j,k}\chi_{ijk}^{(2)}E_jE_k + \sum_{j,k,l}\chi_{ijkl}^{(3)}E_jE_kE_l + \dots\right). \tag{10.17}$$

Skupiając się na klasycznym ujęciu częstotliwościowym (skalarnym), z równania (10.16) wynika dla składowej liniowej:

$$P^{(1)}(\omega_j) = \varepsilon_0\chi^{(1)}(\omega_j)E(\omega_j), \tag{10.18}$$

gdzie symbol ω_j oznacza częstość kołową j-tej składowej pola elektrycznego. W procesach nieliniowych, gdzie w ośrodku współistnieje wiele fal o różnych częstościach, indeks j pozwala jednoznacznie przypisać odpowiedź ośrodka (podatność) do konkretnej składowej widmowej pola. Z drugiej strony, na podstawie mikroskopowej definicji fizycznego wychylenia ładunków wiemy, że:

$$P^{(1)}(\omega_j) = Nqx^{(1)}(\omega_j). \tag{10.19}$$

Przyrównując oba równania otrzymujemy wyrażenie na podatność liniową z modelu Lorentza:

$$\chi^{(1)}(\omega_j) = \frac{Nq^2}{\varepsilon_0 m D(\omega_j)}. \tag{10.20}$$

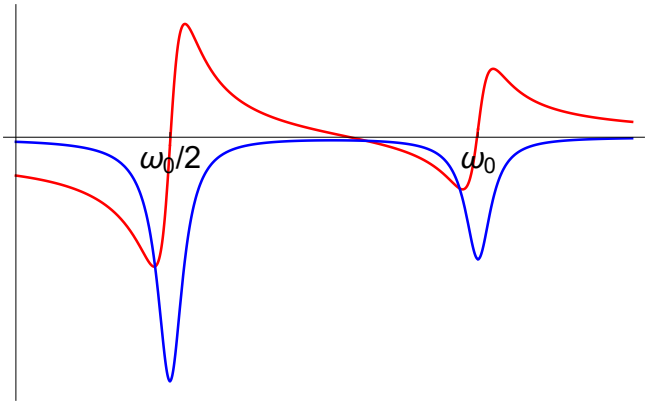
Rozszerzając to rozwiązanie o uwzględnienie siły anharmonicznej drugiego rzędu (wprowadzając zredukowany do masy parametr siły anharmonicznej $a^{(2)} = k_q^{(2)}/m$), można analitycznie wyznaczyć wychylenie $x^{(2)}(2\omega_1)$, a z niego podatność w drugim rzędzie dla częstości $2\omega_1$. Po matematycznym pogrupowaniu czynników, podatność drugiego rzędu można niezwykle elegancko powiązać z podatnościami liniowymi, co nosi nazwę **reguły Millera** (Robert Charles Miller 1925–2016):

$$\chi^{(2)} = -\frac{Nq^3 a^{(2)}}{\varepsilon_0 m^2 D(2\omega_1) D^2(\omega_1)} = -\frac{m a^{(2)} \varepsilon_0^2}{N^2 q^3} \chi^{(1)}(2\omega_1) \left[\chi^{(1)}(\omega_1)\right]^2. \tag{10.21}$$

Przeanalizujmy otrzymany wynik. Jak widać, w podatności nieliniowej $\chi^{(2)}$ mamy do czynienia z dwiema częstotliwościami. Pierwszą z nich jest częstość fali padającej ω_1, na którą reaguje ośrodek. Obok niej pojawia się wyindukowana w materiale wyższa częstotliwość – w tym przypadku podwojona ($2\omega_1$). Powyższy układ równań [równania (10.20) oraz (10.21)] ma dalekosiężne fizyczne konsekwencje, ponieważ oznacza występowanie dwóch rezonansów dla nieliniowej funkcji odpowiedzi. Pierwszy rezonans, wymuszony przez mianownik $D(\omega_1)$, występuje, gdy częstotliwość fali wymuszającej jest równa częstotliwości własnej oscylatora ($\omega_1 = \omega_0$). Drugi rezonans, narzucony przez mianownik $D(2\omega_1)$, występuje dla fali o dwukrotnie mniejszej częstości, tzn. gdy $2\omega_1 = \omega_0$, co implikuje $\omega_1 = \omega_0/2$. Stosując do tego klasycznego układu obrazową nomenklaturę kwantową, procesy te można by interpretować odpowiednio jako **rezonans jednofotonowy** (gdzie energia pojedynczego fotonu rezonuje z przejściem atomowym) oraz **rezonans dwufotonowy** (gdzie do uzyskania energii rezonansowej w ujęciu kwantowym potrzebna jest jednoczesna, silnie nieliniowa absorpcja dwóch fotonów). Proces ten został obrazowo przedstawiony na rys. 10.2.

Jeśli na ośrodek nieliniowy padać będą jednocześnie dwie wiązki światła o różnych częstościach ω_1 oraz ω_2, w pełni analogiczne rozważania można przeprowadzić dla wszystkich możliwych kombinacji tych częstotliwości. Wynikiem takiego sprzężenia nieliniowego będzie pojawienie się nowych fal nie tylko na częstotliwościach harmonicznych ($2\omega_1$, $2\omega_2$), ale również na częstotliwościach będących ich kombinacją liniową: częstotliwości sumarycznej $\omega_1 + \omega_2$ oraz różnicowej $|\omega_1 - \omega_2|$.

Zastanówmy się nad globalnymi konsekwencjami wynikającymi z istnienia nieliniowości w układzie. Po pierwsze, z równań wprost wynika, że zarówno efektywna podatność, a w konsekwencji także współczynnik załamania ośrodka, stają się funkcją amplitudy (natężenia) padającego światła. W praktyce laboratoryjnej to właśnie zależność parametrów reemitowanego światła od natężenia wiązki pompującej jest koronnym dowodem na nieliniowy charakter obserwowanych zjawisk.



Rysunek 10.2: Część rzeczywista (czerwony) i urojona podatności elektrycznej w drugim rzędzie.

Drugą, niezwykle istotną konsekwencją jest fundamentalne załamanie się zasady superpozycji. O ile bowiem w reżimie liniowym oświetlenie materiału dwiema wiązkami sprawia, że ośrodek opuszczają fale o dokładnie tych samych częstotliwościach (choć o ewentualnie zmodyfikowanych amplitudach i fazach), o tyle w przypadku oddziaływania nieliniowego pole elektryczne staje się wewnątrz ośrodka źródłem zupełnie nowych częstotliwości (sumy, różnicy czy wyższych harmonicznych). Oznacza to, że nieliniowa polaryzacja działa jako nowe, wewnętrzne źródło promieniowania. Fakt, że właściwości optyczne ośrodka dla jednej fali mogą być modyfikowane przez obecność innej fali (np. w zjawisku optycznego efektu Kerra), otwiera fascynującą możliwość „kontroli fotonów za pomocą fotonów", bez pośrednictwa układów elektronicznych.

Zjawiska te stanowią fundament współczesnej **fotoniki nieliniowej**, umożliwiając budowę przestrajalnych źródeł laserowych, ultraszybkich przełączników w pełni optycznych oraz kryształów generujących splątane pary fotonów, które napędzają dzisiejszy rozwój **optyki kwantowej**.

Dygresja

Generacja wyższych harmonicznych jest szeroko wykorzystywana w technikach laserowych do generowania fal o nowych, często niedostępnych wprost, długościach fali. Przykładowo, proces ten zachodzi wtedy, kiedy wiązka intensywnego światła laserowego, emitowanego np. z lasera na ciele stałym, jakim jest domieszkowany neodymem granat itrowo-aluminiowy (Nd:YAG), generującego promieniowanie o długości fali 1064 nm, oddziałuje z nieliniowym materiałem. W przypadku potrojenia czy czterokrotnego zwiększenia częstości tego światła, możemy w ten sposób wytworzyć promieniowanie spójne w zakresie ultrafioletu, co ma m.in. zastosowanie w mikroskopii i litografii. Generacja wyższych harmonicznych jest również wykorzystywana

w badaniach nad strukturą atomową i molekularną, umożliwiając uzyskanie ważnych informacji spektroskopowych.

10.4 Optyczne zjawiska nieliniowe

Przyjrzyjmy się teraz polaryzacji ośrodka w dwóch najniższych rzędach rozwinięcia nieliniowego, czyli $P^{(2)}$ oraz $P^{(3)}$. Przyczynki te są opisywane dwiema podatnościami elektrycznymi $\chi^{(2)}$ oraz $\chi^{(3)}$ i z tego powodu w żargonie naukowym często mówi się o nieliniowości $\chi^{(2)}$ oraz $\chi^{(3)}$.

Nieliniowość $\chi^{(2)}$

Jak wspomniano wcześniej, z nieliniowością proporcjonalną do podatności $\chi^{(2)}$ mamy do czynienia w ośrodkach, które nie są centrosymetryczne (nie posiadają środka inwersji). Przykładem takich ośrodków są niektóre kryształy. Wyróżniamy w nich następujące podstawowe efekty nieliniowe:

- **efekt elektrooptyczny** (efekt Pockelsa)

  Zjawisko to, pokrótce omówione w rozdziale 8, polega na tym, że zmiana współczynnika załamania zachodzi na skutek przyłożenia zewnętrznego, stałego pola elektrycznego do kryształu.

- **generacja drugiej harmonicznej**

  Efekt, w którym, w myśl optyki kwantowej, dwa fotony o tej samej częstotliwości ω oddziałują z ośrodkiem nieliniowym, ulegając anihilacji, a w ich miejsce powstaje jeden foton o podwojonej częstotliwości 2ω. Zjawisko to zostało szczegółowo omówione w poprzednim podrozdziale.

- **generacja sumy** i **różnicy częstości**

  Zjawiska polegające na nieliniowym mieszaniu dwóch wiązek światła o częstotliwościach ω_1 i ω_2, w wyniku czego powstaje nowa fala o częstotliwości równej odpowiednio ich sumie ($\omega_3 = \omega_1 + \omega_2$) lub różnicy ($\omega_3 = |\omega_1 - \omega_2|$). Wynikają one bezpośrednio z nieliniowych właściwości ośrodka w drugim rzędzie rozwinięcia podatności elektrycznej. Zjawiska te są bezpośrednią manifestacją zasady zachowania energii (odpowiednio: fuzji dwóch fotonów w jeden lub rozpadu fotonu wysokoenergetycznego na dwa o niższej energii).

- **optyczne wzmocnienie parametryczne**

  Zjawisko to polega na przeniesieniu części energii z jednej wiązki (tzw. pompy) do drugiej (tzw. sygnału), przy jednoczesnym wygenerowaniu trzeciej wiązki (tzw. fali jałowej), tak aby zachowana była zasada zachowania energii i pędu fotonów. W ten sposób można wzmacniać słabe wiązki światła, których generacja innymi metodami albo jest bardzo trudna, albo wręcz nie istnieją wąskie spektralnie źródła światła (lasery), które pozwalałyby na wytworzenie tego typu wiązki.

Nieliniowość $\chi^{(3)}$

W przeciwieństwie do nieliniowych zjawisk optycznych drugiego rzędu, zjawiska nieliniowe trzeciego rzędu mogą być generowane w ośrodkach, które posiadają środek symetrii (ciecze izotropowe, gazy, szkła, a także kryształy centrosymetryczne). Wśród najważniejszych zjawisk rządzonych podatnością $\chi^{(3)}$ wymienić należy:

- **optyczny efekt Kerra**

  Zjawisko polegające na zmianie współczynnika załamania ośrodka pod wpływem zmiennego pola elektrycznego fali świetlnej. Wywołana zmiana proporcjonalna jest do natężenia przechodzącego przez ośrodek światła, a przez to może prowadzić do zmian warunków propagacji światła w różnych miejscach przekroju poprzecznego wiązki (zależnie od lokalnego natężenia). Skutkuje to często tzw. **samoogniskowaniem wiązki**. Ważnym zastosowaniem efektu Kerra jest generowanie ultrakrótkich impulsów laserowych (rozdział 11.4.4).

- **zjawisko samomodulacji fazy**

  Czasową konsekwencja optycznego efektu Kerra jest zjawisko samomodulacji fazy. Zjawisko to polega na zmianie współczynnika załamania ośrodka pod wpływem chwilowego natężenia propagującego się impulsu światła $I(t)$. W efekcie różne części impulsu (zbocze narastające, szczyt, zbocze opadające) doznają różnego opóźnienia fazowego. Prowadzi to do rozsunięcia przestrzennego częstotliwości w obrębie impulsu (powstaje tzw. **świergot częstotliwości**) i znacznego poszerzenia jego widma. Samomodulacja fazy jest kluczowym zjawiskiem wykorzystywanym m.in. w procesie generacji **superkontinuum** oraz w budowie laserów generujących impulsy femtosekundowe.

- **mieszanie czterech fal**

  Zjawisko to polega na oddziaływaniu trzech fal w ośrodku nieliniowym, co prowadzi do powstawania czwartej składowej częstotliwościowej. Podczas mieszania czterech fal, częstotliwość nowo powstałej fali jest kombinacją liniową częstotliwości fal wzbudzających (np. $\omega_4 = \omega_1 + \omega_2 - \omega_3$) z zachowaniem energii i pędu. Oznacza to, że z fal początkowych powstaje nowa wiązka o różnej częstotliwości. Intensywność tej nowej fali zależy od siły oddziaływania między falami początkowymi oraz od ich amplitud i faz. Co więcej, kierunek emisji fali w tym zjawisku jest zdeterminowany przez tzw. **warunek dopasowania fazowego** (zasadę zachowania pędu), przez co może być precyzyjnie kontrolowany.

- **generacja trzeciej harmonicznej**

  Proces analogiczny do generacji drugiej harmonicznej, polegający na nieliniowej zamianie trzech fotonów o częstotliwości ω na jeden foton o częstotliwości 3ω. Z racji znacznie mniejszych wartości podatności $\chi^{(3)}$ względ-

dem $\chi^{(2)}$, zjawisko to wymaga bardzo wysokich natężeń światła oraz precyzyjnego dopasowania fazowego.

- **wymuszone rozpraszanie Ramana**

  Nieliniowy proces rozpraszania nieelastycznego, w którym część energii padających fotonów wiązki pompującej przekazywana jest do ośrodka w postaci wzbudzeń oscylacyjno-rotacyjnych molekuł (lub optycznych fononów w sieci krystalicznej). Skutkuje to wygenerowaniem silnej fali o niższej częstotliwości, zwanej **falą Stokesa**. Zjawisko to jest powszechnie stosowane jako mechanizm wzmocnienia w światłowodowych laserach ramanowskich.

- **elektromagnetycznie indukowana przezroczystość**

  To ww. zjawisko jest efektem kwantowej interferencji. W zjawisku tym silna wiązka pompująca modyfikuje własności optyczne nieprzezroczystego ośrodka atomowego (zazwyczaj gazu) w taki sposób, że staje się on całkowicie przezroczysty dla słabej wiązki próbnej o częstotliwości rezonansowej. Zjawisko to wiąże się z gigantyczną nieliniowością optyczną $\chi^{(3)}$ oraz efektem „wolnego światła".

