15.01.2019

Temat

XIII

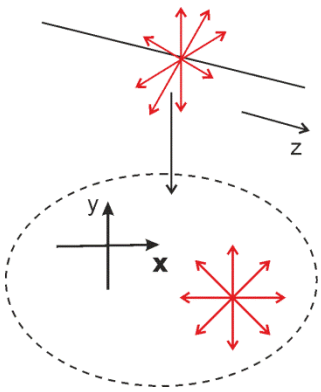
POLARYZACJA

1. Polaryzacja światła ♦

Z tematu (§TX 1) wiemy, że fale możemy podzielić na dwie grupy: fale poprzeczne i fale podłużne. Przykładem fali podłużnej jest fala dźwiękowa. Innym przykładem jest fala podłużna, która rozchodzi się w układzie ciężarków pokazanych na rysunku (TIX 4.4.1). Przykładem fali poprzecznej jest fala świetlna (ogólniej elektromagnetyczna). Do uzasadnienia tego faktu dojdziemy przy tematach związanych z elektrodynamiką. Przy omawianiu optyki falowej (§TX) zaniedbywaliśmy efekty związane z poprzecznym charakterem fali świetlnej. W tym temacie musimy to nadrobić.

W fali elektromagnetycznej oscylacjom podlegają wektor natężenia pola elektrycznego **E** oraz wektor natężenia pola magnetycznego **H**. W fali harmonicznej oba wektory drgają w kierunku prostopadłym do kierunku rozchodzenia się fali świetlnej (rys. TX 1.2). Ponieważ oba wektory są wzajemnie do siebie prostopadłe i powiązane równaniami Maxwella wystarczy, gdy skupimy się na jednym z nich – powszechnym wyborem jest wektor natężenia pola elektrycznego **E**, który będą w skrócie nazywał wektorem pola elektrycznego.

W przypadku fal poprzecznych mamy dodatkowy stopień swobody związany z tym, że kierunków prostopadłych do kierunku rozchodzenia się fali jest nieskończenie wiele (rys. 1.1).



Rysunek 1.1. Dla fal poprzecznych drgania zachodzą prostopadle do kierunku propagacji. Istnieje wiele możliwych kierunków prostopadłych – stąd drgania poprzeczne wykazują zjawisko polaryzacji.

Sytuacja jest nawet bardziej ciekawa. Kto powiedział, że wektor pola elektrycznego musi trzymać się wybranego kierunku drgań? Nie możemy wskazać powodów, dla których tak miałoby być. Oznacza to, że wektor pola elektrycznego może się kręcić wokół kierunku propagacji światła. Jaki tor zakreśla, w wybranej płaszczyźnie prostopadłej do kierunku propagacji światła, koniec wektora pola elektrycznego? Pytamy o następującą rzecz. Wyobraź sobie, że światło biegnie dokładnie prostopadle do ekranu. Jaki ślad rysuje na ekranie cień końca wektora pola elektrycznego **E**? Ograniczamy się do monochromatycznych fal płaskich czyli prostych drgań harmonicznych składowych wektora pola elektrycznego wyznaczonych w płaszczyźnie

prostopadłej (x, y) do kierunku (z) propagacji fali świetlnej. Niech współrzędne x końca wektora pola elektrycznego opisane są wzorem

$$E_x = E_{0x}\cos(\omega t - kz + \delta_x) \tag{1.1a}$$

E_{0x} jest wartością największego wychylenia wektora pola elektrycznego wzdłuż osi x, czyli amplitudą drgań wzdłuż tej osi. Dla osi y powinniśmy zapisać

$$E_y = E_{0y}\cos\left(\omega t - kz + \delta_y\right) \tag{1.1b}$$

Tutaj E_{0y} jest wartością amplitudy drgań wzdłuż osi y. Cofając obie fazy o δ_x otrzymamy prostsze wyrażenia

$$E_x = E_{0x}\cos(\omega t) \tag{1.2a}$$

$$E_y = E_{0y}\cos(\omega t + \Delta) \tag{1.2b}$$

$$\Delta = \delta_y - \delta_x \tag{1.2c}$$

Widać, że Δ jest względnym przesunięciem fazowym drgań harmoniczny między osią x i y. Wyrażenia (1.1) możemy zapisać w postaci zespolonej (TIX 1.1.2)

$$E_x = E_{0x}e^{i(\omega t+\delta_x)}e^{-ikz} \tag{1.3a}$$

$$E_y = E_{0y}e^{i(\omega t+\delta_y)}e^{-ikz} \tag{1.3b}$$

Podobnie jak wyrażenia (1.2), (1.3) może mieć postać zespoloną

$$E_x = E_{0x}e^{i\omega t}e^{-ikz} \tag{1.4a}$$

$$E_y = E_{0y}e^{i(\omega t+\Delta)}e^{-ikz} \tag{1.4b}$$

Część zależna od współrzędnej z-towej jest taka sama dla obu współrzędnych i nic nie wnosi do względnych relacji pomiędzy składowymi E_x i E_y. Oznacza to, że nie wpływa ona na stan polaryzacji światła. Dlatego nagminnie będę o niej „zapominał”. Na przykład wyrażenia (1.4) będę zwykle pisał w postaci

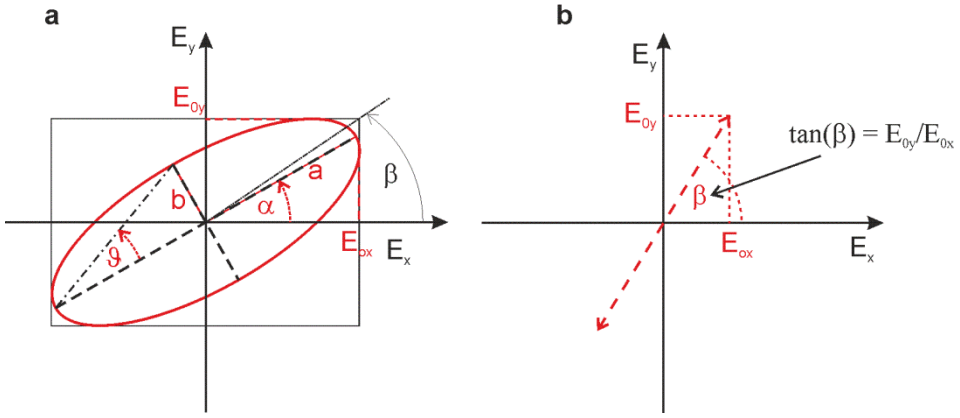
$$E_x = E_{0x}e^{i\omega t} \tag{1.5a}$$

$$E_y = E_{0y}e^{i(\omega t+\Delta)} \tag{1.5b}$$

Widać, że mamy zagadnienie składania prostopadłych drgań harmonicznych, o tej samej częstości, które zostało rozwiązane w (§TVIII 2.2). Zgodnie z uzyskanymi rozwiązaniami koniec wektora pola elektrycznego zakreśla w płaszczyźnie prostopadłej do kierunku ruchu fali elipsę. Co możemy powiedzieć na temat tej elipsy? Po pierwsze nadamy jej nazwę: elipsa stanu polaryzacji światła.

Definicja 1.1: Elipsa stanu polaryzacji światła

Elipsa wykreślona przez koniec wektora elektrycznego elektromagnetycznej fali harmonicznej, w rzucie na płaszczyznę prostopadłą do kierunku propagacji tej fali świetlnej, nazywamy elipsą stanu polaryzacji światła



Rysunek 1.2. a) Koniec wektora pola elektrycznego zakreśla elipsę; b) W szczególnym przypadku koniec wektora elektrycznego porusza się po linii prostej.

Układ równań (1.1) możemy sprowadzić, przez eliminację parametru *t*, (czyli czasu) do postaci

$$\frac{E_x^2}{E_{0x}^2}+\frac{E_y^2}{E_{0y}^2}-2\frac{E_xE_y}{E_{0x}E_{0y}}\cos(\Delta)=\sin^2(\Delta) \qquad 1.6$$

Co daje nam równanie elipsy w postaci (Dxxxx) bardziej znanej ze szkoły.

Stwierdziłem wyżej, że ograniczam się do fal płaskich, co wymaga komentarza. Zgodnie z (def. TIX 1.1.1) dla fali płaskiej zbiór punktów, dla których faza fali jest taka sama, w ustalonej chwili czasu, tworzy układ płaszczyzn oddalonych o długość fali (rys. TIX 1.1.9). Teraz jednak opisujemy falę za pomocą układu dwóch równań (1.2a-b). Każde z nich z osobna stanowi równanie fali harmonicznej. Wynikające stąd dwa układy powierzchni równej fazy nie pokrywają się, poza przypadkami szczególnymi. Wynika stąd, że harmonicznej fali elektromagnetycznej nie możemy, w ogólnym przypadku, przypisać jednej płaskiej powierzchni falowej. Przyjmiemy zatem, że spolaryzowana fala jest harmoniczna (płaska), gdy obie składowe fali wyrażają się przez równania fali harmonicznej (płaskiej). Wtedy, w każdym punkcie przestrzeni, w której rozchodzi się fala elektromagnetyczna, możemy narysować elipsę stanu polaryzacji opisującą stan polaryzacji światła w tym punkcie.

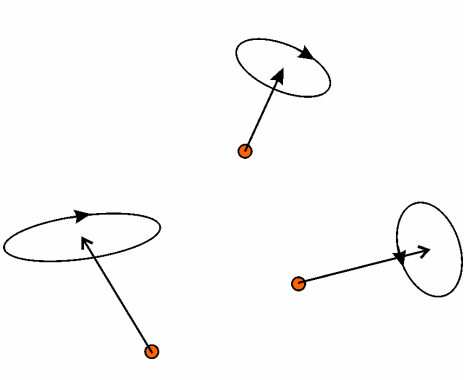
W ogólnym przypadku tak łatwo nie jest. Jednak, gdy światło jest spolaryzowane, lokalnie oscylacje fali świetlnej podobne są do oscylacji harmonicznych. To oznacza, że w konkretnym punkcie możemy wykreślić elipsę stanu polaryzacji, tyle że od punktu do punktu mogą to być różne elipsy (rys. 1.3). W przypadku gdy zaniedbujemy polaryzację fali, wracamy do opisu skalarnego przedstawionego w (§TIX).

Definicja 1.1: skalarny model fali świetlnej

Skalarny model fali świetlnej to model, w którym zaniedbujemy efekty polaryzacyjne, a falę opisujemy funkcją skalarną.

W szczególnych przypadkach elipsa polaryzacji może stać się linią prostą; mówimy wtedy, że światło jest spolaryzowane jest liniowo. Aby to nastąpiło musi być:

$$\Delta = 0 \text{ lub } \Delta = \pi \qquad 1.7$$



Rysunek 1.3. W ogólnym przypadku, gdy światło jest spolaryzowane każdym punkcie przestrzeni polaryzację światła opisuje odpowiednio zorientowana elipsa polaryzacji.

Kąt nachylenia tej prostej zdefiniowany jest przez stosunek E_{0y}/E_{0x}, tak jak jest to pokazane na rysunku (1.2). Jeżeli spełnione są warunki

$$\Delta = \pm\frac{1}{2}\pi \qquad 1.8a$$

$$E_{0x} = E_{0y} \qquad 1.8b$$

to koniec wektora pola elektrycznego zakreśla okrąg i mamy światło o polaryzacji kołowej lewo lub prawoskrętnej. Zauważ, że dla polaryzacji kołowej kąt azymutu jest nieokreślony. Jest to przykład innego, po fazie (§TX 1.1.1) parametru opisującego światło, które dla pewnych stanów jest źle określony. Kołowy stan polaryzacji jest przykładem nieciągłości polaryzacyjnej. Nieciągłości polaryzacyjne są przedmiotem wielu prac teoretycznych i aplikacyjnych. Jednak w tym temacie nie będę ich omawiał. Ustalimy jeszcze kiedy obieg elipsy jest lewo, a kiedy prawoskrętny. Wielkości E_x i E_y we wzorach (1.1) reprezentują współrzędne punktu na płaszczyźnie. Obliczając pochodne tych wielkości po czasie obliczymy składowe wektora prędkości tego punktu.

$$\frac{\mathrm{d}E_x}{\mathrm{d}t} = -E_{0x}\omega\sin(\omega t) \qquad 1.9a$$

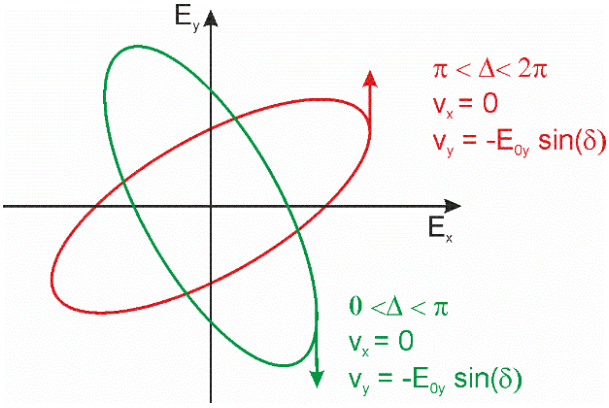
$$\frac{\mathrm{d}E_y}{\mathrm{d}t} = -E_{0y}\omega\sin(\omega t + \Delta) \qquad 1.9b$$

Wystarczy teraz ustalić kierunek wektora tej prędkości dla wybranej chwili. Niech tą chwilą będzie t=0. Wtedy równania (1.9) przyjmują postać:

$$\frac{\mathrm{d}E_x}{\mathrm{d}t} = 0 \qquad 1.10a$$

$$\frac{\mathrm{d}E_y}{\mathrm{d}t} = -E_{0y}\omega\sin(\Delta) \qquad 1.10b$$

Zatem w chwili $t = 0$ składowa x-owa wektora prędkości jest równa zeru, a składowa y-owa skierowana jest w kierunku ujemnym osi y, jeżeli $0 < \Delta < \pi$ oraz dodatnim jeżeli $\pi < \Delta < 2\pi$. Jeżeli składowa y-owa wektora prędkości skierowana jest dodatnio, to obieg elipsy jest prawoskrętny, a przy kierunku ujemnym obieg elipsy jest lewoskrętny, tak jak to pokazuje rysunek (1.4). Dla $\Delta = 0$ lub $\Delta = \pi$ otrzymujemy polaryzację liniową, dla której nie możemy zdefiniować orientacji stanu polaryzacji światła. Musimy ponadto założyć, że



Rysunek 1.4. Czerwona elipsa obiegana jest w lewą stronę na co wskazuje kierunek wektora prędkości w punkcie, gdzie $v_x = 0$. Z tego samego powodu wnioskujemy, że zielona elipsa obiegana jest w prawą stronę.

Mamy jeszcze jeden problem zilustrowany na rysunku (1.5). Gdy popatrzymy na elipsę polaryzacji z dwóch stron osi z, to kierunek jej obiegu zmienia się z lewo na prawo skrętny i lub odwrotnie. Konsekwentnie będę się trzymał prawoskrętnego układ współrzędnych (rys. TV 3.2.7). W takim wypadku oś z wychodzi z płaszczyzny rysunku (1.4) do czytelnika. Fala świetlna w naszych przykładach propaguje się w kierunku osi +z. Przyjmiemy zatem, że

Konwencja 1.1: Kierunek obiegu elipsy polaryzacji

Kierunek obiegu elipsy polaryzacji definiujemy patrząc się, w kierunku nadbiegającego światła, na ruch końca wektora pola elektrycznego, w prawoskrętnym układzie współrzędnych

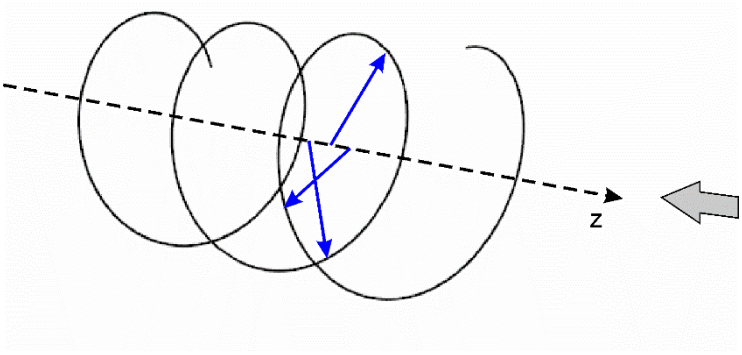
Tak więc w przykładzie z rysunku (1.5) patrzymy się w kierunku przeciwnym do kierunku osi z. Gdy światło biegnie w przeciwną stronę musimy patrzeć zgodnie z kierunkiem osi z.

Do tej pory opisywaliśmy elipsę stanu polaryzacji posługując się trzema parametrami: E_{0x}, E_{0y} i Δ. Do tego celu wystarczą tylko dwa parametry. Poniżej podaję definicję dwóch takich parametrów często wykorzystywanych do opisu stanu polaryzacji światła. Wartości E_{0x} i E_{0y} odpowiadają za tak zwaną eliptyczność ε elipsy to jest stopień jej spłaszczenia w stosunku do okręgu. Eliptyczność ε definiujemy w następujący sposób

Definicja 1.2: Eliptyczność

Eliptyczność definiujemy jako iloraz długości małej i dużej osi elipsy stanu polaryzacji światła

$$\varepsilon = \frac{b}{a} \qquad 1.11$$



Rysunek 1.5. Fala świetlna biegnie zgodnie z kierunkiem osi *z*. Patrzymy się w kierunku przeciwnym (szara strzałka). W tym kierunku widać, że spirala kręci się prawoskrętnie. Niebieskie strzałki pokazują wektory pola elektrycznego w trzech wybranych punktach. Spirala pokazuje tor końca wektora elektrycznego.

Obok eliptyczność używa się równoważnej wielkości nazywanej kątem eliptyczności (rys. 1.3)

Definicja 1.3: Kąt eliptyczności

Kątem eliptyczności θ nazywamy wartość arcustangens obliczony z eliptyczność ε. Przyjmujemy przy tym, że kąt ten jest dodatni dla prawej skrętności polaryzacji a ujemny dla lewej skrętności polaryzacji.

$$\theta = \pm\mathrm{atan}(\varepsilon) \qquad 1.12$$

Zobaczmy jaki jest zakres zmienności kąta eliptyczności. Gdy mała oś elipsy jest równa zeru, to i kąt eliptyczności równy jest zeru; zeru jest również równa eliptyczność elipsy. Mamy wtedy do czynienia z linią prostą. Zwiększając rozmiar małej osi dochodzimy do wartości $\theta = 45°$, dla której obie osie mają taką samą długość. Dalsze zwiększanie długości małej osi nie ma sensu, gdyż oś ta staje się wtedy osią większą. Pełny opis stanu polaryzacji wymaga rozszerzenia tego zakresu na kąty ujemne, co zostanie wyjaśnione nieco poniżej. Zatem kąt eliptyczności zmienia się w zakresie

$$-\frac{\pi}{4} \le \theta \le \frac{\pi}{4} \qquad 1.13$$

Jak już wiemy, orientacja dużej osi elipsy określana jest przez tak zwany kąt azymutu α (rys. 1.2).

Definicja 1.4: Kąt azymutu

Kątem azymutu nazywamy kąt jaki duża oś elipsy polaryzacji tworzy z wybraną osią odniesienia

Musimy zastanowić się nad zakresem zmienności kąta azymutu. Z rysunku (1.2) widać, że wszystkie możliwe orientacje elipsy stanu polaryzacji wyczerpują kąty azymutu z przedziału

$$-\frac{\pi}{2} \le \alpha \le \frac{\pi}{2} \qquad 1.14$$

Możemy znaleźć różne relacje pomiędzy wprowadzonymi tu parametrami opisującymi stan polaryzacji światła. Wymaga to nieco rachunków z użyciem tożsamości trygonometrycznych, czego robić tu nie będę ograniczając się do wypisania przykładowych wzorów[1]

$$a^2 + b^2 = E_{0x}^2 + E_{0y}^2 \qquad 1.15a$$

$$ab = E_{0x}E_{0y}\sin(\Delta) \qquad 1.15b$$

$$\sin(2\theta) = \sin(2\beta)\sin(\Delta) \qquad 1.15c$$

Ostatnie wyrażenie wyjaśnia dlaczego wartości kąta eliptyczności θ, została poszerzona o zakres wartości ujemnych (1.13). Dla prawoskrętnej orientacji polaryzacji przesunięcie fazowe Δ mieści się w zakresie od π do 2π. Dla takiego zakresu $\sin(\Delta)$ jest ujemny, przez co $\sin(2\theta)$ jest też ujemny i kąt θ musi być ujemny. Zatem ujemne wartości kąta θ wyznaczają tą samą eliptyczności co dodatnie, ale wskazują na polaryzację prawoskrętną.

Wprowadzę powszechnie przyjęte oznaczenia

Definicja 1.5: Stan polaryzacji H

Przy wyróżnionej osi x-ów, światło spolaryzowane liniowo zgodnie z kierunkiem tej osi oznaczamy jak światło w stanie polaryzacji H

Definicja 1.6: Stan polaryzacji V

Przy wyróżnionej osi x-ów, światło spolaryzowane liniowo prostopadle do kierunku tej osi oznaczamy jak światło w stanie polaryzacji V

Oznaczenia H i V pochodzą od angielskiego *horizontal* (poziomy, horyzontalny) i *vertical* (pionowy, wertykalny). To, że oś *x*-ów jest wyróżniona mówi nam, że od tej osi liczymy kąty.

[1] Przykłady takich obliczeń można znaleźć w F. Ratajczak Optyka ośrodków anizotropowych, PWN, Warszawa, 1994

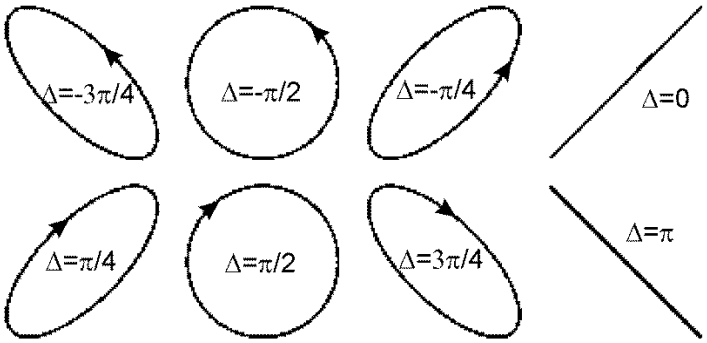
Definicja 1.7: Stan polaryzacji R

Światło spolaryzowane kołowo prawoskrętnie oznaczamy jak światło w stanie polaryzacji R

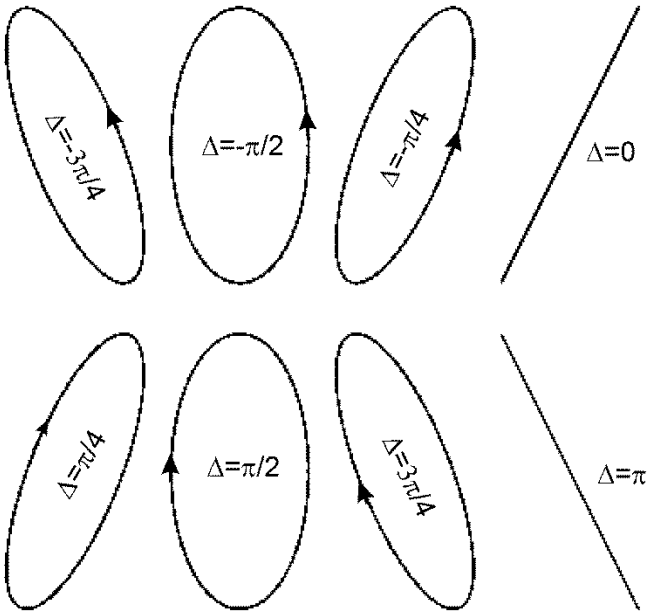
Definicja 1.8: Stan polaryzacji L

Światło spolaryzowane kołowo lewoskrętnie oznaczamy jak światło w stanie polaryzacji L

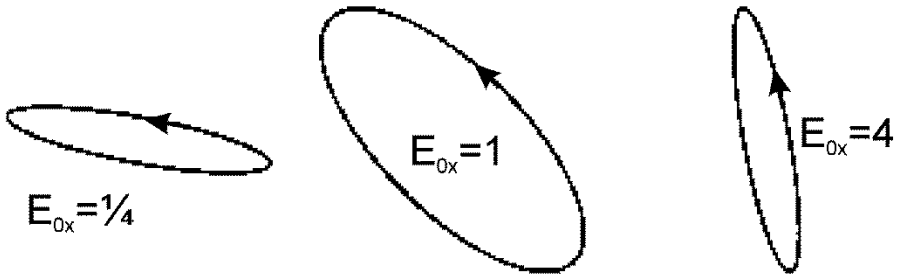
Oznaczenia R i L pochodzą od angielskiego *right* (prawo) i *left* (lewo). Różne stany polaryzacji przy różnych wartościach parametrów elipsy polaryzacji pokazują rysunki (1.6-8)



Rysunek 1.6. Przykładowe elipsy polaryzacji dla różnych wartości parametru Δ, przy czym $E_{0x}=E_{0y}$.



Rysunek 1.7. Przykładowe elipsy polaryzacji dla różnych wartości parametru Δ, przy czym $2E_{0x}=E_{0y}$.



Rysunek 1.8. Przykładowe elipsy polaryzacji dla różnych wartości amplitudy E_{0x}, przy czym E_{0y}=1 i Δ=-3π/4.

1.1. Metody otrzymywania światła spolaryzowanego

Zjawisko polaryzacji światła nie ma jednego odkrywcy. Efekty polaryzacyjne były obserwowane już w zamierzchłych czasach choć nikt ich wtedy nie postrzegał jako coś, co wymaga szczególnej uwagi. Według niektórych historyków Wikingowie używali faktu, że światło słoneczne jest częściowo spolaryzowane do nawigacji już około roku 700, ale jest to ciągle dyskutowana teza[2]. Wiadomo natomiast, że polaryzację wykorzystują w nawigacji niektóre owady (np. pszczoły[3]). Pewną informacją jest również, to że Wikingowie odkryli Islandię, a że klimat był wówczas łaskawszy, szybko zaczęli ją kolonizować. Na Islandii odnaleziono minerał nazywany szpatem islandzkim, który odegrał ważną rolę w badaniach nad polaryzacją. Szpat islandzki, to nic innego jak kryształy kalcytu, czyli węglan wapnia o wzorze $CaCO_3$. W 1669 roku duński matematyk Erasmus Bartholinus wykonał serię eksperymentów z kalcytem, których wyniki spisał w sześćdziesięciostronicowej pracy. Patrząc przez kryształ kalcytu widzimy rozdwojony obraz, co wynika z podziału wiązki światła na dwie (rys. 1.1.1).

Rysunek 1.1.1. Podwójny obraz widziany przez kryształ szpatu islandzkiego; źródło Wikipedia

Wiązki te są wzajemnie prostopadle spolaryzowane. Efekt powstawania podwójnego obrazu znany był przed Bartholinusem, ale jego praca, opisująca dokładnie sam efekt i przeprowadzone przy tym eksperymenty dała początek

[2] Nie mniej dziś opracowane zostały metody nawigowania z wykorzystaniem polaryzacji światła (zobacz np. Guixia Guan et. al., *The novel method of North-finding based on the skylight polarization*, Journal of Engineering Science and Technology Review, **6** 107-110 (2013)

[3] Rossel S, Wehner R., *The bee's map of the e-vector pattern in the sky*, Nature, 1982, 79, pp: 4451-4455.

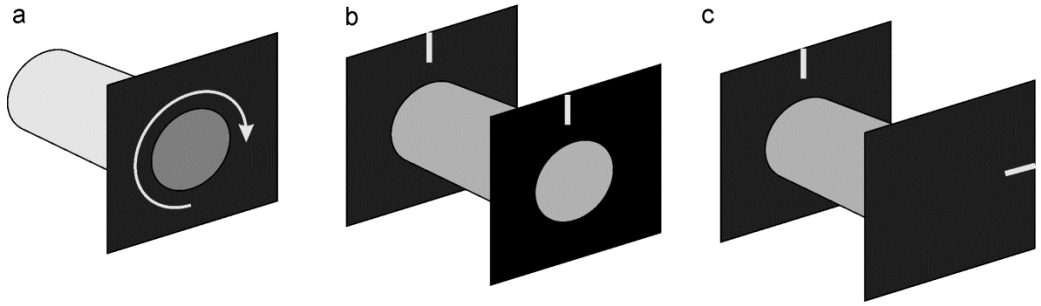
szerszemu zainteresowaniu temu zjawisku. O tym, że obie wiązki są prostopadle spolaryzowane Bartholinus nie wiedział. Nie wiedział również i o tym, że światło można opisywać jako falę świetlną. Christian Huygens zaprzągł do opisu polaryzacji sformułowaną przez siebie zasadę, którą dziś nazywamy zasadą Huygensa (§TIX 2). Zdał sobie sprawę, że efekt nazywany dziś efektem dwójłomności, który odpowiada za powstawanie podwójnych obrazów, może powstać, gdy fala, ze źródła punktowego, rozchodzi się w ośrodku nie jako sfera ale jako elipsoida, był to pogląd bliski dzisiejszemu. Etienne Malus eksperymentując z kryształem szpatu islandzkiego odkrył zjawisko polaryzacji światła przez odbicie od szyby. Prowadząc kolejne eksperymenty sformułował prawo, które dziś nazywamy prawem Malusa (okr. 2.1.). Wyniki Malusa nie ucieszyły ani Younga ani Fresnela. Obaj fizycy walnie przyczynili się do rozwoju teorii falowej światła (§TX 2), ale obaj rozumieli światło jako skalarne fale eteru, na wzór fali dźwiękowej. Skalarne fale nie powinny wykazywać zjawiska polaryzacji. Po pewnym czasie obaj panowie doszli do wniosku, że światło musi być falą poprzeczną. Od tego momentu możemy mówić, że zakończyła się prehistoria rozwoju badań nad zjawiskiem polaryzacji światła.

Przedstawię kilka najbardziej znanych metod generowania światła o określonej polaryzacji, bez ich szerszego teoretycznego uzasadnienia. Teoretyczne uzasadnienie można zbudować na gruncie praw elektrodynamiki, a to dopiero przed nami. Niemniej większość tych metod odkryto i badano jeszcze przed kompletnym sformułowaniem praw elektrodynamiki klasycznej. A dopiero komplet tych praw (równania Maxwella) pozwolił na sformułowanie bardziej podstawowej teorii zjawiska polaryzacji światła.

1.1.1. Polaryzacja przez odbicie

Wśród metod otrzymywania światła spolaryzowanego najprostszego oprzyrządowania wymaga metoda polaryzowania przez odbicie. Jeżeli światło niespolaryzowane pada na materiał dielektryczny (np. szybę), to światło odbite ulega częściowej polaryzacji liniowej. Kierunek polaryzacji jest równoległy do powierzchni płytki. Co to znaczy, że światło ulega częściowej polaryzacji? Powiedzmy, że mamy taki przyrząd, który przepuszcza światło o polaryzacji liniowej, o zadanym kierunku, a pozostałą część wiązki absorbuje. Przyrząd ten nazwę analizatorem, gdyż może służyć do analizy stanu polaryzacji światła. Nie jest w tym momencie istotne jak taki przyrząd zrobić. Powiedzmy, że ustawiamy analizator na wprost w stosunku do nadbiegającej wiązki światła. Jeżeli zmiana kąta orientacji analizatora nie zmienia natężenia światła przechodzącego, to mówimy, że światło nie ma polaryzacji liniowej (rys. 1.1.2). Jeżeli dla pewnego kąta ustawienia analizatora wiązka padająca jest całkowicie absorbowana, to światło jest spolaryzowane liniowo, a kierunek polaryzacji jest prostopadły do kierunku analizatora. W takim przypadku aby przepuścić światło przez analizator, bez obniżenia jego natężenia, musimy go obrócić o kąt prosty. Gdy natężenie światła przechodzącego przez analizator jest zależne od kąta

analizatora, a przy tym dla żadnego z tych kątów natężenie światła przechodzącego nie spada do zera, to mówimy, że światło jest częściowo spolaryzowane. Oznacza to również, że światło przechodzące przez nasz przyrząd (dalej będę mówił płytkę) jest spolaryzowane liniowo, zgodnie z kierunkiem analizatora. Płytka może więc równie dobrze służyć za polaryzator. To czy daną płytkę traktujemy jak analizator, czy jak polaryzator zależy od sposobu jej użycia. Na rysunku (1.1.2b i c), mamy użycie dwóch płytek, pierwsza pracuje jako polaryzator a druga jako analizator.