Dygresja

Jednym z najbardziej spektakularnych zjawisk optyki nieliniowej jest generacja **superkontinuum**. Gdy wprowadzimy ultrakrótki impuls laserowy (np. z lasera tytanowo-szafirowego) do specjalnego światłowodu fotonicznego, silne pole elektryczne impulsu uruchamia całą kaskadę zjawisk nieliniowych trzeciego rzędu (głównie samomodulację fazy oraz mieszanie czterech fal). W efekcie pierwotnie wąskie widmo lasera (np. w podczerwieni) ulega drastycznemu poszerzeniu, pokrywając cały zakres widzialny i bliską podczerwień. Promieniowanie opuszczające światłowód wygląda dla ludzkiego oka jak zwykłe białe światło (np. z żarówki), ale zachowuje pełną spójność przestrzenną i czasową lasera. Za wykorzystanie tego zjawiska do budowy tzw. optycznych grzebieni częstotliwości przyznano w 2005 roku Nagrodę Nobla z fizyki.

10.5 Podsumowanie rozdziału

W niniejszym rozdziale nakreślono fizyczne podstawy optyki nieliniowej, odchodząc od klasycznego paradygmatu proporcjonalności polaryzacji ośrodka materialnego do amplitudy pola elektrycznego światła. Wykazano w nim, że przy odpowiednio dużych natężeniach światła zasada superpozycji załamuje się, a ośrodki materialne stają się aktywnymi środowiskami potrafiącymi generować nowe częstotliwości promieniowania. Powyższe zestawienie dowodzi, jak bogatym i fascynującym obszarem fizyki jest optyka nieliniowa drugiego i trzeciego

rzędu. Należy jednak pamiętać, że wszystkie te zjawiska (a w szczególności efekty rzędu trzeciego) wymagają do swojej efektywnej obserwacji znacznie wyższych natężeń światła niż klasyczne zjawiska optyki liniowej. Wspólnie stanowią one nie tylko potężne narzędzie badawcze, ale również fundament współczesnej fotoniki nieliniowej i technologii ultraszybkich.

Zadania

10.1. Rozważ model elektronu w nieliniowym potencjale opisanym siłą przywracającą $F_{\mathrm{ah}} = -k_q x - k_q^{(2)} x^2$. Oblicz, dla jakiego wychylenia x człon nieliniowy siły staje się równy 1% członu liniowego. Zakładając, że wychylenie x jest rzędu promienia atomu wodoru (promień Bohra, $a_0 \approx 0{,}5 \times 10^{-10}$ m), oszacuj stosunek nieliniowej stałej sprężystości $k_q^{(2)}$ do liniowej stałej sprężystości k_q.

10.2. Laser neodymowy (Nd:YAG) emituje promieniowanie o długości fali $\lambda_1 = 1064$ nm. Wiązka ta zostaje rozszczepiona. Jedna z części ulega procesowi generacji drugiej harmonicznej. Następnie wygenerowana w ten sposób wiązka ulega procesowi generacji sumy częstości z drugą (niezmienioną) częścią wiązki pierwotnej. Oblicz długość fali promieniowania opuszczającego układ. Jak inaczej nazywa się ten sumaryczny proces nieliniowy?

10.3. W krysztale nieliniowym zachodzi proces optycznego wzmocnienia parametrycznego. Kryształ pompowany jest wiązką lasera o długości fali $\lambda_p = 355$ nm (trzecia harmoniczna lasera Nd:YAG). Na wejście układu podawany jest również słaby sygnał o długości fali $\lambda_s = 600$ nm. Korzystając z zasady zachowania energii, oblicz długość fali wiązki jałowej λ_i, która zostanie wygenerowana w tym procesie.

10.4. Proces generacji drugiej harmonicznej dla fal płaskich wymaga spełnienia warunku zachowania pędu fotonów: $\hbar\mathbf{k}_{2\omega} = 2\hbar\mathbf{k}_{\omega}$. Udowodnij, że w przypadku propagacji współliniowej obu fal, warunek ten sprowadza się do wymogu zrównania współczynników załamania dla obu częstości: $n(2\omega) = n(\omega)$. Dlaczego w typowych ośrodkach izotropowych wykazujących normalną dyspersję ($dn/d\omega > 0$) spełnienie tego warunku jest niemożliwe bez wykorzystania dwójłomności kryształów?

10.5. Optyczny efekt Kerra opisuje się często fenomenologicznym równaniem na zależny od natężenia współczynnik załamania: $n(I) = n_0 + n_2 I$, gdzie n_0 to liniowy współczynnik załamania, a n_2 to nieliniowy współczynnik załamania (proporcjonalny do $\chi^{(3)}$). Dla topionego kwarcu $n_0 = 1{,}45$, a $n_2 \approx 3 \times 10^{-16}$ cm^2/W. Laser impulsowy emituje impulsy o mocy szczytowej $P = 10$ MW, skupione na plamce o średnicy $d = 10$ μm. Oblicz natężenie światła I w ognisku oraz oszacuj, o ile zmieni się współczynnik załamania Δn w centrum wiązki.

10.6. Zgodnie z regułą Millera, wielkość odpowiedzi nieliniowej drugiego rzędu można powiązać z podatnością liniową: $\chi^{(2)}(2\omega;\omega,\omega) \propto \chi^{(1)}(2\omega)\left[\chi^{(1)}(\omega)\right]^2$. Załóżmy, że dysponujemy dwoma różnymi kryształami niedoskonale przezroczystymi w badanym widmie. Dla pierwszego kryształu liniowa podatność wynosi $\chi_A^{(1)} = 1$, a dla drugiego, wykazującego silniejszą polaryzację, $\chi_B^{(1)} = 2$ (zakładamy dla uproszczenia brak dyspersji, $\chi^{(1)}(\omega) \approx \chi^{(1)}(2\omega)$). Przy założeniu tej samej struktury anharmonicznej (ten sam współczynnik proporcjonalności w regule Millera), ile razy silniejszej podatności nieliniowej $\chi^{(2)}$ należy spodziewać się w krysztale B w stosunku do kryształu A?

Rozdział 11

Źródła Światła

11.1 Wprowadzenie

Zrozumienie mechanizmów leżących u podstaw emisji promieniowania elektromagnetycznego w ujęciu elektrodynamiki klasycznej, a następnie wprowadzenie modelu Lorentza, który z dobrym przybliżeniem opisuje oddziaływanie światła z materią, było celem rozdziału 9. W niniejszym rozdziale omówimy źródła światła wykorzystywane współcześnie w nauce i technice. Nierzadko w tych rozważaniach będziemy uciekać się do kwantowego opisu światła i materii, który pozwoli lepiej zrozumieć naturę różnych źródeł oraz statystykę emitowanego przez nie promieniowania.

Kwantowa natura promieniowania pozwala sklasyfikować źródła światła na podstawie **statystyki emitowanych fotonów**. Źródła chaotyczne, czyli większość występujących w naturze źródeł termicznych oraz niespójnych źródeł wyładowczych, charakteryzują się statystyką opartą na **rozkładzie Bosego-Einsteina**. Emisja w takich układach jest sekwencją niezależnych, losowych zdarzeń spontanicznych. Manifestacją tego zjawiska są fluktuacje natężenia światła oraz **zjawisko grupowania fotonów**. **Funkcja korelacji czasowej drugiego rzędu**, definiowana jako:

$$g^{(2)}(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2}, \tag{11.1}$$

gdzie τ jest opóźnieniem czasowym pomiędzy detekcją dwóch fotonów, a $I(t)$ oznacza chwilowe natężenie promieniowania, dla światła termicznego przyjmuje wartość $g^{(2)}(0) = 2$ (co wynika z relacji momentów dla tego rozkładu, gdzie $\langle I^2\rangle = 2\langle I\rangle^2$), po czym zanika do jedności dla opóźnień $\tau \to \infty$. Oznacza to, że prawdopodobieństwo detekcji drugiego fotonu natychmiast po pierwszym jest dwukrotnie wyższe niż dla zdarzeń całkowicie niezależnych.

Zupełnie odmienną statystykę wykazują źródła koherentne, czyli **lasery**. W przeciwieństwie do źródeł termicznych, w tego typu urządzeniach musi być spełniony określony warunek progowy, określany mianem **progu akcji laserowej**.

Tablica 11.1: Porównanie parametrów statystycznych promieniowania ze źródeł termicznych i laserowych.

Parametr	**Źródło termiczne (chaotyczne)**	**Laser (idealny stan koherentny)**
Statystyka emisji fotonów	Bosego-Einsteina	Poissona
Wariancja liczby fotonów $\langle(\Delta n)^2\rangle$	$\langle n\rangle^2 + \langle n\rangle$	$\langle n\rangle$
Funkcja korelacji 2. rzędu $g^{(2)}(0)$	2 (grupowanie fotonów)	1 (brak grupowania)
Szerokość widmowa $\Delta\nu$	Rzędu THz do PHz	Rzędu Hz do MHz

Powyżej progu zaczyna dominować tzw. **emisja wymuszona**, która determinuje proces generacji promieniowania. Emitowane z lasera fotony znajdują się w tzw. **stanie koherentnym**. W stanie tym iloczyn nieoznaczoności pomiędzy fazą a amplitudą emitowanego pola przyjmuje wartość minimalną, a statystyka emisji fotonów opisywana jest rozkładem Poissona. Dla idealnego światła spójnego wariancja liczby fotonów jest równa jej wartości średniej, a funkcja korelacji wynosi $g^{(2)}(\tau) = 1$ niezależnie od czasu opóźnienia. Oznacza to absolutny brak zjawiska grupowania fotonów oraz drastyczną redukcję fluktuacji amplitudy pola w stosunku do światła chaotycznego. To właśnie przekroczenie ograniczeń naturalnej statystyki emisji termicznej zainicjowało rozwój nowoczesnej fotoniki nieliniowej i optyki kwantowej.

Zestawienie właściwości różnych śródeł światła zostało zaprezentowane w Tablicy 11.1.

11.2 Termiczne źródła światła i promieniowanie ciała doskonale czarnego

Każde ciało, którego temperatura jest wyższa od zera bezwzględnego, nieustannie emituje promieniowanie elektromagnetyczne. Spektrum tego promieniowania (czyli to, jakie długości fal są emitowane najsilniej) zależy jednak w rzeczywistości od wielu skomplikowanych cech materiału. Aby zrozumieć fundamentalne prawa rządzące emisją termiczną, w roku 1859 Gustav Kirchhoff zaproponował koncepcję obiektu wyidealizowanego – **ciała doskonale czarnego**.

Ciało doskonale czarne to abstrakcyjny fizyczny układ, który pochłania całe padające na niego promieniowanie, niezależnie od długości fali i kąta padania (stąd przymiotnik „czarne"). Z termodynamiki wynika jednak bardzo ważny wniosek: skoro ciało doskonale pochłania, w stanie równowagi termicznej musi również doskonale emitować, aby nie ulec nieskończonemu nagrzaniu. Stanowi ono zatem idealny emiter promieniowania – w danej temperaturze T emituje ono

maksymalną możliwą moc promieniowania dla każdej długości fali. Dobrym modelem takiego układu jest szczelnie zamknięta wnęka (piec) rozgrzana do temperatury T, z niewielkim otworem, przez który może wydostawać się światło. Promieniowanie wylatujące przez ten otwór ma charakterystykę promieniowania idealnego ciała doskonale czarnego.

11.2.1 Katastrofa w nadfiolecie

Pod koniec XIX wieku fizycy próbowali opisać rozkład promieniowania emitowanego z wnęki, używając klasycznej termodynamiki oraz elektrodynamiki Maxwella. Wyobraźmy sobie promieniowanie we wnęce jako zbiór stojących fal elektromagnetycznych (tzw. **modów wnęki**). Fale o krótszej długości łatwiej „upakować" we wnęce niż fale długie, tzn. więcej fragmentów o połowie długości fali można zmieścić we wnęce o danej długości dla fal krótszych niż dłuższych. Dokładna analiza matematyczna pokazuje, że gęstość dostępnych modów pola (czyli ich liczba na jednostkową objętość i przedział częstości) gwałtownie rośnie wraz z częstością kołową ω i wyraża się wzorem $g(\omega) = \omega^2/\pi^2 c^3$.

W myśl klasycznego **twierdzenia o ekwipartycji energii**, każdy mod wnęki w stanie równowagi termicznej powinien posiadać taką samą średnią energię $\langle E \rangle = k_{\mathrm{B}}T$, niezależnie od swojej częstości, gdzie k_{B} to stała Boltzmanna. Mnożąc gęstość modów $g(\omega)$ przez tę samą średnią energię na mod, $k_{\mathrm{B}}T$, uzyskujemy klasyczne **prawo Rayleigha-Jeansa**. Opisuje ono gęstość energii promieniowania we wnęce $\rho_{\mathrm{RJ}}(\omega)$ w funkcji częstości kołowej ω:

$$\rho_{\mathrm{RJ}}(\omega) = g(\omega)\langle E \rangle = \frac{\omega^2}{\pi^2 c^3} k_{\mathrm{B}}T. \tag{11.2}$$

Wynik ten natychmiast ujawnia fundamentalny problem dziewiętnastowiecznej fizyki. Prawo Rayleigha-Jeansa przewiduje bowiem, że gęstość promieniowania rośnie do nieskończoności wraz ze wzrostem częstości. Oznaczałoby to, że otwarcie drzwiczek domowego pieca skutkowałoby śmiertelnym poparzeniem promieniami γ, a całkowita energia wnęki byłaby nieskończona. Zjawisko to przeszło do historii jako **katastrofa w nadfiolecie** i – jak się okazało – było bezpośrednim impulsem, który pchnął Maxa Plancka na drogę do stworzenia mechaniki kwantowej.

11.2.2 Kwantowanie pola

Rozwiązanie tego paradoksu, zaproponowane w 1900 roku przez Maxa Plancka, wymagało rewolucyjnego odejścia od klasycznych założeń fizyki. Planck zrozumiał, że średnia energia modów o wysokich częstościach nie może być równa $k_{\mathrm{B}}T$, a musi ona drastycznie maleć, by zapobiec katastrofie. Aby to osiągnąć, sformułował słynny postulat (początkowo traktując go zaledwie jako „matematyczną sztuczkę"): energia nie może być wymieniana pomiędzy atomami ścian wnęki a polem elektromagnetycznym w sposób ciągły. Zgodnie z tą tezą energia jest uwalniana i pochłaniana tylko w dyskretnych porcjach zwanych **kwantami**.

Wielkość pojedynczego kwantu zależy od częstości promieniowania i wynosi:

$$E_f = h\nu = \hbar\omega. \tag{11.3}$$

W klasycznej termodynamice prawdopodobieństwo obsadzenia wyższego stanu energetycznego rządzie się **rozkładem Boltzmanna** (Ludwig Boltzmann 1844–1906) i zależy od czynnika $\exp(-E/k_{\mathrm{B}}T)$. Dla małych częstości (fale radiowe, podczerwień), porcja energii $\hbar\omega$ jest malutka w porównaniu do dostępnej energii termicznej $k_{\mathrm{B}}T$. W takim przypadku podukłady z łatwością wymieniają energię, i wszystko zachowuje się zgodnie z klasyczną zasadą ekwipartycji.

Sytuacja zmienia się diametralnie w nadfiolecie i dla promieni o jeszcze większej częstotliwości. Dla tak wysokich częstości kwant energii $\hbar\omega$ jest znacznie większy od średniej energii cieplnej $k_{\mathrm{B}}T$. Z punktu widzenia fizyki statystycznej układu po prostu „nie stać" na to, by wzbudzić taki wysokoczęstotliwościowy mod (prawdopodobieństwo wzbudzenia $\exp(-\hbar\omega/k_{\mathrm{B}}T)$ jest bliskie zera). Można zatem powiedzieć, że wysokoczęstotliwościowe oscylatory „zamarzają".

Stosując sumowanie po statystyce Boltzmanna dla skwantowanych poziomów energii $E_n = n\hbar\omega$, wyznaczamy nową, kwantową średnią energię oscylatora:

$$\langle E\rangle = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) - 1}. \tag{11.4}$$

Mnożąc tę (poprawną już) wartość średniej energii przez gęstość dostępnych modów elektromagnetycznych, uzyskujemy fundamentalne **prawo Plancka**:

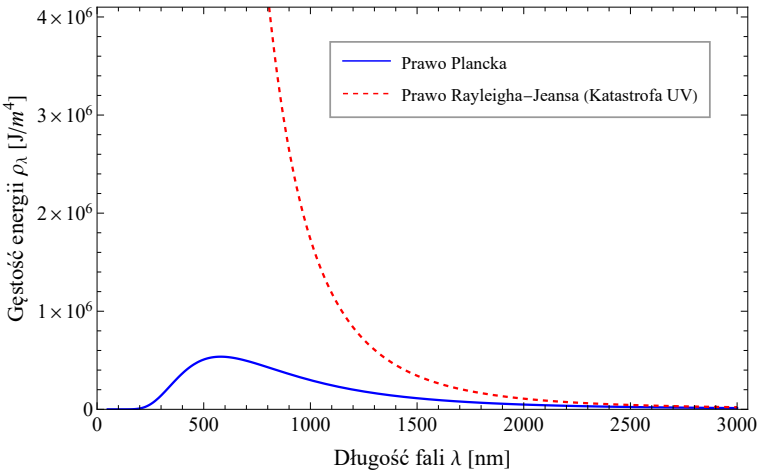
$$\rho(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp\left(\frac{\hbar\omega}{k_{\mathrm{B}}T}\right) - 1}. \tag{11.5}$$

Zauważmy elegancję tego rozwiązania: czynnik eksponencjalny w mianowniku sprawia, że dla dużych ω rozkład gwałtownie zanika do zera (rozwiązując problem katastrofy nadfioletowej). Z kolei dla bardzo małych częstości, mianownik możemy rozwinąć w szereg Taylora ($\exp(x) \approx 1+x$), co sprawia, że wzór Plancka płynnie sprowadza się do klasycznego wzoru Rayleigha-Jeansa. Prawo Plancka stało się historycznym początkiem rewolucji kwantowej.

Różnica pomiędzy widmem przewidywanym dla ciała doskonale czarnego w ramach elektrodynamiki klasycznej oraz mechaniki kwantowej pokazana jest na rys. 11.1.

11.2.3 Rzeczywiste termiczne źródła światła

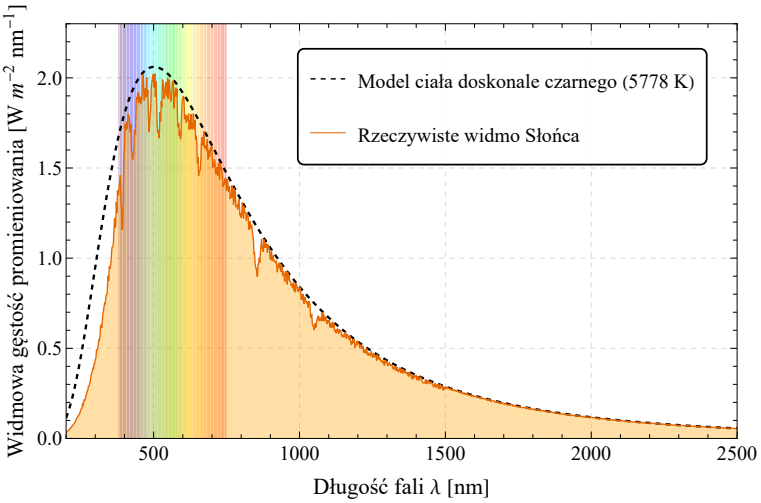
Zanim przejdziemy do dalszych rozważań matematycznych, warto odnieść wyidealizowany model abstrakcyjnego ciała doskonale czarnego do otaczającej nas rzeczywistości. Skoro niniejszy rozdział poświęcony jest badaniu źródeł światła, nasuwa się pytanie: gdzie w przyrodzie i technice odnajdujemy obiekty działające na podstawie opisywanej tu emisji termicznej?



Rysunek 11.1: Rozkład gęstości energii promieniowania ciała doskonale czarnego przewidywany przez teorię klasyczną (prawo Rayleigha-Jeansa) oraz mechanikę kwantową (prawo Plancka). Nieograniczony wzrost energii dla krótkich fal w modelu klasycznym obrazuje absurd fizyczny znany jako katastrofa w nadfiolecie. W rzeczywistości, zgodnie z postulatami Plancka, widmowa gęstość energii dla bardzo krótkich fal (wysokich częstości) gwałtownie dąży do zera.

Najpotężniejszym i najbardziej naturalnym przykładem takich źródeł są gwiazdy. Widmo promieniowania naszego Słońca, którego temperatura powierzchni (fotosfery) wynosi około 5778 K, z ogromną dokładnością pokrywa się z krzywą Plancka (co obrazuje rys. 11.2). Co fascynujące, maksimum emisji Słońca przypada na centralną część widma widzialnego (barwa żółto-zielona), do której ludzkie oko ukształtowało się ewolucyjnie, aby widzieć optymalnie.

Z kolei w technice sztandarowym przykładem jest żarówka z żarnikiem wolframowym. Przepływający prąd elektryczny rozgrzewa w niej cienkie włókno do temperatury około 3000 K. Zgodnie z prawem Plancka, dla takiej temperatury maksimum emisji znajduje się głęboko w paśmie bliskiej podczerwieni. Jedynie niewielki ułamek (około 5%) emitowanej przez żarówkę energii przypada na użyteczny zakres widzialny. Zrozumienie krzywej Plancka tłumaczy zatem bezpośrednio, dlaczego tradycyjne żarówki są tak mało wydajnymi źródłami światła – w istocie stanowią one doskonałe grzejniki na podczerwień, a samo zjawisko widzialnego świecenia jest jedynie ich termicznym „skutkiem ubocznym". Nawet żarzące się drwa w kominku to nic innego jak emitery termiczne, których słaby, czerwonawy blask wprost informuje nas o ich dużo niższej temperaturze (zazwyczaj poniżej 1200 K).



Rysunek 11.2: Porównanie krzywej promieniowania ciała doskonale czarnego dla temperatury 5778 K z rzeczywistym widmem promieniowania słonecznego docierającego do szczytu ziemskiej atmosfery. Niezwykła zgodność obu krzywych stanowi bezsprzeczny empiryczny dowód uniwersalności prawa Plancka. Widoczne „szarpane" spadki w pomierzonym widmie Słońca wynikają ze zjawisk absorpcji światła w chłodniejszych, zewnętrznych warstwach atmosfery gwiazdy.

11.2.4 Prawo Stefana-Boltzmanna i prawo Wiena

Całkowanie rozkładu Plancka po pełnym spektrum pozwala wyznaczyć całkowite natężenie ciała doskonale czarnego. Prowadzi to analitycznie do **prawa Stefana-Boltzmanna** (Josef Stefan 1835–1893):

$$I(T) = \int_0^\infty \frac{c}{4}\rho(\omega, T)\, d\omega = \sigma T^4, \tag{11.6}$$

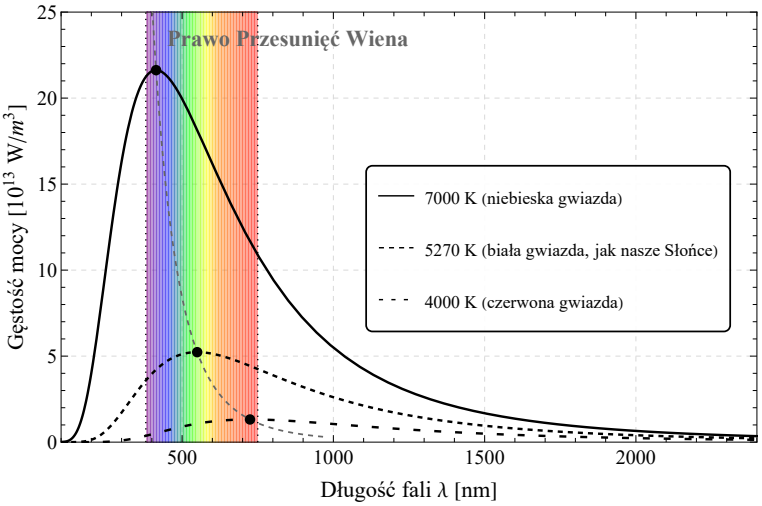
gdzie σ jest stałą Stefana-Boltzmanna. Zależność całkowitej emisji od czwartej potęgi temperatury determinuje wykorzystanie wysokotemperaturowych źródeł plazmowych do generacji niezwykle intensywnych pól radiacyjnych.

Rozkład gęstości promieniowania można przetransformować z domeny częstości ω do domeny długości fali λ. Różniczkowanie tak uzyskanego rozkładu $\rho(\lambda, T)$ względem λ i przyrównanie pochodnej do zera pozwala wyznaczyć wierzchołek krzywej emisyjnej. Rozwiązanie tej relacji transcendentalnej dostarcza **prawa przesunięć Wiena** (Wilhelm Wien 1864–1928):

$$\lambda_{\max} T = b, \tag{11.7}$$

gdzie b jest stałą Wiena. Fizyczna interpretacja tego prawa pozwala ilościowo określić zmianę barwy promieniowania temperaturowego. Wzrost temperatury skutkuje przesunięciem maksimum emisji ze strefy podczerwieni w kierunku

widma optycznego i ultrafioletu. Dlatego rozgrzewany metal najpierw zaczyna świecić na czerwono, następnie na żółto, a najgorętsze gwiazdy wydają się błękitne, co znajduje szerokie zastosowanie w pirometrii i astrofizyce. Rysunek 11.3 przedstawia widma emisji ciała doskonale czarnego dla kilku przykładowych temperatur.



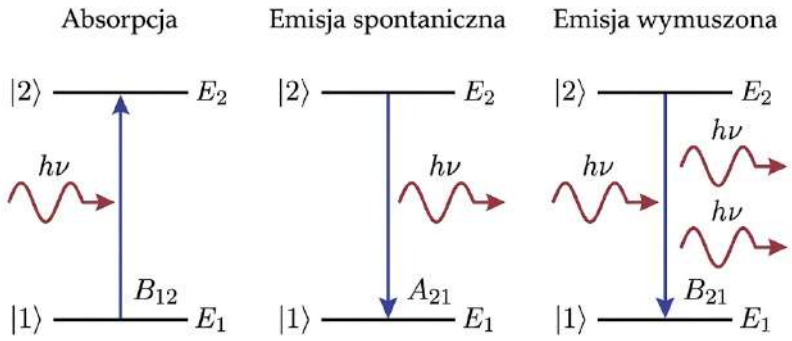
Rysunek 11.3: Widma promieniowania ciała doskonale czarnego (krzywe Plancka) wykreślone dla różnych temperatur. Pole powierzchni pod każdą z krzywych obrazuje całkowitą emitancję, która zgodnie z prawem Stefana-Boltzmanna rośnie proporcjonalnie do czwartej potęgi temperatury (T^4). Z kolei prawo przesunięć Wiena manifestuje się poprzez przesuwanie się wierzchołków krzywych (maksimów emisji) w stronę coraz krótszych fal w miarę nagrzewania się obiektu.

11.3 Niespójne źródła atomowe i półprzewodnikowe

Przejście od makroskopowych źródeł termicznych do selektywnych źródeł o węższym widmie wymaga opisu układów kwantowych na poziomie mikroskopowym. Emisja w takich układach jest pozbawiona fazowej koordynacji, stanowiąc superpozycję zdarzeń niezależnych, stąd wynikowe promieniowanie charakteryzuje się bardzo krótką drogą koherencji.

11.3.1 Emisja spontaniczna

Wyobraźmy sobie układ kwantowy posiadający jedynie dwa stany energetyczne: **stan podstawowy** i **stan wzbudzony**. W stanie równowagi termodyna-



Rysunek 11.4: Oddziaływanie promieniowania z materią: absorpcja, emisja spontaniczna i emisja wymuszona. Strzałki pionowe wskazują kierunek przejścia kwantowego, a faliste reprezentują zaabsorbowane lub wyemitowane fotony.

micznej proporcja obsadzeń tych stanów opisana ww. jest **rozkładem Boltzmanna**:

$$\frac{N_2}{N_1} = \exp\left(-\frac{\Delta E}{k_{\mathrm{B}}T}\right), \tag{11.8}$$

gdzie N_1 i N_2 są obsadzeniami odpowiednio stanu podstawowego i wzbudzonego, ΔE jest różnicą energii pomiędzy tymi stanami, zaś T temperaturą układu. Warto zwrócić uwagę, że obsadzenia te bardzo silnie zależą od różnicy energii ΔE. Przykładowo, w temperaturze pokojowej energia termiczna wynosi zaledwie $k_{\mathrm{B}}T \approx 25$ meV, podczas gdy różnica energii stanów związanych z przejściami elektronowymi w atomach wynosi zwykle kilka elektronowoltów. Sprawia to, że czynnik eksponencjalny jest niezwykle bliski zera ($\Delta E \gg k_{\mathrm{B}}T$), a więc 100% populacji znajduje się w stanie podstawowym, zaś termiczne wzbudzenie elektronu jest niemal całkowicie niemożliwe. Inaczej jest w przypadku przejść pomiędzy poziomami oscylacyjnymi lub rotacyjnymi, gdzie odległości między poziomami są znacznie mniejsze (odpowiednio dziesiątki meV oraz pojedyncze meV). W ich przypadku nawet w temperaturze pokojowej termiczne wzbudzenie wyższych poziomów staje się wysoce prawdopodobne, a obsadzenia sąsiadujących poziomów są do siebie zbliżone.