Rysunek 1.1.2. a) gdy na analizator (czarna płytka) pada światło niespolaryzowane liniowo, to wiązka przechodząca ma o połowę mniejsze natężenie w stosunku do padającej. Ponadto natężenie wiązki przechodzącej nie zależy od kąta orientacji polaryzatora; b) ta część wiązki, która przechodzi przez polaryzator jest spolaryzowana liniowo zgodnie z osią polaryzatora. Na rysunku oś polaryzatora zaznaczona jest krótką jasną kreską. Po przejściu przez drugi polaryzator (który pełni funkcję analizatora), o takim samym ustawieniu osi natężenie wiązki przechodzącej jest takie same jak wiązki padającej na ten drugi polaryzator (pod warunkiem, że polaryzator jest idealny); c) jeżeli oś drugiego polaryzatora (analizatora) ustawimy prostopadle do osi pierwszego, to wiązka zostanie całkowicie zatrzymana.

Po tej krótkiej dygresji wrócę do odbicia światła od powierzchni dielektryka. Przy odbiciu od dielektrycznej płytki światło ulega częściowej lub całkowitej polaryzacji linowej. Jeżeli ośrodek, na który pada światło jest przezroczysty, to światło przechodzące jest częściowo spolaryzowane. W 1815 David Brewster odkrył, że jeżeli światło pada na płytkę w taki sposób, że kąt załamania tworzy z kątem padania kąt prosty (rys. 1.1.3), to światło odbite jest całkowicie spolaryzowane. Kąt padania dla którego spełniony jest powyższy warunek nazywamy kątem Brewstera. Łatwo jest go wyznaczyć. Z rysunku (1.1.3) i prawa Snella mamy

$$\sin(\mathrm{i}) = n\sin(\mathrm{i}') \qquad 1.1.1$$

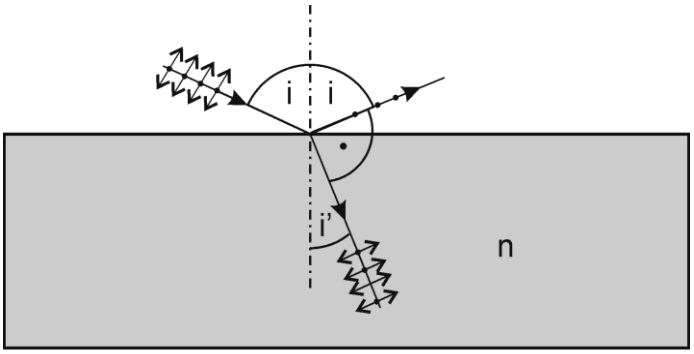
Oraz

$$i' = \frac{\pi}{2} - i \qquad 1.1.2$$

Wstawiając (2.1.2) do (2.1.1) i porządkując tak otrzymane wyrażenie mamy

$$\tan(i) = n \qquad 1.1.3$$

Równanie to wyraża warunek na kąt Brewstera



Rysunek 1.2.3. Jeżeli na powierzchnię dielektryka pada światło niespolaryzowane to światło odbite jest spolaryzowane liniowo w płaszczyźnie równoległej do powierzchni płytki. Stopień polaryzacji zależy od kąta padania światła. Jeżeli kąt padania jest taki, że kąt pomiędzy promieniami odbitym i załamanym jest równy kątowi prostemu, to światło odbite jest całkowicie spolaryzowane. Strzałki pokazuję orientację polaryzacji w płaszczyźnie kartki, a kropki wskazują na polaryzację prostopadłą do kartki.

Dlaczego całkowitej polaryzacji, pod kątem Brewstera, ulega światło odbite, a nie przechodzące? Natężenie światła w wiązce odbitej to zwykle kilka procent natężenia wiązki padającej. Zjawisko odbicia usuwa tylko część światła spolaryzowanego równolegle do powierzchni płytki. Pozostała część przechodzi w wiązce ugiętej, która przez to nie może być całkowicie spolaryzowana. Niemniej stos płytek ułożonych jedna nad drugą może dać światło przechodzące o wysokim stopniu polaryzacji. Na przykład 45 płytek ze szkła BK7 ułożonych jedna nad drugą może dać światło spolaryzowane liniowo, po przejściu przez stos, o stopniu polaryzacji około 0.9. Układanie 45 płytek może się wydać kłopotliwym przedsięwzięciem. Jednak gdy płytki staną się cienkimi warstwami o grubości liczonych w mikrometrach, to duża liczba takich warstw ma ciągle niewielką grubość i może zostać wykorzystana jako dobrej jakości polaryzator.

Co to znaczy, że światło jest spolaryzowane liniowo w stopniu 0.9? Wstawmy analizator w badaną wiązkę światła i zacznijmy nim kręcić tak, by znaleźć maksymalną wartość natężenia światła przechodzącego I_{max}. Następnie szukamy położenia, dla którego natężenie światła przechodzącego jest minimalne I_{min}. Stopień polaryzacji liniowej jest równy

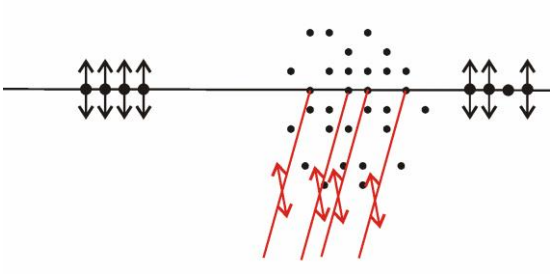
$$P = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad 1.1.4.$$

Jeżeli światło padające na analizator jest całkowicie spolaryzowane liniowo, to znaczy istnieje takie położenie analizatora, że przechodzi całe światło $I_{max}=I_0$, to w położeniu prostopadłym natężenie świata przechodzącego jest równe zeru $I_{min}=0$ i ze wzoru (1.1.4) znajdujemy $P=1$. Gdy światło jest niespolaryzowane lub spolaryzowane jest kołowo, to ilość światła przechodzącego przez analizator liniowy jest dla każdego kąta osi tego analizatora taka sama, zatem $I_{max}=I_{min}$ i $P=0$. Stopień polaryzacji liniowej jest równy zeru. Nie oznacza to jeszcze, że

światło nie ma uporządkowanego stanu polaryzacji; może być spolaryzowane kołowo. Przy polaryzacji kołowej obrót analizatora nie zmienia natężenia światła przechodzącego. Światło kołowo spolaryzowane ma stopień polaryzacji liniowej równej zeru, pomimo że jest to światło całkowicie spolaryzowane. Widać z tego, że polaryzator liniowy nie wystarcza do stwierdzenia braku polaryzacji światła. Istnieją jednak metody, które pozwalają taki test przeprowadzić. Wystarczyłoby zbudować polaryzator kołowy, który badał by polaryzację światła kołową lewą i prawo skrętną. Wzór (1.1.4) możemy rozumieć również w szerszym sensie. Gdy P=0, to nie ma żadnej wyróżnionej polaryzacji światła w badanej wiązce. Gdy P=1 istnieje jedna taka polaryzacja. Może być to polaryzacja liniowa, kołowa lub jakakolwiek inna eliptyczna.

1.1.2 Polaryzacja przez rozpraszanie

Kolejnym mechanizmem prowadzącym do polaryzacji światła jest jego rozpraszanie na cząsteczkach mniejszych od długości fali. Przykładem zbioru takich cząsteczek jest powietrze. Kiedy obserwujemy niespolaryzowane światło przechodzące przez zbiór takich cząsteczek, a kierunek obserwacji jest prostopadły do kierunku propagacji światła, to rozproszona wiązka jest spolaryzowana w kierunku prostopadłym do płaszczyzny utworzonej przez kierunek propagacji światła i kierunek wzdłuż, którego obserwujemy rozproszone światło (rys. 1.1.4). Fakt ten odnotował w 1811 roku Arago. W prosty sposób mechanizm polaryzowania przez rozpraszanie można sobie wyjaśnić odwołując się do faktu, że składowa pola elektrycznego w kierunku propagacji musi być równa zeru. Zatem, gdy kierunek propagacji jest prostopadły do kierunku padania rozchodzić się może składowa, która jest prostopadła i do nowego kierunku i do kierunku padania wiązki świetlnej. Niestety rozumowanie takie jest mocno uproszczone, ale pozwala łatwo zapamiętać, w jakim kierunku spolaryzowane jest światło rozproszone.



Rysunek 1.1.4 Polaryzowanie światła przez rozpraszanie na małych cząsteczkach

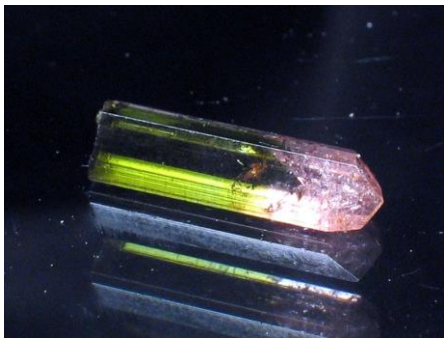
Przechodzące przez atmosferę światło słoneczne ulega częściowej polaryzacji w kierunku prostopadły do kierunku rozchodzenia się promieni świetlnych. Polaryzacja jest częściowa, gdyż wiązka promieni słonecznych nie jest równoległa, zatem nie można się do całej wiązki ustawić w kierunku prostopadłym. Zjawisko to łatwo sprawdzić, patrząc na promienie słoneczne

w odpowiednim kierunku (to jest jak najbardziej prostopadle) przez płytkę polaroidu. Obracają płytkę zaobserwujesz zmiany natężenia światła przechodzącego przez płytkę.

Niektóre zwierzęta wykrywają polaryzację słonecznego światła. Należą do nich między innymi pszczoły, które wykorzystują tą zdolność do nawigacji. Co więcej część z nas jest w stanie wykryć polaryzację światła słonecznego gołym okiem. Widzą oni tzw. szczotkę Haidingera; efekt jest słaby i wymaga uwagi. Opis tego efektu jak i wielu innych związanych z naturalnym światłem można znaleźć w doskonałej książce Minnareta[4].

1.1.3. Polaryzacja przez selektywną absorpcję

Niektóre kryształy anizotropowe optycznie wykazują skrajnie różne współczynnik absorpcji dla światła o różnych stanach polaryzacji; nazywamy to zjawiskiem pleochroizmu. Zjawisko to dostrzegł i opisał angielski lekarz William Herapath (1852). Jeden z jego uczniów zauważył, że dodanie jodyny do moczu psa żywionego chininą powoduje wytrącenie zielonych kryształów (wodosiarczan chininy, jako minerał nazywany herapatytem). Herapath oglądając kryształy pod mikroskopem zauważył, że polaryzują one światło. Innym przykładem kryształu o takich własnościach jest turmalin rysunek (1.1.5). Podobny efekt można uzyskać wytwarzając siatkę cienkich, gęsto upakowanych drucików (rys. 1.1.6)

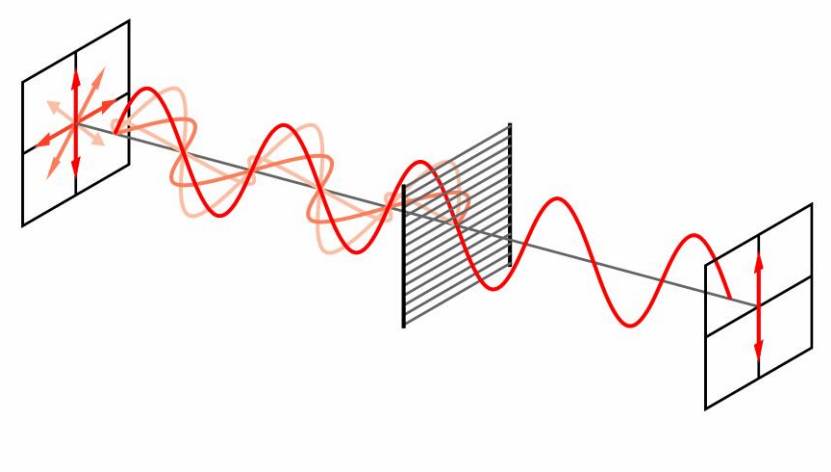


Rysunek 1.1.5. Kryształu turmalinu; źródło Wikipedia, pod licencją GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation.

Światło o kierunku polaryzacji prostopadłej do ułożenia drucików przechodzi przez kryształ jak przez ośrodek przezroczysty, a światło o polaryzacji równoległej jest tłumione. Zjawisko to nazywamy zjawiskiem *selektywnej absorpcji* światła. Teoretyczny opis tego zjawiska jest możliwy na gruncie elektrodynamiki klasycznej. Elementy polaryzacyjne wycięte z kryształów turmalinu są drogie (turmalin zaliczany jest do kamieni jubilerskich). Wykorzystanie drobniejszych kryształków herapatytu pozwoliło na opracowanie tańszych polaryzatorów – *polaroidów*. Polaroid został opatentowany przez Amerykanina Edwina Landa. Pierwsze polaroidy były

[4] M. Minnaert, Światło i barwa w przyrodzie, PWN, Warszawa, 1961

płytkami z nitrocelulozy, w których zatopiono zostały podłużne jak igły kryształy jodosiarczanu chininy. Aby kryształy ułożyły się równolegle (inaczej każdy przepuszczałby światło o innej polaryzacji liniowej), podkład był na gorąco i w obecności silnego pola elektrycznego, rozciągany. Tak otrzymana płytka absorbowała światło o polaryzacji poprzecznej do kierunku ułożenia kryształków. Od 1938 roku polaroid produkowany jest z polialkoholu winylowego z domieszką jodku potasu. Jodek potasu łączy się z polialkoholem przez wiązania wodorowe, tak że podczas rozciągania, mikrodomeny jodku potasu orientują się wzdłuż jednego kierunku bez potrzeby przykładania pola elektrycznego. Jest to metoda wykorzystywana do dziś, gdyż pozwala na produkcję polaryzatorów o niskiej cenie. Niestety niska cena polaroidów okupiona jest ich niską jakością optyczną, co eliminuje je z bardziej wymagających zastosowań.



Rysunek 1.1.6. Niespolaryzowane światło pada na układ równoległych drucików. Drgania wektora pola elektrycznego w kierunku równoległym do drucików są silnie tłumione. Drgania prostopadłe propagują się przez sieć drutów; w efekcie otrzymujemy światło spolaryzowane liniowo.

1.1.4. Działanie polaroidu

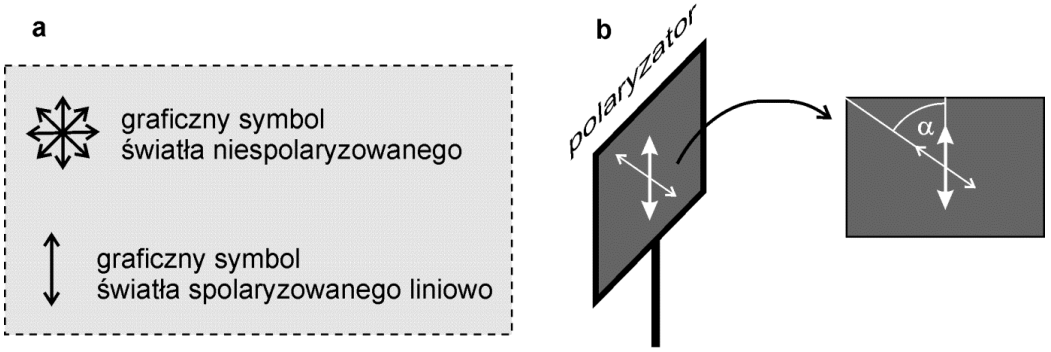
Jak działa polaroid na padające światło? Wiemy już, że przepuszcza światło spolaryzowane liniowo zgodnie z własnym kierunkiem przepuszczania, który nazwę osią polaryzatora. Regułę pozwalającą na obliczenie względnej ilości światła o danej polaryzacji liniowej, przechodzącego przez polaryzator o danej orientacji (rys. 1.2.7) odkrył Etienne Malus.

Określenie 1.2: Prawo Malusa

Jeżeli na polaryzator liniowy pada światło, prostopadle do płaszczyzny polaryzatora i linowo spolaryzowane, o natężeniu I_0 a kąt między osią polaryzatora, a kierunkiem polaryzacji światła wynosi α, to natężenie światła przechodzącego wyraża się wzorem

$$I = I_0 \cos^2(\alpha) \qquad 1.2.5$$

Gdy kąt α=0 polaryzator idealny przepuszcza całe światło (rys. 1.2.2b), a gdy $\alpha=\pi/2$ polaryzator idealny tłumi całe światło (rys. 1.2.2c). A co się stanie, jeżeli oświetlimy polaryzator wiązką spolaryzowaną eliptycznie? Każdą elipsę możemy opisać jako złożenie dwóch, prostopadłych drgań harmonicznych (tab. TVIII 2.2.1, wzory (1.2 i 1.3)). Wybierzmy zatem oś *x*-ów układu współrzędnych tak, aby pokrywała się z osią polaryzatora i rozłóżmy drgania pola elektrycznego na dwie składowe prostopadłe. Polaryzator przepuści tylko składową *x*-ową blokując prostopadłą do niej składową *y*-ową. Stąd natężenie światła przepuszczonego przez polaryzator będzie równe



Rysunek 1.2.7. a) graficzne symbole dla światła niespolaryzowanego i spolaryzowanego liniowo; b) polaryzator liniowy z zaznaczonym poprzez podwójną (grubą) strzałkę kierunkiem osi. Na polaryzator pada światło spolaryzowane liniowo (cienka strzałka) tak, że kąt między kierunkiem osi polaryzatora liniowego, a kierunkiem polaryzacji światła jest równy α

$$I = \langle E_x^2 \rangle_T = E_{0x}^2 \langle \cos^2(\omega t) \rangle_T = \frac{1}{2} E_{0x}^2 \sim E_{0x}^2 \qquad 1.2.6$$

Zgodnie z dyskusją (TX 1.4), w przypadku światła prędkość oscylacji jest tak duża, że mierzymy nie chwilowe wartości natężenia światła, ale jego średnie po czasie. Wzór (1.2.6) podpowiada nam skąd się wzięło prawo Malusa. Powiedzmy, że światło padające jest liniowo spolaryzowane pod kątem α do osi polaryzatora. Niech natężenie światła (już to uśrednione po czasie, czyli mierzone) wynosi

$$I_0 = \langle E_0^2 \rangle_T \langle \cos^2(\omega t) \rangle_T = \frac{1}{2} E_0^2 \qquad 1.2.7$$

Rzut wektora o długości E_0 na oś polaryzatora jest równy

$$E_{0x} = E_0 \cos(\omega t)\cos(\alpha) \qquad 1.2.8$$

Tylko ta część przechodzi przez polaryzator, stąd natężenie światła przechodzącego

$$I = \langle E_{0x}^2 \rangle_T = E_0^2 \cos^2(\omega t)\langle \cos^2(\omega t)\rangle_T = \frac{1}{2} E_0^2 \cos^2(\alpha) \qquad 1.2.9$$

Korzystając z (1.2.7) mamy

$$I = I_0 \cos^2(\alpha) \qquad 1.2.10$$

Czyli prawo Malusa. To co dla Malusa było prawem empirycznym, dla nas jest prostą konsekwencją wektorowej teorii falowej.

Przedstawiony na rysunku (1.2.2c) układ dwóch polaroidów o wzajemnie prostopadłych osiach jest podstawowym układem do badania próbek zmieniających stan polaryzacji światła (rys. 1.2.8). Układ do badania stanu polaryzacji światła nazywa się polarymetrem. Przypomnę, że pierwszy polaroid przygotowuje światło w zadanym stanie polaryzacji liniowej (stąd jego nazwa polaryzator), a drugi polaroid pozwala na określenie zmian tego stanu polaryzacji (stąd jego nazwa analizator). Układ do analizy światła spolaryzowanego złożony z dwóch liniowych polaryzatorów o wzajemnie ortogonalnych orientacjach osi nazywa się „krzyżem polaryzacyjnym".

Definicja 1.2.1: Polarymetr

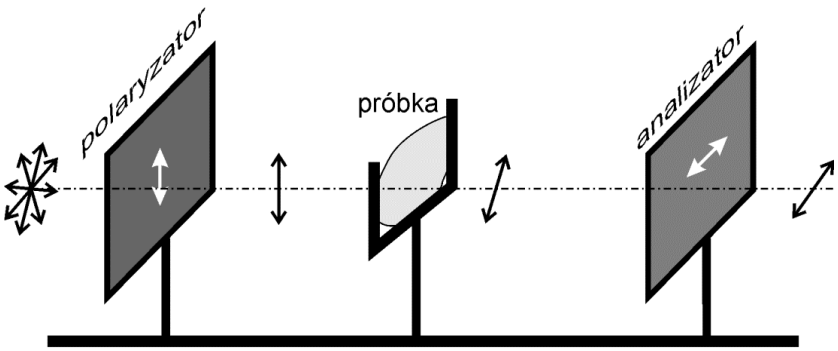
Przyrząd pozwalający na określenie stanu światła spolaryzowanego, lub wpływu próbki na stan polaryzacji światła nazywamy polarymetrem

W temacie (§TV 3.2) omówiłem zjawisko chiralności cząstek chemicznych. Cząstki chiralne o określonej skrętności skręcają kierunek polaryzacji światła. Nie możemy na tym etapie omówić mechanizmu mikroskopowego tego zjawiska i nie musimy tego mechanizmu znać, aby samo zjawisko praktycznie wykorzystać. Tak to też wyglądało od strony historycznej. Jeszcze przed stworzeniem modelu teoretycznego, nauczono się praktycznie wykorzystywać zjawisko skręcenia polaryzacji przez roztwory cząstek chiralnych. Zjawisko to odkrył Jean Baptist Biot w 1815 roku w roztworach cukru i kwasu winowego. Zawartość cząstek chiralnych w roztworze można określić z pomocą krzyża polaryzacyjnego. Pomiędzy polaryzator i analizator wkładamy badaną próbkę. Po przejściu przez próbkę (może to być roztwór) zmienia się orientacja światła, przez co część światła przechodzi przez analizator. Obracając analizator szukamy takiego jego położenia, przy którym światło zostanie ponownie wygaszone. Kąt tego obrotu wyznacza nam kąt skręcenia polaryzacji światła przez badaną próbkę. Na tej podstawie możemy wyznaczyć na przykład stężenie substancji chiralnej w roztworze. Mierzy się w ten sposób stężenie cukru w wodzie, korzystając z prostej zależności

$$c = \frac{\alpha}{\alpha_w D} \qquad 1.2.11$$

Gdzie c jest mierzonym stężeniem cukru (lub innej substancji) α zmierzonym kątem skręcenia polaryzacji, D grubością warstwy roztworu przez, który przeszło światło, α_w skręcalnością właściwą. Skręcalność właściwa określa zdolność do skręcania kierunku polaryzacji przez daną substancję. Wymiarem skręcalności właściwej jest

$$[\alpha_w] = [\text{miara kąta}]\frac{[\text{długość}]^2}{[\text{masa}]} \qquad 1.2.12$$



Rysunek 1.2.8. Układ polarymetru typu krzyż polaryzacyjny. Polaryzator przepuszcza światło o liniowym stanie polaryzacji. Analizator to liniowy polaryzator, którego oś jest ustawiona prostopadle do osi polaryzatora. Gdy między analizatorem a polaryzatorem jest wolna przestrzeń światło jest przez układ blokowane. Gdy między polaryzatorem a analizatorem wstawimy próbkę, która zmienia stan polaryzacji światła, to przejdzie ono częściowo przez analizator. Mierząc natężenie światła przechodzącego względem światła padającego na próbkę, można określić pewne własności optyczne próbki.

Jak widać z tych zależność kąt skręcenia polaryzacji rośnie liniowo z grubością roztworu.

Definicja 1.2.2: Skręcalność właściwa

Skręcalność właściwa wyraża liczbowo kąt skręcenia płaszczyzny polaryzacji światła przez roztwór o stężeniu jednostkowym i jednostkowej grubości warstwy.

Fakt 1.2.1.

W układzie SI wymiarem skręcalności właściwej jest

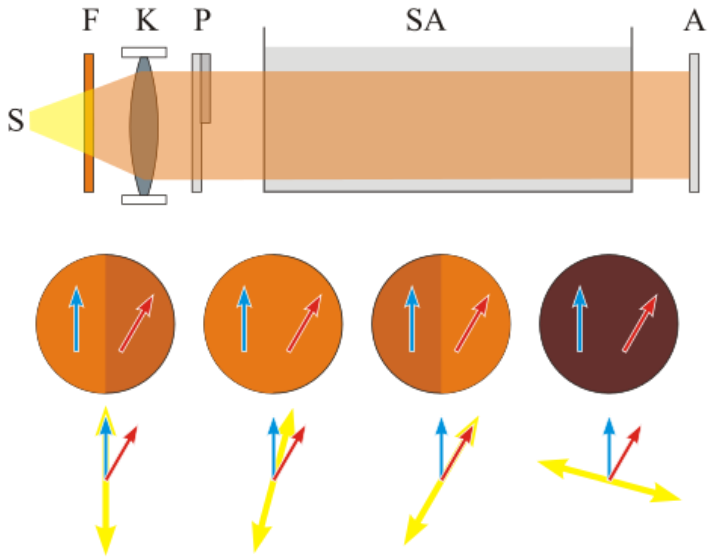
$$\left[\frac{\text{stopień m}^2}{\text{kg}}\right] \qquad 1.2.13$$

Z powodów historycznych w praktyce używa się innych jednostek skręcalności właściwej. Wszystkie są jednak definiowane względem jednostek układu SI (§TI 6).

Fakt 1.2.2.

Wartość skręcalności właściwej dla danej substancji jest zależna od długości fali światła.

Opisany wyżej polarymetr ustępuje dokładnością polarymetrowi półcieniowemu (rys. 1.2.9). W polaryzatorze półcieniowym Laurenta na część polaryzatora nałożona jest półfalówka[5], która skręca kąt polaryzacji światła wychodzącego z polaryzatora. W efekcie po przejściu przez próbkę i analizator widzimy dwa obszary o różnym natężeniu światła. Kręcimy polaryzatorem tak długo aż zniknie różnica jasności miedzy tymi dwoma obszarami. Występuje to w dwóch przypadkach, gdy obie połówki są tak samo jasne lub gdy obie połówki są tak samo ciemne. Do pomiaru wybieramy sytuację gdy obie połówki są tak samo ciemne, co wynika z faktu, że przy mniejszym natężeniu światła precyzyjniej potrafimy ocenić różnicę jasności dwóch sąsiadujących obszarów (prawo Webera-Fechnera). Polarymetr półcieniowy wykorzystuje fakt, że z większą precyzją określamy różnicę kontrastu między dwoma powierzchniami niż moment, w którym natężenie światła przechodzącego przez analizator jest najmniejsze (jak to ma miejsce w najprostszym polarymetrze). Polarymetry są przykładem prostego ale bardzo efektywnego narzędzia pomiarowego. Stosowane są w przemyśle spożywczym, farmaceutycznym i w chemii.

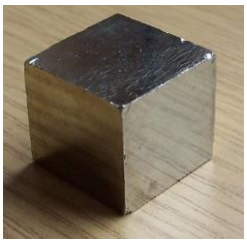


Rysunek 1.2.9. U góry schemat budowy polarymetru półcieniowego. S – źródło światła, F – filtr monochromatyczny, K – kolimator, P – polaryzator z nałożoną półfalówką, SA kuweta z badaną substancją, A – analizator; u dołu strzałka niebieska pokazuje kierunek polaryzacji światła po przejściu przez polaryzator i próbkę, strzałka czerwona kierunek polaryzacji światła po przejściu przez polaryzator półfalówkę i próbkę, strzałka żółta kierunek orientacji analizatora. Drugi rysunek pokazuje sytuację gdy oba obszary są tak samo jasne, czwarty gdy są tak samo ciemne; źródło Wikipedia

[5] Półfalówkę omówię już w (§3.2.1) tu wystarczy powiedzieć, że skręca ona polaryzację światła liniowo spolaryzowanego o kąt $\pi/2$.

2. Zjawisko dwójłomności ♦

Opis przejścia światła przez ośrodki materialne wymaga analizy oddziaływania impulsu elektromagnetycznego z cząsteczkami tego ośrodka. Zawierające elektrony cząsteczki są elektrycznie aktywne i mogą absorbować energię fali elektromagnetycznej a następnie ją reemitować. W efekcie ośrodki przezroczyste spowalniają falę świetlną i powodują jej ugięcie. Oddziaływanie pola elektrycznego fali świetlnej z cząsteczkami ośrodka zależy nie tylko od własności tych cząsteczek. Gdy cząsteczki tworzą stałą w czasie uporządkowaną strukturę (kryształy), to przejście fali świetlnej będzie również zależało od geometrii tego uporządkowania. Najprostszymi pod tym względem są kryształy regularne (rys. 2.1). W ten sposób krystalizuje około 12% wszystkich minerałów; na przykład sól kuchenna, diament, fluoryt, granaty, piryt (rys. 2. 2). W kryształach regularnych (inaczej, izotropowych) zjawisko załamania światła zachodzi tak samo jak w ośrodkach niekrystalicznych.

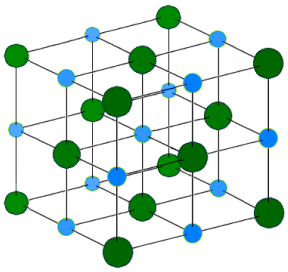


a) kryształ pirytu; FeS_2

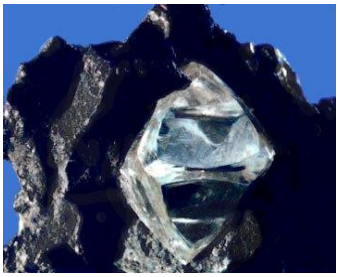
b) kryształ piropu; $Mg_3Al_2(SiO_4)_3$

c) kryształ piropu po oszlifowaniu

d) kryształ fluorytu; CaF_2



e) komórka kryształu soli; NaCl



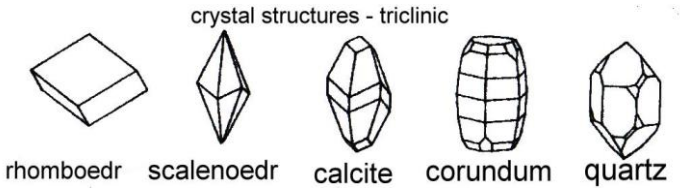
f) kryształ diamentu; C

Rysunek 2.1. Przykłady kryształów regularnych; źródło Wikipedia

Inną klasę stanowią kryształy optycznie anizotropowe (nieregularne). Kryształy, które nie należą do kryształów regularnych wykazują zjawisko dwójłomności. Mówimy, że są to ośrodki optycznie anizotropowe. W ośrodkach anizotropowych prędkość rozchodzenia się światła zależy od kierunku jego propagacji. Zależność prędkości rozchodzenia się światła od kierunku dobrze reprezentuje powierzchnia równej fazy. Niestety w kryształach anizotropowych kształt powierzchni równej fazy zależy od stanu polaryzacji światła. Okazuje

się, że istnieją dwa takie, wzajemnie ortogonalne (prostopadłe) stany liniowej polaryzacji, dla której powierzchnie falowe dla źródła punktowego dadzą się prosto określić. W przyszłości istnienie tych szczególnych stanów polaryzacji wykażemy rachunkiem, wychodząc z równań Maxwella. Na razie jednak nie dysponujemy niezbędną ku temu wiedzą. Stawia to nas na pozycji badaczy działających przed połową XIX wieku, którzy nie mieli do dyspozycji równań Maxwella. Nie przeszkodziło im to w dotarciu, do wielu głębokich konkluzji. Z opisem propagacji światła w ośrodkach anizotropowych jednoosiowych poradził sobie już w XVII wieku Huygens adaptując do nowej sytuacji swoją zasadę opisującą propagację fal ((§TIX 2).

Kryształy optycznie anizotropowe dzielą się na dwie klasy: jednoosiowe i dwuosiowe. Opis propagacji światła w kryształach dwuosiowych jest bardziej złożony niż w jednoosiowych, dlatego zostawimy go sobie na późniejszy temat – związany z równaniami Maxwella. Tu zajmiemy się ośrodkami jednoosiowymi, tym bardziej, że podstawowe elementy optyki polaryzacyjnej bazują na takich kryształach. Rysunek (2.2) przedstawia przykłady kryształów jednoosiowych.



Przykłady kryształów anizotropowych jednoosiowych (w układzie trygonalnym)

kryształ górski; SiO_2

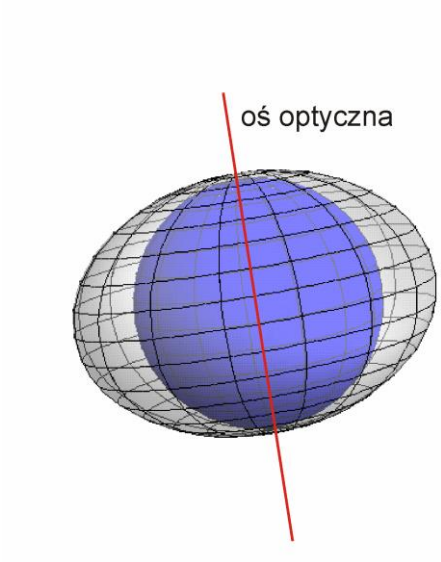
akwamaryn; Be

kryształ kalcytu; $CaCO_3$

Rysunek 2.2. Przykłady kryształów wykazujących anizotropię optyczną jednoosiową; źródło Wikipedia

Załóżmy, że promienie świetlne wychodzą z wyimaginowanego, punktowego, źródła światła znajdującego się wewnątrz kryształu. Jeżeli istnieje taki stan polaryzacji liniowej fali, dla którego powierzchnia falowa jest powierzchnią sferyczną, a dla polaryzacji do niej ortogonalnej jest powierzchnią elipsoidalną, to mamy do czynienia z ośrodkiem anizotropowym optycznie, jednoosiowym (rys.2.3). Jak widać z rysunku sfera i elipsoida w ośrodku anizotropowym jednoosiowym mają jeden wspólny promień. Wzdłuż kierunku wyznaczonego przez ten promień światło rozchodzi się z tą samą prędkością

niezależnie od stanu polaryzacji. Kierunek ten nazywamy osią optyczną kryształu. Fale wychodzące z punktu będę również nazywał falami cząstkowymi. W zasadzie Huygensa fale cząstkowe są tożsame z falami wtórnymi, które generowane są przez każdy punkt do którego dotarło zaburzenie. Rysunek (2.3) przestawia powierzchnie falowe fal cząstkowych w ośrodku anizotropowym jednoosiowym.



Rysunek 2.3. Dwie powierzchni falowe dla światła propagującego się ze źródła punktowego wewnątrz kryształu dwójłomnego. W przypadku ośrodka jednoosiowego istnieje stan polaryzacji, dla którego powierzchnia falowa dla źródła znajdującego się wewnątrz kryształu jest sferą. Dla polaryzacji prostopadłej do tej jednej szczególnej powierzchnia falowa jest elipsoidę obrotową (o przekroju kołowym). Sfera i elipsa mają wzdłuż jednej osi te same promienie. Oś skierowaną zgodnie z tym promieniem nazywamy osią optyczną kryształu. Wzdłuż osi optycznej promień nadzwyczajny rozchodzi się tak jak promień zwyczajny. Wzdłuż osi prostopadłej do osi optycznej różnica prędkości promienia zwyczajnego i nadzwyczajnego jest największa.

Definicja 2.1: Oś optyczna ośrodka anizotropowego

Kierunek wzdłuż, którego światło rozchodzi się z tą samą prędkością niezależnie od stanu polaryzacji nazywamy osią optyczną ośrodka anizotropowego.

Widać z tego również, że oś optyczna jest osią symetrii elipsoidy.

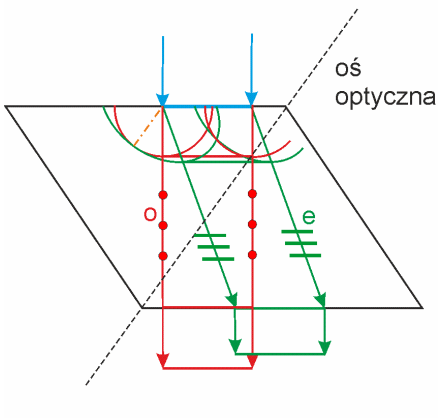
Definicja 2.2: Przekrój główny kryształu anizotropowego

Płaszczyzna wyznaczona przez oś optyczną kryształu anizotropowego i kierunek padania promienia nazywamy przekrojem głównym tego kryształu

Promienie reprezentujące propagację światła dla stanu polaryzacji, dla której powierzchnia falowa jest sferyczna nazywamy „promieniami zwyczajnymi”, a odnośna fala nazywana jest falą zwyczajną. Polaryzacja wiązki światła reprezentowanej przez promienie zwyczajne jest prostopadła do osi optycznej kryształu. Uzasadnię to w jednym z późniejszych tematów. Symbole wielkości związane z promieniami zwyczajnymi są zwykle wzbogacone o indeks „o” od łacińskiego „ordinarius”; na przykład współczynnik załamania dla promienia zwyczajnego oznaczamy jako n_o. Promienie związane ze stanem

ortogonalnym do „zwyczajnego” nazywamy „promieniami nadzwyczajnymi”, a wielkości z nimi związane piszemy z indeksem „e” (łacińskie „extraordinarius”); i tak współczynnik załamania dla promienia nadzwyczajnego oznaczamy jako n_e.

Do tej pory analizowaliśmy źródło punktowe emitujące falę o szczególnych polaryzacjach liniowych, tak że rozchodziła się fala albo sferyczna albo elipsoidalna. Jeżeli z punktowego źródła światła wyemitujemy światło niespolaryzowane, to fala świetlne rozdzieli się na te dwie szczególne składowe, zwyczajną o kształcie sferycznym i nadzwyczajną o kształcie elipsoidalnym. Obie fale będą liniowo i wzajemnie ortogonalnie spolaryzowane. Przy czym polaryzacja fali nadzwyczajnej „leży” w płaszczyźnie głównej (symbolizują zielone kreski, na rysunku (2.4)), a polaryzacja fali zwyczajnej jest prostopadła do płaszczyzny głównej (symbolizują ją czerwone kropki).



Rysunek 2.4. W ośrodku anizotropowym jednoosiowym światło można rozłożyć na dwa ortogonalne stany polaryzacji liniowej. Przy czym kierunki tych stanów można wybrać tak, że dla jednego z nich światło będzie się rozchodziło jak w ośrodku izotropowym, a powierzchnie falowe dla źródła punktowego będą sferami. Dla drugiego stanu polaryzacji powierzchnie falowe będą elipsoidami. Na rysunku czerwonym kolorem narysowane są fale wtórne i promienie dla fali normalnej, a zielonym kolorem dla fali nadzwyczajnej. czerwona kropki wskazują na kierunek polaryzacji fali zwyczajnej, a zielona kreski na kierunek polaryzacji fali nadzwyczajnej.

Narysowane tu powierzchnie elipsoidalne i sferyczne można interpretować jako powierzchnie reprezentujące prędkości rozchodzenia się fal wzdłuż różnych kierunków. W ośrodku jednorodnym powierzchnie falowe są punktami, do którego dotarło zaburzenie z punktowego źródła światła, w zadanej chwili czasu. Jeżeli czas przyjmiemy za jednostkowy, to długość wektora zaczepionego w punkcie źródłowym, o końcu na powierzchni falowej będzie taka sama, jak długość odnośnego wektora prędkości. Przy takiej interpretacji wykreślone powierzchnie nazywamy „powierzchniami prędkości fali”.

Definicja 2.3: Powierzchnie prędkości fali

Powierzchnią prędkości fali wyznaczoną jest przez punkty, do których światło dotarło, ze źródła punktowego, po ustalonym czasie

W przypadku kryształów jednoosiowych powierzchnia prędkości fali jest powierzchnią dwupowłokową złożoną ze sfery i elipsoidy. Dla kryształów jednoosiowych przyjmujemy ponadto następującą konwencję:

Definicja 2.4: Kryształy optycznie dodatnie i ujemne

Gdy prędkość promienia nadzwyczajnego jest większa od prędkości promienia zwyczajnego to kryształ jest optycznie ujemny, w przeciwnym razie kryształ jest optycznie dodatni.

Wprowadzę również następujące pojęcie

Definicja 2.5: Czoło fali w ośrodku anizotropowym

Czołem fali w ośrodku anizotropowym nazywamy powierzchnię styczną do powierzchni prędkości promienia zwyczajnego lub nadzwyczajnego, w danej chwili czasu

Dla fali cząstkowej czoło fali jest styczne do powierzchni tej fali. Na koniec jeszcze jedna definicja

Definicja 2.6: Ośrodek anizotropowy optycznie

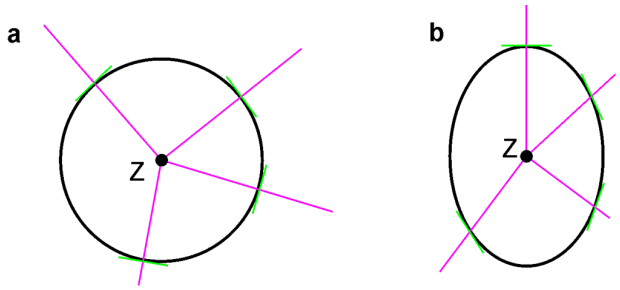
Mówimy, że ośrodek jest anizotropowy optycznie jeżeli jego cechy optyczne (takie jak np. prędkość rozchodzącej się w nim fali świetlnej) wykazują zależność od kierunku

Mikroskopowa analiza rozchodzenia się światła w kryształach jednoosiowych wymaga odwołania się do równań Maxwella. Kilka istotnych faktów można jednak wydedukować korzystając z zasady Huygensa (§TIX 2), co teraz uczynimy. Rysunek (2.4) przedstawia płaską falę padającą na kryształ jednoosiowy. Kryształ jest zeszlifowany w kształt rombu, a kierunek biegu fali padającej jest prostopadły do jednej z jego powierzchni. Zgodnie z tym co zostało powiedziane światło o ustalonej polaryzacji możemy rozłożyć na dwie składowe o polaryzacji liniowej wzajemnie prostopadłej. Na rysunku (2.4) fala płaska (rysowana na niebiesko) reprezentowana jest na dwa sposoby, przez fragment powierzchni falowej oraz przez promienie odpowiadające tym fragmentom powierzchni falowych. Skupimy uwagę na punktach na końcach niebieskiego odcinka. Zgodnie z zasadą Huygensa możemy je traktować jako źródła fal wtórnych. Dla fali zwyczajnej są to fale kuliste (rysowane kolorem czerwonym). Rysując styczną do dwóch takich fal kulistych wychodzących z punktów na brzegu odcinka niebieskiego widać, że fala zwyczajna rozchodzi się wzdłuż kierunku padania fali wejściowej. Jest to zgodne z prawem Snella. Niezależnie od współczynnika załamania kryształu, jeżeli kąt padania jest równy

zeru to i kąt załamania jest równy zeru. Ponadto, tak jak to było do tej pory, kierunek biegu promienia zgadza się z kierunkiem biegu fali (promień jest prostopadły do powierzchni falowej). Obok fali zwyczajnej rozchodzi się fala nadzwyczajna. Tu falami wtórnymi są elipsoidy, dla których jeden z promieni jest równoległy do osi optycznej kryształu i równy promieniowi odpowiedniej sfery dla promienia zwyczajnego (rys. 2.2). Na przykładzie z rysunku (2.4) widać, że prędkość promienia nadzwyczajnego jest większa od prędkości promienia zwyczajnego, co oznacza, że kryształ jest kryształem ujemnym (def. 2.4). Z rysunku widać, że fala nadzwyczajna wyprzedziła falę zwyczajną. W kryształach optycznie dodatnich kolejność fal byłaby odwrotna. Fale i promienie nadzwyczajne rysowane są na zielono. Obwiednia fal wtórnych będzie również linią prostą równoległą do fali padającej. Zatem kierunek biegu fali nie ulegnie zmianie. Jednak promienie nadzwyczajne zmienią kierunek biegu (biegną wzdłuż osi elipsy (rys. (2.3)))! Ponieważ, promienie wyznaczają kierunek biegu energii, to wektor falowy dla fali nadzwyczajnej nie jest prostopadły do jej powierzchni falowej[6], a ponadto po przejściu przez nasz kawałek kryształu powstaną dwa obrazy fali padającej wzajemnie między sobą przesunięte. Skutki tego efektu możesz obejrzeć na rysunku (1.1.1). Uwaga: do tej pory przyjmowaliśmy, że promienie są prostopadłe do powierzchni falowych, teraz jednak musimy sprawę sprecyzować. Tak jest dla ośrodka izotropowego lub fali zwyczajnej. W ośrodku optycznie anizotropowym i w przypadku fali nadzwyczajnej tak już zwykle nie jest. Sprawę ilustruje rysunek (2.5).

Fakt 2.1:

Dla fali nadzwyczajnej normalna do czoła fali w danym punkcie zwykle nie ma kierunku promienia przechodzącego przez ten punkt

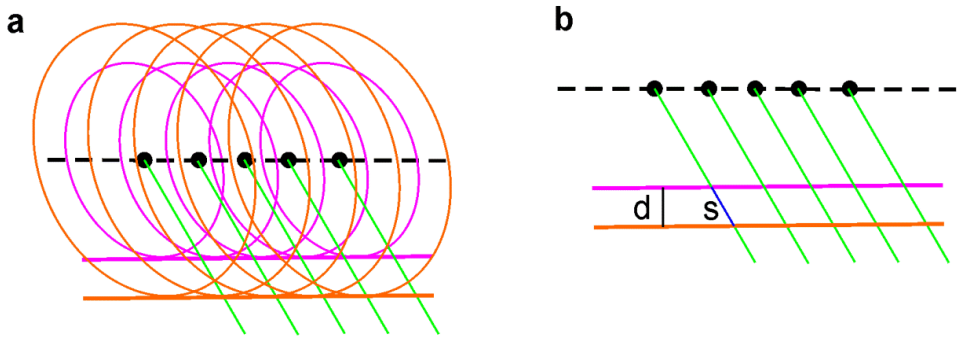


Rysunek 2.5. a) źródło punktowe *Z* emituje falę kulistą. Promienie są prostopadłe do powierzchni falowej. Jest tak w przypadku ośrodka optycznie izotropowego, lub fali zwyczajnej w ośrodku anizotropowym; b) źródło punktowe *Z* emituje falę elipsoidalną. Większość promieni nie jest prostopadła do powierzchni falowej. Wyjątkiem są promienie idące wzdłuż osi elipsoidy.

Należy pamiętać, że współczynnik załamania światła odnosi się do fali, a nie do promieni (rys. TIX 2.7). Współczynnik załamania mówi nam ile razy

[6] Wektor falowy pokazuje kierunek biegu fali

prędkość fazowa fali świetlnej jest mniejsza w ośrodku materialnym w stosunku do jej prędkości w próżni. Dla fali nadzwyczajnej prędkość promienia (prędkość z jaką przenosi się energia) nie jest ogólnie rzecz biorąc taka sama jak prędkość fazowa fali, co ilustruje rysunek (2.6).



Rysunek 2.6. a) na przerywanej linii czarnej (granica ośrodków o różnych współczynnikach załamania) zaznaczone jest kilka punktów, z których rozchodzą się eliptyczne fale wtórne. Przyjmijmy, że fala zostaje wyemitowana w chwili t_0. W chwili t_1 fale wtórne narysowane są na różowo, a ich obwiednia zaznaczona grubszą, różową kreską, w chwili t_2 fale wtórne narysowane są na pomarańczowo, a ich obwiednia zaznaczona grubszą kreską pomarańczową. Promienie nadzwyczajne zaznaczone są na zielono; b) rysunek przedstawia sytuację z części (a) ale bez fal wtórnych. Różowa i pomarańczowa kreska to front falowy fali nadzwyczajnej w chwili t_1 i t_2. Jak widać w czasie $\Delta t=t_2-t_1$ front ten przebył drogę d. W tym samym czasie promienie nadzwyczajne przeszły dłuższa drogę s. Promienie dla fali nadzwyczajnej nie są prostopadłe do powierzchni falowych, a prędkość promienia różni się od prędkości fazowej fali.

Fakt 2.2:

Dla fali nadzwyczajnej prędkość promienia jest w ogólnym przypadku różna od prędkości fazowej fali

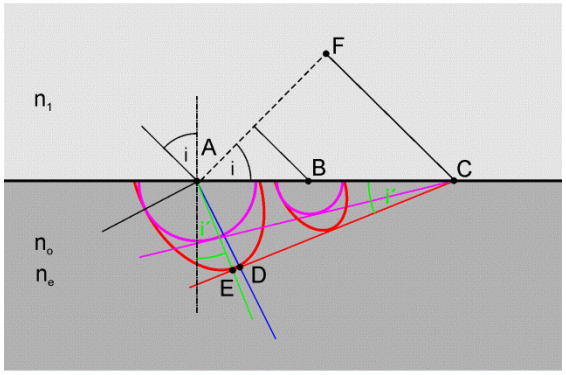
W (§TIX 2) wyprowadziłem prawo załamania w przypadku płaskiej granicy między dwoma ośrodkami optycznie izotropowymi. W podobny sposób można wyprowadzić prawo załamania dla fali nadzwyczajnej w ośrodkach anizotropowych (rys. 2.7). Niech fala płaska AF pada na powierzchnię kryształu anizotropowego jednoosiowego. Kolorem czerwonym wykreślone są dwie fale wtórne rozchodzące się z punktu A i B, w czasie, gdy światło z punktu F dojdzie do punktu C. Przyjmijmy, że droga FC jest jednostkowa, gdyż przyjęcie chwilowo takiej jednostki długości ułatwia rachunki. Wtedy długość odcinka FC jest równa co do wartości prędkości światła v_1 fali w ośrodku o współczynniku n_1. Ponadto odcinek AE będzie równy co do wartości prędkości v_e fali nadzwyczajnej w ośrodku o współczynniku załamani n_e. Możemy zatem zapisać

$$\frac{v_1}{v_e}=\frac{\overline{\mathrm{F\,C}}}{\overline{\mathrm{AE}}}=\frac{\overline{\mathrm{AC}}\sin(i)}{\overline{\mathrm{AC}}\sin(i_e)}\Longrightarrow\frac{v_1}{v_e}=\frac{\sin(i)}{\sin(i'_e)} \qquad 2.2.1$$

Biorąc pod uwagę, że $n_1=c/v_1$ i $n_e=c/v_e$, mamy

$$\frac{n_e}{n_1} = \frac{\sin(i)}{\sin(i_e)} \qquad 2.2.2$$

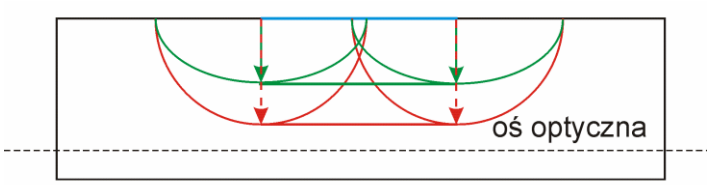
Co jest prawem załamania dla promienia nadzwyczajnego.



Rysunek 2.7. Na granicy ośrodków, izotropowego o współczynniku załamania n_1 i anizotropowego o współczynnikach załamania n_o i n_e zachodzi zjawisko załamania fali. Na rysunku przerywaną linią pokazany jest front falowy padającej fali płaskiej. Gdy fala dociera do punktu C, w punktach A i B wygenerowane już zostały dwie fale wtórne, zwyczajna (różowe okręgi) i nadzwyczajne (czerwone elipsy). Zgodnie z zasadą Huygensa (§TX 2) różowa linia wyznacza front falowy fali zwyczajnej, czerwona linia wyznacza front falowy fali nadzwyczajnej. Zielona linia wyznacza linię prostopadłą do nadzwyczajnego frontu falowego, a linia niebieska wskazuje bieg promienia nadzwyczajnego.

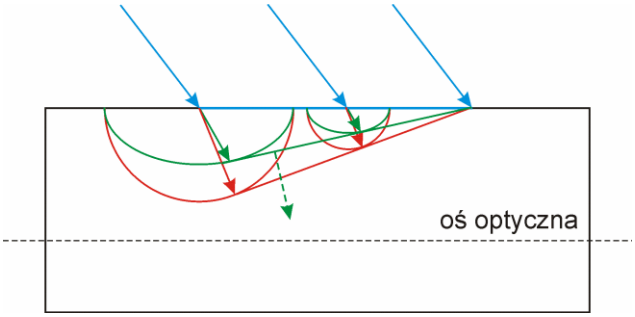
Jak widać postać prawa załamania dla fali nadzwyczajnej jest taka sama jak dla fali zwyczajnej. Wynika z tego, że jeżeli kąt padania fali jest równy zeru, to zeru jest też równy kąt załamania. Z rysunku (2.4) widać, że dla kąta padania zero promienie nadzwyczajne załamuję się pod kątem różnym od zera. Problem tkwi w tym, że jesteśmy przyzwyczajeni do wiązania prawa załamania z biegiem promieni. Ale jak już zaznaczyłem odnosi się ono do biegu fali. W ośrodkach optycznie izotropowych zapomnienie tego faktu nie prowadzi do błędu. Jednak w ośrodkach optycznie anizotropowych przypisanie prawa załamania promieniom może prowadzić do błędów. Przy kącie padania różnym od zera powierzchnia fali nadzwyczajnej również zostanie ugięta.

Gdy powierzchnia, na którą pada fala płaska jest równoległa do osi optycznej, to fala zwyczajna i nadzwyczajna podobnie jak odpowiednie promienie mają takie same kierunki (rys. 2.8). Rysunek (2.9) obrazuje przypadek, gdy tak jak na rysunku (2.8) powierzchnia, na którą pada światło jest ścięta równolegle do osi optycznej ale padająca fala płaska jest pochylona względem tej powierzchni. W takim przypadku powierzchnia falowa fali zwyczajnej (kolor czerwony) jest nachylona w stosunku do powierzchni falowej fali nadzwyczajnej (kolor zielony).



Rysunek 2.8. Fala płaska pada normalnie do kryształu anizotropowego jednoosiowego, którego oś biegnie prostopadle do kierunku padania światła. Wewnątrz kryształu wiązka rozdziela się na zwyczajną (kolor czerwony) i nadzwyczajną (kolor zielony). Jednak w tym szczególnym przypadku promień nadzwyczajny jest prostopadły do nadzwyczajnej powierzchni falowej.

Kierunek biegu promienia nadzwyczajnego (strzałka rysowana linią ciągłą; zielona) jest inny niż kierunek rozchodzenia się fali nadzwyczajnej (zielona strzałka rysowana linią przerywaną).



Rysunek 2.9. Sytuacja podobna do tej z rysunku (2.8), teraz jednak kierunek propagacji płaskiej fali padająca jest nachylony do normalnej. Kierunek propagacji czoła fali (przerywana strzałka) jest inny niż kierunek propagacji promienia (zielona strzałka ciągła).

Tabela (2.1) zawiera wartości współczynnika załamania zwyczajnego i nadzwyczajnego dla wybranych kryształów anizotropowych dla fali o długości λ=590nm. Dla większości dwójłomnych kryształów różnice wartości pomiędzy współczynnikiem załamania nadzwyczajnym i zwyczajnym nie są duże. W tabeli wyróżniają się pod tym względem kalcyt i rutyl. Niestety odpowiednia wartość współczynnika załamania to nie jedyna cecha kwalifikująca dany kryształ jako materiał do wyrobu przyrządów wykorzystujących zjawisko dwójłomności. Liczą się również inne cechy, jak wytrzymałość mechaniczna, łatwość obróbki, trwałość chemiczna, toksyczność, zakres fal, dla których kryształ jest przezroczysty. Z tych innych powodów nie wykorzystuje się rutylu. Szeroko używany jest kalcyt, choć jest minerałem silnie higroskopijnym i wymaga powłok ochronnych.

kryształ	n_o	n_e	n_e-n_o
kalcyt ($CaCO_3$)	1,658	1,486	-0,172
kwarc (SiO_2)	1,544	1,553	0,009
rutyl (TiO_2)	2,616	2,903	0,287
szafir (Al_2O_3)	1,768	1,760	-0,008
turmalin	1,669	1,638	-0,031

Tabela 2.1. Współczynniki załamania zwyczajny n_o i nadzwyczajny n_e dla wybranych kryształów wykazujących anizotropię jednoosiową.

Podsumuję jeszcze kwestie polaryzacji fali nadzwyczajnej i zwyczajnej. Wiemy już, że obie fale są spolaryzowane liniowo i wzajemnie ortogonalnie. Nadto kierunek polaryzacji fali nadzwyczajnej leży w płaszczyźnie głównej i jest prostopadły do normalnej do powierzchni fali (nie do kierunku promienia). Kierunek polaryzacji fali zwyczajnej jest prostopadły do płaszczyzny głównej i do normalnej do powierzchni fali i zarazem do kierunku promienia (bo kierunek promienia jest prostopadły do powierzchni falowej).

Fakt 2.3:

Fala zwyczajna i nadzwyczajna jest spolaryzowana liniowo, przy czym obie polaryzacje są wzajemnie ortogonalne, tak że kierunek polaryzacji fali nadzwyczajnej jest zgodna z płaszczyzną główną i prostopadły do normalnej do powierzchni fali nadzwyczajnej, a kierunek polaryzacji fali zwyczajnej jest prostopadły to tej płaszczyzny i do kierunku normalnego do powierzchni fali zwyczajnej.

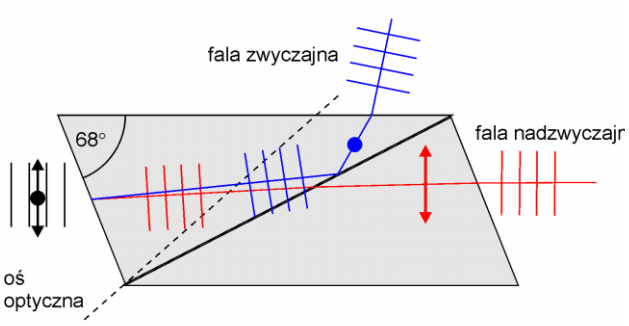
3. Dwójłomne elementy polaryzacyjne ♦

Kryształy wykazujące dwójłomność pozwalają na budowę wysokiej klasy elementów polaryzacyjnych. Wadą tych elementów jest ich wysoka cena. W tej części omówię najważniejsze elementy optyki polaryzacyjnej wykorzystujące zjawisko anizotropii optycznej

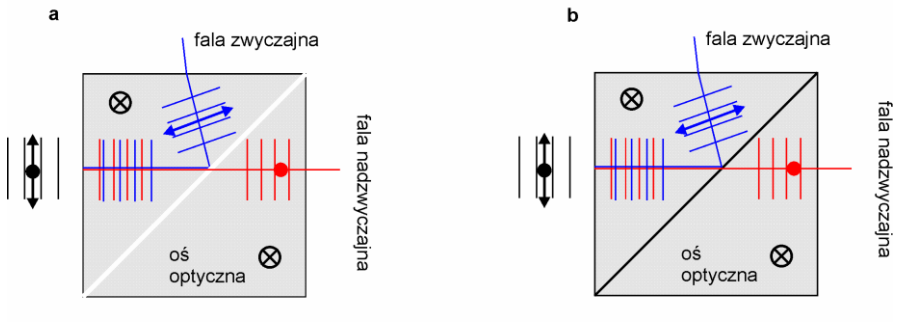
3.1. Pryzmat Nicole

Pierwszym elementem optycznym wykorzystującym zjawisko dwójłomności, który omówię, jest pryzmat Nicole. Wynaleziony w 1828 roku przez Szkota Williama Nicole składa się z dwóch sklejonych pryzmatów wyciętych z materiału optycznie dwójłomnego. Nicole wykorzystał kryształ kalcytu (kalcyt jest kryształem jednoosiowym ujemnym). Pryzmaty wycięte pod kątem 68° (rys. 3.1.1) zostały następnie sklejone balsamem kanadyjskim[7]. Dla przykładu dla światła o długości fali 589nm współczynnik załamania promienia zwyczajnego wynosi 1.658, dla promienia nadzwyczajnego i 1.486, a dla balsamu kanadyjskiego 1.545. Fala świetlna padająca w kierunku równoległym do osi optycznej układu, ulega rozdzieleniu na falę zwyczajną i nadzwyczajną. Pośrednia wartość współczynnika załamania balsamu kanadyjskiego pozwala tak dobrać geometrię sklejonych kryształów, że jedna z fal (na rysunku reprezentowana przez promień zwyczajny, który jest zawsze prostopadły do powierzchni fali zwyczajnej) pada na granicę klejenie pod kątem większym od kąta granicznego. W wyniku zjawiska całkowitego wewnętrznego odbicia (rys. TX 1.1.2) fala zostaje skierowana na ściankę boczną gdzie ulega pochłonięciu (ścianka boczna jest pokryta materiałem absorbującym). Przez pryzmat przechodzi spolaryzowana liniowo fala nadzwyczajna. Pryzmat Nicole działa jak polaryzator liniowy. Ze względu na to, że materiałem jest kryształ powierzchnie pryzmatu Nicole można bardzo dokładnie wyszlifować, co pozwala na jego zastosowanie w interferometrach, gdy wymagane są precyzyjne pomiary frontów falowych. Jedną z wad pryzmatu Nicole są skośne ścianki wejściowe. Obecnie pryzmaty Nicole zostały zastąpione przez inne, doskonalsze konstrukcje. Przykładami są pryzmaty Glana-Foucalta i Glana-Thompsona (rys. 3.1.2).

[7] Balsam kanadyjski to ekstrakt z żywicy jodły balsamicznej o współczynniku załamania zbliżonym do najpopularniejszych szkieł optycznych. Wykorzystywany jest do klejenia ze sobą elementów optycznych. Spoina z balsamu kanadyjskiego ma niską wytrzymałość termiczną. Obecnie zastępowany jest przez kleje syntetyczne.



Rysunek 3.1.1. Pryzmat Nicole wykonanego z kalcytu. Dwa pryzmaty z materiały dwójłomnego są ścięte tak, że po sklejeniu fala nadzwyczajna przechodzi przez warstwę klejącą a fala zwyczajna ulega całkowitemu wewnętrznemu odbiciu.



Rysunek 3.1.2. a) pryzmat Glana-Foucaluta składa się z dwóch pryzmatów wyciętych z materiału dwójłomnego (zwykle jest to kalcyt), tak jak pokazano na rysunku. Pomiędzy pryzmatami znajduje się warstwa powietrza, na której fala zwyczajna ulega zjawisku całkowitego wewnętrznego odbicia; b) pryzmat Glana-Taylora nie ma warstwy powietrznej. Pryzmaty są sklejone przez co są łatwiejsze w wykonaniu. Jednak pryzmat Glana-Foucaulta wytrzymuje pracę z wiązkami laserowymi o dużej mocy (do ok. 100W/cm^2), przy których warstwa klejąca w pryzmatach Glana-Taylora ulega degradacji.