Jeżeli z uwagi na procesy zewnętrzne (na przykład pochłonięcie energii w wyniku zderzenia niesprężystego) układ kwantowy (np. atom) znajdzie się w stanie wzbudzonym o energii E_2, układ zostaje wytrącony z równowagi termodynamicznej. Ponieważ stan taki jest niestabilny, to po pewnym czasie atom ulega relaksacji, powracając do stanu podstawowego E_1. Proces ten, zwany **emisją spontaniczną**, odbywa się poprzez wyemitowanie kwantu energii (fotonu) o częstości kołowej $\omega = (E_2 - E_1)/\hbar$ (rys. **??**). Co ciekawe z punktu widzenia fundamentów fizyki, relaksacja ta indukowana jest przez oddziaływa-

nie atomu z tzw. **zerowymi fluktuacjami próżni** pola elektromagnetycznego, jest więc sztandarowym przykładem procesu opisywanego w ramach elektrodynamiki kwantowej.

Dla zespołu wielu atomów, szybkość powracania układu do równowagi termodynamicznej (czyli tempo ubywania populacji stanu wzbudzonego N_2) opisywana jest fenomenologicznym **współczynnikiem Einsteina** A_{21} (Albert Einstein 1879–1955):

$$\frac{dN_2}{dt} = -A_{21}N_2. \tag{11.9}$$

Aby wyznaczyć analitycznie wartość tego współczynnika z zasad pierwszych, konieczne jest zastosowanie aparatu rachunku zaburzeń i wywodzącej się z niego **złotej reguły Fermiego**. Reguła ta pozwala precyzyjnie obliczyć prawdopodobieństwo przejścia układu ze stanu początkowego do końcowego pod wpływem sprzężenia z polem elektromagnetycznym. Według niej tempo tego przejścia jest wprost proporcjonalne do dwóch odrębnych, kluczowych składowych. Po pierwsze, do gęstości dostępnych stanów końcowych (w tym przypadku mowa o liczbie dostępnych modów próżni pola elektromagnetycznego, w które może zostać wyemitowany foton). Po drugie, do kwadratu elementu macierzowego, który opisuje wewnętrzną siłę sprzężenia między atomem a polem. Ograniczając się do najsilniejszego członu tego oddziaływania (tzw. **przybliżenie dipolowe**), kluczowym parametrem staje się elektryczny dipolowy moment przejścia $\hat{\mathbf{d}}_{12} = e\langle 1|\hat{\mathbf{r}}|2\rangle$ (daszki nad symbolami oznaczają tu operatory mechaniki kwantowej, zaś $\langle \dots |$ oraz $| \dots \rangle$ to wektory stanu w notacji Diraca).

Fizycznie, element macierzowy momentu dipolowego przejścia $\hat{\mathbf{d}}_{12}$ wynika z kwantowej superpozycji stanu początkowego i końcowego. Zależy on wprost od tego, w jakim stopniu funkcje falowe obu tych stanów przekrywają się w przestrzeni. Rozkład ładunku w takim stanie superpozycji wykazuje oscylującą asymetrię, dzięki czemu atom sprawnie sprzęga się z polem elektromagnetycznym (działając niczym miniaturowa antena). Im większa wartość momentu przejścia, tym szybciej zachodzi emisja. Ostatecznie, na podstawie złotej reguły Fermiego, współczynnik emisji spontanicznej w próżni wyraża się wzorem:

$$A_{21} = \frac{\omega^3 |\hat{\mathbf{d}}_{12}|^2}{3\pi\epsilon_0 \hbar c^3}. \tag{11.10}$$

Widać zatem, że szybkość emisji spontanicznej jest wprost proporcjonalna do trzeciej potęgi częstości emitowanego promieniowania (ω^3) oraz do kwadratu wartości oczekiwanej momentu dipolowego przejścia. Ten silny związek z częstością tłumaczy, dlaczego układy tak chętnie i błyskawicznie emitują promieniowanie wysokoenergetyczne (np. rentgenowskie czy optyczne), podczas gdy stany wzbudzone w reżimie mikrofalowym czy radiowym mogą charakteryzować się niezwykle długim czasem życia.

Z powyższej dyskusji wyłania się najważniejsza cecha emisji spontanicznej: jest to proces na wskroś stochastyczny i niezależny. Każdy atom emituje swój foton w zupełnie losowym momencie, z przypadkową fazą początkową i w wysoce

przypadkowym kierunku w przestrzeni trójwymiarowej. Wypadkowe promieniowanie zliczane z ogromnej liczby takich atomów przypomina szum akustyczny w zatłoczonym pomieszczeniu – jest całkowicie pozbawione globalnej koherencji. Właśnie z tego powodu klasyczne źródła światła, oparte na wyładowaniach elektrycznych czy rekombinacji nośników w półprzewodnikach, o których piszemy w kolejnym rozdziale, generują wyłącznie światło niespójne. Podlega ono **statystyce światła chaotycznego** (termicznego), zdominowanej przez rozkład Bosego-Einsteina, o czym wspominaliśmy na samym początku tego rozdziału.

11.3.2 Lampy wyładowcze i diody elektroluminescencyjne

Klasycznym przykładem wykorzystania opisanego powyżej mechanizmu są lampy wyładowcze wypełnione rozrzedzonymi gazami szlachetnymi (np. neon, argon) lub parami metali (np. sód, rtęć). Układ taki jest „pompowany" energią kinetyczną ze zderzeń niesprężystych z przyspieszanymi w zewnętrznym polu elektrycznym swobodnymi elektronami (emitowanymi z katody, przez którą płynie prąd). Powstające linie widmowe nie są nieskończenie wąskie. Czas życia stanu wzbudzonego $\tau_{\mathrm{sp}} = 1/A_{21}$ wymusza **poszerzenie naturalne linii** (lorentzowskie) do szerokości $\Delta\omega_{\mathrm{nat}} = A_{21}$, co wynika bezpośrednio z zasady nieoznaczoności Heisenberga.

W rzadkich gazach dominuje jednak **poszerzenie dopplerowskie** o charakterze niejednorodnym. Z uwagi na termiczny ruch atomów, częstość emitowanego promieniowania podlega **zjawisku Dopplera**, które sprawia, że foton emitowany w układzie własnym poruszającego się atomu z częstością ω_0, dociera do obserwatora z przesuniętą wartością $\Delta\omega = \omega_0(v/c)$, gdzie v jest prędkością atomu względem obserwatora, mogącą przyjmować zarówno wartości dodatnie, jak i ujemne. Ponieważ prędkości atomów w gazie opisywane są statystycznym rozkładem Maxwella-Boltzmanna, wypadkowe widmo emisji z całego gazu ulega poszerzeniu i przyjmuje kształt profilu Gaussa o pełnej szerokości połówkowej:

$$\Delta\omega_{\mathrm{D}} = \omega_0 \sqrt{\frac{8k_{\mathrm{B}}T\ln 2}{mc^2}}, \tag{11.11}$$

gdzie m jest masą pojedynczego atomu.

Zwiększenie ciśnienia w gazie (co wymusza znacznie częstsze elastyczne zderzenia między atomami) prowadzi do kolejnego efektu dekoherencji, poszerzając dodatkowo emitowane linie widmowe (tzw. jednorodne **poszerzenie zderzeniowe**).

Znaczący postęp w elektronice i optoelektronice przyniósł rozwój **diod elektroluminescencyjnych** (LED). Urządzenia te nie opierają się na emisji termicznej w wysokiej temperaturze (jak np. żarówki żarnikowe), lecz na **elektroluminescencji**, która jest procesem nietermicznej emisji światła wywołanej przepływem prądu przez **złącze pn**. Złącze takie składa się z dwóch obszarów półprzewodnika: typu p (gdzie dominują dodatnie dziury) oraz typu n (gdzie dominują swobodne elektrony). Ich fizyczne zetknięcie prowadzi do swobodnej dyfuzji ładunków i powstania na granicy wąskiego **obszaru zubożonego**

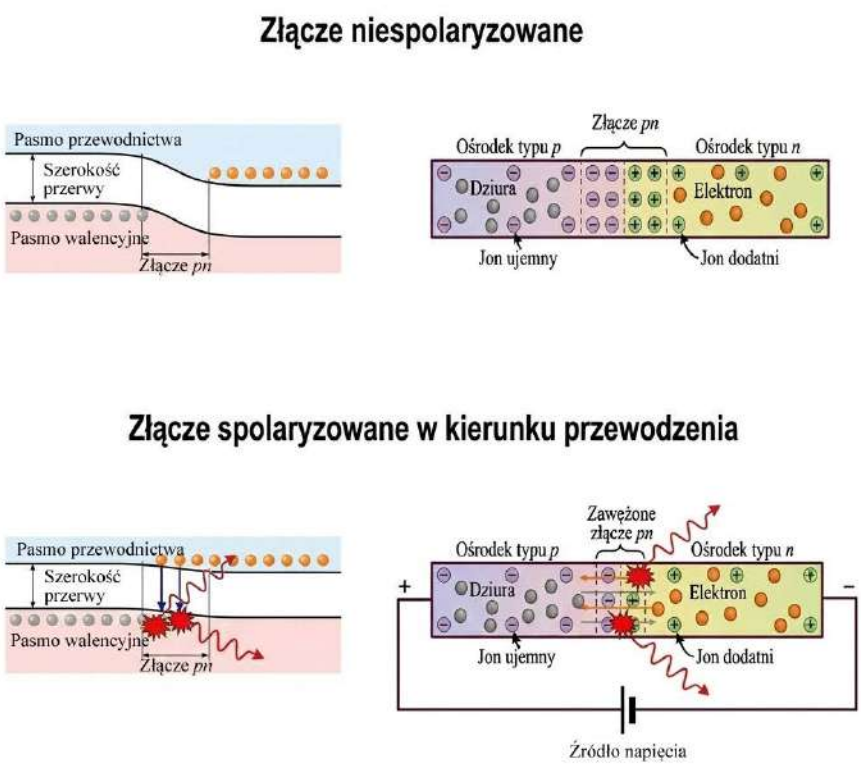
z wolnych nośników. W stanie równowagi termodynamicznej (bez przyłożonego napięcia) w obszarze tym wciąż dochodzi do pojedynczych, rzadkich aktów spontanicznej rekombinacji i emisji fotonów [lewa strona rys. 11.5]. Proces ten jest jednak ściśle równoważony przez termiczną generację nowych par elektron-dziura, w związku z czym złącze makroskopowo nie emituje światła.

Aby dioda LED mogła sprawnie emitować promieniowanie, buduje się ją z materiałów półprzewodnikowych posiadających tzw. **prostą przerwę energetyczną** E_{pw} (np. arsenek galu GaAs, azotek galu GaN). Oznacza to, że na wykresie pasmowym wierzchołek **pasma walencyjnego** (gdzie przebywają dziury) i dno **pasma przewodnictwa** (gdzie swobodnie płyną elektrony) znajdują się w tym samym punkcie przestrzeni pędów (mają ten sam wektor falowy $\mathbf{k}$). Jest to kluczowe zjawisko fizyczne: ponieważ uciekający foton niesie ze sobą znikomo mały pęd, elektron spadający z pasma przewodnictwa do walencyjnego może zrekombinować z dziurą oddając całą swoją energię w postaci światła, bez naruszania zasady zachowania pędu.

Sytuacja zmienia się drastycznie po spolaryzowaniu złącza pn w kierunku przewodzenia — czyli po przyłożeniu zewnętrznego napięcia o polaryzacji przeciwnej do wewnętrznego pola elektrycznego obszaru zubożonego [prawa strona rys. 11.5]. Dochodzi wówczas do istotnego obniżenia dyfuzyjnej bariery potencjału. Umożliwia to gwałtowny przepływ większościowych nośników ładunku przez granicę złącza: elektrony z obszaru n dyfundują do obszaru p, natomiast dziury z obszaru p przenikają do obszaru n. Nośniki te stają się w tych rejonach nadmiarowymi nośnikami mniejszościowymi, stąd zjawisko to nosi nazwę **wstrzykiwania nośników mniejszościowych**. Ta lokalna, masowa akumulacja ładunków całkowicie zaburza stan równowagi układu, wywołując kaskadę procesów rekombinacji spontanicznej i związanej z nią makroskopowej emisji promieniowania, z której na co dzień korzystamy.

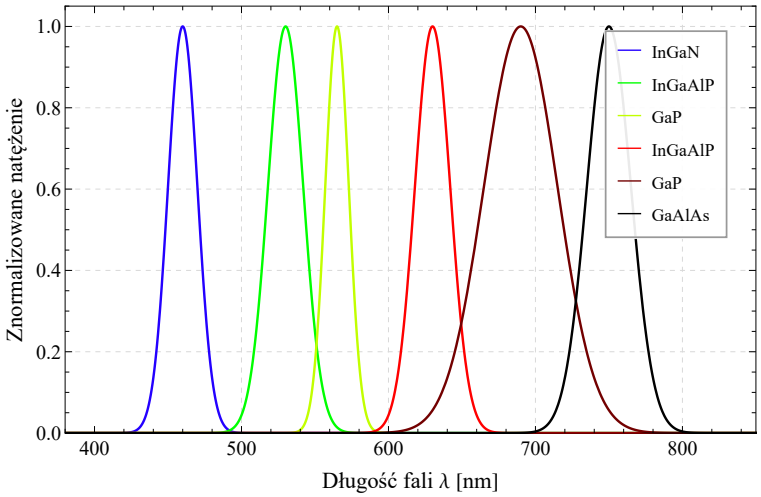
W pierwszym przybliżeniu długość fali emitowanej z diody LED zależy od szerokości przerwy wzbronionej. Jednocześnie kształt profilu widmowego emitowanego promieniowania jest bezpośrednim odzwierciedleniem rozkładu energetycznego wstrzykniętych nośników w pasmach. Pełna szerokość połówkowa linii wynosi w przybliżeniu $\Delta E \approx 1{,}8 k_{\mathrm{B}} T$. W temperaturze pokojowej ($T \approx 300$ K) energia termiczna wynosi $k_{\mathrm{B}} T \approx 25$ meV, co implikuje szerokość energetyczną linii na poziomie około 45 meV. W dziedzinie optycznej przekłada się to bezpośrednio na szerokość widmową od 10 nm do 50 nm, zależnie od centralnej długości fali emisyjnej. Taka charakterystyka plasuje diody LED jako źródła o właściwościach kompromisowych pomiędzy szerokopasmową, ciągłą emisją termiczną a wysoce monochromatycznymi liniami laserowymi. Przykłady widm emisji różnych diod LED zostały przedstawione na rys. 11.6.