3.2. Płytki falowe

Polaryzatory oparte o efekt dwójłomności są elementami wysokiej jakości ale wykonują tą samą pracę co polaroidy. Nowego typu elementami są płytki falowe. Płytka falowa wycięta jest z dwójłomnego jednoosiowego kryształu. Światło wprowadzane jest prostopadle do jej powierzchni, przy czym kierunek prostopadły do powierzchni jest zarazem kierunkiem prostopadłym do osi optycznej kryształu (rys. 3.2.1). Oznacza, to że wzdłuż tego kierunku różnica faz po przejściu przez płytkę między promieniem zwyczajnym i nadzwyczajnym jest największa. Różnicę tę łatwo obliczyć wykorzystując różne wartości współczynnika załamania dla promienia zwyczajnego (współczynnik załamania n_o) i nadzwyczajnego (współczynnik załamania n_e), przechodzącego przez płytkę o grubości d wynosi

$$\Delta s = d(n_e - n_o) \qquad 3.2.1$$

Wiąże się to z różnicą faz

$$\Delta\varphi = kd(n_e - n_o) \qquad 3.2.2$$

Różnica faz, poprzez liczbę falową k zależna jest od długości fali użytego światła. Wyróżnia się następujące rodzaje płytek falowych:

3.2.1. Półfalówka

Dla danej długości fali płytka półfalowa wnosi różnicę dróg optycznych równą

$$\Delta\varphi = kd(n_e - n_o) = m2\pi + \pi \qquad 3.2.3$$

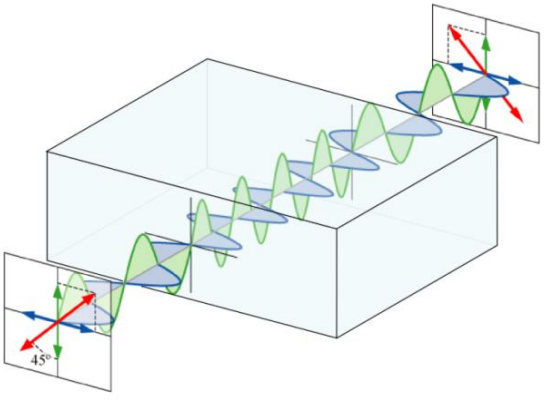
W zależności od tego czy fala nadzwyczajna jest szybsza czy zwyczajna, półfalówka powoduje, że szybsza z nich wyprzedza wolniejszą o π. Rysunek (3.2.2) ilustruje co się stanie gdy na półfalówkę puścimy światło spolaryzowane liniowo pod kątem $\pi/4$ do osi kryształu. Analizę tego przypadku możemy sprowadzić do składania dwóch harmonicznych drgań prostopadłych (tab. TVIII 2.2.1). Na tej samej zasadzie można pokazać, że półfalówka zmieni skrętność polaryzacji światła kołowo spolaryzowanego (z lewej na prawą i na odwrót). Gdy na wejściu pada światło spolaryzowane liniowo ale pod innym kątem niż $\pi/4$ lub światło o polaryzacji eliptycznej półfalówka zmieni stan polaryzacji światła na jakiś inny eliptyczny. Formalny opis takiej transformacji możliwy jest w formalizmie macierzy Jonesa (§4.1).

3.2.2. Ćwierćfalówka

Dla danej długości fali płytka ćwierćfalowa wnosi różnicę dróg optycznych równą

$$\Delta\varphi = kd(n_e - n_o) = m2\pi + \frac{\pi}{2} \qquad 3.2.4$$

W zależności od tego czy fala nadzwyczajna jest szybsza czy zwyczajna, ćwierćfalówka powoduje, że szybsza z nich wyprzedza wolniejszą o $\pi/2$. Gdy światło liniowo spolaryzowane pada na ćwierćfalówkę pod kątem $\pi/4$, to na wyjściu otrzymamy światło kołowo spolaryzowane (lewo lub prawoskrętnie) i na odwrót – światło kołowo spolaryzowane transformowane jest na światło kołowo liniowo spolaryzowane pod kątem $\pi/4$ do osi ćwierćfalówki.



Rysunek 3.2.3. Gdy światła padające na półfalówkę jest spolaryzowane liniowo pod kątem $\pi/4$ do osi kryształu to składowe H i V mają tą samą amplitudę i tą samą fazę na wejściu do kryształu. Obie składowe reprezentują fale harmoniczne, oscylujące wzdłuż dwóch prostopadłych osi. Fale te są zaznaczone kolorem zielonym i niebieskim. Na wyjściu jedna z fal jest opóźniona w fazie o $2\pi m+\pi$, co po ich złożeniu daje polaryzację liniową obróconą o π w stosunku do wejściowej; źródło rysunku Wikipedia

Dla danej długości fali falówka wnosi różnicę dróg optycznych równą

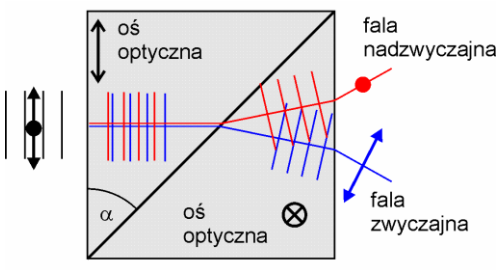
$$\Delta\varphi = kd(n_e - n_o) = m2\pi \qquad 3.2.5$$

3.2.3. Falówka

W zależności od tego czy szybsza jest fala nadzwyczajna czy zwyczajna, falówka powoduje, że szybsza z nich wyprzedza wolniejszą o 2π. Stan polaryzacji światła na wyjściu falówki jest taki sam jak na wejściu. Działanie płytki falowej wydaje się pozbawione sensu. Tak jednak nie jest. Światło białe spolaryzowane liniowo przechodząc przez falówkę obliczoną na daną długość fali λ_f pozostaje, dla tej długości fali, liniowo spolaryzowane (jeżeli kierunek polaryzacji jest zgodny z kierunkiem H lub V kryształu). Dla pozostałych długości fali, światło staje się eliptycznie spolaryzowane. Odpowiednio ustawiony analizator wycina składową λ_f, przepuszczając częściowo składowe o innych długościach fal. Analiza barwy światła na wyjściu pozwala na określenie własności polaryzacyjnych ośrodka (np. minerału) i jego identyfikacji. Efekt ten jest wykorzystywany w mikroskopii polaryzacyjnej.

3.2.4. Pryzmat Wollastona

Często stosowanym elementem polaryzacyjnym jest pryzmat Wollastona, który przepuszcza obie wiązki, zwyczajną i nadzwyczajną, przy czym na wyjściu z pryzmatu wiązki te są względem siebie przesunięte i rozchodzą się pod różnymi kątami jak to pokazuje rysunek (3.2.4).



Rysunek 3.2.4. Pryzmat Wollastona składa się z dwóch pryzmatów prostokątnych wykonanych tak, że po ich sklejeniu ich osie optyczne są wzajemnie prostopadłe i prostopadłe do założonego biegu wiązki padającej. Przez pryzmat przechodzą obie wzajemnie prostopadłe fale zwyczajna i nadzwyczajna. Kąt odchylenia tych fal na wejściu jest zależny od kąta α pryzmatów.

Widać, że w płytce półfalowej fala świetlna rozdziela się na dwie, które będziemy nazywali dalej modami.

Fakt 3.2.1: Mody polaryzacyjne

W płytce dwójłomnej światło rozdziela się na dwa wzajemnie ortogonalnie spolaryzowane mody

3.3. Dwójłomność wymuszona

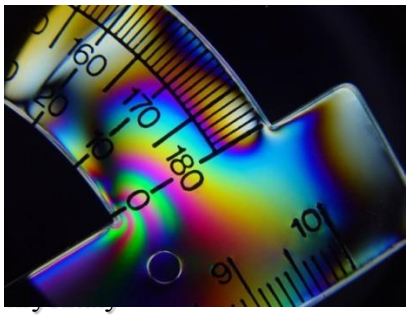
Dwójłomność naturalnie występuje w kryształach o ograniczonej symetrii. Można ją również wymusić na ciałach optycznie izotropowych poprzez zadziałanie czynnika zewnętrznego. Prostym tego przykładem jest wymuszenie

dwójłomności na izotropowym ciele stałym poprzez przyłożenie zewnętrznej siły.

Definicja 3.3.1: Fotoelastyczność

Fotoelestyczność jest efektem wymuszenia dwójłomności na izotropowym ciele stałym poprzez przyłożenie zewnętrznego nacisku mechanicznego

Najprostszy układ do badania pomiaru efektu fotoelastycznego oparty jest o krzyż polaryzacyjny (rys. 1.1.8). Włożenie wewnątrz krzyża polaryzacyjnego (między polaryzator a analizator) elementu, na który wywieramy nacisk pozwala obserwować barwne prążki na jego powierzchni (obrazy rejestrowane są oczywiście za analizatorem). Pojawienie się dwójłomności nie musi być związane z naciskiem. Wewnątrz materiału takiego jak szkło czy plastyk mogą się pojawić naprężenia „zamrożone" na skutek procesu produkcyjnego; na przykład zbyt szybkiego stygnięcia materiału. Przykład takich naprężeń widzianych w krzyżu polaryzacyjnym przedstawia rysunek (3.3.1). Pojawienie się barwnych rozkładów związane jest ze zmianą stanu polaryzacji światła przy przejściu przez naprężony element badany. W efekcie przez analizator przechodzi część światła. Ponieważ zmiana stanu polaryzacji światła jest zależna od długości światła, niektóre barwy są, w danym obszarze, bardziej tłumione przez analizator niż inne. Stąd mamy barwne obrazy za analizatorem. Metoda ta służy do pomiaru rozkładu naprężenia w materiałach optycznie izotropowych na skutek nacisku, wzrostu temperatury, ciśnienia lub naprężeń zamrożonych.



Rysunek 3.3.1. Plastikowy kątomierz widziany w krzyżu polaryzacyjnym. Barwne wzory świadczą o naprężeniach wewnątrz materiału; źródło Wikipedia, autor Nevit Dilmen, Creative Commons Attribution-Share Alike 3.0 Unported

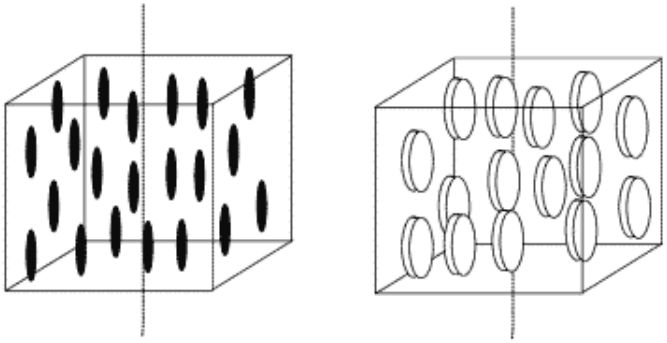
3.4. Ciekłe

Wielką karierę we współczesnej elektronice użytkowej zrobiły ciekłe kryształy. Zaczęło się od badań austriackiego fizjologa Friedricha Reinitzera (1888), który zajmował się własnościami pochodnych cholesterolu. Reinitzer odkrył dwustopniowe topnienie benzoesanu cholesterylu. W temperaturze 145.5°C związek ten topił się do postaci mętnej cieczy, a następnie w temperaturze 178,5°C przechodził do stanu przezroczystej cieczy. Otto Lehmann doszedł do wniosku, że faza mętna ma uporządkowane ułożenie cząsteczek, charakterystyczne dla kryształu, które jednak mogą się względem siebie przemieszczać, co jest charakterystyczne dla cieczy – stąd obecna nazwa „ciekłe

kryształy”. W następnych badaniach Reinitzer odkrył, że benzoesan cholesterolu skręca polaryzację światła (wykazuje aktywność optyczną). Po tych odkryciach zaprzestał badań nad ciekłymi kryształami. Przez kolejne osiemdziesiąt lat nie widziano dla nich istotnych zastosowań. Niemniej, pojawiały się w tym czasie prace donoszące o odkryciu kolejnych związków wykazujących fazę ciekłokrystaliczną. Ciekłe kryształy przez wiele dziesięcioleci pozostawały tematem badań czystej nauki, cieszącym się raczej niską popularnością.

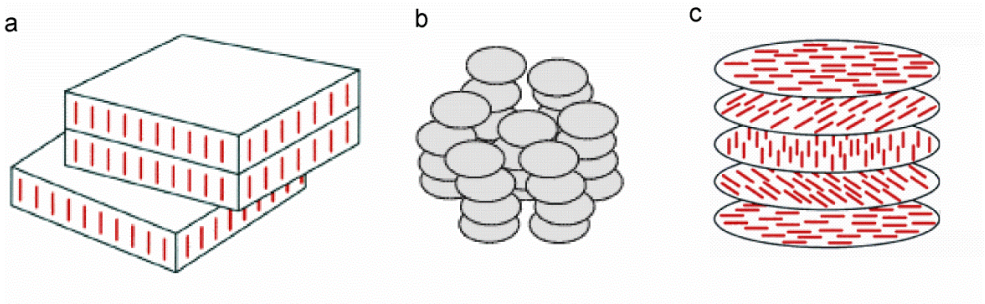
Uporządkowanie wewnątrz ciekłego kryształu związane jest z asymetrią w budowie cząsteczek. Są to cząsteczki albo bardzo wyciągnięte (długie sztywne pręty) albo w kształcie płaskich dysków. W pewnych warunkach cząsteczki te układają się w określony sposób powodując, że cała makroskopowa struktura ma złamaną symetrię. Stąd wynika anizotropia własności optycznych, mechanicznych, elektrycznych i magnetycznych takich ciał.

Badania nad ciekłymi kryształami doprowadziły do podziału ich na dwie duże grupy, nazywane fazami. Są to faza termotropowa i faza liotropowa. Fazy liotropowe są specyficznymi rodzajami emulsji, w której porządek wymuszany jest albo przez cząsteczki ciekłego kryształu, albo przez cząsteczki rozpuszczalnika. Ciekłe kryształy należące do fazy liotropowej nie znalazły do tej pory szerszego zastosowania i nie będziemy się nimi dalej zajmować. Kryształy fazy termotropowej przechodząc w stan ciekłokrystaliczny pod wpływem ogrzewania (stan ten nazywa się mezofazą termotropową). Benzoesanu cholesterylu, w którym Friedricha Reinitzera odkrył fazę ciekłokrystaliczną należy do tej grupy. Wśród kryształów termotropowych wyróżniamy następujące podgrupy: fazę smektyczną (3.4.1), nematyczną (rys. 3.4.2a) i kolumnową (rys. 3.4.2b) oraz cholesteryczną (rys. 3.4.2c). W 1962 roku Richard Williams, pracujący wówczas w laboratoriach RCA odkrył, że pod wpływem zewnętrznego pola elektrycznego przyłożonego do cienkiej warstwy nematyka, tworzą się regularne struktury, które nazwał domenami (dziś są to domeny Williamsa). Otworzyło to drogę do budowy prototypu wyświetlacza ciekłokrystalicznego. Pierwsze prace w tym kierunku prowadzone były przez Georga Heilmeiera. Jednak używany przez niego nematyk para-Azoxyanisole wykazywał fazę ciekłokrystaliczną w temperaturze wyższej niż 116°C – w efekcie pierwsze prototypy wyświetlaczy ciekłokrystalicznych nie miały szans na upowszechnienie.



Rys. 3.4.1. Faza smektyczna – długie cząsteczki lub cząsteczki w kształcie dysku są ułożone wzdłuż tego samego kierunku. Mogą się jednak względem siebie przesuwać; źródło Wikipedia, autor, Creative Commons Attribution-Share Alike 3.0 Unported;

Nematyki wykazujące efekt ciekłokrystaliczny w temperaturze pokojowej i niższej zostały odkryte przez J. Goldmachera i A. Castaleno (pracujących w RCA), w 1966 roku. To odkrycie otworzyło drogę ciekłym kryształom do praktycznych zastosowań. Na początku lat 70-tych ubiegłego wieku, wyświetlacze ciekłokrystaliczne pojawiły się w naręcznych zegarkach elektronicznych. Wraz z pierwszymi masowymi aplikacjami szybko rosła liczba prac nad własnościami i technologią ciekłych kryształów.

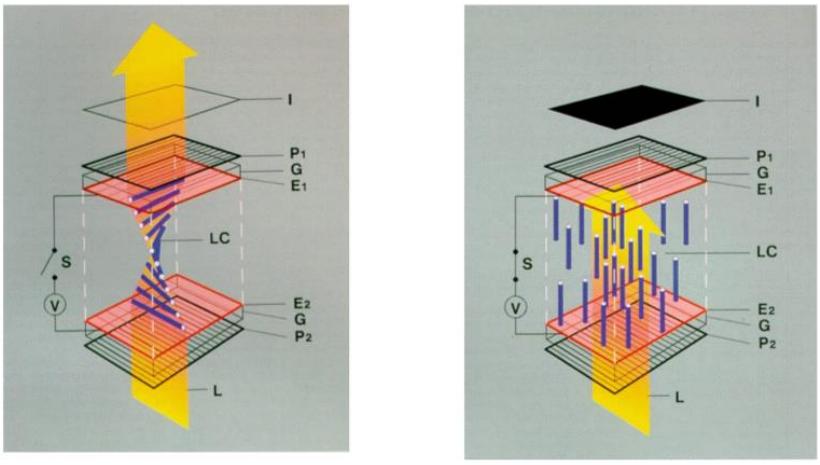


Rysunek 3.4.2. a) w fazie nematycznej cząsteczki ułożone są w warstwy. W każdej warstwie układają się wzdłuż jednego kierunku, prostopadłego do warstwy, jednak poszczególne warstwy mogą być względem siebie obrócone i przemieszczone; b) w fazie kolumnowej cząstki w kształcie dysku ułożone są wzdłuż kolumn; c) w fazie cholesterolowej cząstki układają się w warstwy równolegle do powierzchni tej warstwy. Kierunek orientacji cząstek obraca się od warstwy do warstwy; źródło Wikipedia, Creative Commons Attribution-Share Alike 3.0 Unported

Na początku lat 90-tych zaczęły rozpowszechniać się wyświetlacza ciekłokrystaliczne do przenośnych komputerów najpierw czarnobiałe a następnie kolorowe. Potem rozwój techniki poszedł bardzo szybko, tak że w 2006 roku wolumen sprzedaży telewizorów z wyświetlaczami LCD (od angielskiego (liquid crystal display; wyświetlacz ciekłokrystaliczny) był większy niż telewizorów z lampą kineskopową (CRT – ang. cathode ray tube). Rysunek (3.4.3) przedstawia schemat pojedynczej komórki wyświetlacza ciekłokrystalicznego.

Długa droga od laboratorium do przemysły ciekłych kryształów dobrze ilustruje sens badań podstawowych. Odkrycie ciekłych kryształów datuje się na 1888 r. W tym czasie była to w zasadzie ciekawostka naukowa. Przez kolejne lata badacze zajmowali się ciekłymi kryształami na zasadzie badania ciekawego

i złożonego zjawiska. Powoli budowała się baza danych związana z wiedzą na temat właściwości ciekłych kryształów i metod badawczych. Rozbudowywała się również teoria opisująca ich właściwości optyczne. Na tym etapie rygorystyczni praktycy mogliby stwierdzić, że finansowanie tych prac to strata czasu i pieniędzy. Dopiero w połowie lat 60-tych XX wieku pojawiła się wyraźna idea ich zastosowania. Jednak mało kto wówczas przypuszczał, że ciekłe kryształy mogą wyprzeć klasyczne lampy kineskopowe (CRT). Patrząc na pierwsze czarno- białe wyświetlacze (rys. 3.4.4) trudno się temu dziwić. A jednak, z pozoru nic nie mające praktycznego znaczenia odkrycie z 1888 roku, kiedy jeszcze nie było telewizji i komputerów, ponad sto lat później zrewolucjonizowało rynek mediów elektronicznych. Bez stopniowej akumulacji wiedzy przez uczonych, cierpliwie badających z pozoru nieużyteczne zjawisko ta rewolucja mogłaby nie nadejść lub mocno się opóźnić.



Rysunek 3.4.3. Schemat pojedynczej komórki wyświetlacza ciekłokrystalicznego transparentnego, w stanie wyłączonym (odłączone napięcie) – z lewej, oraz załączonej – z prawej. E_1 i E_2 przezroczyste elektrody, P_2 i P_1 układ polaryzator – analizator, G płytki szklane. W stanie wyłączonym polaryzacja światła zostaje skręcona o $\pi/2$ i światło jest całkowicie przepuszczane przez układ polaryzator – analizator. W stanie włączonym jednorodnie uporządkowane cząsteczki ciekłego kryształu nie mają wpływu na stan polaryzacji przechodzącego światła i światło jest blokowane; źródło Wikipedia, Creative Commons Attribution-Share Alike 3.0 Unported.

Rysunek 3.4.4. Segmentowy czarno biały wyświetlacz jaki był montowany w zegarkach od pierwszej połowy lat 70-tych XX wieku. Patrząc na taki wyświetlacz trudno było przewidzieć erę kolorowych telewizorów o dużych ekranach zrealizowanych z wykorzystaniem ciekłych kryształów.

3.5. Aktywność optyczna

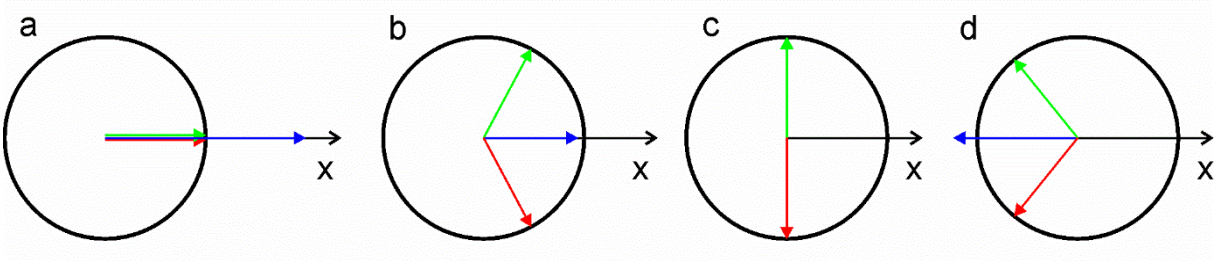
O aktywności optycznej już wspomniałem pisząc o polarymetrach (rys. 1.1.9). Przy przejściu przez niektóre substancje, kierunek światła spolaryzowanego liniowo ulega skręceniu. Efekt ten nazywamy aktywnością optyczną. Wiemy już, że aktywność optyczna związana jest z chiralnością cząstek (def. TV_3.2.5). Cukier wytwarzany biologicznie (przez burak cukrowy lub trzcinę cukrową) składa się z jednego enancjomeru (def. TV_3.2.6), dlatego spolaryzowane liniowo światło przy przejściu przez wodny roztwór cukru ma skręcony kierunek polaryzacji. Racemat cukru (mieszanina lewo i prawo skrętnego enancjomeru cukru w równych proporcjach) nie skręca kierunku polaryzacji. Przyjrzymy się bliżej zjawisku aktywności optycznej.

Aktywność optyczną odkrył w 1811 roku francuski fizyk Dominique Arago, który zaobserwował skręcenie kierunku polaryzacji światła przy przejściu przez płytkę wyciętą z kryształu kwarcu. W tym samym czasie J. Biot zaobserwował podobne efekty w niektórych cieczach i ich parach. Biot rozróżnił również materiały prawo i lewo skrętne (skrętność określamy patrząc w kierunku źródła). W 1822 angielski astronom John Herschel doszedł do wniosku, że w przypadku kryształów skrętność optycznie aktywnych materiałów zależy od orientacji sieci krystalicznej. Część kryształów występuje w dwóch formach lewo lub prawo skrętnej (rys. TV_3.2.12).

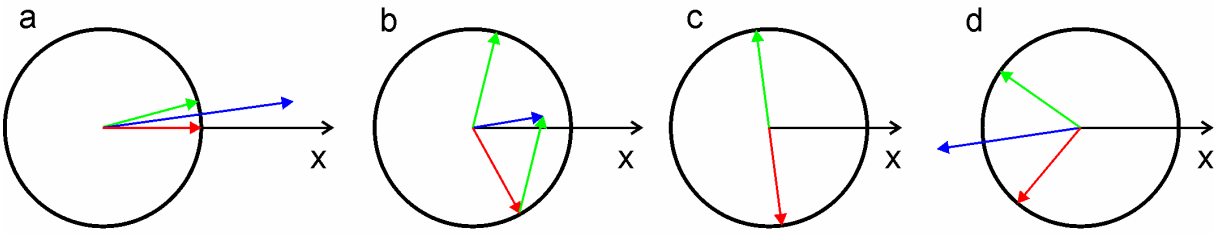
Podejście mikroskopowe do zjawiska aktywności optycznej wymaga odwołania się do elektrodynamiki klasycznej, a pełna teoria opiera się na mechanice kwantowej. Jednak jak już zapewne się zorientowałeś zanim dojdzie do opisania tego czy innego zjawiska za pomocą wielkich teorii ludzie radzą sobie za pomocą mniejszych konstrukcji, które pozwalają na budowanie mniej uniwersalnych, ale całkiem skutecznych modeli. Wielkie teorie powstają jako synteza wielu drobniejszych kroczków. Zbudujemy teraz małą teorię aktywności optycznej. Nie będziemy jednak szli drogą historyczną. Nasza teoria ma wartość dydaktyczną. Nie ma się na co tu obrażać, dobra teoria dydaktyczna jest cennym nabytkiem na drodze poznawania reguł przyrody a wielu uczonych, po cichu, dalej korzysta z teorii dydaktycznych, gdyż są one dla danej grupy problemów przyjaźniejsze od tych ogólnych. Pamiętasz co pisałem o Archimedesie (§DB 4). Najpierw obliczał pola za pomocą metody niepodzielnych (która ma charakter teorii dydaktycznej), a następnie, kiedy już wiedział jak się rzeczy mają, to dowodził swych wyników z pomocą metody wyczerpywania, która jest trudniejsza w użyciu ale za to ścisła. Nie oznacza to jednak, że operowanie teorią ogólną jest niepotrzebne. Bywa, że teoria dydaktyczna myli się, w pewnych aspektach badanego zjawiska. Analiza z punktu widzenia teorii ogólnej uwidacznia takie pułapki, jak również pokazuje możliwości ukryte przed teorią dydaktyczną.

Teoria, którą tu przedstawię została zaproponowana przez Fresnela w 1825 roku. Wiemy, że liniowo spolaryzowane światło możemy przedstawić

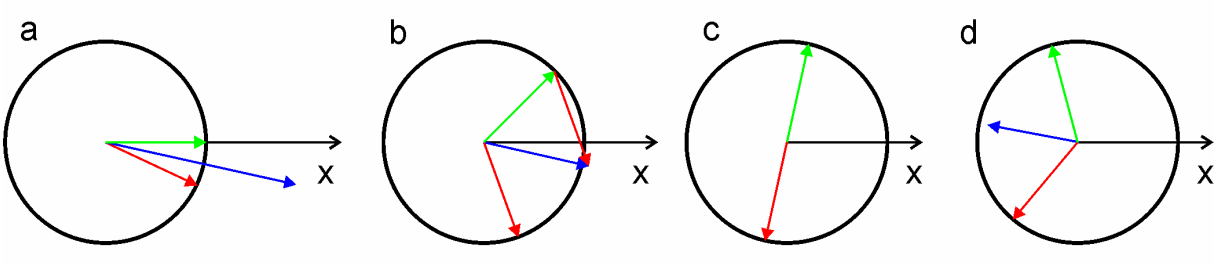
jak superpozycję dwóch ortogonalnych polaryzacji kołowych lewo i prawo skrętnej L i R. Wynika to z teorii składania drgań harmonicznych (TVIII 2.2), a sprawę omówię dokładnie w rozdziale (V). Fresnel przyjął, że w substancjach aktywnych optycznie, obie składowe poruszają się z różną prędkością. Rysunek (3.5.1) pokazuje, że suma dwóch przeciwbieżnych polaryzacji kołowych daje, jako wypadkową polaryzację liniową. Z rysunków (3.5.2) i (3.5.3) widać również, że jeżeli jedna ze składowych polaryzacji opóźnia się fazowo w stosunku do drugiej (to znaczy jedna składowa propaguje się wolniej niż druga), to kierunek polaryzacji liniowej zmienia się.



3.5.1. Zielony wektor (fazor) reprezentuje składową spolaryzowaną lewoskrętnie, czerwony prawoskrętnie, a niebieski fazory wypadkowy: a) stan początkowy, oba fazory są w fazie; b) chwila później oba fazory odeszły od osi x w przeciwne strony. Fazor wypadkowy zachowuje kąt, ale uległ skróceniu; c) kolejny moment, w który fazory składowe sumują się do zera; d) kolejny moment, fazor wypadkowy zachowuje kierunek ale ma teraz przeciwny zwrot w stosunku do sytuacji przedstawionej na rysunku (a) i (b). Widać, że wypadkowy fazor reprezentuje polaryzację liniową.



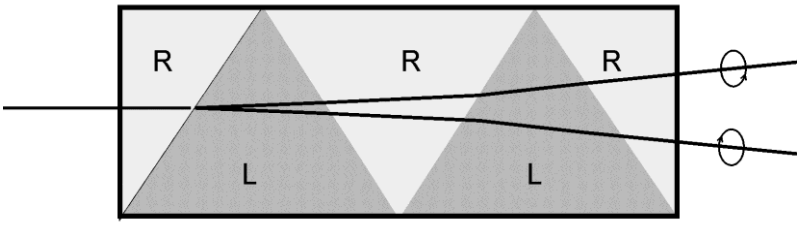
Rysunek 3.5.2. Na ośrodek aktywny optycznie pada światło spolaryzowane liniowo. W tym ośrodku składowa lewoskrętna odczuwa większy współczynnik załamania niż składowa prawoskrętna $n_R>n_L$. Oznacza to, że w pewnej odległość z od granicy tego ośrodka, gdy fazor czerwony powróci do punktu wyjścia fazor zielony będzie go wyprzedzał o jakiś kąt α (a). Dalej niech światło leci poza ośrodkiem aktywnym optycznie (np. w powietrzu), co oznacza, że $z=d$, gdzie d jest grubością ośrodka. W każdej następnej chwili kąt o jaki fazor zielony wyprzedza fazor czerwony będzie równy kątowi α. W efekcie w kolejnych chwilach, pokazanych na rysunkach (b, c, d) wypadkowy fazor będzie miał ten sam kierunek, a wypadkowa polaryzacja będzie polaryzacją liniową. Jednak kierunek wyjściowej polaryzacji liniowej będzie różny o kąt α, w stosunku do kierunku wejściowego. Dla $n_R>n_L$, kierunek polaryzacji liniowej obróci się w prawo.



Rysunek 3.5.3. Przypadek analogiczny do pokazanego na rysunku (2.2.8), z tym że $n_L>n_R$. Kierunek polaryzacji liniowej obraca się w lewo.

Dla przypadku, gdy współczynnik załamania dla składowej polaryzacji prawoskrętnej jest większy niż dla składowej polaryzacji lewoskrętnej $n_R>n_L$, kierunek polaryzacji światła będzie się obracał przeciwnie do biegu wskazówek zegara (rys. 3.5.2). Gdy $n_L>n_R$, kierunek polaryzacji będzie światła będzie się obracał zgodnie z ruchem wskazówek zegara (rys. 3.5.3).

Aby pokazać, że rozkład liniowej polaryzacji na dwa kołowe komponenty ma sens fizyczny, a nie tylko matematyczny, Fresnel rozdzielił obie składowe polaryzacji kołowej wykorzystując przyrząd pokazany na rysunku (3.5.4)



Rysunek 3.5.4. Przyrząd Fresnela składa się z pryzmatów wyciętych z lewo i prawo skrętnych kryształów kwarcu, sklejonych tak jak na rysunku. Przy takie konstrukcji na granicach pryzmatów lewo i prawo skrętna składowe światła spolaryzowanego liniowo (na wejściu do pryzmatu) odchylają się w przeciwne strony, co prowadzi do rozdzielenia obu wiązek na wyjściu. Przyrząd ten nazywamy „złożonym pryzmatem Fresnela".

Dotarliśmy do miejsca, w którym dalszy sensowny wykład z polaryzacji, wymaga więcej narzędzi matematycznych. Następny rozdział poświęcony będzie teorii macierzy. Przy okazji pokażę, prawdziwie możliwości metody ABCD, którą wprowadziłem w (§TXI 3.4).

4. Matematyczna przerwa na macierze ♣

Mam nadzieję, że z pojęciem macierzy się już oswoiłeś, bo będziemy się z nimi często spotykać. W (§TXI 3) przedstawiłem wykorzystanie macierzy do rozwiązywania problemów z optyki geometrycznej. Tu chcę zaprząc je do pracy przy opisie propagacji światła spolaryzowanego. Zanim do tego przystąpię muszę uzupełnić nasze zasoby wiedzy o rachunku macierzowym. W tym uzupełnianiu pomocne będzie przypomnienie sobie tego co o macierzach już wiemy, z dodatku (DE) oraz z wstępu do teorii przestrzeni wektorowych (§TIV 3.1).

Macierze kwadratowe możemy traktować jako operatory przekształcające jeden wektor w drugi (oba wektory należą do tej samej przestrzeni wektorowej V). Powiedzmy, że mamy przestrzeń wektorów V, o wymiarze 3 nad ciałem liczb rzeczywistych. Zapisując współrzędne wektora w postaci kolumny możemy przekształcić go w inny wektor za pomocą macierzy 3x3

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \\ a_{31}v_1 + a_{32}v_2 + a_{33}v_3 \end{bmatrix} = \begin{bmatrix} v'_1 \\ v'_2 \\ v'_3 \end{bmatrix} \qquad 4.1$$

Jak widać z powyższego primowane współrzędne nowego wektora można otrzymać z współrzędnych starego wektora poprzez działanie macierzy **A**.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad 4.1a$$

Ogólnie fakt ten zapisujemy

$$\mathbf{v'} = \mathbf{Av} \qquad 4.2$$

Otrzymane przekształcenie jest liniowe (co łatwo udowodnić), to znaczy, że jeżeli α i β są liczbami (mogą to być liczby zespolone), to

$$\mathbf{A}(\alpha\mathbf{v} + \beta\mathbf{w}) = \alpha\mathbf{Av} + \beta\mathbf{Aw} \qquad 4.3$$

Bez trudu można powyższe wyrażenia uogólnić na przestrzeń wektorową dowolnego wymiaru. Oczywiście dla przestrzeni o wymiarze N macierz przekształcenia jest macierzą o wymiarze NxN.

Przypominam, że w (§TIV 3.1) trafiliśmy na podobne wyrażenia. Na przykład wzór (TIV 3.1.26) opisywał relację między współrzędnymi wektora w dwóch różnych bazach. Formalnie jest to ten sam kształt wzoru co (3.1), ale jego znaczenie matematyczne jest inne. Wzór (4.1) opisuje zmianę współrzędnych jednego wektora na współrzędne innego wektora. Oba wektory muszą mieć wyrażone współrzędne w tej samej bazie. Wzór (TIV 3.1.26)

opisuje współrzędne tego samego wektora w dwóch różnych bazach. Teraz skupiamy się na pierwszym przypadku, to jest na operacji przekształcania jednego wektora w inny wektor poprzez działanie na jego współrzędne.

Macierz przekształcenia liniowego **A** może działać na dowolny wektor z danej przestrzeni wektorowej. Powiedzmy zatem, że w wyniku działania macierzy **A** na dwa wektory **v** i **w** otrzymujemy ten sam wektor **c**, wtedy, korzystając z liniowości (4.3) mamy

$$\mathbf{Av}-\mathbf{Aw}=\mathbf{A}\left(\mathbf{v}-\mathbf{w}\right)=\mathbf{c}-\mathbf{c}=0\Rightarrow\mathbf{v}-\mathbf{w}=0\Rightarrow\mathbf{v}=\mathbf{w} \qquad 4.4$$

Wynika z tego, że dwa różne wektory nie mogą, poprzez przekształcenie liniowe (4.2) przekształcić się w ten sam wektor. Słowem operator macierzowy **A** tasuje wektory przestrzeni wektorowej, na którą działa. Jest jednak jedno zastrzeżenie. Niech macierz **A** ma postać

$$\mathbf{A}=\begin{bmatrix}0 & 0 & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix} \qquad 4.5$$

Wtedy

$$\begin{bmatrix}0 & 0 & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}=\begin{bmatrix}0\\ a_{21}v_1+a_{22}v_2+a_{23}v_3\\ a_{31}v_1+a_{32}v_2+a_{33}v_3\end{bmatrix}=\begin{bmatrix}v'_1\\ v'_2\\ v'_3\end{bmatrix} \qquad 4.6$$

Jakakolwiek będzie współrzędna v_1 wektora **v**, to zawsze w wyniku otrzymamy $v'_1=0$. Możemy teraz wziąć dwa wektory różniące się współrzędną v_1 i po przekształceniu macierzą (4.5) otrzymamy ten sam wektor. Z (TIV 3.1.26) wiemy, że tak się dzieje gdy wyznacznik macierzy **A** jest równy zeru. Oznacza to, że jednoznaczności przekształcenia liniowego wektorów ma miejsce wtedy, gdy macierz tego przekształcenia jest nieosobliwa (jej wyznacznik jest różny od zera). Gdy jeden z wierszy macierz jest wypełniony zerami tak jak to ma miejsce w przykładzie (4.5), to wszystkie przekształcane wektory gubią jedną współrzędną. Macierz osobliwa redukuje wymiar przestrzeni wektorowej. W naszym przykładzie wszystkie wektory, po przekształceniu macierzą osobliwą (4.5) zamieszkują przestrzeń dwuwymiarową.

Przekształcenie odwrotne do przekształcenia A definiujemy w oczywisty sposób

Definicja 4.1: przekształcenie odwrotne do operatora liniowego A

Przekształcenie liniowe przestrzeni wektorowych $\mathcal{A}^{-1}$ nazywamy przekształceniem odwrotnym do przekształcenia liniowego przestrzeni wektorowych $\mathcal{A}$, gdy

$$\mathbf{Av}=\mathbf{w}\Rightarrow\mathbf{A}^{-1}\mathbf{w}=\mathbf{v} \qquad 4.7$$

Nie każde przekształcenie liniowe ma przekształcenie odwrotne

Twierdzenie 4.1: Istnienie przekształcenia odwrotnego do przekształcenia liniowego A

Przekształcenie liniowe A ma przekształcenie odwrotne gdy

$$\mathbf{v} \neq \mathbf{w} \Rightarrow \mathbf{A}\mathbf{v} \neq \mathbf{A}\mathbf{w} \qquad 4.8$$

Inaczej rzecz ujmując wyznacznik macierzy **A** musi być różny od zera. Z powyższego wynika bezpośrednio że

$$\mathbf{A}\mathbf{A}^{-1}\mathbf{v} = \mathbf{v} \Rightarrow \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \qquad 4.9$$

I jest macierzą jednostkową, to znaczy, że macierz przekształcenia odwrotnego do przekształcenia reprezentowanego przez macierz **A** jest macierzą odwrotną do macierzy **A**.

Macierz **A** interpretujemy jako macierz przekształcenia wektorów w wektory. Ogólnie (nie odwołując się do współrzędnych) możemy posługiwać się wzorem (4.2). Wtedy o obiekcie **A** nie mówimy macierz, tylko operator liniowy. Macierz jest reprezentacją operatora liniowego w danej bazie. To znaczy, że żeby napisać operator liniowy **A** w postaci macierzowej (4.1a) musimy w danej przestrzeni wektorowej wybrać bazę.

Zbadamy jak zmienia się reprezentacja macierzowa operatora przekształcenia liniowego przy zmianie bazy w przestrzeni wektorowej V. Rozumowanie będę prowadził dwutorowo na wzorach ogólnych i dla przypadku dwuwymiarowego. Niech zbiory wektorów $\{\mathbf{e}_i\}$ i $\{\mathbf{e}'_i\}$ są dwiema bazami przestrzeni wektorowej o wymiarze *N*. Macierz **A** działa na wybrany wektor bazowy w następujący sposób

$$\begin{bmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1N} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{N1} & \vdots & \vdots & \vdots & a_{NN} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ki} \\ \vdots \\ a_{Ni} \end{bmatrix} \qquad 4.10$$

W przypadku dwuwymiarowym mamy dwa wektory bazy, zatem

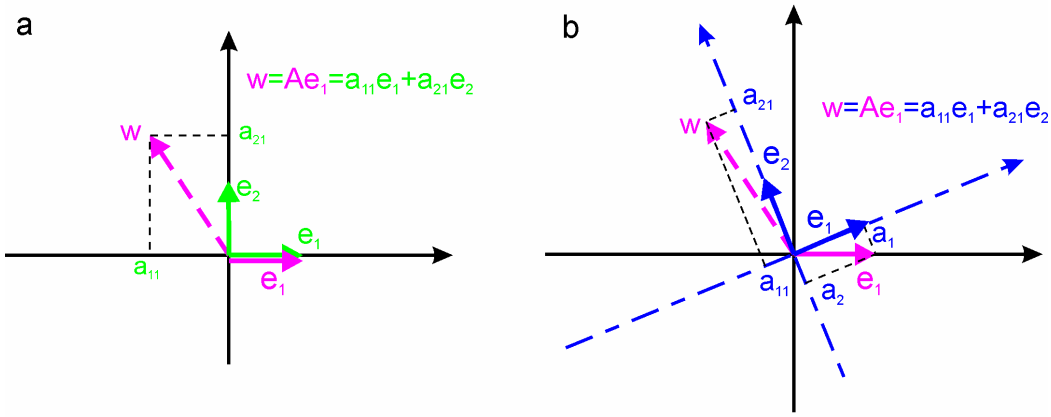
$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \text{ lub } \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} \qquad 4.10a$$

Możemy ogólnie zapisać

$$\mathbf{A}\mathbf{e}_i = \sum_{k=1}^{N} a_{ki}\mathbf{e}_k \qquad 4.11$$

Działanie macierzy **A** na wektor bazowy ilustruje rysunek (4.1.1a). W bazie primowanej współrzędne wektora $\mathbf{e}_i$ i macierzy operatora liniowego **A** ulegną

zmianie (oznaczamy je primami), choć zmiana bazy nie zmienia ani wektora ani operatora, tylko jego reprezentację.



Rysunek 4.1.1. a) Ilustracja działania operatora **A** na wektor $\mathbf{e}_1$. Wektor $\mathbf{e}_1$ przechodzi w wektor **w**. Wektor $\mathbf{e}_1$ jest tożsamy z elementem bazy jednostkowej $\mathbf{e}_1$ i jako taki ma współrzędne (1,0). Działanie operatora **A** w tej bazie dane jest przez liczby a_{11} i a_{21}. Jak widać z rysunku liczby a_{11} i a_{21} są jednocześnie współrzędnymi wektora **w**. Zatem kolumna macierzy **A** pokazuje jak wyglądają współrzędne wektora bazy po przekształceniu liniowym danym macierzą **A**, przy czym *n*-ta kolumna reprezentuje współrzędne *n*-tego wektora bazy; b) W innym (niebieskim układzie współrzędnych) ten sam wektor bazy $\mathbf{e}_1$ ma inne współrzędne (a_1,a_2). Nie jest on już wektorem bazy. Wektory nowej bazy narysowane są na niebiesko. W nowej bazie ten sam operator **A** działa na $\mathbf{e}_1$ w taki sam sposób; ten sam operator działając na ten sam wektor da ten sam wektor **w**. Ale w nowych współrzędnych jego działanie jest opisane przez nowe współrzędne a_{11} i a_{21}, które na rysunku wyróżnione są innym kolorem.

W nowej bazie działanie tego samego operatora na ten sam wektor wyrazi się wzorem (ograniczę się do zapisu dwuwymiarowego)

$$\begin{bmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\begin{bmatrix} a'_{11} \\ a'_{21} \end{bmatrix} \text{ lub } \begin{bmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix}=\begin{bmatrix} a'_{12} \\ a'_{22} \end{bmatrix} \qquad 4.12$$

Co przedstawia rysunek (4.1.1b). W skrócie wzór (4.12) zapiszę tak

$$\mathbf{A}'\mathbf{e}'_j=\sum_{k=1}^{N} a'_{kj}\,\mathbf{e}'_k \qquad 4.13$$

Podkreślę, że wzory (4.11 i 4.13) opisują przekształcenie tego samego $\mathbf{e}_j$ wektora bazy przez ten sam operator liniowy **A**, ale w różnych bazach. Niech **B** będzie macierzą przejścia od bazy $\{\mathbf{e}_i\}$ do bazy $\{\mathbf{e}'_i\}$. We wzorze (TIV 3.1.21) oznaczałem taką macierz przez **A**, ale tu litera **A** zarezerwowana jest dla macierzy reprezentujących operatory liniowe, dlatego użyłem litery **B**. Zgodnie z wzorem (TIV 3.1.21), możemy każdy wektor bazowy $\mathbf{e}'_j$ wyrazić przez wektory bazowe $\{\mathbf{e}_i\}$

$$\mathbf{e}'_j = b_{1j}\mathbf{e}_1 + \cdots + b_{Nj}\mathbf{e}_N = \sum_{i=1}^{N} b_{ij}\mathbf{e}_i \qquad 4.14$$

Prawa strona (4.14) ma taką samą postać jak prawe strony (4.11) i (4.13). Oznacza to, że macierz przejścia między bazami $\{\mathbf{e}_i\}$ i $\{\mathbf{e}'_i\}$ może być również interpretowana jako macierz reprezentująca operator liniowy działający na przestrzeni wektorowej V, zgodnie z wzorem

$$\mathbf{e}'_j = \sum_{i=1}^{N} b_{ij}\mathbf{e}_i = \mathbf{B}\mathbf{e}_i \qquad 4.15$$

Wstawiając (4.15), za $\mathbf{e}'_j$ do lewej strony (4.13) mam

$$\mathbf{A}'\mathbf{e}'_j = \sum_{k=1}^{N} a'_{kj}\,\mathbf{e}'_k = \sum_{k=1}^{N} a'_{kj}\,\mathbf{B}\mathbf{e}_k = \sum_{k=1}^{N} a'_{kj} \sum_{i=1}^{N} b_{ik}\mathbf{e}_i = \sum_{i=1}^{N}\left(\sum_{k=1}^{N} b_{ik}a'_{kj}\right)\mathbf{e}_i \qquad 4.16a$$

Z drugiej strony, ponownie korzystając z (4.15) mam

$$\mathbf{A}\mathbf{e}'_j = \sum_{k=1}^{N} b_{kj}\mathbf{A}\mathbf{e}_k = \sum_{k=1}^{N} a_{kj} \sum_{i=1}^{N} b_{ik}\mathbf{e}_i = \sum_{i=1}^{N}\left(\sum_{k=1}^{N} a_{ik}b_{kj}\right)\mathbf{e}_i \qquad 4.16b$$

Ponieważ wektory bazy są liniowo niezależne, z porównania (4.16a) i (4.16b) mamy

$$\sum_{k=1}^{N} a_{ik}b_{kj} = \sum_{k=1}^{N} b_{ik}a'_{kj} \qquad 4.17$$

W postaci macierzowej ostatnia równość przyjmie kształt

$$\mathbf{AB} = \mathbf{BA}' \Rightarrow \mathbf{A}' = \mathbf{B}^{-1}\mathbf{AB} \qquad 4.18$$

Wyrażenie (4.18) określa szukany wzór na postać macierzy przekształcenia liniowego $\mathbf{A}'$ w bazie primowanej $\{\mathbf{e}'_i\}$, gdy znana jest jego postać $\mathbf{A}$ w bazie nieprimowanej $\{\mathbf{e}_i\}$, a macierz $\mathbf{B}$ jest macierzą przejścia od bazy nieprimowanej do bazy primowanej.

Definicja 4.2: Macierze podobne

Mówimy, że macierz $\mathcal{A}'$ jest podobna do macierzy $\mathcal{A}$, jeśli istnieje taka macierz odwracalna $\mathcal{B}$, że zachodzi (4.18).

Podobieństwo macierzy jest pojęciem ogólnym. W kontekście przekształcenia liniowego to, że dwie macierze są podobne oznacza, że są reprezentacją tego samego przekształcenia liniowego wektorów w dwóch różnych bazach.

Fakt 4.1:

Macierz podobne reprezentują, w różnych bazach, ten sam operator liniowy

Dla macierzy podobnych zachodzi użyteczne twierdzenie

Twierdzenie 4.2: O wyznacznikach macierzy podobnych

Jeżeli $\mathbf{B}$ i $\mathbf{C}$ są macierzami podobnymi to $\det\mathbf{B}=\det\mathbf{C}$

Łatwo jest to udowodnić. Ponieważ **B** i **C** są podobne, to istnieje taka nieosobliwa macierz **A**, że $\mathbf{C}=\mathbf{A}^{-1}\mathbf{B}\mathbf{A}$. Korzystając z faktu, że wyznacznik iloczynu macierzy jest równy iloczynowi wyznaczników oraz wzoru mamy

$$\det\mathbf{C}=\det\mathbf{A}^{-1}\mathbf{B}\mathbf{A}=\det\mathbf{A}^{-1}\det\mathbf{A}\det\mathbf{B}=\det\mathbf{B} \tag{4.19}$$

Użyteczną wielkością charakteryzującą macierze jest ich ślad.

Definicja 4.3: Ślad macierzy A

Śladem macierzy kwadratowej A nazywamy sumę wartości jej elementów diagonalnych

Dla macierzy kwadratowej o *N* wierszach zapisujemy to tak

$$\mathrm{Tr}\mathbf{A}=\sum_{i=1}^{N}a_{ii} \tag{4.20}$$

Oznaczenie Tr pochodzi od angielskiego słowa „trace” (ślad). Dla śladu macierzy podobnych zachodzi

Twierdzenie 4.3: Ślad macierzy podobnych

Jeżeli $\mathbf{B}$ i $\mathbf{C}$ są macierzami podobnymi to $\mathrm{Tr}\mathbf{B}=\mathrm{Tr}\mathbf{C}$

Stosunkowo prosty dowód tego twierdzenia można znaleźć w każdym podręczniku algebry liniowej. Z twierdzeń (4.2) i (4.3) wynika, że ślad i wyznacznik nie zależą od wyboru bazy, w której obliczamy współczynniki macierzy reprezentującej przekształcenie liniowe wektorów. Zatem i wyznacznik i ślad reprezentują właściwości odwzorowania liniowego.

Fakt 4.2.

Ślad i wyznacznik macierzy nieosobliwej reprezentującej dany operator liniowy A nie zależy od wyboru bazy, co oznacza, że i ślad i wyznacznik charakteryzują ten operator

Masz może już dość matmy? O nie mój drogi za dobrze nam idzie. Proponuję przerwę na pięć przysiadów i jeszcze trochę matematycznego wysiłku. Wprowadzony tu formalizm będzie bardzo użyteczny.

4.1. Wartości własne

Wartości własne to jedno z podstawowych pojęć rachunku macierzowego. Krótkie wprowadzenie do tego tematu zacznę od definicji wartości własnych, ale od razu dla operatora liniowego. Ponieważ operatory liniowe reprezentowane są przez macierze od razu widać jak przestawić tą definicję z operatorów na macierze

Definicja 4.1.1: Wartość własna operatora liniowego A

Wartością własną operatora liniowego $\mathcal{A}$ jest taka liczba (skalar), że dla pewnego niezerowego wektora $\mathbf{v}$ spełnione jest równanie

$$\mathbf{Av} = \lambda \mathbf{v} \qquad 4.1.1.$$

Idąc za ciosem zdefiniuję wektory własne

Definicja 4.1.2: Wektor własny operatora liniowego A

Niech λ jest wartością własną operatora liniowego $\mathcal{A}$, wtedy wektor $\mathbf{v}$ spełniający równanie (4.1.1) nazywamy wektorem własnym tego operatora

Równanie (4.1.1) oznacza, że dla danego przekształcenia liniowego **A** wyszukujemy takie wektory z przestrzeni liniowej V, że działanie tego przekształcenia na te wektory sprowadza się do mnożenia tegoż wektora przez liczbę. Słowem wektory nie zmieniają kierunku, pod wpływem działania operatora **A** (ale mogą zmienić zwrot gdy wartość własna jest ujemna).

Powiedzmy, że dla danego odwzorowania liniowego i danej wartości własnej λ istnieje N wektorów własnych, $\mathbf{v_1}, \ldots, \mathbf{v_N}$. Wtedy kombinacja liniowa tych wektorów jest też wektorem własnym

$$\mathbf{A}\left(\alpha_1\mathbf{v}_1 + \cdots + \alpha_N\mathbf{v}_N\right) = \alpha_1\mathbf{A}\mathbf{v}_1 + \cdots + \alpha_N\mathbf{A}\mathbf{v}_N = \lambda\left(\alpha_1\mathbf{v}_1 + \cdots + \alpha_N\mathbf{v}_N\right) \qquad 4.1.2$$

Definicja 4.1.3: Geometryczna krotność wartości własnej

Geometryczną krotnością wartości własnej nazywamy liczbę liniowo niezależnych wektorów własnych należących do tej wartości własnej

Skoro zajmujemy się wartościami własnymi, to muszą mieć one jakieś istotne znaczenie dla fizyki. Zatem nie powinny zależeć od wyboru bazy w jakiej reprezentujemy, przez odpowiednią macierz, dany operator liniowy. Jak można się domyśleć zachodzi następujące twierdzenie

Twierdzenie 4.1.4: Wartości własne macierzy podobnych

Macierze podobne mają te same wartości własne. Wartości własne macierzy podobnych mają tą samą geometryczną krotność

Oczywiście na domysłach w matematyce nie kończymy, a prosty dowód tego twierdzenie dostępny jest w każdym podręczniku algebry liniowej.

Obliczę wartość własne macierzy diagonalnej

$$\begin{bmatrix} \beta_{11} & 0 & 0 \\ 0 & \beta_{22} & 0 \\ 0 & 0 & \beta_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \qquad 4.1.3$$

Po wymnożeniu tej macierzy przez wektor mamy trzy równania na wartości własne

$$\begin{cases} \beta_{11}v_1 = \lambda v_1 \\ \beta_{22}v_2 = \lambda v_2 \\ \beta_{33}v_3 = \lambda v_{33} \end{cases} \qquad 4.1.4$$

Stąd jasno wynika, że elementy macierzy diagonalnej są jej wartościami własnymi

$$\begin{cases} \lambda_1 = \beta_{11} \\ \lambda_2 = \beta_{22} \\ \lambda_3 = \beta_{33} \end{cases} \qquad 4.1.5$$

Tym wartościom własnym odpowiadają trzy wektory własne o współrzędnych

$$\lambda_1 \to \begin{bmatrix} v_1 \\ 0 \\ 0 \end{bmatrix};\ \lambda_2 \to \begin{bmatrix} 0 \\ v_2 \\ 0 \end{bmatrix};\ \lambda_3 \to \begin{bmatrix} 0 \\ 0 \\ v_2 \end{bmatrix} \qquad 4.1.6$$

Fakt 4.1.1:

Wartości własne macierzy diagonalnej są wartościami jej diagonalnych współczynników

Widać z tego, że jeżeli znajdziemy metodę diagonalizacji macierzy, to znajdziemy łatwy sposób obliczania jej wartości własnych. Dlaczego? Pomyśl diagonalizacja oznacza zmianę wartość współczynników macierzy, czyli zmianę bazy, w której tą macierz obliczamy. Intepretując macierze jako reprezentacje operatora liniowego mamy ten sam operator w dwóch bazach, przy czym w jednej z nich macierz operatora jest diagonalna. Obie macierze reprezentują ten sam operator liniowy, muszą być więc macierzami podobnymi, które na mocy twierdzenia (4.1.4) mają te same wartości własne. Przekształcenie macierzy operatora liniowego do postaci diagonalnej oznacza, że istnieje baza, w której działanie macierzy tego operatora na wektor sprowadza się do mnożenia współrzędnych tego wektora przez kolejne wartości własne tego operatora. Odpowiada to znalezieniu takiej bazy, dla której dany wektor własny jest równoległy do jednego z wektorów bazowych. Znalezienie postaci diagonalnej macierzy danego operatora (jeżeli taka postać istnieje) znacznie upraszcza rachunki. Przykładem takiego uproszczenia jest następujące zadanie. Mamy operator liniowy **A**. Jak wygląda *n*-krotne działanie tego operatora na wektor? Takie *n*-krotne działanie operatora możemy reprezentować jako *n*-krotne mnożenie macierzy **A** reprezentującej ten operator. Możemy to potraktować jak obliczenie *n*-tej potęgi macierzy **A**.

$$\mathbf{A}^n = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}^n \qquad 4.1.7$$

Dla dużego *n* i dużej macierzy jest to poważne zadanie obliczeniowe, chyba że uda się nam znaleźć macierz diagonalną $\mathbf{\Lambda}$ i macierz przekształcenia **B**, taką że zachodzi (4.19), wtedy mamy

$$\mathbf{A}^n = \left(\mathbf{B}^{-1}\mathbf{\Lambda}\mathbf{B}\right)^n = \mathbf{B}^{-1}\mathbf{\Lambda}\mathbf{B}\,\mathbf{B}^{-1}\mathbf{B}\cdots\mathbf{B}^{-1}\mathbf{\Lambda}\mathbf{B} = \mathbf{B}^{-1}\mathbf{\Lambda}^n\mathbf{B} \qquad 4.1.8$$

Kiedy znamy macierz **B** zmiany wektorów bazy, to obliczenie *n*-tej potęgi macierzy **A** staje się proste. Zauważ, że

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix} \qquad 4.1.9$$

Potęgowanie macierzy diagonalnej sprowadza się do obliczenia potęg jej diagonalnych wyrazów. Po obliczeniu potęgi macierzy diagonalnej, korzystając z ostatniego wyrazu w (4.1.8) możemy obliczyć *n*-tą potęgę macierzy **A**.

Jak znajdować wartości własne macierzy **A**? Równanie (4.1.1) przepiszę w postaci

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \qquad 4.1.10$$

I jest macierzą jednostkową. Wyrażenie to można intepretować jako układ równań liniowych. Układ ten ma rozwiązanie gdy

$$\det|(\mathbf{A} - \lambda\mathbf{I})| = 0 \qquad 4.1.11$$

W postaci jawnej wyrażenie (4.1.11) jest dla macierzy **A** o wymiarze *n*×*n* wielomianem stopnia *n*. Tak otrzymane równanie wielomianowe nazywa się równaniem charakterystycznym (wiekowym). Rozwiązanie równanie charakterystycznego ze względu na λ, pozwala na obliczenie wartości własnych

Definicja 4.1.4: Równanie charakterystyczne (wiekowe)

Równanie charakterystyczne macierzy A zdefiniowane jest wzorem (4.1.11)

Dygresja 4.1.1.

Ograniczyłem się tu do macierzy 3x3. Ale te same rachunki można wykonać dla macierzy o dowolnym rozmiarze N×N, więc traktuję powyższe wyniki jako ogólne.

Dla przykładu obliczę wartości własne macierzy

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix} \qquad 4.1.12$$

Równanie charakterystyczne ma postać

$$\det\begin{bmatrix}1-\lambda & 0 & -1\\ 1 & 2-\lambda & 1\\ 2 & 2 & 3-\lambda\end{bmatrix}=0 \Longrightarrow (3-\lambda)(2-3\lambda+\lambda^2)=0 \qquad 4.1.13$$

Równanie to ma trzy rozwiązania

$$\lambda_1 = 1,\ \ \lambda_2 = 2,\ \ \lambda_3 = 3 \qquad 4.1.14$$

Dla λ_1 równanie (4.1.1) przyjmie postać

$$\begin{bmatrix}1 & 0 & -1\\ 1 & 2 & 1\\ 2 & 2 & 3\end{bmatrix}\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}=1\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix} \qquad 4.1.15$$

Stąd mamy układ trzech równań

$$\begin{cases} v_1 - v_3 = v_1 \Rightarrow v_3 = 0\\ v_1 + 2v_2 + v_3 = v_2 \Rightarrow v_1 + v_2 + v_3 = 0\\ 2v_1 + 2v_2 + 3v_3 = v_3 \Rightarrow v_1 + v_2 + v_3 = 0 \end{cases} \qquad 4.1.16$$

Rozwiązaniem tego układu jest wektor o współrzędnych $\mathbf{v_1}\{-a,a,0\}$, gdzie a jest niezerową liczbą rzeczywistą. Liczba a nie może być zerem, gdyż wektor $\mathbf{v}$ w równaniu (4.1.1) nie może być wektorem zerowym. Dwie pierwsze współrzędne wektora własnego nie są konkretna liczbą. Otrzymaliśmy warunek, że współrzędne te nie mogą być zerowe i muszą być liczbami o przeciwnych znakach. Dwa ostatnie równanie układu (4.1.16) nie chcą nic więcej na ten temat powiedzieć. Dla λ_2, współrzędne wektora własnego otrzymujemy w podobny sposób; wynoszą one $\mathbf{v_2}\{-2,1,2\}$, a dla λ_3 – $\mathbf{v_3}\{-1,1,2\}$.

Pozostaje nam odpowiedzieć na pytanie: Czy wyliczenie wartości własnych jest równoważne ze stwierdzeniem, że macierz można zdiagonalizować. Niestety, nie jest to takie proste. Istnieje ważna klasa macierzy, które można zawsze zdiagonalizować, są to macierze symetryczne, czyli takie, które są równe własnej transpozycji.

Definicja 4.1.5: Macierz symetryczna

Macierz kwadratowa jest macierzą symetryczną, jeżeli wyrazy położone symetrycznie względem przekątnej są równe

$$a_{ij} = a_{ji} \qquad 4.1.17$$

Łatwo pokazać, że jeżeli macierz $\mathbf{A}$ jest symetryczna, to jest równa swojej transpozycji

$$\mathbf{A} = \mathbf{A^T} \qquad 4.1.18$$

Twierdzenie 4.1.5: Diagonalizowalność macierzy symetrycznych

Macierz symetryczna jest macierzą diagonalizowalną

Gdy macierz ma współczynniki zespolone wyróżniamy nadto klasę macierzy hermitowskich

Definicja 4.1.5: Macierz hermitowska

Macierz kwadratowa jest macierzą hermitowską, gdy jej współczynniki spełniają związek

$$a_{ij} = a_{ji}^{*} \tag{4.1.19}$$

Gdzie gwiazdka oznacza wyraz sprzężony (def. DD 1.8). Symetryczne macierze rzeczywiste są również macierzami hermitowskimi. Zatem macierze hermitowskie to macierze, dla których spełniona jest zależność

$$\mathrm{A} = \left(\mathrm{A}^{*}\right)^{T} = \mathrm{A}^{\dagger} \tag{4.1.20}$$

To znaczy, że macierz jest hermitowska gdy jest równa swojemu sprzężeniu zespolonemu i transpozycji. Te dwie operacje ujmujemy jednym symbolem krzyżyka.

Twierdzenie 4.1.6: Diagonalizowalność macierzy hermitowskich

Macierz hermitowska jest macierzą diagonalizowalną

Zauważ, że ze wzoru (4.1.19) i faktu (4.1.1) wynika, że wyrazy diagonalne muszą być rzeczywiste. Oznacza to, że

Twierdzenie 4.1.7: Wartości własne macierzy hermitowskich

Wartości własne macierzy hermitowskich są rzeczywiste

Pokazać można również, że macierze symetryczne o współczynnikach zespolonych, w postaci diagonalnej mają na przekątnej wyrazy rzeczywiste.

Tyle na razie, jeżeli chodzi o teorię operatorów liniowych i macierzy. Czas na praktykę.