Osiągnięcie pełnego spektrum barw, ze szczególnym uwzględnieniem kluczowego obszaru fioletu i błękitu, stało się możliwe dzięki wprowadzeniu struktur studni kwantowych opartych na mieszankach azotku galu i indu (InGaN). Przełom ten stanowił bezpośredni fundament rewolucji we współczesnej technice oświetleniowej i energooszczędnej. Niemniej jednak, z punktu widzenia inżynierii optycznej i fizyki urządzeń, kluczowym wyzwaniem pozostaje optymalizacja **wydajności ekstrakcji światła** z wnętrza struktury. Wysoki współ-



Rysunek 11.5: Zasada działania diody elektroluminescencyjnej (LED). W górnej części widoczny jest schemat struktury złącza *pn* obrazujący przestrzenne rozmieszczenie nośników ładunku. W dolnej części przedstawiono uproszczony diagram pasmowy dla materiału z prostą przerwą energetyczną. Widoczne po lewej stronie złącze niespolaryzowane cechuje się równowagą termodynamiczną (pojedyncze akty rekombinacji). Polaryzacja złącza w kierunku przewodzenia (poniżej) obniża barierę potencjału, co umożliwia potężne wstrzykiwanie nośników i lawinową rekombinację skutkującą jasną emisją fotonów.

czynnik załamania światła materiałów półprzewodnikowych ($n \approx 2{,}5 - 3{,}5$) sprawia, że kąt krytyczny dla całkowitego wewnętrznego odbicia na płaskiej granicy półprzewodnik-powietrze jest skrajnie mały i wynosi zazwyczaj poniżej 20°. W efekcie przeważająca większość wygenerowanych stochastycznie fotonów zostaje uwięziona wewnątrz kryształu i ulega bezpromienistej reabsorpcji oraz termizacji. W celu przełamania tego ograniczenia współczesne przyrządy



Rysunek 11.6: Znormalizowana charakterystyka widmowa emisji diod LED wykonanych z różnych związków w temperaturze pokojowej.

poddaje się zaawansowanym procesom technologicznym, takim jak chemiczne teksturowanie powierzchni, stosowanie soczewek kapsułkujących o dopasowanym współczynniku załamania czy integracja z kryształami fotonicznymi, co pozwala skutecznie wyprowadzić promieniowanie na zewnątrz diody.

11.4 Źródła koherentne – lasery

Laser jest źródłem światła, które w odróżnieniu od emiterów omówionych poprzednio, bazuje nie na dyssypacji energii poprzez rozpad spontaniczny, lecz na powielaniu fotonów w procesie koherentnego oddziaływania fali elektromagnetycznej z ośrodkiem aktywnym, wykorzystując zjawisko **emisji wymuszonej**. Dzięki temu lasery są źródłami światła o niezwykle wysokiej gęstości spektralnej energii (choć niekoniecznie wielkości samej dostarczanej energii), które charakteryzuje się doskonałą spójnością przestrzenną i czasową. Budowa każdego lasera opiera się na trzech fundamentalnych filarach: **ośrodku optycznie czynnym** zapewniającym wzmocnienie fali, **układzie pompowania** generującym stan **inwersji obsadzeń**, oraz **wnęce rezonansowej** (rezonatorze optycznym) odpowiedzialnej za selekcję modową i formowanie geometrycznych parametrów wiązki.

11.4.1 Inwersja obsadzeń i emisja wymuszona

W 1916 roku Albert Einstein zrewolucjonizował kwantową teorię światła i materii, postulując obok absorpcji i emisji spontanicznej istnienie trzeciego procesu – **emisji wymuszonej** (rys. 11.4). Proces ten zachodzi, gdy na atom znajdujący

się w stanie wzbudzonym o energii E_2 pada zewnętrzny, rezonansowy foton o energii odpowiadającej dokładnie odległości między poziomami: $\hbar\omega = E_2 - E_1$. Pole elektryczne fali wywołuje natychmiastowe przejście atomu ze stanu wzbudzonego do stanu podstawowego, czemu towarzyszy emisja wtórnego fotonu. Co kluczowe, foton wtórny jest doskonałą kopią fotonu inicjującego cały proces, tzn, posiada tę samą fazę, częstotliwość, polaryzację oraz kierunek propagacji. Proces ten jest zatem procesem przeciwnym do procesu absorpcji, podczas którego atom znajdujacy się w stanie podstawowym absorbuje foton i przechodzi o stanu wzbudzonego.

Prawdopodobieństwo zajścia procesu absorpcji oraz emisji wymuszonej na jednostkę czasu jest wprost proporcjonalne do spektralnej gęstości energii pola promieniowania $\rho_p(\omega)$ wewnątrz ośrodka. Stałymi proporcjonalności są odpowiednie współczynniki Einsteina B_{12} oraz B_{21}. Rozważając warunki globalnej równowagi termodynamicznej materii z promieniowaniem ciała doskonale czarnego, Einstein wykazał, że szanse na zajście obu tych procesów sterowanych zewnętrznym polem są tożsame, co prowadzi do relacji:

$$B_{12} = B_{21}. \tag{11.12}$$

Taka forma równania osiągana jest przy założeniu, że poziomy energetyczne nie są zdegenerowane (wagi statystyczne stanów są równe jedności).

Żądanie jednoczesnej równowagi termodynamicznej z rozkładem Plancka narzuca również ścisły pomost matematyczny pomiędzy współczynnikiem emisji wymuszonej B_{21} a współczynnikiem emisji spontanicznej A_{21}:

$$\frac{A_{21}}{B_{21}} = \frac{\hbar\omega^3}{\pi^2 c^3}. \tag{11.13}$$

Zależność ta, jawnie proporcjonalna do trzeciej potęgi częstości (ω^3), niesie fundamentalne wyzwanie technologiczne. Wynika z niej bowiem, że w miarę przechodzenia do fal coraz krótszych (dalekiego nadfioletu i fal rentgenowskich) tempo niekoherentnej emisji spontanicznej rośnie, co drastycznie utrudnia konstrukcję wydajnych laserów dla tego typu fal.

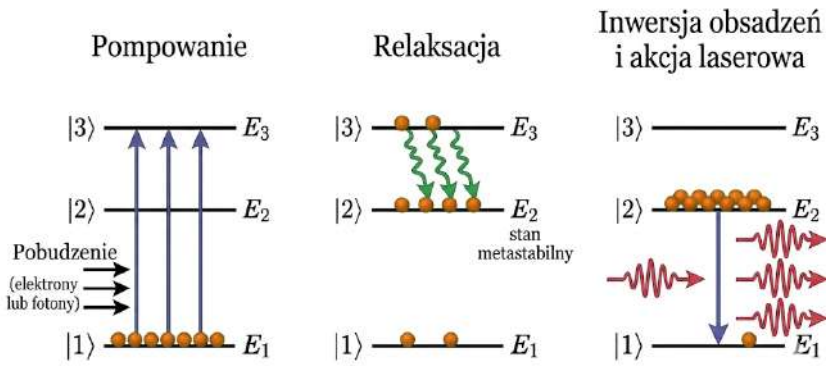
Przejdźmy teraz od opisu pojedynczego atomu do makroskopowego ujęcia fali o spektralnym natężeniu I_ω propagującej się wzdłuż osi z przez ośrodek o gęstościach obsadzeń poziomów wynoszących odpowiednio N_1 i N_2. Zmiana natężenia wiązki na jednostkowej drodze zależy od bilansu konkurujących ze sobą procesów wzmocnienia (emisji wymuszonej) i tłumienia (absorpcji):

$$\frac{dI_\omega}{dz} = \sigma_w(\omega)\left[N_2 - N_1\right] I_\omega \equiv \gamma_e(\omega) I_\omega, \tag{11.14}$$

gdzie $\sigma_w(\omega)$ oznacza przekrój czynny na emisję wymuszoną, natomiast wielkość $\gamma_e(\omega)$ definiujemy jako ułamkowy współczynnik wzmocnienia ośrodka czynnego.

Równanie (11.14) pod względem struktury matematycznej jest identyczne z klasycznym prawem absorpcji Beer-Lamberta (9.27) z tą różnicą, że czynnik stojący przy natężeniu może teraz przyjmować wartości ujemne (odpowiadające absorpcji), ale także dodatnie (odpowiadające emisji wymuszonej).

Aby przechodzące promieniowanie ulegało rzeczywistemu, wykładniczemu wzmocnieniu ($\gamma_e > 0$), czynnik w nawiasie kwadratowym musi być dodatni, co implikuje konieczność spełnienia warunku $N_2 > N_1$. Stan, w którym liczba atomów na poziomie wzbudzonym przewyższa liczbę atomów na poziomie podstawowym, nosi nazwę **inwersji obsadzeń** (rys. 11.7). Ponieważ w warunkach



Rysunek 11.7: Trzy etapy generacji promieniowania w laserze trzypoziomowym. Zewnętrzne pobudzenie przenosi populację atomów z poziomu podstawowego E_1 na poziom E_3. Szybka, bezpromienista deekscytacja na poziom E_2 prowadzi do nagromadzenia tam nośników. Gdy obsadzenie stanu metastabilnego przewyższy obsadzenie stanu podstawowego, osiągnięta zostaje inwersja obsadzeń, a inicjujący foton wywołuje lawinową emisję wymuszoną spójnego promieniowania.

równowagi termodynamicznej (zgodnie z rozkładem Boltzmanna) stan podstawowy jest zawsze najbardziej obsadzony, inwersja jest stanem głęboko nierównowagowym. Jej wytworzenie wymaga dostarczenia z zewnątrz energii, co nazywamy pompowanie układu.

Z równania (11.14) wypływa jeszcze jedna fundamentalna konsekwencja. Próba uzyskania inwersji obsadzeń w prostym układzie dwupoziomowym skazana jest na niepowodzenie. Nawet nieskończenie silne pole pompujące może w takim reżimie doprowadzić jedynie do idealnego wyrównania populacji obu poziomów ($N_1 = N_2$). Wówczas procesy absorpcji i emisji wymuszonej dokładnie się równoważą, współczynnik wzmocnienia spada do zera ($\gamma_e = 0$), a ośrodek staje się optycznie transparentny (jest to zjawisko nasycenia absorpcji). Z tego powodu stabilnie działające systemy laserowe muszą opierać się na bardziej złożonych, asymetrycznych układach kinetycznych (trój- lub czteropoziomowych),

w których poziomy zaangażowane w pompowanie są odseparowane od poziomów roboczych akcji laserowej. Schemat tego procesu został przedstawiony na rys. 11.7.

11.4.2 Rezonatory optyczne i własności wiązki

Sprzężenie zwrotne, absolutnie niezbędne do powstania oscylacji (akcji laserowej), realizowane jest przez wprowadzenie ośrodka optycznie czynnego do wnęki rezonansowej. Najczęściej jest to wnęka typu Fabry'ego-Pérota, zbudowana z dwóch zwierciadeł o współczynnikach odbicia R_1 oraz R_2. Warunek progowy akcji laserowej osiągany jest w takim układzie wtedy, gdy wzmocnienie fali na pełnym dystansie w tę i z powrotem kompensuje całkowite straty rezonatora:

$$R_1 R_2 \exp\left[2(\gamma_{\mathrm{pr}} - \alpha_{\mathrm{str}})L\right] = 1 \implies \gamma_{\mathrm{pr}} = \alpha_{\mathrm{str}} + \frac{1}{2L} \ln\left(\frac{1}{R_1 R_2}\right), \qquad (11.15)$$

gdzie α_{str} oznacza logarytmiczne straty wewnątrzwnękowe (np. rozpraszanie, absorpcja na defektach), γ_{pr} to progowy współczynnik wzmocnienia, a L jest fizyczną długością wnęki.

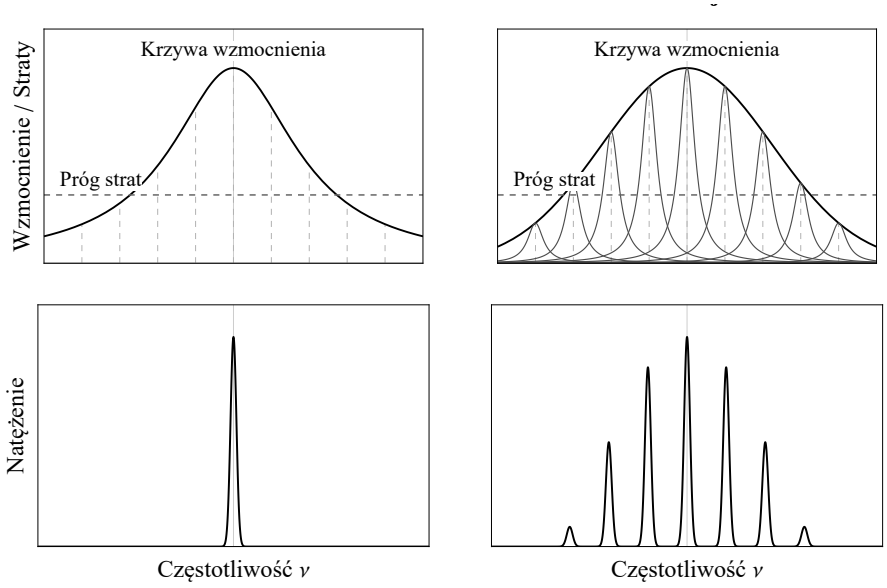
Fala krążąca we wnęce ulega wielokrotnej interferencji. Jak szczegółowo omówiono w rozdziale poświęconym interferometrowi Fabry'ego-Pérota, zjawisko to prowadzi do selekcji tzw. **modów podłużnych**. Oznacza to, że we wnęce powstają tylko takie fale stojąca, dla których całkowita wielokrotność połowy długości fali odpowiada długości wnęki. W konsekwencji w laserze emitowane mogą być tylko takie mody podłużne, dla których częstotliwość ν_j dana jest zależnością:

$$\nu_j = j\frac{c}{2L}, \qquad (11.16)$$

gdzie j jest liczbą całkowitą. W realnym laserze wzbudzają się tylko te mody podłużne, które trafiają w profil wzmocnienia ośrodka $\gamma(\nu)$ i przewyższają próg strat. Odstęp pomiędzy sąsiednimi modami podłużnymi laser zależy wyłącznie od geometrii rezonatora i jest równy FSR rezonatora.

To, jak wiele z modów podłużnych będzie generowanych jednocześnie, zależy w dużej mierze od natury poszerzenia linii emisyjnej ośrodka czynnego. Jeśli ośrodek charakteryzuje się **poszerzeniem jednorodnym** (np. w kryształach w ciele stałym), wszystkie atomy są nierozróżnialne i wspólnie odpowiadają za wzmocnienie światła. W takim układzie zachodzi silna **konkurencja modów**: mod najbliższy środka krzywej narasta najszybciej, zużywa niemal całą inwersję obsadzeń ośrodka i dławi rozwój pozostałych modów. Skutkuje to przeważnie **pracą jednomodową** z wąską linią emisyjną. W ośrodkach a **poszerzeniu niejednorodnym** (np. rozszerzenie dopplerowskie w gazach), poszczególne częstotliwości są emitowane przez zbiory atomów o różnych prędkościach termicznych. Wówczas jeden oscylujący mod „zużywa" jedynie atomy rezonansowe względem niego (zjawisko **wypalania dziur widmowych**), pozostawiając inwersję obsadzeń dostępną dla innych modów podłużnych, co w naturalny sposób skutkuje wielomodową emisją z lasera. Zasadnicza różnica w kształtowaniu się

ostatecznego widma emisji laserów dla obu tych przypadków została schematycznie przedstawiona na rys. 11.8.