4.1.1. Mathematica

Dla pakietów matematycznych obliczanie wartości własnych macierzy należy do standardowych opcji. Zagadnienie obliczania wartości i wektorów własnych w wysokiej klasy pakietach, takich jak Mathematica, potraktowane jest z należytą uwagą. Oferowane procedury są szybkie i dokładne. Ta druga cecha jest ważna, ze względu na fakt, że przy wielu zagadnieniach związanych z obliczaniem wartości własnych błędy numeryczne mogą okazać się istotnie duże. Jak zwykle skorzystam tu z okazji do uwagi dydaktycznej. Kiedy schodzi się z dobrze wydeptanych ścieżek na jakich bazują szkolne lub studenckie przykłady można się natknąć na pułapki, o których mało (jeżeli w ogóle) mówi

się w szkole lub na studiach. W takiej sytuacji nie wystarczy oprogramowanie. Trzeba jeszcze dobrze znać rozwiązywany problem i ograniczenia używanego oprogramowania. W przypadku obliczania wartości własnych macierzy trzeba mieć na uwadze potencjalne błędy numeryczne, co wiąże się z jakością używanych pakietów, lub algorytmów użytych do napisania własnych procedur.

Dla przykładu obliczę wartości własne i wektory własne dla macierzy (3.1.12). Zauważ, że wartości własnej 1, Mathematica przypisała wektor o współrzędnych {-1,1,0}, zamiast {-*a*,*a*,0}, jak to uczyniłem z (3.1.16). Możemy jednak wektor własny przemnożyć przez dowolną liczbę i dalej pozostanie wektorem własnym, choć odpowiadającym innej wartości własnej. Mathematica wyciąga z rozwiązania wspólny czynnik przed nawias przez co otrzymujemy wektor {-1,1,0} i podaje wynik bez czynnika przed nawiasem. Komputer szybko liczy, ale użytkownik musi rozumieć wynik.

Eigensystem[{{1,0,-1},{1,2,1},{2,2,3}}] // MatrixForm	Eigensystem jest instrukcją do obliczania wartości własnych i wektorów własnych; //MatrixForm nakazuje napisać wynik w postaci macierzowej
$\begin{pmatrix} 3 & 2 & 1 \\ \{-1,1,2\} & \{-2,1,2\} & \{-1,1,0\} \end{pmatrix}$	Wynik – w pierwszym rzędzie wartości własne, w drugim odpowiadające im wektory własne

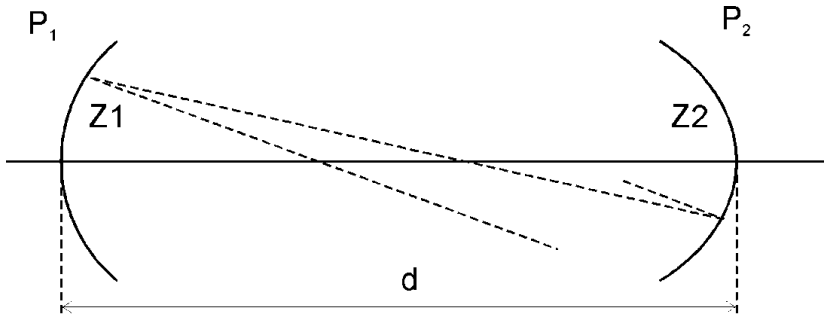
M.4.1.1. Przykładowe obliczanie wartości własnych w pakiecie Mathematica

4.2. Przykład

Powrócę do optyki geometrycznej w ujęciu macierzowym. Jak już wspominałem w (§TXI 3) nie wykorzystałem pełnej mocy metody macierzowej. Mając narzędzia omówione wyżej mogę wam tą moc pokazać. Zrobię to na ważnym przykładzie.

Lasery gazowe potrzebują rezonatorów, które potrafią uwięzić w swym wnętrzu światło. W najprostszym ujęciu rezonator laserowy to dwa zwierciadła *Z*1 i *Z*2 ustawione powierzchniami odbijającymi do siebie (rys. 4.2.1). Zwierciadło *Z*2 jest półprzepuszczalne, tak że niewielka część energii światła wypływa na zewnątrz, a pozostała część odbija się i jest uwięziona w objętości rezonatora. Jeżeli ośrodek aktywny (zwykle mieszanina gazów) wewnątrz rezonatora zostaje pobudzony do świecenia to wyemitowane światło jest magazynowane przez rezonator przez co zwiększa się gęstość promieniowania wewnątrz wnęki rezonatora, aż do momentu kiedy akcja laserowa może zachodzić stabilnie. Po osiągnięciu punktu stabilnej pracy gęstość energii wewnątrz rezonatora pozostaje stała. Wtedy energia emitowana przez ośrodek

czynny jest równa energii wyświecanej na zewnątrz przez zwierciadło *Z2* oraz absorbowanej przez elementy konstrukcyjne rezonatora. To tyle co musimy wiedzieć o rezonatorach laserowych. A teraz nasz problem. Jak zaprojektować rezonator tak, aby światło z jego wnętrza nie uciekało na zewnątrz? Inaczej mówiąc: załóżmy, że z pewnego punktu, pod małym kątem biegnie promień świetlny (te, które biegną pod dużym kątem i tak nie trafią w zwierciadła więc się nimi nie przejmujemy). Chcemy zaprojektować rezonator tak, aby tenże promień odbijając się od zwierciadeł pozostał w jego wnętrzu, a nie wyskoczył na zewnątrz (rys. 4.2.1.). Problem spróbujemy rozwiązać na bazie najprostszego wartościowego modelu (kulista krowa), czyli w przybliżeniu paraksjalnym. To przybliżenie realizuje metoda macierzy ABCD przedstawiona w (§TXI 3). Załóżmy, że mamy promień, który wychodzi od zwierciadła *Z*2 (odbił się od tego zwierciadła) i kieruje się do zwierciadła *Z*1, następnie odbija się od zwierciadła *Z*1 i kieruje się do zwierciadła *Z*2, a po odbiciu od zwierciadła *Z*2 ponownie kieruje się do zwierciadła *Z*1 i tak dalej.



Rysunek 4.2.1. Prosty rezonator laserowy składa się z dwóch zwierciadeł *Z*1 i *Z*2 o mocy optycznej odpowiednio P_1 i P_2. Wierzchołki zwierciadeł są w odległości *d*. Analizujemy biegu promienia pomiędzy zwierciadłami.

Widać, że mamy tu cykliczność kolejnych etapów propagacji promienia. Aby ująć jeden cykl musimy zdefiniować cztery macierze: macierz odbicia $\mathbf{R}_2$ (def. TXI 3.2.3) od zwierciadła *Z*2, macierz $\mathbf{T}_2$ (def. TXI 3.1.3) przejścia od zwierciadła *Z*2 do zwierciadła *Z*1, macierz $\mathbf{R}_1$ odbicia od zwierciadła *Z*1 i macierz przejścia $\mathbf{T}_1$ od zwierciadła *Z*1 do zwierciadła *Z*2. Kolejne macierze mają postać. Macierz odbicia od zwierciadła *Z2*

$$\mathbf{R}_2 = \begin{bmatrix} 1 & 0 \\ -P_2 & 1 \end{bmatrix} \qquad 4.2.1$$

P_2 jest mocą optyczną zwierciadła *Z2* (def. TXI 3.2.1). Macierz przejścia od *Z*2 do *Z*1

$$\mathbf{T}_2 = \begin{bmatrix} 1 & \dfrac{-d}{-n(n=1)} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \qquad 4.2.2$$

Minusy przy *d* i *n* przypominają, że jesteśmy po odbiciu i światło propaguje się przeciwnie do wyróżnionego kierunku propagacji (kon. TXI 3.1). Macierz odbicia od zwierciadła *Z*1 ma postać

$$\mathbf{R}_1 = \begin{bmatrix} 1 & 0 \\ -P_1 & 1 \end{bmatrix} \qquad 4.2.3$$

P_1 jest tu mocą optyczną pierwszego zwierciadła Macierz przejścia od *Z1* do *Z2*

$$\mathbf{T}_1 = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} \qquad 4.2.4$$

Pojedyncze przejście promienia pełnego cyklu możemy opisać wzorem

$$\begin{bmatrix} y_o \\ V_o \end{bmatrix} = \mathbf{T}_1\mathbf{R}_1\mathbf{T}_2\mathbf{R}_2 \begin{bmatrix} y_p \\ V_p \end{bmatrix} = \mathbf{M} \begin{bmatrix} y_p \\ V_p \end{bmatrix} \qquad 4.2.5$$

Macierz **M** jest macierzą ABCD układu odpowiadającą jednemu pełnemu cyklowi. Po przemnożeniu składowych otrzymamy postać tej macierz

$$\mathbf{M} = \begin{bmatrix} 1 - P_1T - 2P_2T + P_1P_2T^2 & T(2 - P_1T) \\ -P_1 - P_2 + P_1P_2T & 1 - P_1T \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \qquad 4.2.6$$

Macierz **M** możemy traktować jako reprezentację operatora **A**, który przekształca dwuwymiarowy wektor (y_p, V_p) w dwuwymiarowy wektor (y_o, V_o). Jeżeli chcemy zobaczyć co się stanie po *N* cyklach to musimy macierz **M** pomnożyć przez siebie *N* razy, co jest mało przyjemną operacją, szczególnie przy dużych *N*. Chyba, że odwołamy się do teorii z podrozdziału (4.1). Gdyby udało się zdiagonalizwać macierz **M**, to wtedy podnoszenie do *N* tej potęgi takiej macierzy byłoby dziecinnie proste. Szukamy zatem takiej nieosobliwej macierz **B**, oraz takiej diagonalnej macierz $\mathbf{\Lambda}$, że spełnione jest równanie.

$$\mathbf{M} = \mathbf{B}\mathbf{\Lambda}\mathbf{B}^{-1} \qquad 4.2.7$$

Wtedy macierze **M** i $\mathbf{\Lambda}$, jako macierze podobne reprezentują ten sam operator **A** i mają te same wartości własne. Współczynniki szukanej macierzy diagonalnej $\mathbf{\Lambda}$, są równe wartościom własnym tej macierzy, czyli rozwiązaniom równania (4.1.11).

$$\det(\lambda\mathbf{I} - \mathbf{M}) = \begin{vmatrix} \lambda - A & -B \\ -C & \lambda - D \end{vmatrix} = 0 \qquad 4.2.8$$

Przez **I** oznaczyłem macierz jednostkową. Obliczając powyższy wyznacznik otrzymamy równanie charakterystyczne (def. 4.1.4)

$$\lambda^2 - (A + D)\lambda + 1 = 0 \qquad 4.2.9$$

Rozwiązania tego równania mają postać

$$\lambda_{1,2} = \frac{1}{2}\left(A + D \mp \sqrt{(A + D)^2 - 4}\right) \qquad 4.2.10$$

Widać, że w wyrażeniu tym dwukrotnie powtarza się wyrażenia *A*+*D*, czyli ślad macierzy (def. 4.3), który jest niezmiennikiem, to znaczy dla wszystkich macierzy podobnych jest taki sam. Możemy zatem stwierdzić, że ślad macierzy charakteryzuje nasz operator **A** i jest przez to ważny. Z tego tytułu zasługuje na własne oznaczenie.

$$\chi = A + D \qquad 4.2.11$$

Rozwiązania na wartości własne mogę teraz zapisać w postaci

$$\lambda_{1,2} = \frac{1}{2}\left(\chi \mp \sqrt{\chi^2 - 4}\right) \qquad 4.2.12$$

Możemy teraz znaleźć postać macierzy **B**, Postępując zgodnie z (DE)

$$\mathbf{F} = \begin{bmatrix} \lambda_1 - D & \lambda_2 - D \\ C & C \end{bmatrix} \qquad 4.2.13a$$

$$\mathbf{F}^{-1} = \frac{1}{\mathrm{C}(\lambda_1 - \lambda_2)} \begin{bmatrix} \mathrm{C} & \mathrm{D} - \lambda_2 \\ -C & \lambda_1 - \mathrm{D} \end{bmatrix} \qquad 4.2.13b$$

Łatwo możecie sprawdzić, że $\mathbf{M}=\mathbf{B}\Lambda\mathbf{B}^{-1}$. Wykonanie tych mnożeń prowadzi do wyniku.

$$\mathbf{B}\Lambda\mathbf{B}^{-1} = \begin{bmatrix} A & \dfrac{-1 + AD}{C} \\ C & D \end{bmatrix} \qquad 4.2.14$$

Właściwą postać uzyskamy korzystając z warunku, że wyznacznik macierzy ABCD **M** jest równy jeden: *AD*-*BC*=1 (fakt TXI 3.4.1), skąd otrzymujemy

$$AD = 1 + \mathrm{BC} \qquad 4.2.15$$

Wstawienie tego wyrażenie do (3.2.17) daje nam wyjściową macierz **M**.

Korzystając z postaci diagonalnej obliczę *N*-tą potęgę macierzy **M**.

$$\mathbf{M}^N = \mathbf{B}\Lambda^N\mathbf{B}^{-1} = \begin{bmatrix} \dfrac{\lambda_1^{N+1} - \lambda_2^{N+1} + D(-\lambda_1^N + \lambda_2^N)}{\lambda_1 - \lambda_2} & \dfrac{(D - \lambda_1)(D - \lambda_2)(\lambda_1^N - \lambda_2^N)}{C} \\ \dfrac{C(\lambda_1^N - \lambda_2^N)}{\lambda_1 - \lambda_2} & \dfrac{D(\lambda_1^N - \lambda_2^N) + \lambda_1\lambda_2(\lambda_2^{N-1} - \lambda_1^{N-1})}{\lambda_1 - \lambda_2} \end{bmatrix} \qquad 4.2.16$$

Wstawiając macierz $\mathbf{M}^{\mathrm{N}}$ do równania (TXI 3.4) otrzymamy

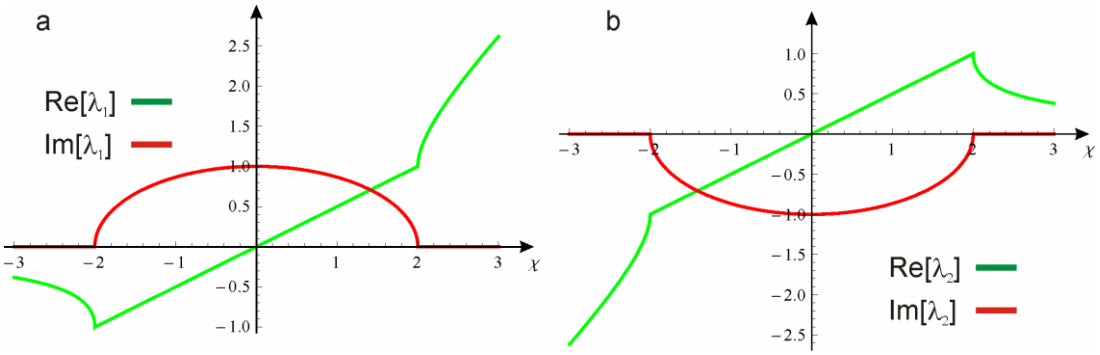
$$y_o = \frac{\lambda_1^{N+1} - \lambda_2^{N+1} + D(-\lambda_1^N + \lambda_2^N)}{\lambda_1 - \lambda_2} y_p - \frac{(D - \lambda_1)(D - \lambda_2)(\lambda_1^N - \lambda_2^N)}{C(\lambda_1 - \lambda_2)} V_p \qquad 4.2.17a$$

$$V_o = \frac{C(\lambda_1^N - \lambda_2^N)}{\lambda_1 - \lambda_2} y_p + \frac{D(\lambda_1^N - \lambda_2^N) + \lambda_1\lambda_2(\lambda_2^{N-1} - \lambda_1^{N-1})}{\lambda_1 - \lambda_2} V_p \qquad 4.2.17b$$

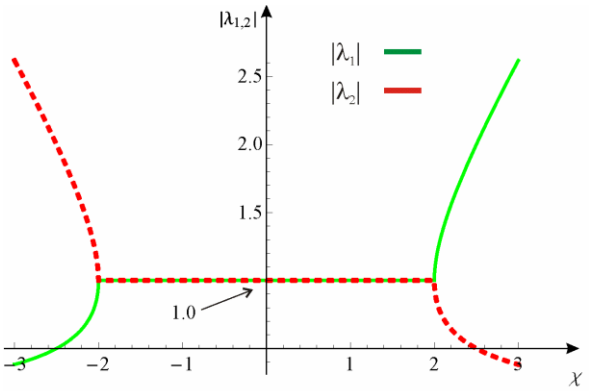
W ten sposób znaleźliśmy współrzędne promienie po N cyklach w układzie rezonatora. A teraz oblicz to samo (np. po dwudziestu cyklach) nie korzystając z macierzowy (czyli korzystając ze wzorów jakie zwykle się podaje w szkole lub na kursach fizyki), to znaczy oblicz położenie promienia po na przykład 20 cyklach wewnątrz rezonatora, a od razu docenisz potęgę metod macierzowych. Ale to bynajmniej nie wszystko co możemy wycisnąć z metody macierzowej.

Przyjrzyjmy się jak zmieniają się wartości własne (4.2.12) λ_1 i λ_2 jako funkcje śladu χ macierzy **M** (rys. 4.2.2). Z rysunku (4.2.2) widać, że na przedziale $\chi\in$[-2,2] wartości λ_1 i λ_2 są zespolone. Wewnątrz tego przedziału część urojona jest różna od zera i wartości własne są liczbami zespolonymi. Rysunek (4.2.3) przedstawia moduł z wartości własnych jako funkcję śladu χ. Zauważ, że na przedziale $\chi\in$[-2,2] moduł obu wartości własnych jest stały i równy jeden. Co jest bardzo szczęśliwą okolicznością. Wielkość zespolona, której moduł jest równy jedności, wyznacza na płaszczyźnie zespolonej punkty na okręgu o promieniu jeden (rys. DD 1.1.5). Możemy więc spróbować dla przedziału $\chi\in$[-2,2] zdefiniować wartości własne λ_1 i λ_2, tak jak to się robi dla okręgu jednostkowego. Wprowadzę następujące podstawienie

$$A + D = \chi = 2\cos(\varphi) \qquad 4.2.18$$



Rysunek 4.2.2. Część rzeczywista i urojona wartości własnych λ_1 (a) i λ_2 (b) jako funkcji śladu χ macierzy (3.2.6). Widać, że poza przedziałem $\chi\in$[-2,2] część urojona jest równa zeru;



Rysunek 4.2.3. Przebieg wartości modułu z wartości własnych λ_1 i λ_2, jako funkcji śladu macierzy (4.2.6). Jak widać na przedziale $\chi\in[-2,2]$ moduł ten jest równy jeden.

Zauważ, że tak przedstawiony ślad macierzy przebiega przedział wartości [-2,2], czyli nie wchodzi w zakres, gdzie ślad ma wartości zespolone. Wstawiając to wyrażenie do równania (4.2.12) mamy

$$\begin{aligned}\lambda_{1,2} &= \frac{1}{2}\left(2\cos(\varphi) \mp \sqrt{4\cos^2(\varphi)-4}\right)\\ &= \frac{1}{2}\left(2\cos(\varphi) \mp 2\sqrt{\cos^2(\varphi)-1}\right)\\ &= \cos(\varphi) \mp i\sqrt{1-\cos^2(\varphi)} = \cos(\varphi) \mp i\sin(\varphi)\\ &= e^{\mp i\varphi}\end{aligned} \qquad 4.2.19$$

Na przedziale $\chi\in[-2,2]$ wartość własne λ_1 i λ_2 dały się wyrazić w bardzo prosty sposób. Co więcej możemy teraz ten wynik zinterpretować. Jak widać ze wzoru (4.2.16) macierz $\mathbf{M}^N$ opisująca N-cykli przejścia promienia zależy od wartości własnych λ_1 i λ_2 w rosnących potęgach. To znaczy, że kolejny cykl zwiększa występujące w tym wyrażeniu wykładniki potęg przy λ_1 i λ_2 o jeden. Potęgowanie sprowadza się teraz do potęgowania wyrażenia typu $e^{\pm i\varphi}$. Mamy przy tym zależności

$$\left(e^{\mp i\varphi}\right)^N = e^{\mp iN\varphi} \qquad 4.2.20$$

Na płaszczyźnie zespolonej, potęgowanie przesuwa tylko λ_1 i λ_2 po jednostkowym okręgu. W efekcie macierz $\mathbf{M}^N$ dla kolejno rosnących N przekształca współrzędne promienia o jeden cykl odbić i przejść między zwierciadłami tak, że promienie te mieszczą się w ograniczonym obszarze przestrzeni. Współrzędna y-owa punktu trafienia promienia nie ucieka gdzieś do nieskończoności, ale zmienia się cyklicznie w ramach określonego skończonego przedziału wartości. A o to nam chodziło; żeby promienie po kolejnych odbiciach nie wychodziły poza ograniczony obszar rezonatora. Pozostaje pytanie, czy ograniczony obszar przemiatania promieni jest mniejszy od obszaru rezonatora? Aby odpowiedzieć na to pytanie musimy znać parametry konkretnego rezonatora. Na razie znaleźliśmy warunek, dla którego wiemy, że promienie trzymają się pewnego obszaru. Warunek ten brzmi: Ślad χ macierzy

ABCD opisujący jeden cykl odbić i przejść w rezonatorze musi się mieścić w przedziale $\chi\in[-2,2]$.

Podstawmy zatem wzór (4.2.19) do wzoru (4.2.16), po uproszeniach (zastosuj wzór Eulera (DD 1.1.3) do wyrażenia $e^{\pm iN\varphi}$) mamy

$$\mathbf{M}^N = \begin{bmatrix} \dfrac{-D\sin(\mathrm{N}\varphi)+\sin(\varphi+N\varphi)}{\sin(\varphi)} & \dfrac{(2\,D\cos(N\,\varphi)-D^2-1)\sin(N\varphi)}{\sin(\varphi)} \\ \dfrac{C\sin(N\varphi)}{\sin(\varphi)} & \dfrac{D\sin(\varphi)+\sin(\varphi-n\varphi)}{\sin(\varphi)} \end{bmatrix} \quad 4.2.21$$

Ze wzorów (4.2.18) i (4.2.15) łatwo jest wyprowadzić zależność

$$B = 2\,D\cos(N\varphi) - D^2 - 1 \quad 4.2.22$$

która po podstawieniu do wzoru (4.2.21) daje

$$\mathbf{M}^N = \begin{bmatrix} \dfrac{-D\sin(\mathrm{N}\varphi)+\sin(\varphi+N\varphi)}{\sin(\varphi)} & \dfrac{B\sin(N\varphi)}{\sin(\varphi)} \\ \dfrac{C\sin(N\varphi)}{\sin(\varphi)} & \dfrac{D\sin(\varphi)+\sin(\varphi-n\varphi)}{\sin(\varphi)} \end{bmatrix} \quad 4.2.23$$

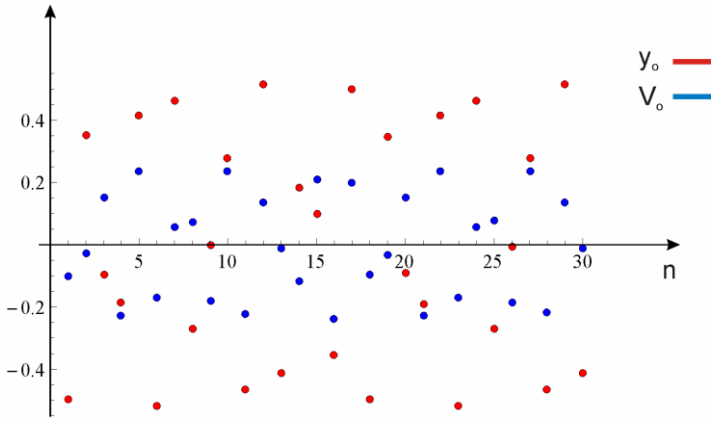
Na marginesie dodam, że mogliśmy również skorzystać z twierdzenia Sylwestra, które mówi, że jeżeli macierz taka jak macierz ABCD ma wartości własne w postaci (4.2.19), to jej *N*-ta potęga ma postać (4.2.23).

Wykorzystując postać *N*-tej potęgi macierzy **M** możemy obliczyć współrzędne punktu po *N* cyklach w rezonatorze.

$$y_o = \frac{-D\sin(\mathrm{N}\varphi)+\sin(\varphi+N\varphi)}{\sin(\varphi)} y_p + \frac{B\sin(N\varphi)}{sin(\varphi)} V_p \quad 4.2.24a$$

$$V_o = \frac{C\sin(N\varphi)}{\sin(\varphi)} y_p + \frac{D\sin(\varphi)+\sin(\varphi-n\varphi)}{\sin(\varphi)} V_p \quad 4.2.24b$$

Rysunek (4.2.4) pokazuje jak zmieniają się współrzędne y_o i V_o dla kolejnych cykli przejścia przez rezonator dla którego ślad macierzy *ABCD* opisującej jednej cykl mieści się w przedziale $\chi\in[-2,2]$. Widać, że obliczone współrzędne trzymają się pewnego skończonego przedziału wartości.



Rysunek 4.2.4. Współrzędna y_o i V_o promienia po n-tym cyklu; liczone ze wzorów (4.2.25) Wyraźnie widać, że obie współrzędne zmieniają się prawie cyklicznie wewnątrz pewnego pasa wartości. Przy takich parametrach promień nie będzie uciekał poza rezonator. Przyjęte parametry mają wartości: B=0.5; C=0; D=-0.5; χ=-1.7; y_p=0.5; V_p=0.2.

Poza tym przedziałem wartości własne są rzeczywiste. Przy pewnej dozie intuicji matematycznej szybko można dojść do podstawienia, które stawiają sprawę stabilności rezonatora w tym przypadku jasnym świetle. Nie możemy stosować teraz podstawienia (4.2.18). Możemy jednak skorzystać z usług funkcji hiperbolicznych

$$\begin{cases} \chi = 2\cosh(t) \text{ dla } \chi > 2 \\ \chi = -2\cosh(-t) \text{ dla } \chi < 2 \end{cases} \qquad 4.2.25$$

Po wstawieniu tych podstawień do wzoru (4.2.12) i skorzystaniu z (DG xx) otrzymamy

$$\lambda_1 = \begin{cases} e^t \text{ dla } \chi > 2 \\ -e^t \text{ dla } \chi < 2 \end{cases} \qquad 4.2.26a$$

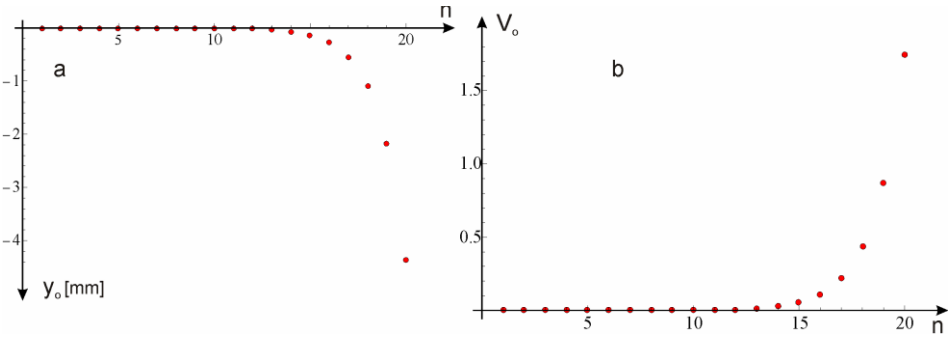
$$\lambda_2 = \begin{cases} e^{-t} \text{ dla } \chi > 2 \\ -e^{-t} \text{ dla } \chi < 2 \end{cases} \qquad 4.2.26b$$

Wstawmy teraz te wyrażenia do wzorów (4.2.17)

$$y_o = \frac{D(e^{-Nt} - e^{Nt}) + e^{(1+n)t} - e^{-(1+n)t}}{e^t - e^{-t}} y_p + \frac{(D - e^{-t})(D - e^t)(e^{Nt} - e^{-Nt})}{e^t - e^{-t}} V_p \qquad 4.2.27a$$

$$V_o = \frac{C(e^{Nt} - e^{-Nt})}{e^t - e^{-t}} y_p + \frac{D(e^{Nt} - e^{-Nt}) + e^{t-Nt} - e^{-t+Nt}}{e^t - e^{-t}} V_p \qquad 4.2.27b$$

We wzorach (4.2.27) pojawiły się rozbieżne wyrazy typu e do rosnącej potęgi rzeczywistej. Fakt ten źle wróży stabilności rezonatora, a nasze obawy potwierdza rysunek (4.2.5)



Rysunek 4.2.5. Wykres y_o (a) i V_o (b) dla kolejnych N=20 cykli. Rezonator ma parametry χ=-2,5; C=0; D=-0,5, czyli jest niestabilny. Analizowany promień „startuje" z punktu na osi y_p=0 i jest prawie równoległy do osi optycznej. V_p=-0.0000025. Z rysunku widać, że wartości y_o i V_o uciekają dla rosnących n.

4.2.1. Konkretny przykład

Przeanalizować rezonator złożony z dwóch sferycznych zwierciadeł. Moc optyczna tych zwierciadeł wynosi P_1=2D i P_2=4D. Odległość między zwierciadłami wynosi d=0.75m, a współczynnik załamania ośrodka n=1.

Zaczniemy od wypisania macierzy ABCD dla odbić i przejścia między zwierciadłami Macierz odbicia od pierwszego zwierciadła ma postać

$$R_1 = \begin{bmatrix} 1 & 0 \\ -P_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \qquad 4.2.28$$

Macierz odbicia od drugiego zwierciadła ma postać

$$R_2 = \begin{bmatrix} 1 & 0 \\ -P_2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -4 & 1 \end{bmatrix} \qquad 4.2.29$$

Macierz przejścia między zwierciadłami ma postać

$$T = \begin{bmatrix} 1 & T = \dfrac{d}{n} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0{,}75 \\ 1 & 0 \end{bmatrix} \qquad 4.2.30$$

Macierz ABCD dla pojedynczego cyklu ma postać (4.2.6)

$$\mathbf{M} = \begin{bmatrix} 1 - P_1T - 2P_2T + P_1P_2T^2 & T(2 - P_1T) \\ -P_1 - P_2 + P_1P_2T & 1 - P_1T \end{bmatrix} = \begin{bmatrix} -2 & 0{,}375 \\ 0 & -0{,}5 \end{bmatrix} \qquad 4.2.31$$

Ślad tej macierzy jest równy -2.5 i leży poza przedziałem [-2,2]. Widać, że parametry rezonatora są źle dobrane i należy się spodziewać szybkiej ucieczki promieni z jego obszaru. Spróbujmy poprawić konstrukcję rezonatora. W tym celu możemy na przykład tak zmienić wzajemne położenia zwierciadeł, aby ślad macierzy **M** był mniejszy od dwóch i większy od minus dwóch.

$$\mathrm{tr}(\mathbf{M}) = 2 - 2(P_1 + P_2)T + P_1P_2T^2 \qquad 4.2.32$$

Przyrównując tą wartość kolejno do dwóch i minus dwóch mamy

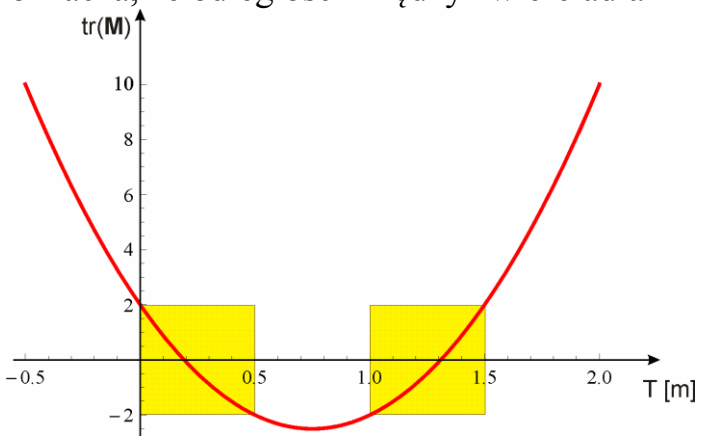
$$0 = -2(P_1 + P_2)T + P_1P_2T^2 \Longrightarrow T = 2\frac{P_1 + P_2}{P_1P_2} = \frac{3}{2} \vee T = 0 \qquad 4.2.33a$$

$$0 = 4 - 2(P_1 + P_2)T + P_1P_2T^2 \Longrightarrow T = \frac{2}{P_1} = 1 \vee T = \frac{2}{P_2} = \frac{1}{2} \qquad 4.2.33b$$

Otrzymujemy zatem dwa przedziały, w których wartość śladu macierzy **M** mieści się w granicach [-2,2], co jest pokazane na rysunku (4.2.6). Przyjmijmy dla przykładu, że T=1.25. Macierz **M** układu ma teraz postać,

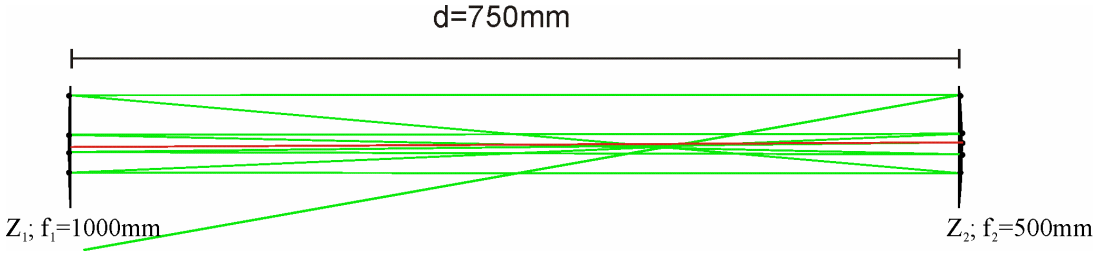
$$\mathbf{M} = \begin{bmatrix} 1 & -0.625 \\ 4 & -1.5 \end{bmatrix} \qquad 4.2.34$$

A jej ślad wynosi tr(**M**)=-0.5, co mieści się w z zadanym przedziale wartości. Oznacza to, że tak zmodyfikowany rezonator powinien pracować stabilnie. Wartość T=1.25, oznacza, że odległości między zwierciadłami wynosi d=1.25m.