Rysunek 11.8: Wpływ charakteru poszerzenia linii emisyjnej na widmo generowane przez laser. Górne panele przedstawiają krzywą wzmocnienia na tle progu strat (pozioma linia przerywana) oraz dozwolonych modów wnęki rezonansowej (pionowe linie). Dolne panele ukazują wynikowe widmo natężenia promieniowania. (Po lewej) W ośrodku o poszerzeniu jednorodnym wszystkie atomy wspólnie odpowiadają za wzmocnienie. Silna konkurencja modów sprawia, że mod centralny zużywa dostępną inwersję, co prowadzi do dławienia pozostałych i pracy jednomodowej. (Po prawej) W ośrodku o poszerzeniu niejednorodnym całkowite wzmocnienie jest sumą wkładów od niezależnych grup atomów (np. o różnych prędkościach). Zjawisko wypalania dziur widmowych zapobiega dławieniu, umożliwiając jednoczesną, wielomodową akcję laserową dla wszystkich modów przekraczających próg strat.

Rozkład przestrzenny pola w płaszczyźnie prostopadłej do osi rezonatora opisują z kolei **mody poprzeczne**. Fundamentalnym rozwiązaniem przyosiowego równania falowego jest wiązka gaussowska (omówiona w rozdziale 3). Warto przypomnieć w tym miejscu, że jej nadrzędną cechą jest zminimalizowana rozbieżność, ograniczona wyłącznie prawami dyfrakcji.

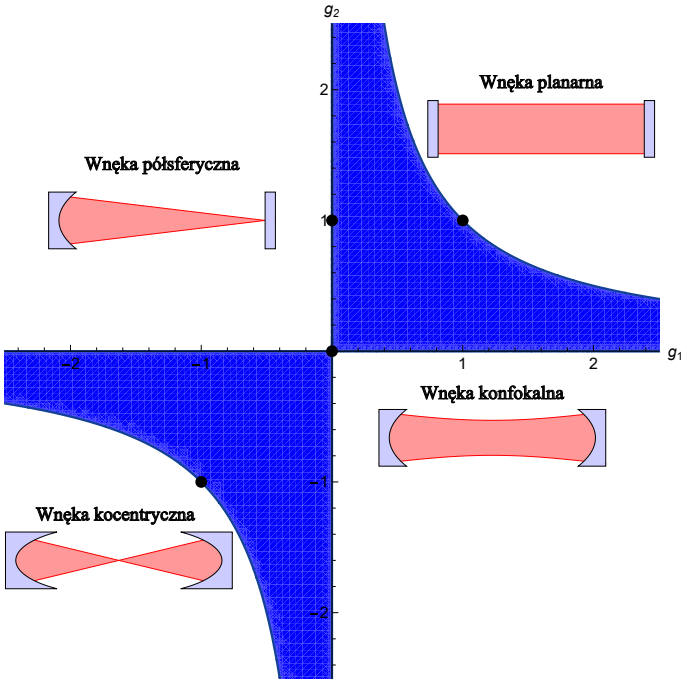
Aby wnęka zbudowana ze sferycznych zwierciadeł mogła stabilnie podtrzymywać propagację takiej wiązki, krzywizny frontów falowych wiązki Gaussa na zwierciadłach muszą idealnie pokrywać się z promieniami krzywizny samych powierzchni odbijających (R_1, R_2). Wykorzystując zaprezentowany uprzednio formalizm optyki macierzowej (w tym celu wprowadzono macierze ABCD), mo-

żemy sformułować fundamentalne kryterium stabilności rezonatora. Poprzez wprowadzenie znormalizowanych parametrów konfiguracyjnych wnęki $g_1 = 1 - L/R_1$ oraz $g_2 = 1 - L/R_2$, warunek stabilności geometrycznej przyjmuje postać:

$$0 \leq g_1 g_2 \leq 1. \tag{11.17}$$

Tak skonstruowany diagram stabilności rezonatora (w osiach g_1, g_2) stanowi niezastąpione narzędzie inżynieryjne przy analizie i projektowaniu systemów laserowych, ułatwiając selekcję modów poprzecznych poprzez wprowadzanie kontrolowanych strat dyfrakcyjnych zależących od fazy Gouya.

Na rys. 11.9 pokazano różne geometrie wnęk wykorzystywane do generacji światła laserowego w zestawieniu z opisanym kryterium stabilności.



Rysunek 11.9: Diagram stabilności wnęk laserowych (w układzie współrzędnych g_1, g_2) oraz najpopularniejsze geometrie rezonatorów typu Fabry'ego-Pérota. Zaznaczony zacieniony obszar, ograniczony ramionami hiperbol ($0 \leq g_1 g_2 \leq 1$), wyznacza parametry stabilnej pracy wnęki – promienie świetlne są wewnątrz okresowo ogniskowane i skutecznie pułapkowane, co pozwala na narastanie akcji laserowej. Dla układów wykraczających poza ten obszar (np. gdy długość wnęki jest zbytnio niedopasowana do promieni krzywizny zwierciadeł) praca lasera staje się niemożliwa z powodu gigantycznych strat dyfrakcyjnych.

11.4.3 Przegląd rodzajów laserów

Istnieje szereg klasyfikacji laserów. Jedna z nich dzieli lasery zgodnie z typem substancji aktywnej wykorzystanej do wytworzenia akcji laserowej:

- **Lasery gazowe**

  Ośrodkiem aktywnym w tego typu układach jest gaz (atomowy, molekularny lub zjonizowany). Zaletą środowiska gazowego jest brak drgań sieci krystalicznej, co w połączeniu ze stosunkowo małą gęstością ośrodka skutkuje niezwykle wąskimi liniami emisyjnymi (poszerzanymi w głównej mierze jedynie przez zjawisko Dopplera). Klasycznym i najlepiej poznanym przykładem jest laser helowo-neonowy (He-Ne). Proces pompowania zachodzi tu wieloetapowo: wyładowanie elektryczne w rurze gazowej przyspiesza elektrony, które poprzez zderzenia wzbudzają atomy helu do długożyciowych **stanów metastabilnych** (długożyjących stanów wzbudzonych). Ponieważ stany te energetycznie niemal idealnie pokrywają się z wyższymi poziomami atomu neonu, energia jest wysoce wydajnie transferowana na atomy neonu w trakcie zderzeń. Ostateczna emisja laserowa (np. popularna czerwona linia o długości 632,8 nm) następuje podczas relaksacji elektronów w atomach neonu na niższe poziomy. Wśród innych powszechnych laserów gazowych warto wymienić lasery jonowo-argonowe (Ar^+) oraz potężne molekularne lasery na dwutlenku węgla (CO_2), generujące duże moce w paśmie dalekiej podczerwieni.

- **Lasery barwnikowe**

  Jako ośrodek czynny wykorzystują złożone cząsteczki barwników organicznych (np. rodamina 6G) rozpuszczone w odpowiednim rozpuszczalniku. Ze względu na występowanie ogromnej liczby gęsto ułożonych poziomów oscylacyjno-rotacyjnych w dużych molekułach, barwniki charakteryzują się niezwykle szerokimi, ciągłymi pasmami fluorescencji. Ta unikalna właściwość sprawia, że lasery barwnikowe są silnie **przestrajalne**. Jest to możliwe przez wprowadzenie do wnęki rezonansowej elementu dyspersyjnego (np. siatki dyfrakcyjnej), pozwalają na płynną zmianę długości emitowanej fali świetlnej w szerokim zakresie. Historycznie odgrywały one kluczową rolę w zaawansowanej spektroskopii.

- **Lasery na ciele stałym**

  Stanowią klasę urządzeń, które jako ośrodek czynny wykorzystują dielektryczne ciała stałe, najczęściej kryształy lub szkła domieszkowane odpowiednimi jonami lantanowców bądź metali przejściowych. Przy ich pomocy możliwe jest emitowanie promieniowania o ogromnej mocy ciągłej oraz impulsowej, co czyni je niezastąpionymi narzędziami w przemysłowej obróbce materiałów, medycynie oraz zaawansowanych badaniach naukowych.

  Sztandarowym i powszechnie stosowanym rozwiązaniem jest laser Nd:YAG (granat itrowo-aluminiowy domieszkowany jonami neodymu Nd^{3+}). Ma-

tryca krystaliczna tego ośrodka działa jako wysoce efektywny układ czteropoziomowy, ułatwiający uzyskanie silnej inwersji obsadzeń przy stosunkowo niskim progu pompowania. Jony neodymu absorbują promieniowanie wzbudzające (pochodzące najczęściej z diod półprzewodnikowych) w wąskim paśmie wokół 808 nm, natomiast stabilna i wydajna akcja laserowa zachodzi w zakresie bliskiej podczerwieni, generując fali o długości 1064 nm.

Osobną, rewolucyjną klasą emiterów krystalicznych jest laser tytanowo-szafirowy (kryształ szafiru domieszkowany jonami tytanu Ti^{3+}). Jego wyjątkowość w kontekście współczesnej fizyki eksperymentalnej wynika z ekstremalnie szerokiego pasma wzmocnienia, rozciągającego się od około 650 nm do ponad 1100 nm. Ta unikalna cecha pozwala no tylko na płynne przestrajanie długości fali generowanego światła w ogromnym zakresie spektralnym, ale stanowi przede wszystkim fizyczny fundament generacji najkrótszych na świecie impulsów promieniowania. Dzięki temu lasery tytanowo-szafirowe stały się podstawowym narzędziem w optyce ultraszybkich impulsów, umożliwiając badanie dynamiki reakcji chemicznych oraz śledzenie procesów wewnątrzatomowych w czasie rzeczywistym.

- **Lasery półprzewodnikowe**

  Choć pod względem fizycznym wykorzystują one materię w stanie stałym, ze względu na fundamentalnie odmienny mechanizm pompowania stanowią zupełnie niezależną klasę urządzeń. W przeciwieństwie do kryształów (które pompujemy optycznie), lasery półprzewodnikowe są pompowane elektrycznie prądem przepływającym przez złącze pn. Zasadniczo, laser tego typu to silnie domieszkowana dioda LED, w której wymuszono inwersję obsadzeń i dodano sprzężenie zwrotne w postaci rezonatora optycznego.

  Podczas gdy w standardowej diodzie LED wstrzykiwane nośniki (elektrony i dziury) ulegają rekombinacji spontanicznej, w diodzie laserowej gęstość prądu jest na tyle wysoka, że w złączu osiągana jest inwersja populacji. Naturalną wnękę rezonansową tworzą tu najczęściej same gładko odłupane równoległe ścianki kryształu (np. GaAs czy stopów InGaN), które dzięki wysokiemu współczynnikowi załamania działają jak lustra częściowo przepuszczalne rezonatora. W ten sposób spełnione są wszystkie warunki do powstania akcji laserowej i emisji promieniowania spójnego.

Dygresja

Bardzo ciekawym przykładem układu generującego światło spójne jest laser na wolnych elektronach. W urządzeniu tym, które w praktyce stanowi potężną, kilometrową infrastrukturę badawczą, ośrodkiem czynnym nie jest gaz, kryształ ani złącze półprzewodnikowe. Funkcję tę pełnią elektrony poruszające się swobodnie w próżni z prędkościami bliskimi prędkości światła. Wiązka takich elektronów zostaje wstrzyknięta do wnętrza specjalnej, pe-

riodycznej struktury magnetycznej zwanej undulatorem. Przechodząc przez ten układ naprzemiennie ułożonych magnesów elektrony wykonują swego rodzaju szybki slalom w próżni.

Przejście przez undulator sprawia, że przyspieszane poprzecznie na każdym zakręcie elektrony zaczynają emitować silne promieniowanie synchrotronowe. Co kluczowe, emitowana fala świetlna zaczyna oddziaływać z gęstością samej paczki elektronów, powodując ich mikro-grupowanie w warstwy odległe o długość fali. W rezultacie elektrony zaczynają emitować promieniowanie całkowicie spójnie i w zgodnej fazie, a natężenie światła narasta lawinowo wzdłuż sekcji undulatorowej w procesie samowzmocnienia.

Kluczową zaletą laserów na wolnych elektronach jest niedostępna dla innych systemów przestrajalność długości fali. Zmieniając energię kinetyczną wstrzykiwanych elektronów lub modyfikując odstęp między magnesami undulatora, można płynnie przestrajać generowane promieniowanie od fal milimetrowych, przez zakres widzialny i ultrafiolet, aż po twarde promieniowanie rentgenowskie – i to bez wymiany jakiegokolwiek elementu optycznego. Najnowocześniejsze współczesne instalacje tego typu generują ekstremalnie jasne błyski rentgenowskie o czasie trwania rzędu zaledwie kilku femtosekund (10^{-15} s). Pozwala to naukowcom dosłownie „fotografować" pojedyncze molekuły białek w trakcie reakcji chemicznej, śledzić ruchy atomów w czasie rzeczywistym oraz rekonstruować trójwymiarową strukturę wirusów z próbek zbyt małych dla klasycznej krystalografii.

11.4.4 Lasery impulsowe

Jak wykazano we wcześniejszych rozważaniach, fundamentalną cechą promieniowania laserowego pracującego w układzie ciągłym jest jego jego monochromatyczność. Istnieje jednak niezwykle ważna klasa urządzeń – **lasery impulsowe** – które zamiast stałego strumienia światła emitują regularne błyski (impulsy) promieniowania o ogromnej mocy chwilowej (choć niekoniecznie dużej energii). Spektrum takiego promieniowania charakteryzuje się bardzo dużą szerokością widmową, co z pozoru stoi w sprzeczności z klasyczną ideą monochromatyczności. Natura tego zjawiska jest jednak zupełnie inna niż w przypadku szerokopasmowych źródeł termicznych czy diod LED.

Aby zrozumieć fizykę generacji impulsów, należy odwołać się do aparatu matematycznego transformaty Fouriera, która nierozerwalnie wiąże ze sobą dziedzinę czasu i częstotliwości. Zgodnie z relacją nieoznaczoności $\Delta t\,\Delta\nu \geq 1$, im krótszy jest czas trwania danego impulsu (Δt), tym szersze musi być jego widmo częstotliwości ($\Delta\nu$). Oznacza to bezpośrednio, że krótkiego impulsu światła nie da się uformować z pojedynczej fali o idealnie określonej częstotliwości. W przeciwieństwie jednak do źródeł klasycznych, W laserach impulsowych, impuls powstaje w wyniku jednoczesnej, koherentnej superpozycji wielu różnych modów podłużnych (wielu dobrze określonych częstotliwości).

Kluczem do wymuszenia jednoczesnej oscylacji tak wielu częstotliwości jest wykorzystanie ośrodka czynnego o odpowiednio szerokim paśmie wzmocnienia

(takiego jak wspomniany wcześniej kryształ tytanowo-szafirowy), które pokrywa tysiące dyskretnych modów wnęki rezonansowej. W klasycznym laserze mody te oscylują niezależnie, a ich fazy początkowe fluktuują w sposób losowy. W rezultacie wypadkowe pole elektromagnetyczne ulega przypadkowej interferencji, przypominając chaotyczny szum o stałej średniej mocy. Sytuacja zmienia się diametralnie, gdy do układu wprowadzony zostanie mechanizm wymuszający stacjonarną, sztywną relację między fazami wszystkich sąsiadujących modów. Fale te zaczynają wówczas konstruktywnie interferować tylko w jednym, precyzyjnie określonym punkcie przestrzeni i czasu, ulegając całkowitemu wygaszeniu w pozostałych miejscach. W ten sposób ciągłe promieniowanie lasera zostaje przekształcone w okresowy ciąg ultrakrótkich impulsów. Technika ta nosi nazwę **synchronizacji modów**.

Dla wygenerowania impulsów piko-, a nawet femtosekundowych przy użyciu synchronizacji modów, pole elektryczne fali przyjmuje matematyczną postać sumy:

$$E(t) = \sum_{j=-N/2}^{N/2} E_j e^{-i(\omega_0 + j\Delta\omega)t}. \tag{11.18}$$

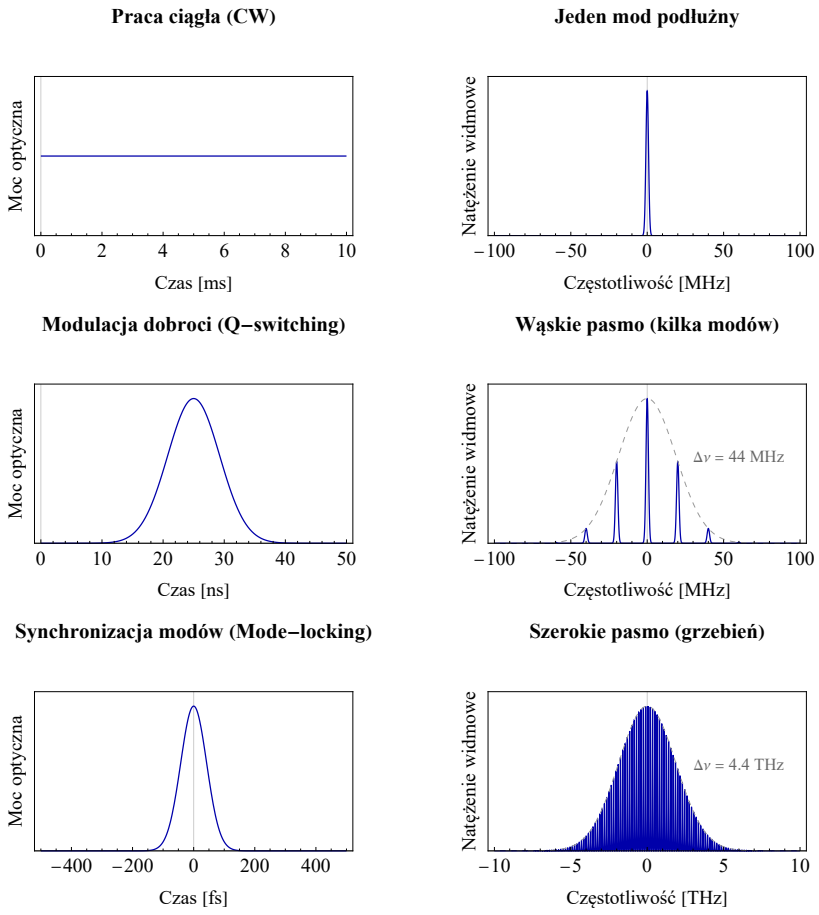
Minimalny czas trwania tak wygenerowanego impulsu jest ściśle limitowany szerokością dostępnego pasma wzmocnienia ośrodka ($\Delta t \propto 1/\Delta\nu$).

Odrębną metodą kompresji energii w czasie, pozwalającą na generację impulsów o nanosekundowym czasie trwania i ekstremalnie wysokich pikach mocy, jest **modulacja dobroci rezonatora**. Technika ta polega na celowym, okresowym blokowaniu optycznego sprzężenia zwrotnego (stan niskiej dobroci), podczas gdy układ pompujący nieustannie zasila ośrodek. Brak możliwości zajścia lawinowej emisji wymuszonej pozwala na skumulowanie gigantycznej inwersji obsadzeń, znacznie przekraczającej klasyczny próg akcji laserowej. Następnie szybki modulator elektrooptyczny (np. komórka Pockelsa) lub pochłaniacz nasycalny gwałtownie przywraca pełną transmisję wnęki. Skutkuje to natychmiastowym uwolnieniem zmagazynowanej energii w postaci pojedynczego, lawinowego impulsu o mocy sięgającej nierzadko dziesiątek megawatów.

Unikalne właściwości światła z laserów impulsowych – przede wszystkim gigantyczna moc szczytowa przy zachowaniu relatywnie niskiej mocy średniej – otworzyły zupełnie nowe horyzonty aplikacyjne w nauce i przemyśle. Zjawisko to znalazło rewolucyjne zastosowania u podstaw diagnostyki układów biologicznych, gdzie potężne zagęszczenie fotonów w czasie pozwala na obrazowanie za pomocą mikroskopii wielofotonowej bez termicznego uszkadzania żywych tkanek. Z kolei w precyzyjnej mikrochirurgii laserowej (np. w korekcji wzroku) oraz w obróbce materiałów, femtosekundowe impulsy pozwalają na tzw. zimną ablację – materiał odparowuje tak szybko, że ciepło nie zdąży rozprzestrzenić się do otaczającego obszaru, co niemal całkowicie eliminuje strefę wpływu ciepła.

Dygresja

Jednym z najbardziej fascynujących i technologicznie zaawansowanych przed-



Rysunek 11.10: Porównanie reżimów pracy lasera w dziedzinie czasu (lewa kolumna) oraz odpowiadających im widm promieniowania (prawa kolumna.

sięwzięć współczesnej fizyki eksperymentalnej jest próba okiełznania energii gwiazd za pomocą zaawansowanych technologii laserowych. Sercem tych badań jest potężna instalacja National Ignition Facility (NIF), zlokalizowana w Lawrence Livermore National Laboratory w Stanach Zjednoczonych. Infrastruktura NIF, zajmująca powierzchnię trzech boisk piłkarskim, kryje w sobie gigantyczny system, który ze wspólnego generatora początkowego formuje i rozbudowuje 192 niezależne, perfekcyjnie zsynchronizowane linie potężnych wzmacniaczy laserowych.

Nadrzędnym celem tej gigantycznej maszyny optycznej jest wywołanie kontrolowanej fuzji termojądrowej i otwarcie ludzkości drogi do bezpiecznego, czystego oraz praktycznie niewyczerpanego źródła energii. Zasada działania opiera się na koncepcji bezwładnościowego uwięzienia plazmy. Wszystkie 192 wiązki laserowe, po przejściu przez setki metrów szkieł neodymowych i konwersji do zakresu ultrafioletu, zostają jednocześnie, z dokładnością do pojedynczych pikosekund, skierowane do wnętrza milimetrowej komory. Znajduje się tam kapsułka zawierająca termojądrowe paliwo – izotopy wodoru (deuter i tryt).

Gwałtowne uderzenie potężnego impulsu o łącznej energii przekraczającej 2 MJ wywołuje natychmiastową, ekstremalnie symetryczną implozję kapsułki. Materia zostaje ściśnięta do gęstości setki razy większej od gęstości ołowiu, a temperatura w punkcie centralnym przekracza 100 milionów stopni Celsjusza. Na ułamek sekundy na Ziemi powstają warunki panujące w samym jądrze Słońca. Historyczne sukcesy zespołu NIF, polegające na osiągnięciu tzw. zapłonu termojądrowego (uzyskaniu dodatniego bilansu energetycznego z samej fuzji), stanowią empiryczny dowód, że zaawansowane systemy optyczne mogą w nadchodzących dekadach zrewolucjonizować światową energetykę.

11.5 Źródła światła nieklasycznego

Wszystkie omówione do tej pory źródła promieniowania – od makroskopowych emiterów termicznych po zaawansowane układy laserowe – znajdują pełny opis w ramach klasycznej elektrodynamiki falowej lub optyki półklasycznej, w której pole elektromagnetyczne traktuje się jako ciągłą falę. Nawet idealne światło laserowe, mimo swojej spójności, wykazuje fluktuacje natężenia ograniczone przez tzw. **szum śrutowy**, opisany statystyką Poissona. Współczesna optyka kwantowa i technologie informatyczne wymagają jednak źródeł o pełnej kwantowej charakterystyce – tzw. **źródeł światła nieklasycznego**, których właściwości statystyczne i korelacje nie mają żadnego odpowiednika w fizyce klasycznej.

Sztandarowym przykładem takiego układu są **źródła pojedynczych fotonów**. W idealnym przypadku emitują dokładnie jeden foton na jedno impulsowe wzbudzenie zewnętrzne, wykazując zjawisko tzw. **antygrupowania fotonów**, gdzie funkcja korelacji czasowej drugiego rzędu przyjmuje wartość $g^{(2)}(0) \to 0$. Realizacja takich źródeł opiera się najczęściej na izolowanych układach dwupoziomowych, takich jak pojedyncze kropki kwantowe, centra barwne w diamencie

(np. centra NV) czy pojedyncze jony uwięzione w pułapkach magnetycznych. Światło to stanowi fundament bezpiecznej kryptografii kwantowej, uniemożliwiając podsłuch oparty na podziale intensywności wiązki.

Kolejną kluczową klasą są źródła generujące **fotony splątane**. Wykorzystują one procesy nieliniowe, takie jak **spontaniczny podział częstotliwości**, zachodzące w kryształach anizotropowych. Foton pompy o wysokiej energii ulega w nich rozpadowi na parę fotonów o niższych energiach, których stany kwantowe (np. polaryzacja) są ze sobą nierozerwalnie skorelowane, niezależnie od odległości dzielącej cząstki. Układy te są kluczowe w eksperymentach nad **teleportacją kwantową** oraz w architekturze komputerów kwantowych.

Warto również wspomnieć o **świetle ściśniętym**, w którym na mocy relacji nieoznaczoności Heisenberga udaje się zredukować fluktuacje kwantowe (szum) jednej ze składowych pola (np. amplitudy lub fazy) poniżej poziomu szumu próżni, kosztem zwiększenia nieoznaczoności w składowej ortogonalnej. Światło ściśnięte pozwala na przekroczenie standardowego limitu kwantowego w pomiarach interferometrycznych, co znalazło spektakularne zastosowanie m.in. w detektorach fal grawitacyjnych LIGO, radykalnie zwiększając ich zasięg obserwacyjny kosmosu.

Podsumowanie rozdziału

W niniejszym rozdziale dokonano szczegółowego przeglądu fizycznych mechanizmów leżących u podstaw generacji fali elektromagnetycznej, przeprowadzając czytelnika od makroskopowych emiterów klasycznych po zaawansowane źródła kwantowe. Analiza promieniowania ciała doskonale czarnego oraz fundamentalnych praw Kirchhoffa, Plancka i Wiena pozwoliła zrozumieć naturę emisji termicznej, stanowiącej podłoże działania gwiazd czy klasycznych żarników wolframowych. Pokazano, że ograniczenia statystyki chaotycznej i związany z nią problem katastrofy w nadfiolecie stały się historycznym katalizatorem narodzin mechaniki kwantowej.

Przejście na poziom mikroskopowy umożliwiło zdefiniowanie procesów emisji spontanicznej w izolowanych układach atomowych oraz w złączach półprzewodnikowych, co stanowi fizyczny fundament działania lamp wyładowczych oraz współczesnych diod LED. Kwintesencją rozdziału było jednak wprowadzenie pojęcia emisji wymuszonej, sformułowanej przez Einsteina, która stała się podwaliną technologii laserowej. Zaprezentowane kryteria progowe wzmocnienia, selekcja modów podłużnych i poprzecznych we wnękach rezonansowych oraz techniki pracy impulsowej – modulacja dobroci Q oraz synchronizacja modów – zobrazowały, w jaki sposób współczesna inżynieria optyczna potrafi formować wiązki o skrajnie wysokiej spójności i potężnej mocy chwilowej. Rozdział zamknęło zarysowanie perspektyw optyki kwantowej poprzez wprowadzenie źródeł światła nieklasycznego, co dowodzi, że ewolucja źródeł światła nieustannie przesuwa granice możliwości poznawczych fizyki i współczesnej technologii informatycznej.

Podsumowanie

Głównym celem niniejszego podręcznika było zapoznanie Czytelnika z fundamentalnymi pojęciami optyki. Zaprezentowany materiał skupiał się przede wszystkim na prawach i zjawiskach optyki falowej, które stanowią rzetelną bazę dla tej dziedziny. Choć opisywane mechanizmy opierają się na wiedzy rozwijanej w dużej mierze do końca XIX wieku, starano się je systematycznie poszerzać o dygresje ukazujące ich rolę we współczesnej fizyce. Stąd w tekście znalazły się odwołania do optyki kwantowej (m.in. splątania kwantowego), fotoniki, fizyki laserowej oraz nowoczesnych układów eksperymentalnych, które pozwalają testować fundamentalne prawa przyrody i poszukiwać nowej fizyki, wykraczającej poza Model Standardowy.

Nadrzędnym celem tych zabiegów było ukazanie, że optyka nie jest zamkniętym i ugruntowanym historycznie rozdziałem nauki. Przeciwnie — to niezwykle dynamicznie rozwijająca się dyscyplina, która nieustannie przesuwa granice ludzkiego poznania i prowadzi do opracowania wielu przełomowych technologii. Wskazanie tego szerszego, aplikacyjnego i badawczego horyzontu ma stanowić dla Czytelnika zachętę do dalszego zgłębiania tematu i sięgnięcia po literaturę bardziej zaawansowaną. Zrozumienie tych nowoczesnych, fascynujących zjawisk nie byłoby jednak możliwe bez uprzedniego opanowania klasycznych podstaw, którym z powodzeniem poświęcono tę książkę.