Rysunek 4.2.6. Na żółto zaznaczone są obszary, dla których wartość T odpowiada stabilnej konfiguracji pracy rezonatora.

Działanie rezonatorów w obu konfiguracjach, niestabilnej i stabilnej ilustrują rysunki (4.2.6) i (4.2.7).



Rysunek 4.2.6. Przykładowy bieg promienia obliczony dla pierwszego układu rezonatora. Promień został wypuszczony przy zwierciadle Z_1 w lewą stronę. Punkt początkowy promienia znajduje się 0.1mm nad osią optyczną układu, a jego kąt nachylenia to 0,3° w stosunku do osi optycznej. Pierwszy przebieg promienia pomiędzy zwierciadłami Z_1 i Z_2 zaznaczony jest na czerwono. Średnica zwierciadeł wynosi 10cm. Po siódmym odbiciu promień opuszcza obszar rezonatora.

Rysunek 4.2.7. Przykładowy bieg promienia obliczony dla poprawionego układu rezonatora. Promień został wypuszczony przy zwierciadle Z_1 w lewą stronę. Punkt początkowy promienia znajduje się 1mm nad osią optyczną układu, a jego kąt nachylenia to 0,6° w stosunku do osi optycznej. W tym przykładzie, na początku, promień jest bardziej wychylony od osi układu niż w przykładzie z rysunku (3.2.6). Pierwszy przebieg promienia pomiędzy zwierciadłami Z_1 i Z_2 zaznaczony jest na czerwono. Średnica zwierciadeł wynosi 10cm. Mimo gorszych parametrów na starcie po 31 odbiciach promień ciągle znajdował się wewnątrz układu rezonatora. Po 31 odbiciu obliczenia zostały zatrzymane.

Czas na uwagę natury dydaktycznej. Analizę biegu promienia świetlnego moglibyśmy prowadzić, krok po kroku, z szkolnych wzorów na odbicie od zwierciadła sferycznego. Byłaby to jednak droga przez mękę. Za pomocą techniki macierzowej możemy szybko wyprowadzić wzór (4.2.6) na macierz opisującą bieg promienia w pojedynczym cyklu. Obliczanie współrzędnych po *n*-cyklach jest ciągle złożone, gdy robi się to w najprostszy sposób. Jest jednak inna droga. Zapominamy o optyce i zanurzamy się w abstrakcyjny świat macierzy reprezentujących przekształcenie liniowe. Tam znajdujemy liczne bardzo użyteczne twierdzenia pozwalająca znaleźć zbawiennie prosty sposób na podnoszenie zdiagonalizowanej macierzy do *n*-tej potęgi. Nasza macierz ABCD jest na szczęście diagonalizowalna, więc szybko dochodzimy do wzorów *n*-tą potęgę. Wyobraź sobie sytuację, gdy promotor twojej pracy dyplomowej każe ci policzyć położenie promienie po jedenastu cyklach. Tak się składa, że nie zna siły metody macierzowej i myśli sobie, że to ciężka robota. Ty masz jednak w ręku wzór na *n*-tą potęgę i w pół godziny kończysz zadanie (obliczasz oczywiście na komputerze). Potem twój promotor dochodzi do wniosku, że trzeba przyjąć nieco inne dane początkowe i dobrze byłoby przeliczyć

dwadzieścia jeden cykli – katorga; ale dla wtajemniczonych w magię macierzy to kilka minut roboty (wzory na komputerze masz już wpisane).

Wzór na *n*-tą potęgę macierzy to dopiero początek naszych sukcesów. Mając nieco szerszą wiedzę matematyczną znajdujemy przekształcenia (4.2.18 i 4.2.25). Przekształcenie (4.2.18) w sposób jawny pokazuje nam, że tam gdzie spełnione są warunki dla tego przekształcenia promień jest uwięziony w ograniczonym obszarze. A warunek jest prosty – ślad macierzy **M** musi należeć do przedziału [-2,2]. Bajka - nie tylko że łatwo znajdujemy położenia promienia po *n*-cyklach, ale potrafimy, dodając do siebie dwie współrzędne macierzy ABCD (χ=A+D), powiedzieć, czy rezonator będzie stabilny czy nie. W ręce mamy efektywne narzędzie poprawy konstrukcji rezonatora. Trzeba tak zmienić jego konstrukcję, aby ślad macierzy mieścił się w zadanym przedziale.

Zwróć uwagę na jedną rzecz, większość tych wniosków to czysta matematyka, której potem przypinamy odpowiednią interpretację fizyczną. Kto zna matematykę może mieć ogromną przewagę na tym, który jej nie zna.

Nasze rozwiązanie zostało znalezione w przybliżeniu paraksjalnym (TXI 3.1.3), czyli na poziomie modelu kulistej krowy. Zwykle jest to model niewystarczający i trzeba potem przejść do bardziej dokładnego i złożonego modelu. Ale praktyka projektanta jest zawsze taka – zaczynamy od modelu kulistej krowy, wyciskamy z niego co się da, a jak nie da się wystarczająco dużo to przechodzimy do bardziej złożonego modelu. W naszym przypadku mógłby to być model uwzględniający trzeciorzędowe aberracje układów optycznych. Zwykle jest tak, że z bardziej złożonymi modelami pracuje się lżej, jeżeli wiemy, w którym obszarze szukać rozwiązania. Tak więc prosty model pozwala nam znaleźć obszar, w którym znajduje się dobre rozwiązanie, a dokładniejszy model pozwala na przeczesanie tego obszaru. Pomimo, że model paraksjalny jest prosty, to ciągle jest bardzo użyteczny i wykorzystywany jako pierwszy krok do rozwiązywania jak najbardziej poważnych problemów.

Ten sukces macierzy na w obszarze optyki geometrycznej jasno wskazuje, że im prędzej przejdziemy do macierzowego opisu polaryzacji tym będzie dla nas lepiej.

5. Operatory i wektory Jonesa ♣

W teorii polaryzacji mamy dwa ogólnie znane podejścia macierzowe, metoda macierzy Jonesa i Müllera. Nie jest to bynajmniej nadmiar bogactwa, oba podejścia mają swoje zalety. By nam jednak od macierzy nie zakręciło się w głowie w tym temacie skupię się na bardziej popularnej metodzie Jonesa. O macierzach Müllera wspomnę na końcu.

5.1. Wektory Jonesa

Wzory (1.4) opisywały polaryzację fali harmonicznej, dla wygody tutaj je przepiszę

$$E_x = E_{0x} e^{i\delta_x} e^{i(\omega t - kz)} \qquad 5.1.1a$$

$$E_y = E_{0y} e^{i\delta_y} e^{i(\omega t - kz)} \qquad 5.1.1b$$

Oba wyrażenia na składowe wektora pola elektrycznego możemy zapisać w postaci kolumny

$$\mathbf{E} = \begin{bmatrix} E_{0x} e^{i\delta_x} \\ E_{0y} e^{i\delta_y} \end{bmatrix} e^{i(\omega t - kz)} \qquad 5.1.2$$

W dalszej części pominę człon propagacyjny (wykładniczy). Zatrzymamy się nad następującym zagadnieniem. Co się dzieje, gdy na drodze światła spolaryzowanego ustawimy element zmieniający stan polaryzacji światła? Po przejściu przez ten element otrzymamy nowy wektor pola elektrycznego

$$\mathbf{E}' = \begin{bmatrix} E'_{0x} e^{i\delta'_x} \\ E'_{0y} e^{i\delta'_y} \end{bmatrix} \qquad 5.1.3$$

Formalizm, który zastosujemy powie nam jak zmienia stan polaryzacji fali harmonicznej po przejściu przez ten element. W takim przypadku nie działamy na czynnik propagacyjny, stąd jego pominięcie.

Wiemy, że wektory transformują się poprzez macierze (§4). Możemy więc oczekiwać, że spełniona będzie zależność

$$\begin{bmatrix} E'_{0x} e^{i\delta'_x} \\ E'_{0y} e^{i\delta'_y} \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} E_{0x} e^{i\delta_x} \\ E_{0y} e^{i\delta_y} \end{bmatrix} \qquad 5.1.4$$

Macierz **A** jest matematyczną reprezentacją działanie elementu zmieniającego stan polaryzacji, czyli operatora $\hat{\mathrm{A}}$, który reprezentujące działanie fizycznego element zmieniającego stan polaryzacji światła. Przesuwając obie składowe fazy δ_x i δ_y o $-\delta_x$ (1.2c), podobnie czyniąc dla faz primowanych możemy zapisać

$$\begin{bmatrix} E_{0x} \\ E_{0y}e^{i\Delta'} \end{bmatrix} = \mathbf{A}\begin{bmatrix} E_{0x} \\ E_{0y}e^{i\Delta} \end{bmatrix} \quad 5.1.5$$

$$\Delta = \delta_y - \delta_x \quad 5.1.5a$$

$$\Delta' = \delta'_y - \delta'_x \quad 5.1.5b$$

Przejście światła przez element polaryzacyjny opisałem, od strony formalnej, zupełnie tak samo jak opisywaliśmy, z użyciem macierzy ABCD, przejście światła, przez element optyczny (§TXI 3). Jednak macierzy **A** nie nazywamy macierzą ABCD, tylko macierzą Jonesa, która reprezentuje w danym układzie współrzędnych operator Jonesa.

Definicja 5.1.1: Wektor Jonesa

Wektor opisujący stan polaryzacji światła nazywamy wektorem Jonesa

Definicja 5.1.2: Kolumna Jonesa

Kolumna w postaci (4.1.2) lub (4.1.5) reprezentująca wektor Jonesa, w danym układzie współrzędnych nazywamy kolumną Jonesa

Definicja 5.1.3: Operator Jonesa

Operator transformujący wektory Jonesa nazywamy operatorem Jonesa. Operator Jonesa opisuje działanie liniowych elementów polaryzacyjnych

Definicja 5.1.4: Macierz Jonesa

Macierz Jonesa jest macierzą reprezentującą, w danym układzie współrzędnych, operator Jonesa

Dlaczego elementy liniowe? Macierze reprezentują odwzorowania liniowe, wobec tego można z ich pomocą opisywać te elementy optyczne, których działanie ma charakter liniowy (wyraża się przez zależności liniowe). Na szczęście w optyce polaryzacyjnej zakres działania opisu liniowego jest bardzo duży.

Chociaż macierze Jonesa to macierze 2x2, to nie nazywamy ich macierzami ABCD. Nazwa macierz ABCD nie wyróżnia całej klasy macierzy o wymiarze 2x2, tylko wskazuje na ich interpretację fizyczną - są to macierze opisujące, w przybliżeniu liniowym, przejście promienia przez układ optycznych. Nazwa „macierz Jonesa" oznacza, że macierz używana jest do opisu zmiany stanu polaryzacji światła przez elementy liniowe. Istnieje również różnica formalna. Macierze Jonesa i stowarzyszone z nimi wektory mają współczynniki zespolone. Będę się musiał nad tym faktem pochylić. Przedtem podam kilka przykładów kolumny Jonesa reprezentującej wybrane wektory Jonesa. Wszystkie przykłady podane są w postaci (5.1.5).

Powiedzmy, że faza początkowa jest równa zeru, a światło jest spolaryzowane liniowa zgodnie z osią *x* (osią *y*). Kolumna Jonesa opisująca stan polaryzacji ma postać

$$\mathbf{E} = \begin{bmatrix} E \\ 0 \end{bmatrix} \text{ polaryzacja liniowa wzdłuż osi } x \qquad 5.1.6a$$

$$\mathbf{E} = \begin{bmatrix} 0 \\ E \end{bmatrix} \text{ polaryzacja liniowa wzdłuż osi } y \qquad 5.1.6b$$

Czasem przyjmujemy, że natężenie światła E^2 jest równe jeden (i tak mierzymy zwykle względne różnice w natężeniu między różnymi punktami). Wtedy wzory (4.1.6) przyjmują szczególnie prostą formę

$$\mathbf{E} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ polaryzacja liniowa wzdłuż osi } x \qquad 5.1.7a$$

$$\mathbf{E} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ polaryzacja liniowa wzdłuż osi } y \qquad 5.1.7b$$

Światło o amplitudzie E=1 i polaryzacji liniowej pod kątem $\pm\pi/4$ do osi *x* będzie miało wektor o współrzędnych

$$\mathbf{E} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ \pm 1 \end{bmatrix} \qquad 5.1.8$$

Światło o amplitudzie E=1 spolaryzowane kołowo odpowiednio prawo i lewo skrętnie jest reprezentowane przez kolumnę Jonesa

$$\mathbf{E} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad 5.1.9a$$

$$\mathbf{E} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix} \qquad 5.1.9b$$

Sprawdzenie poprawności wzorów (5.1.9) zaczniemy od obliczenia amplitudy

$$|E|^2 = |\mathbf{E}\mathbf{E}^*| = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = 1 \qquad 5.1.10$$

Amplituda jest rzeczywiście jednostkowa. Aby być zgodnym z formą (5.1.5) stany (5.1.9a) i (5.1.9b) zapiszę w postaci

$$\begin{bmatrix} E_{0x} \\ E_{0y}e^{i\delta} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ e^{i\frac{\pi}{2}} \end{bmatrix}; \quad \text{bo } e^{i\frac{\pi}{2}} = i \qquad 5.1.11a$$

$$\begin{bmatrix} E_{0x} \\ E_{0y}e^{i\delta} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ e^{-i\frac{\pi}{2}} \end{bmatrix}; \quad \text{bo } e^{-i\frac{\pi}{2}} = -i \qquad 5.1.11b$$

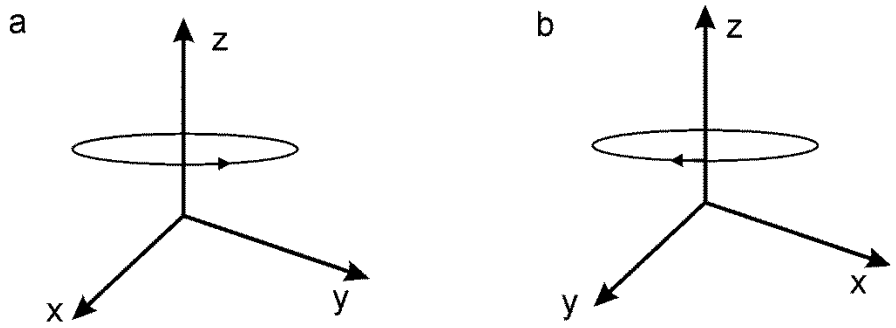
Zatem wzory (4.1.1) dla tego przypadku mają postać

$$\begin{cases} E_x = \dfrac{1}{\sqrt{2}} e^{i(\omega t - kz)} \\ E_y = \dfrac{1}{\sqrt{2}} e^{i\frac{\pi}{2}} e^{i(\omega t - kz)} \end{cases} \qquad 5.1.12a$$

$$\begin{cases} E_x = \dfrac{1}{\sqrt{2}} e^{i(\omega t - kz)} \\ E_y = \dfrac{1}{\sqrt{2}} e^{-i\frac{\pi}{2}} e^{i(\omega t - kz)} \end{cases} \qquad 5.1.12b$$

W przypadku (5.1.12a) mamy złożenie dwóch ortogonalnych drgań harmonicznych o tej samej amplitudzie, przy czym drganie wzdłuż osi *y* wyprzedza drganie wzdłuż osi *x* o $\pi/2$. Zgodnie z tym co było powiedziane o składaniu drgań harmonicznych (tab. TVIII 2.2.1), wypadkowy ruch jest ruchem po okręgu zgodnie z kierunkiem ruchu wskazówek zegara (ruch prawoskrętny). W przypadku (5.1.21b) otrzymujemy ruch po okręgu przeciwnie do wskazówek zegara (ruch lewoskrętny).

Przy opisie stanu polaryzacji światła trzeba uważać na orientację układu współrzędnych (rys. TV 3.2.10). Zmiana orientacji powoduje, że zmienia się skrętność polaryzacji dla światła opisanego wzorami (5.1.9), co ilustruje rysunek (5.1)



Rysunek 5.1. Powiedzmy, że kolumna Jonesa opisuje w prawoskrętnym układzie współrzędnych światło spolaryzowane prawoskrętnie (a). Punkt który mija dodatnią półoś *x*-ów kieruje się do dodatniej osi *y*-ów. Gdy zmienimy skrętność układu (b) nie zmieniając ułożenia współrzędnych w kolumnie Jonesa, to punkt, który mija dodatnią półoś *x*-ów dalej będzie się kierował do dodatniej półosi *y*-ów, co w efekcie da światło spolaryzowane lewoskrętnie.

Uwaga 5.1:

W literaturze nie posługujemy się nazwą „kolumna Jonesa", tylko „wektorem Jonesa". Mnie jednak zależy na tym, byście sobie utrwalili różnicę między wektorem a jego reprezentacją we współrzędnych. Wektor nie jest zależny od wyboru współrzędnych, ale

jego reprezentacja zależna jest. Podobnie operator nie jest zależny od współrzędnych ale jego reprezentacja jest. Dlatego tam, gdzie jawnie posługuję się współrzędnymi będę mówił kolumna lub macierz, a nie wektor lub operator. Co ciekawe w tej samej literaturze zamiast konsekwentnie mówić „operator Jonesa" mówi się „macierz Jonesa".

5.1.1. Wektory nad ciałem liczb zespolonych

Kolumny Jonesa można traktować jako współrzędne wektorów Jonesa należących do przestrzeni wektorowej zdefiniowanej nad ciałem liczb zespolonych (§TIV 3.1). W efekcie współrzędne tych wektorów wyrażą się przez liczby zespolone. Przestrzeń wektorowa nad ciałem liczb zespolonych jest w pełni uprawnioną przestrzenią wektorową. Wszystkie ogólne własności przestrzeni wektorowych są w niej takie same jak w każdej innej. Bardziej specyficzne własności mogą się jednak różnić. Jeżeli chcemy wprowadzić iloczyn skalarny do przestrzeni wektorowej nad ciałem liczb zespolonych, to definicja znana z przestrzeni wektorowej nad ciałem liczb rzeczywistych (def. TV 3.4.9) jest kłopotliwa.

$$\langle \mathbf{v}|\mathbf{w}\rangle = \sum_{i=1}^{N} v_i w_i \qquad 5.1.13$$

Ponieważ współrzędne v_i i w_i są liczbami zespolonymi, to tak zdefiniowany iloczyn odwzorowuje dwa wektory w liczbę zespoloną. Chcielibyśmy aby wyrażenie na iloczyn wektora przez siebie było kwadratem długości tego wektora. Definicja iloczynu w postaci (5.1.13) tego nie zapewnia

$$\langle \mathbf{v}|\mathbf{v}\rangle = \sum_{i=1}^{N} v_i v_i \qquad 5.1.14$$

Sytuacja się zmieni, gdy iloczyn skalarny dwóch zespolonych wektorów zdefiniujemy tak

$$\langle \mathbf{v}|\mathbf{w}\rangle = \sum_{i=1}^{N} v_i w_i^* \qquad 5.1.15$$

Przypominam, że gwiazdka oznacza liczbę zespoloną sprzężoną (def. DD 1.8). Przy takiej definicji iloczyn wektora przez siebie ma wartość rzeczywistą i daje kwadrat długość wektora rozumianą jako suma kwadratów jego części rzeczywistej v_r i urojonej v_i.

$$\langle \mathbf{v}|\mathbf{v}\rangle = \sum_{j=1}^{N} \left(v_{r;j}^2 + v_{i;j}^2 \right) \qquad 5.1.16$$

$$\mathbf{v}(v_{r;1} + iv_{i;1}, \ldots, v_{r;N} + iv_{i;N}) \qquad 5.1.16a$$

Tak zdefiniowany iloczyn skalarny pewne własności ma inne niż w przypadku standardowego iloczynu zdefiniowanego dla przestrzeni wektorowej nad ciałem liczb rzeczywistych, w szczególności nie jest on przemienny

$$\langle \mathbf{v}|\mathbf{w}\rangle = \overline{\langle \mathbf{w}|\mathbf{v}\rangle} \qquad 5.1.17$$

Tutaj sprzężenie zespolone oznaczyłem przez kreskę nad iloczynem skalarnym, co jest, często spotykane, szczególnie w notacji braketów (DD 1.15a). Dla dowolnej liczby zespolonej α mamy ponadto

$$\langle \alpha \mathbf{v} | \mathbf{w} \rangle = \alpha \langle \mathbf{v} | \mathbf{w} \rangle \qquad 5.1.18$$

$$\langle \mathbf{v} | \alpha \mathbf{w} \rangle = \overline{\alpha} \langle \mathbf{v} | \mathbf{w} \rangle \qquad 5.1.18b$$

Widać z tego, że w drugim wyrazie iloczyn skalarny nie jest liniowy, wyciągamy przed iloczyn sprzężenie skalara a nie sam skalar. Możemy zresztą to uznać ze przejaw zmodyfikowanej liniowości w przestrzeniach zespolonych. Dalej jednak dwa wektory są do siebie ortogonalne, gdy ich iloczyn skalarny jest równy zeru

$$\mathbf{v} \perp \mathbf{w} \Leftrightarrow \langle \mathbf{v} | \mathbf{w} \rangle = 0 \qquad 5.1.19$$

Gdy wektor ma współrzędne rzeczywiste, to nowy iloczyn skalarny zachowuje się tak jak ten stary, zdefiniowany dla wektorów rzeczywistych. Możemy więc przyjąć, że ogólna definicja iloczynu skalarnego ma postać (5.1.15). Gdy wektory są rzeczywiste definicja ta jest równoważna definicji (TV 3.4.9).

Definicja 5.1.3: iloczyn skalarny wektorów zespolonych

Iloczyn skalarny dla wektorów zespolonych w układzie bazy ortonormalnej zdefiniowany jest wzorem (5.1.15)

Wykorzystując nowy iloczyn skalarny możemy stwierdzić, że

Fakt 4.1.1:

Dwa wektory Jonesa $\mathbf{E_1}$ i $\mathbf{E_2}$ reprezentują ortogonalne stany polaryzacji, gdy ich iloczyn skalarny jest równy zeru

$$\langle \mathbf{E_1} | \mathbf{E_2} \rangle = 0 \qquad 5.1.20$$

Do każdego stanu polaryzacji światła dobierzemy stan do niego ortogonalny. Poniżej zapisałem w postaci ogólnej dwa wektory Jonesa opisujące wzajemnie ortogonalne stany polaryzacji.

$$\mathbf{E} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} \qquad 5.1.21$$

$$\mathbf{E_{ort}} = \begin{bmatrix} E_x^* \\ -E_y^* \end{bmatrix} \qquad 5.1.21b$$

Łatwo możesz sprawdzić, że dwa wektory Jonesa reprezentujące polaryzację V i H (def. 1.1.5 i 1.1.6) są ortogonalne. Podobnie wektory Jonesa reprezentujące polaryzację L i R (def. 1.1.7 i 1.1.8). Sprawdzę to korzystając z (5.1.9) i (5.1.15) mamy

$$\langle \mathbf{L} | \mathbf{R} \rangle = \frac{1}{2}\left(1 + i^2\right) = \frac{1}{2}(1-1) = 0 \qquad 5.1.22$$

5.1.2. Notacja braketowa dla wektorów Jonesa

Wektory Jonesa będziemy również zapisywać w notacji braketowej (TIX 4.3.23). Stosując wprowadzone oznaczenia stanów polaryzacji możemy zapisać wektory Jonesa dla polaryzacji liniowej poziomej i pionowej jako odpowiednio

$$|H\rangle \text{ lub } |V\rangle \qquad 5.1.23$$

Dla polaryzacji kołowej prawo i lewoskrętnej, odpowiednio

$$|R\rangle \text{ lub } |L\rangle \qquad 5.1.24$$

Ponieważ mamy tylko dwa ortogonalne stany polaryzacji więc odpowiadająca im przestrzeń wektorowa jest dwuwymiarowa.

Fakt 5.1.2:

Przestrzeń wektorów Jonesa jest przestrzenią dwuwymiarową

Ortogonalne stany polaryzacji mogą stanowić bazę takiej przestrzeni (def. TIV 3.1.4), wobec tego układ wektorów Jonesa (5.1.23) i (5.1.24) możemy traktować jako dwie możliwe bazy dla przestrzeni wektorów Jonesa. Oznacza to, że wszystkie inne wektory Jonesa możemy wyrazić jako kombinację liniową wektorów należących do jednej z tych dwóch par. Na przykład wektory dla stanów kołowych można wyrazić przez wektory dla stanów liniowych

$$|R\rangle = \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle) = \frac{1}{\sqrt{2}}\left(\begin{bmatrix}1\\0\end{bmatrix} - i\begin{bmatrix}0\\1\end{bmatrix}\right) = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} \qquad 5.\,1.25$$

$$|L\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle) = \frac{1}{\sqrt{2}}\left(\begin{bmatrix}1\\0\end{bmatrix} + i\begin{bmatrix}0\\1\end{bmatrix}\right) = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix} \qquad 5.1.25b$$

Różnych układów wektorów bazowych jest nieskończenie wiele. Nieskończenie wiele jest również baz ortogonalnych. Mając wzory (5.1.21) każdemu stanowi polaryzacji możemy dobrać stan ortogonalny. Czyli każdy stan polaryzacji jest odpowiedni do konstrukcji bazy ortogonalnej. Należy również pamiętać, że iloczyny skalarne (5.1.20) musimy obliczać dla wektorów Jonesa przedstawionych w tej samej bazie. Zresztą wszystkie operacje na współrzędnych dwóch wektorów należy przeprowadzić w tej samej bazie, inaczej zmieniając bazę dla jednego z wektorów możemy praktycznie dowolnie zmienić wynik sumy czy iloczynu wektorów.

5.2. Macierze Jonesa

Wektory Jonesa opisują stan polaryzacji światła. Aby opisać transformację tego stanu przez liniowe elementy polaryzacyjne potrzebujemy macierzy 2x2 nazywanej macierzą Jonesa (def. 5.2.4). Macierz ta reprezentuje operator, który nazywamy operatorem Jonesa (def. 5.2.3). Macierz Jonesa działa w następujący sposób

$$\begin{pmatrix} E_x^{wyj} \\ E_y^{wyj} \end{pmatrix} = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix} \begin{pmatrix} E_x^{wej} \\ E_y^{wej} \end{pmatrix} \Rightarrow \begin{cases} E_x^{wyj} = J_{xx}E_x^{wej} + J_{xy}E_y^{wej} \\ E_y^{wyj} = J_{yx}E_x^{wej} + J_{yy}E_y^{wej} \end{cases} \qquad 5.2.1$$

Elementy diagonalne opisują przeskalowanie wejściowego stanu polaryzacji. Elementy pozadiagonalne wpływ składników ortogonalnych do danego kierunku. Od strony matematycznej macierz Jonesa to po prostu macierz. Dla fizyka jednak dodatkowe określenie przez słowo „Jonesa" informuje jak należy intepretować fizycznie tą macierz.

Macierz Jonesa musi opisywać działanie liniowych elementów polaryzacyjnych. Jak znaleźć postać macierzy Jonesa dla konkretnego elementu polaryzacyjnego? W podręcznikach do optyki opisane są metody eksperymentalnego wyznaczania macierzy Jonesa dla nieznanego elementu liniowego przekształcającego stan polaryzacji światła. Wymaga to zwykle trzech lub więcej pomiarów z użyciem znanych elementów polaryzacyjnych (polaryzatorów, ćwierć lub półfalówek). Ja nakreślę tu drogę czysto teoretyczną dla przypadku płytek falowych (4.2). Powiedzmy, że płaskorównoległa płytka falowa o grubości d wprowadza do płaskiej fali świetlnej padającej na nią normalnie, różnicę faz między modem (fakt. 4.2.1) szybkim (współczynnik załamania n_s) a modem wolnym (współczynnik załamania n_w) równą Γ

$$\Gamma = k\left(n_s - n_w\right)d \qquad 5.2.2$$

Wprowadzę wielkość φ_s będącą średnią przesunięcia fazowego modu szybkiego i wolnego.

$$\varphi_s = \frac{1}{2}\left(n_s + n_w\right)\frac{2\pi}{\lambda}d \qquad 5.2.3$$

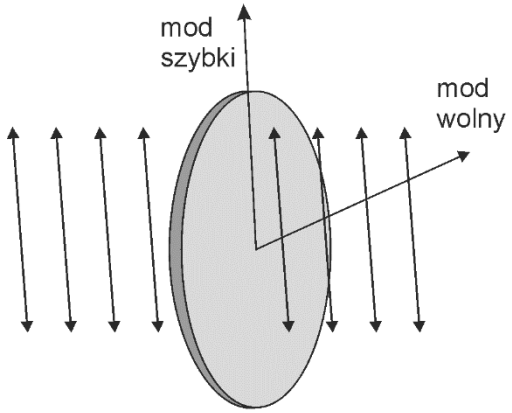
Ponieważ interesuje nas względna różnica faz wprowadzona przez płytkę między modem szybkim a modem wolnym rozłożę tą różnicę równomiernie na oba mody płytki dwójłomnej; wtedy macierz Jonesa przyjmie postać

$$\mathbf{J} = e^{i\varphi_s}\begin{pmatrix} e^{-i\frac{\Gamma}{2}} & 0 \\ 0 & e^{i\frac{\Gamma}{2}} \end{pmatrix} \qquad 5.2.4$$

Zauważ jak działa ta macierz. Pierwszą współrzędną wektora Joensa przemnoży przez czynnik opóźniający fazę o $\Gamma/2$, a drugą przez czynnik przyspieszający fazę o $\Gamma/2$. W efekcie względna różnica faz między modami wynosić będzie Γ, przy czym każdy mod zostanie przesunięty o właściwą wartość fazy. W dalszych rozważaniach czynnik zawierający φ_s nie będzie miał znaczenia więc go pominę. Nie zawsze jest to możliwe. Gdy trzeba uwzględnić interferencję różnych wiązek, np. spowodowanych kolejnymi odbiciami od powierzchni płytki fazowej, to czynnika tego nie wolno zaniedbywać.