Indeks alfabeczny

aberracja
    chromatyczna, 52
    sferyczna, 51, 52
aberracja soczewek, 49
aberracja układu optycznego, 51
aktywność optyczna, 216
amplitudowe współczynniki Fresnela, 85
amplitudowe współczynniki Fresnela dla składowej prostopadłej, 85
analiza widmowa światła, 141
antygrupowanie fotonów, 281
apertura numeryczna, 44, 49
apochromat, 52
astygmatyzm, 53

barwy strukturalne, 151
bipryzmat Fresnela, 125

całkowite wewnętrzne odbicie, 24, 43, 89, 90
    udaremnione, 92
chiralność, 217
chromatyczna zdolność rozdzielcza, 188
ciało doskonale czarne, 259
ciśnienie promieniowania, 13
czas koherencji, 126
czoło fali, 20
czynnik nachylenia, 164, 169
częstotliwość fali, 7
częstość kołowa, 7
częstość plazmowa, 238

destruktywna interferencja, 116
diagram promieni, 38
dichroizm, 208
dioda elektroluminescencyjna, 267
dioda LED, 267, 277
dipol elektryczny, 207, 230, 231
dobroć wnęki, 145
droga koherencji światła, 126
droga optyczna, 23
droga spójności światła, 126
dualizm korpuskularno-falowy, xiv
dudnienia, 110
dwójłomność, 209
    kołowa, 217, 220
dyfrakcja Fraunhofera, 173
dyfrakcja Fresnela, 173
dyfrakcja na fali akustycznej – reżim Bragga, 193
dyfrakcja na fali akustycznej – reżim Ramana-Natha, 193
dyfrakcja na jednowymiarowej szczelinie, 175
dysk Airy'ego, 179
dyspersja, 41, 42, 240
dyspersja anomalna, 115, 237, 240
dyspersja kątowa, 42, 187
dyspersja normalna, 115, 237, 240
dzielnik wiązki, 117
długość fali, 7
długość Rayleigha, 65

efekt Faradaya, 220
    nieliniowy, 222
efekt Kerra, 219
efekt naskórkowy, 101
efekt Pockelsa, 219

eksperyment Younga, 123
ekstynkcja polaryzatora, 211
elektrodynamika klasyczna, 226
elektrodynamika kwantowa, 226, 266
elektromagnetycznie indukowana przezroczystość, 255
element dyspersyjny, 187
elipsoida współczynnika załamania, 211
elipsometria, 103
emisja spontaniczna, 265
emisja wymuszona, 151, 270
etalon, 149

fala cylindryczna, 63
fala elektromagnetyczna, 1
fala monochromatyczna, 5
fala płaska, 57, 66
fala sferyczna, 60, 66, 155
fala stojąca, 109
fala Stokesa, 255
fala zanikająca, 92
fale wtórne, 155
finezja interferometru, 140
fluktuacje natężenia światła, 258
fonon, 192
formalizm macierzowy ABCD, 26
formalizm Stokesa, 203
formalizm wektorów Jonesa, 200
fotoelastyczność, 218
foton, 226, 265
fotoniczna przerwa wzbroniona, 45
fotonika nieliniowa, 252
fotony splątane, 282
front falowy, 57, 155
funkcja korelacji czasowej drugiego rzędu, 258, 281
funkcje Hankela, 62

generacja drugiej harmonicznej, 253
generacja różnicy częstości, 253
generacja sumy, 253
generacja trzeciej harmonicznej, 254
granica rozdzielczości, 180
grupowanie fotonów, 258
grupowy współczynnik załamania, 115
grzebień częstotliwości, 255
gęstość energii pola, 10
głębokość wnikania w przewodnikach, 100

holografia, 127

indukcja pola elektrycznego, 1
indukcja pola magnetycznego, 1
interferencja przez podział amplitudy, 118
interferencja przez podział frontu falowego, 118
interferencja wielowiązkowa, 134
interferogram, 145
interferometr Fabry'ego-Pérota, 273
interferometr Fabry'ego-Perota, 134
interferometr Jamina, 123
interferometr konfokalny Fabry'ego-Perota, 147
interferometr Macha-Zehndera, 220
interferometr Michelsona, 118, 121
inwersja obsadzeń, 45, 270, 272
iryzacja, 151
izolator optyczny, 220

katastrofa w nadfiolecie, 260
koma, 52
kompensator Babineta, 216
komputer kwantowy, xiv, 282
komórka Pockelsa, 279
konkurencja modów, 273
konstruktywna interferencja, 116
kontrast prążków interferencyjnych, 138
konwencja kartezjańska, 25
konwencja znaków, 25
kropka kwantowa, 281
kryształ dwuosiowy, 211
kryształ elektrooptyczny, 121

kryształ fotoniczny, 45
kryształ izotropowy, 211
kryształ jednoosiowy, 211
kryterium Abbego, 181
kryterium Rayleigha, 180, 188
kryterium szerokości połówkowej, 141
kwant energii, 260
kwant promieniowania, xiv
kąt akceptacji, 44, 181
kąt Brewstera, 87, 206
kąt graniczny, 24, 43, 89

laser, 258
    przestrajalny, 276
laser impulsowy, 278
laser jednomodowy, 273
lasery barwnikowe, 276
lasery gazowe, 276
lasery na ciele stałym, 276
lasery półprzewodnikowe, 277
lasery światłowodowe, 45
liczba falowa, 19
luminescencja, 267
lustro dielektryczne, 149
lustro Lloyda, 124
lustro półprzepuszczalne, 118

macierz Jonesa
    płytki fazowej, 214
macierz Müllera
    płytki fazowej, 215
macierz obrotu wektora Stokesa, 215
macierze Jonesa, 201
macierze Müllera, 204
magnetometria optyczna, 222
mechanika kwantowa, xiv
metoda wskazów, 136, 157
mieszanie czterech fal, 254
mikroskop, 48
mikroskopia polaryzacyjna, 218
mikroskopia superzdolności rozdzielczej, 182
mikroskopia TIRF, 92
minimum dyfrakcyjne, 179
mod wnęki, 260
model Lorentza, 231
modulacja dobroci rezonatora, 279
modulator akustooptyczny, 191
modulator elektrooptyczny, 220
modulator natężenia światła, 121
mody lasera
    podłużne, 273
    poprzeczne, 274
mody własne ośrodka, 217
moment dipolowy, 231
moment pędu fotonu, 15
multipleksowanie sygnałów z podziałem długości fali, 45

National Ignition Facility, 281
natężenie pola elektrycznego, 1
natężenie pola magnetycznego, 1
natężenie światła, 11
nieliniowy efekt Kerra, 254
nieodwracalność
    efekt Faradaya, 220
niezmienniczość względem symetrii odbicia w czasie, 93

obiektyw mikroskopu, 48
obiektyw teleskopu, 49
obraz pozorny, 39
obraz rzeczywisty, 39
obszar zubożony, 267
ognisko, 32
ogniskowa, 32
ograniczenie/limit fourierowski, 126
okienka Brewstera, 90
oko, 47
okular mikroskopu, 48
okular teleskopu, 49
optyczne wzmocnienie parametryczne, 253
optyka adaptatywna, 50
optyka falowa, xiv
optyka fourierowska, 194
optyka geometryczna, 21, 22
optyka kwantowa, xiv, 252

opóźnienie fazy Gouya, 69
orbitalny moment pędu światła, 74
osie główne ośrodka, 210
oś optyczna, 25
oś szybka, 214, 215
oś wolna, 214
ośrodek anizotropowy, 209
ośrodek liniowy, 106
ośrodek nieliniowy, 106
ośrodek optycznie czynny, 270
ośrodki dielektryczne, 2

pasmo przewodnictwa, 268
pasmo walencyjne, 268
plamka Arago, 160
plamka Poissona, 158, 160
podatność elektryczna, 231, 235, 247
  tensor, 209
pokrycia antyrefleksyjne, 151
polaroid, 208
polaryzacja p, 84
polaryzacja s, 84
polaryzacja eliptyczna, 9, 102, 112, 200
polaryzacja fali, 8
polaryzacja kołowa, 8, 200
polaryzacja liniowa, 8, 200
polaryzacja ośrodka, 209, 247
polaryzacja światła, 200
  odbicie, 206
polaryzacyjny dzielnik wiązki, 207
polaryzator, 203
polaryzator Glana, 211
pole bliskie, 172
pole dalekie, 172
pole elektromagnetyczne, xiv
pompa optyczna, 270
poprzeczna fala elektromagnetyczna, 59
poprzeczny mod pola elektromagnetycznego, 72
poszerzenie jednorodne, 273
poszerzenie niejednorodne, 273
poszerzenie zderzeniowe, 267
powiększenie układu optycznego, 33
prawo ABCD, 70
prawo Ampère'a, 80
prawo Brewstera
  dla naprężeń, 218
prawo Faradaya, 80
prawo Gaussa, 79, 209
prawo Lamberta-Beera, 238
prawo Malusa, 213
prawo odbicia, 23, 83
prawo Ohma, 97
prawo Rayleigha-Jeansa, 260
prawo Snella, 24, 84
prawo Stefana-Boltzmanna, 263
prawo Wiena, 263
prawo załamania, 24, 84
promienie przyosiowe, 64
promieniowanie synchrotronowe, 229
promień gięcia światłowodu, 43
promień nadzwyczajny, 211
promień przewężenia wiązki gaussowskiej, 66
promień zwyczajny, 210
promień świetlny, 22
propagacja nadświetlna, 241
pryzmat, 41
pryzmat odbijający, 42
pryzmat rozpraszający, 41
przedział spektralny, 140, 147
przenikalność elektryczna, 1, 231
przenikalność magnetyczna, 1
przesunięcie Goosa-Hanchena, 92
przesłona komplementarna, 190
przybliżenie adiabatyczne, 98
przybliżenie dipolowe, 266
przybliżenie Kirchhoffa, 168
przybliżenie optyczne, 169
przybliżenie przyosiowe, 25, 28
przybliżenie wolnozmienne obwiedni, 65
przysłona aperturowa, 47
przysłona polowa, 47
próg akcji laserowej, 258
prążki równego nachylenia, 145

prędkość fazowa, 113
prędkość grupowa, 114, 240
punkty główne, 35
punkty węzłowe, 37
płaszcz światłowodu, 43
płaszczyzna główna obrazowa, 35
płaszczyzna główna przedmiotowa, 35
płaszczyzna ogniskowa, 32
płaszczyzna padania, 83
płaszczyzny główne, 34
płytka fazowa, 205, 206
płytka fazowa jednofalowa, 214
płytka fazowa półfalowa, 214
płytka fazowa wielorzędowa, 215
płytka fazowa zerowego rzędu, 215
płytka fazowa złożona zerowego rzędu, 215
płytka fazowa ćwierćfalowa, 214
płytki fazowe, 213

rdzeń światłowodu, 43
reflektancja, 94, 137
reguła Millera, 251
relacja dyspersji, 114
relacje Kramersa-Kroniga, 242
relacje Stokesa, 135, 139
rezonans dwufotonowy, 251
rezonans jednofotonowy, 251
rezonans optyczny, 234
rezonator optyczny, 270
reżim konfokalny, 147
reżim liniowy, 246
romb Fresnela, 215
rozdzielczość kątowa, 181
rozdzielczość liniowa, 180
rozkład Boltzmanna, 261, 265
rozkład Maxwella-Boltzmanna, 267
równania brzegowe, 81
równania materiałowe, 1, 210
równania materiałowe zespolone, 7
równania Maxwella, 1
równania Maxwella zespolone, 6
równania Stokesa, 93
równanie dyspersji w przewodnikach, 99
równanie eikonału, 21
równanie falowe, 93, 246
równanie falowe dla fal przyosiowych, 65
równanie Helmholtza, 19, 57, 165
równanie Maxwella, 16
równanie soczewkowe, 32
równanie soczewkowe Newtona, 37

samoogniskowanie wiązki, 254
sfera Poincarégo, 204
siatka dyfrakcyjna, 184
siatka dyfrakcyjna holograficzna, 188
siatka dyfrakcyjna profilowana, 188
siła anharmoniczna, 247
siła Lorentza, 2, 12
siła oscylatora, 239
skalarna teoria dyfrakcji, 167
soczewka, 33
soczewka achromatyczna, 52
soczewka dwuwypukła, 33
soczewki gradientowe, 38
spektroskopia strat we wnęce, 144
spektroskopia strat we wnęce optycznej, 144
spirala Cornu, 162
spontaniczny podział częstotliwości, 282
stan koherentny światła, 259
stan metastabilny, 276
statystyka Bosego-Einsteina, 258, 267
statystyka fotonów, 258
stała dielektryczna, 2, 231
stała elastooptyczna, 218
stała Kerra, 219
stała sprężystości, 247
stała Verdeta, 220
stopień polaryzacji, 203
stos Arago, 206
straty zgięciowe, 43
strefy Fresnela, 156, 160
strumień fotonów, 12
strzałka fali, 109
superkontinuum, 254, 255

synchronizacja modów, 279
synchronizacja modów lasera, 279
szczypce optyczna, 74
szerokość Dopplera linii, 267
szerokość naturalna linii, 267
szerokość spektralna, 125
szum śrutowy, 281
szybkie światło, 241

teleportacja kwantowa, 282
teleskop, 49
teleskop zwierciadlany, 50
tensor przenikalności elektrycznej, 210
transformacja Lorentza, 228
transmitancja, 95, 137
twierdzenie całkowe Helmholtza-Kirchhoffa, 167
twierdzenie Greena, 165
twierdzenie o ekwipartycji energii, 260
twierdzenie o wzajemności Helmholtza, 172
tłumienie radiacyjne, 233

układ aplanatyczny, 52

warunek dopasowania fazowego, 254
warunek sinusów Abbego, 52
warunki brzegowe, 79
warunki brzegowe Kirchhoffa, 168
wektor Jonesa, 200
wektor Poyntinga, 11
widzialność prążków, 121
widzialność prążków interferencyjnych, 138
wir optyczny, 74
wiązka gaussowska, 65, 274
wnęka optyczna, 133
wnęka rezonansowa, 273
wolde światło, 255
wolne światło, 240
wolnozmienna obwiednia fali, 64
wolny przedział spektralny, 143
współczynnik absorpcji, 99, 236
współczynnik Einsteina
  emisji spontanicznej, 266
współczynnik ekstynkcji, 99
współczynnik finezji, 140
współczynnik załamania, 4, 209, 231, 236
  zespolony, 238, 242
współczynniki Einsteina
  absorpcji, 271
  emisji wymuszonej, 271
współczynniki Fresnela, 134, 206
  odbicia, 202
  transmisji, 202
wstrzykiwanie nośników, 268
wydajność ekstrakcji światła, 268
wymuszone rozpraszanie Ramana, 255
wypalanie dziur widmowych, 273
względna przenikalność elektryczna ośrodka, 5
wzory Fresnela dla składowej równoległej, 87
wzór Airy'ego, 137
wzór szlifierzy soczewek, 34
węzeł fali, 109
włókno światłowodowe, 44

zasada Babineta, 190
zasada Fermata, 22
zasada Huygensa, 155
zasada Huygensa-Fresnela, 155
zasada przyczynowości, 242
zasada superpozycji, 4, 8, 106, 183, 190, 246, 252
zasada zachowania energii, 15, 94
zasada zachowania momentu pędu, 15
zasada zachowania pędu, 15
zdolność rozdzielcza, 143
zdolność rozdzielcza mikroskopu, 48
zdolność rozdzielcza siatki dyfrakcyjnej, 189
zdolność skupiająca, 32
zegary optyczne, 151

zerowe fluktuacje próżni, 266
zespolona funkcja apertury, 194
zespolone prawo Snella, 102
zespolony współczynnik załamania, 99, 236, 246
zjawiska elektrooptyczne, 219, 253
zjawisko Dopplera, 267
zjawisko samomodulacji fazy, 254
zniekształcenie, 53
zredukowana stała Plancka, 12
złota reguła Fermiego, 266
złącza pn, 267
złącze pn, 277

ładunek topologiczny światła, 74

światło chaotyczne, 267
światło nieklasyczne, 281
światło ściśnięte, 282
światłowód, 43
światłowód fotoniczny, 45, 255
światłowód włóknisty, 44
świergot częstotliwości, 254
źrenica wejściowa, 47
źrenica wyjściowa, 47
źródła pojedynczych fotonów, 281
źródła termiczne, 258
źródła wyładowcze, 258
źródła światła nieklasycznego, 281
źródło promieniowania, 252