Jeżeli puścimy na taki element światło spolaryzowane liniowo zgodnie z osią x lub y, to nic się nie zmieni (rys. 5.2.1); na przykład

$$\begin{pmatrix} e^{-i\frac{\Gamma}{2}} & 0 \\ 0 & e^{i\frac{\Gamma}{2}} \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{-i\frac{\Gamma}{2}} \\ 0 \end{pmatrix} = e^{-i\frac{\Gamma}{2}}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad 5.2.5$$



Rysunek 5.2.1. Gdy na płytkę falową pada światło spolaryzowane liniowo zgodnie z kierunkiem polaryzacji modu szybkiego lub wolnego, to światło to przechodzi bez zmiany stanu polaryzacji. Oznacza to, że w formalizmie Jonesa wektor opisujący ten stan polaryzacji światła jest wektorem własnym macierzy Jonesa opisujące płytkę.

Widać, że liniowa polaryzacja zgodna z osią *x* lub osią *y* jest dla macierzy (5.2.4) wektorem własnym, a połowa względnego przesunięcie fazowe modu wolnego i szybkiego jest wartością własną. Powiedzmy, że na wejściu mamy stan polaryzacji opisany wektorem

$$\mathbf{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix} \qquad 5.2.6$$

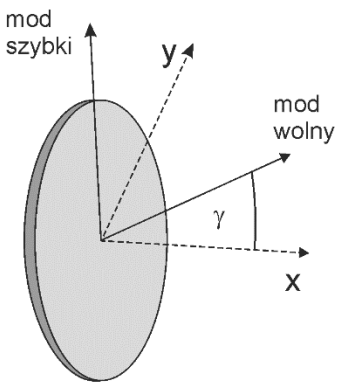
Układ współrzędnych, w których zdefiniowany jest wektor Jonesa nazwę układem laboratoryjnym (rys. 5.2.2). Jeżeli chcemy do opisu stanu światła po przejściu przez płytkę falową użyć macierzy Jonesa postaci (5.2.4) to musimy opisać stan polaryzacji na wejściu w układzie współrzędnych zgodnym z układem, w którym zdefiniowana jest macierz (5.2.4) (nazwę go układem współrzędnych macierzy Jonesa). Pamiętaj, że operator jest pojęciem ogólniejszym od macierzy. Macierz to konkretna reprezentacja operatora w wybranym układzie współrzędnych. Zmiana układu współrzędnych zmienia macierz ale nie operator. Fizyka to operator, który nie może zależeć od wyboru układu współrzędnych. Macierz to matematyczna reprezentacja operatora w danym układzie współrzędnych, czyli wstęp do konkretnych rachunków. Układ, w którym operator Jonesa ma reprezentację macierzową postaci (5.2.4) nazwę układem własnym operatora Jonesa.

Definicja 5.2.1: Układ własny operatora Jonesa

Układ własny operatora Jonesa to taki układ współrzędnych, którego osie są zgodne z kierunkami wektorów własnych tego operatora.

Procedura wygląda zatem tak: przeliczam wejściowy stan polaryzacji do układu własnego danego operatora Jonesa (wtedy macierz tego operatora jest postaci (5.2.4)). Mnożę tak otrzymaną kolumnę Jonesa przez macierz Jonesa, w wyniku

czego otrzymam kolumnę Jonesa w układzie własnym operatora Jonesa. Przeliczam współrzędne tak otrzymanego wektora Jonesa (lub nie, jeżeli nowy układ współrzędnych mi pasuje) z powrotem do starego układu współrzędnych.



Rysunek 5.2.2. Wygodnie jest przejść ze współrzędnymi wektora Jonesa dla danego wejściowego stanu polaryzacji od układu laboratoryjnego do układ własnego macierzy Jonesa reprezentującej dany element polaryzacyjny. Należy również przeliczyć do tego układu współrzędne wektora Jonesa reprezentującego wejściowy stan polaryzacji. Po przeliczeniu działania macierzy, można wrócić do współrzędnych laboratoryjnych.

Współrzędne wektora Jonesa (5.2.6) w obróconym o kąt γ (rys. 5.2.2) układzie współrzędnych (przejście do układu własnego operatora Jonesa) mają postać

$$\mathbf{v}_\gamma = \mathbf{R}(\gamma)\mathbf{v} \qquad 5.2.7$$

$$\mathbf{R}(\gamma) = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ -\sin(\gamma) & \cos(\gamma) \end{pmatrix} \qquad 5.2.7a$$

W nowym układzie współrzędnych na współrzędne wektora Jonesa $\mathbf{v}_\gamma$ możemy podziałać macierzą postaci (4.3.4). Przypominam, że $\mathbf{v}$ i $\mathbf{v}_\gamma$, to te same wektory, tyle że reprezentowane w dwóch obróconych o kąt γ względem siebie układach współrzędnych. Działamy teraz na wektor $\mathbf{v}_\gamma$ macierzą Jonesa (5.2.4)

$$\mathbf{v}'_\gamma = \mathbf{J}\,\mathbf{v}_\gamma = \mathbf{J}\,\mathbf{R}(\gamma)\mathbf{v} \qquad 5.2.8$$

W efekcie otrzymujemy wektor Jonesa reprezentujący stan światła po przejściu przez płytkę falową, ale reprezentowany w układzie własnym operatora Jonesa. Jeżeli chcemy wrócić do układu laboratoryjnego to musimy pomnożyć wynik przez macierz obrotu odwrotnego

$$\mathbf{v}' = \mathbf{R}^{-1}\mathbf{J}\,\mathbf{R}\mathbf{v} \qquad 5.2.9$$

$$\mathbf{R}^{-1} = \mathbf{R}(-\gamma) \qquad 5.2.9a$$

Zdefiniujmy macierz

$$\mathbf{J}_\gamma = \mathbf{R}^{-1}\mathbf{J}\mathbf{R} \qquad 5.2.10$$

Z jej użyciem wzór (5.2.9) przyjmie postać

$$\mathbf{v}' = \mathbf{J}_\gamma\,\mathbf{v} \qquad 5.2.11$$

Wzór (5.2.10) jest wzorem na macierze podobne, które jak wiemy mają te same wektory własne i wartości własne. Ale współrzędne tych wektorów i macierzy są wyliczone w różnych układach współrzędnych. W praktyce procedura obracania współrzędnych wektorów Jonesa do układu zgodnego z układem macierzy Jonesa, a po przemnożeniu macierzy przez taki wektor, dokonania

obrotu odwrotnego jest bardzo wygodna i jest często w obliczeniach stosowana. Zobaczmy na przykładach jak działa formalizm Jonesa.

Na początek zajmę się operatorem Jonesa, który opisuje nic nie robienie, czyli brak zmiany stanu polaryzacji światła. W układzie własnym tego operatora reprezentuje go oczywiście macierz jednostkowa

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \qquad 5.2.12$$

Podziałam tą macierzą na stan polaryzacji liniowej o współrzędnych (a,b). Zakładam przy tym, że polaryzacja jest już wyznaczona w układzie współrzędnych macierzy Jonesa, więc niczego nie muszę obracać.

$$\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=\begin{bmatrix}1\cdot a+0\cdot b\\ 0\cdot a+1\cdot b\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}a\\ 0\end{bmatrix}+\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}0\\ b\end{bmatrix}=\begin{bmatrix}a\\ b\end{bmatrix} \qquad 5.2.12a$$

v1 v2

Łatwo jest zgadnąć, w układzie własnym operatora Jonesa, postać jego macierzy reprezentującego polaryzator liniowy ustawiony zgodnie z osią x lub y.

$$\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} \text{ dla osi } x \qquad 5.2.13a$$

$$\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} \text{ dla osi } y \qquad 5.2.13b$$

Sprawdźmy jak działa macierz Jonesa tego operatora Jonesa dla polaryzatora o osi zgodnej z osią x-ów gdy pada światło spolaryzowane wzdłuż osi x (5.2.14a) i y (5.2.14b); przyjmujemy amplitudę padającego światła jako równią jeden.

$$\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix} \qquad 5.2.14a$$

$$\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix} \qquad 5.2.14b$$

Tak jak tego oczekujemy światło spolaryzowane zgodnie z orientacją polaryzatora jest przepuszczane całkowicie, a to o polaryzacji ortogonalnej jest całkowicie zatrzymane. A co gdy światło jest spolaryzowane pod kątem $\pi/4$ do osi x?

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ 0\end{bmatrix} \qquad 5.2.15a$$

$$\frac{1}{\sqrt{2}}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}0\\ 1\end{bmatrix} \qquad 5.2.15b$$

Tak jak należy otrzymujemy światło spolaryzowane wzdłuż osi x (5.2.15a) lub y (5.2.15b), o połowie mniejszym natężeniu.

To, że macierz Jonesa dla polaryzatora o osi zgodnej z osią *x* nie zmienia stanu polaryzacji światła spolaryzowanego liniowo zgodnie z osią *x*, wynika również faktu, że taki stan polaryzacji jest reprezentowany przez wektor własny tejże macierzy Jonesa (4.3.4). Wiemy, że działanie macierzy na wektor własny sprowadza się do mnożenia tego wektora przez liczbę. Ogólnie (5.2.15a) możemy zapisać w postaci

$$\begin{bmatrix}\alpha & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\alpha\begin{bmatrix}1\\ 0\end{bmatrix} \qquad 5.2.16$$

Co może oznaczać współczynnik α? Gdy jest mniejszy od jeden reprezentuje fakt pochłaniania przez polaryzator światła. Natężenie światła na wyjściu obliczamy jako sumę kwadratów składowych

$$I=\alpha^2+0^2=\alpha^2 \qquad 5.2.16a$$

Każdy polaryzator w jakimś stopniu pochłania światło. Ale my stosujemy prosty model polaryzatora idealnego i dlatego przyjmujemy, że α=1. Gdy α jest większe od jeden to mamy układ, który na wyjściu daje więcej energii niż pojawiło się na wejściu. Takie własności mają elementy aktywne, które kosztem pewnego rezerwuaru energii pompują ją do wiązki świetlnej. Nie jest to jednak przypadek zwykłego polaryzatora.

Co by się stało, gdybym obrócił układ współrzędnych, tak że osie układu nie byłyby osiami polaryzatora? Własności układu nie uległyby zmianie, ale zapis reprezentacji kolumnowej wektorów Jonesa i macierzowej operatorów Jonesa w tych nowych współrzędnych byłby bardziej skomplikowany. Macierze Jonesa miałyby cztery niezerowe współrzędne, a kolumny własne miałyby dwie niezerowe współrzędne i rachunki stałby się bardziej rozwlekłe.

Dla przykładu macierz Jonesa dla idealnego polaryzatora liniowego, ustawionego pod kątem θ do osi *x*, możemy obliczyć ze wzoru (5.2.10).

$$\mathbf{J}_{\mathbf{L;\theta}}=\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{pmatrix} \qquad 5.2.17$$

Po obliczeniu powyższego wyrażenia dostajemy

$$\begin{bmatrix}c_1^2 & c_1s_1\\ c_1s_1 & s_1^2\end{bmatrix} \qquad 5.2.18$$

$$c_1=\cos(\theta) \qquad 5.2.18a$$

$$s_1=\sin(\theta) \qquad 5.2.18b$$

Dalej jest to macierz Jonesa reprezentująca ten sam operator Jonesa, ale zapisana poprzez transformację podobieństwa w mniej wygodnym układzie współrzędnych

Macierz Jonesa dla idealnej ćwierćfalówki, przy kącie orientacji osi szybkiej θ=0 do osi *x*, przyjmie postać

$$\mathbf{J}_{\frac{\lambda}{4}}=\begin{pmatrix} e^{-i\frac{\pi}{4}} & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix} \qquad 5.2.19$$

przy dowolnym kącie orientacji osi szybkiej θ do osi *x*-ów, na podstawie wzoru (5.2.10) otrzymujemy

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1-ic_2 & is_2 \\ is_2 & 1+ic_2 \end{bmatrix} \qquad 5.2.20$$

$$c_2=\cos(2\theta) \qquad 5.2.20a$$

$$s_2=\sin(2\theta) \qquad 5.2.20b$$

Dla przykładu wyprowadzę wzór na pierwszy wyraz (5.2.20). Z mnożenia macierzy $\mathbf{J}_{\lambda/4}$ (5.2.19) przez macierz obrotu **R** mamy

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix}\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix} (1-i)\cos(\theta) & (1-i)\sin(\theta) \\ -(1+i)\sin(\theta) & (1+i)\cos(\theta) \end{pmatrix} \qquad 5.2.21$$

Mnożenie macierzy obrotu odwrotnego $\mathbf{R}^{-1}$ przez tak uzyskaną macierz (5.2.21), na pozycji pierwszego elementu daje

$$(1-i)\cos^2(\theta)+(1+i)\sin^2(\theta)=1-i\cos(2\theta) \qquad 5.2.22$$

Co jest dokładnie pierwszym wyrazem macierzy (5.2.20).

W szczególnych przypadkach macierz (5.2.20) przyjmuje postać

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix},\ \text{gdy}\ \theta=0 \qquad 5.2.23a$$

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1+i & 0 \\ 0 & 1-i \end{bmatrix},\ \text{gdy}\ \theta=\frac{\pi}{2} \qquad 5.2.23b$$

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & \pm i \\ \pm i & 1 \end{bmatrix}\ \text{gdy}\ \theta=\pm\frac{\pi}{4} \qquad 5.2.23c$$

Zobaczmy jak to działa. Przepuśćmy światło spolaryzowane liniowo, zgodnie z osią *x*, przechodzi przez ćwierćfalówkę, której oś szybka zorientowana jest pod kątem $\pi/4$ do osi *x*-ów. Ćwierćfalówkę opisuje wtedy wzór (5.2.23c) ze znakiem plus

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad 5.2.24$$

Przy okazji, w notacji braketów operacja (5.2.24) ma postać

$$\mathbf{J_q}\left(\frac{\pi}{4}\right)|H\rangle \qquad 5.2.25a$$

$$\mathbf{J_q}\left(\frac{\pi}{4}\right)=\frac{1}{\sqrt{2}}\begin{bmatrix}1 & \pm i\\ \pm i & 1\end{bmatrix} \qquad 5.2.25b$$

Zgodnie z ogólną postacią kolumny Jonesa musi być

$$\begin{bmatrix}e^{i\delta_x}\\ e^{i\delta_y}\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix} \qquad 5.2.26$$

Na podstawie (tab. 1.1) wnioskujemy, że na wyjściu otrzymaliśmy światło kołowo spolaryzowane lewoskrętnie. Wynik zgodny z rozważaniami przedstawionymi w (§4.2.2) poświęconym ćwierćfalówce.

Macierz Jonesa dla idealnej półfalówki, przy kącie orientacji osi szybkiej θ=0 do osi x, przyjmie postać

$$\mathbf{J}_{\frac{\lambda}{2}}=\begin{pmatrix}e^{-i\frac{\pi}{2}} & 0\\ 0 & e^{i\frac{\pi}{2}}\end{pmatrix}=\begin{pmatrix}i & 0\\ 0 & -i\end{pmatrix}=i\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \qquad 5.2.27$$

przy dowolnym kącie orientacji osi szybkiej θ do osi x-ów, na podstawie wzoru (5.2.10), postępując podobnie jak w przypadku ćwierćfalówki, otrzymujemy

$$\begin{bmatrix}ic_2 & is_2\\ is_2 & -ic_2\end{bmatrix} \qquad 5.2.28$$

W szczególnych przypadkach mamy

$$\begin{bmatrix}i & 0\\ 0 & -i\end{bmatrix},\ \text{gdy } \theta=0 \qquad 5.2.29a$$

$$\begin{bmatrix}-i & 0\\ 0 & i\end{bmatrix},\ \text{gdy } \theta=\frac{\pi}{2} \qquad 5.2.29b$$

$$\begin{bmatrix}0 & \pm i\\ \pm i & 0\end{bmatrix}\ \text{gdy } \theta=\pm\frac{\pi}{4} \qquad 5.2.29c$$

Zobaczmy jak to działa. Niech na płytkę półfalową zorientowaną pod kątem π/4 pada światło o jednostkowej amplitudzie, spolaryzowane liniowo pod kątem π/4. Kolumna Jonesa (5.1.8) ma postać

$$\mathbf{v}=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ 1\end{pmatrix} \qquad 5.2.30$$

Macierz Jonesa tej płytki falowej obróconej pod kątem $\theta=\pi$/4, na mocy ma postać (5.2.29c) ze znakiem plus.

$$\mathbf{J}=\begin{pmatrix}0 & i\\ i & 0\end{pmatrix} \qquad 5.2.31$$

Zobaczmy jak macierz ta działa na światło o jednostkowej amplitudzie spolaryzowane liniowo pod kątem θ=π/4

$$\frac{1}{\sqrt{2}}\begin{pmatrix}0 & i\\ i & 0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}i\\i\end{pmatrix}=\frac{i}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}=\frac{e^{i\frac{\pi}{2}}}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \qquad 5.2.32$$

Widać, że stan polaryzacji światła nie zmienił się, choć pod wpływem półfalówki faza światła przesunęła się o $\pi/2$. Ponieważ światło jest liniowo spolaryzowane pod kątem $\pi/4$ i półfalówka obrócona jest pod tym samym kątem, to kolumna (5.2.30) reprezentuje wektor własny operatora Jonesa w danym układzie współrzędnych. Słowem układ zachowuje się dokładnie tak samo jak w przypadku gdy światło pod kątem zero pada na nieobróconą półfalówkę.

Na koniec wyznaczę macierz Jonesa dla dowolnej płytki falowej, to jest płytki wprowadzającej przesuniecie fazowe Δ między szybką a wolną osią, o szybkiej osi zorientowanej pod kątem θ do osi x. Dla kąta θ =0 mamy

$$\mathbf{J}_\Delta=\begin{pmatrix}e^{-i\frac{\Delta}{2}} & 0\\ 0 & e^{i\frac{\Delta}{2}}\end{pmatrix} \qquad 5.2.33$$

Po pierwszym przemnożeniu mamy

$$\begin{pmatrix}e^{-i\frac{\Delta}{2}} & 0\\ 0 & e^{i\frac{\Delta}{2}}\end{pmatrix}\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{pmatrix}=\begin{pmatrix}e^{-i\frac{\Delta}{2}}\cos(\theta) & e^{-i\frac{\Delta}{2}}\sin(\theta)\\ -e^{i\frac{\Delta}{2}}\sin(\theta) & e^{i\frac{\Delta}{2}}\cos(\theta)\end{pmatrix} \qquad 5.2.34$$

Drugie mnożenie daje

$$\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}\begin{pmatrix}e^{-i\frac{\Delta}{2}}\cos(\theta) & e^{-i\frac{\Delta}{2}}\sin(\theta)\\ -e^{i\frac{\Delta}{2}}\sin(\theta) & e^{i\frac{\Delta}{2}}\cos(\theta)\end{pmatrix}=$$

$$\begin{pmatrix}e^{-i\frac{\Delta}{2}}\cos^2(\theta)+e^{i\frac{\Delta}{2}}\sin^2(\theta) & \left(e^{-i\frac{\Delta}{2}}-e^{i\frac{\Delta}{2}}\right)\sin(\theta)\cos(\theta)\\ \left(e^{-i\frac{\Delta}{2}}-e^{i\frac{\Delta}{2}}\right)\sin(\theta)\cos(\theta) & e^{i\frac{\Delta}{2}}\cos^2(\theta)+e^{-i\frac{\Delta}{2}}\sin^2(\theta)\end{pmatrix} \qquad 5.2.35$$

Zrobię przykładowe zadanie.

Zadanie 5.2.1.

Płaska fala światła spolaryzowanego kołowo prawoskrętnie o jednostkowym natężeniu pada normalnie na ćwierćfalówkę, a następnie na 1/8 falówkę. Osie szybkie obu płytek falowych zorientowane są wertykalnie (V - to jest wzdłuż osi y). Określ stan światła po przejściu przez pierwszą i drugą płytkę falową.

Kolumna Jonesa dla światła spolaryzowanego prawoskrętnie kołowo dany jest wzorem (4.1.9), a macierz Jonesa dla ćwierćfalówki o orientacji V ma postać

(5.1.21b). Stan polaryzacji światła obliczymy mnożąc przez siebie obie te macierze

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1-i & 0\\ 0 & 1+i\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix}=\frac{1}{2}\begin{bmatrix}1-i\\ i-1\end{bmatrix}= \qquad 5.2.31$$

$$\frac{1-i}{2}\begin{bmatrix}1\\ -1\end{bmatrix}=\frac{e^{-i\frac{\pi}{2}}}{\sqrt{2}}\begin{bmatrix}1\\ -1\end{bmatrix}$$

Otrzymaliśmy światło spolaryzowane liniowo, zorientowane pod kątem - $\pi/4$ do osi x-ów (4.1.8). Zauważ, że ćwierćfalówka dodała do obu składowych H i V przesunięcie fazowe o $-\pi/2$.

Przypadek (b) - ze wzoru (5.1.35) wyznaczę macierz Jonesa dla 1/8 falówki o orientacji V. Kąt opóźnienia między falą szybką i wolną wynosi $\pi/8$. Macierz Jonesa ma postać

$$\mathbf{J}_{\frac{1}{8}}=\begin{pmatrix}1 & 0\\ 0 & \frac{1}{\sqrt{2}}(1+i)\end{pmatrix} \qquad 5.2.32$$

Działamy tą macierzą na stan wejściowy

$$\begin{pmatrix}1 & 0\\ 0 & \frac{1}{\sqrt{2}}(1+i)\end{pmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ i\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ \frac{1}{\sqrt{2}}(i-1)\end{bmatrix}=\frac{1}{\sqrt{2}}\begin{bmatrix}1\\ e^{i\frac{3}{4}\pi}\end{bmatrix} \qquad 5.2.33$$

Stąd mamy

$$\delta=\frac{3}{4}\pi;\quad \sin(\delta)=\frac{1}{\sqrt{2}};\quad \cos(\delta)=-\frac{1}{\sqrt{2}} \qquad 5.2.34$$

Nadto H=V=1. Równanie elipsy polaryzacji przyjmie postać (1.6)

$$x^2-\sqrt{2}xy+y^2=1 \qquad 5.2.35$$

Ponieważ δ leży w przedziale 0-π, elipsa jest obiegana prawoskrętnie.

Formalizm Jonesa dał nam do ręki efektywną metodę obliczania zmiany dowolnego stanu polaryzacji przy przejściu przez element polaryzacyjny, który można opisać macierzą Jonesa. Rachunki mogą na początku zdawać się zagmatwane, szczególnie gdy trzeba przeskakiwać z jednego do drugiego układu współrzędnych, ale gdy ktoś ma potrzebę takich rachunków, to po zrozumieniu wyłożonych tu podstaw szybko może nabyć wprawy w ich robieniu. Wtedy pokazane tu przykłady uzna oczywiście za bardzo proste.

Formalizm Jonesa nie jest jedyną metodą opisu światła spolaryzowanego. Obok niego często używamy formalizmu Stokesa. Ponieważ polaryzacja jest trudnym ale ważnym tematem pozwolę sobie przybliżyć wam w skrócie również ten drugi opis.

5.3. Sfera Poincarego – krótkie wprowadzenie

Wektory i macierze Jonesa mają konkurencję w postaci tzw. wektorów Stokesa i macierzy Müllera. Macierze Müllera mają rozmiar 4x4, co oznacza, że działają na kolumny liczb w czterowymiarowej przestrzeni (tzw. wektory Stokesa). Wydaje się, że cztery współrzędne to za dużo do opisu stanu polaryzacji, ale formalizm macierzy Müllera ma swoje zalety. Pierwszą z nich jest fakt, że współrzędne wektora Stokesa reprezentują wielkości bezpośrednio mierzalne w eksperymencie. Druga, to elegancka interpretacja geometryczna w postaci sfery Poincarego.

Zacznę od przypomnienia, że stan polaryzacji mogę scharakteryzować przez dwa kąty: kąt α określający azymut elipsy polaryzacji kąt θ określający spłaszczenie elipsy. Kiedy mamy dwa kąty i wektor o stałej długości, to sam się narzuca pomysł reprezentowania stanów polaryzacji na powierzchni sfery. Zanim przejdę do sfery, podam jeszcze jeden wzór opisujący dowolny stan spolaryzowanej fali świetlnej

$$\xi = e^{i\Delta}\tan(\beta) \qquad 5.3.1$$

Kąt Δ jest kątem względnego przesunięcia fazowego między składową *y*-ową i *x*-ową (1.2.c). Tangens kąta β to stosunek amplitudy drgań wzdłuż osi *y* - E_{0y} i wzdłuż osi *x* - E_{0x} (rys. 1.2). Określa on zarazem geometrię prostokąta o bokach równoległych do osi współrzędnych, opisanego na elipsie polaryzacji. Z rysunku (1.2) widać, że zakres zmienności kąta β jest określony przez przedział [0, π/2] Wzór (5.3.1) dla danych kątów Δ i β jest liczbą zespoloną, co oznacza, że stan polaryzacji możemy opisać liczbą zespoloną. Związki między kątami Δ i β, a kątami określającymi elipsę polaryzacji α i θ mają postać

$$\tan(\alpha) = \tan(2\beta)\cos(\Delta) \qquad 5.3.2a$$

$$\sin(2\theta) = -\sin(2\beta)\sin(\Delta) \qquad 5.3.2b$$

Kąty Δ i 2β, mogą reprezentować zarówno stan polaryzacji, jak i punkt na sferze. Mnożenie kąta β przez dwa zapewnia jednoznaczne pokrycie całej sfery. Równoważnie możemy posłużyć się podwojonymi kątami azymutu 2α i eliptyczności 2θ. Sprawę ilustruje rysunek (5.3.1).

W celu umożliwienia operacji na sferze wprowadzimy wektor Stokesa o współrzędnych

$$S_0 = I \qquad 5.3.3a$$

$$S_1 = I\cos(2\beta) \qquad 5.3.3b$$

$$S_2 = I\,\sin(2\beta)\cos(\Delta) \qquad 5.3.3c$$

$$S_3 = I\sin(2\beta)\sin(\Delta) \qquad 5.3.3d$$

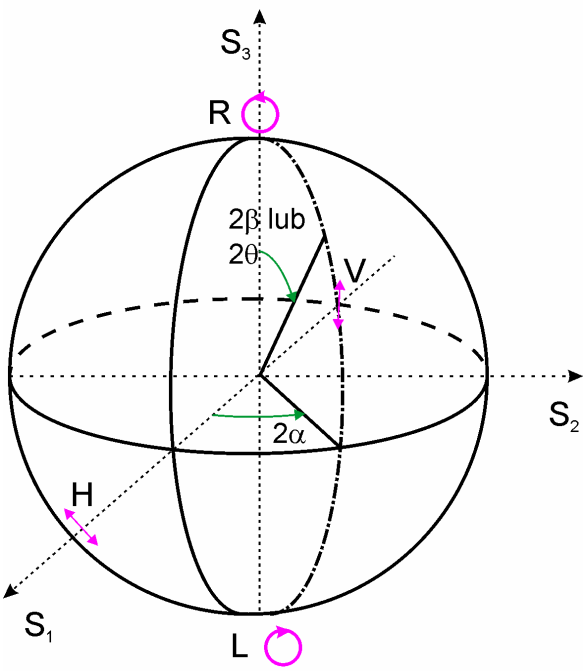
Współrzędne Stokesa w układzie kątów α i θ przyjmą postać

$$S_0 = I \qquad 5.3.4a$$

$$S_1 = I\cos(2\beta) \qquad 5.3.4b$$

$$S_2 = I\,\sin(2\beta)\cos(\Delta) \qquad 5.3.4c$$

$$S_3 = I\sin(2\beta)\sin(\Delta) \qquad 5.3.4d$$



Rysunek 5.3.1. Sfera Poincarego. Punkty na powierzchni sfery identyfikowane są przez dwie współrzędne sferyczne (rys. TIV 5.1.8), które tutaj oznaczane są symbolami przez 2χ i 2γ. Na biegunach sfery znajdują się stany reprezentującą polaryzację kołową prawo i lewo skrętną. Na równiku mamy polaryzacje liniowe. W pozostałych punktach mamy stany eliptyczne. Widać, że wzdłuż równoleżników zmieniają się azymut a wzdłuż południka eliptyczność stanów światła. Antypodalne punkty sfery (punkty leżące po przeciwnej stronie średnicy) wyznaczają stany ortogonalne.

Możemy do sprawy podejść bardziej ogólnie, czyli zrezygnować z założenia całkowitej polaryzacji światła. Jeżeli wiązkę światła uznamy za zbiór paczek falowych, a każda z nich może reprezentować stan o innej polaryzacji, to nie ma możliwości ustalenia stanu polaryzacji światła. Wtedy wzory (5.3.3-4) przestają obowiązywać. Możemy je jednak uogólnić do postaci

$$S_0 = \left\langle E_x^2 \right\rangle + \left\langle E_y^2 \right\rangle \qquad 5.3.5a$$

$$S_1 = \left\langle E_x^2 \right\rangle - \left\langle E_y^2 \right\rangle \qquad 5.3.5b$$

$$S_2 = \left\langle E_x E_y^* \right\rangle + \left\langle E_y E_x^* \right\rangle \qquad 5.3.5c$$

$$S_3 = i\left(\left\langle E_x E_y^* \right\rangle - \left\langle E_y E_x^* \right\rangle\right) \qquad 5.3.5d$$

Ostre nawiasy oznaczają średnią po czasie. Gdy stosujemy zespoloną reprezentację współrzędnych E_x i E_y, to kwadrat tych współrzędnych jest rozumiany jako kwadrat ich modułu. Dla światła całkowicie spolaryzowanego, możemy przejść od wzorów (5.3.5) do wzorów (5.3.3-4), ale wymaga to sporo rachunków, więc sprawy nie będę rozwijał. Związek parametrów Stokesa ze sferą (dla światła spolaryzowanego) pokazany jest na rysunku (5.3.1). Parametr S_0 określa promień sfery.

Jakie jest znaczenie fizyczne parametrów Stokesa? Zacznę od światła niespolaryzowanego. Ponieważ żaden z kierunków polaryzacji nie jest wyróżniony możemy zapisać

$$S_0 = \left\langle E_x^2 \right\rangle + \left\langle E_y^2 \right\rangle = I = 2\left\langle E_x^2 \right\rangle \qquad 5.3.6$$

Zerowy parametr S_0 określa natężenie wiązki I. Wobec równość natężeń na obu składowych polaryzacji liniowej parametr S_1 jest równy zeru, podobnie jak pozostałe dwa parametry, które Zależą od iloczynów mieszanych E_xE_y niezależnie oscylujących harmonik (zakładamy, że każdą falę możemy rozłożyć na harmoniki). Średnie po czasie takich iloczynów są równe zeru. Widać z tego również, że światła niespolaryzowanego nie można reprezentować na sferze o promieniu S_0. Aby można to było zrobić dla światła spolaryzowanego musi być

$$S_0 = \sqrt{S_1^2 + S_2^2 + S_3^2} \qquad 5.3.7$$

Związek (5.3.7) można również wyprowadzić (dla światła spolaryzowanego) przekształcając wzory (5.3.3-4). Wprowadzę wielkości nazywaną stopniem polaryzacji

$$P = \frac{I_{pol}}{I} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \qquad 5.3.8$$

Gdzie I_{pol} to natężenie spolaryzowanej części wiązki świetlnej. Stan wiązki można przedstawić jako sumę części spolaryzowanej i niespolaryzowanej

$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = (1-P)\begin{pmatrix} S_0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + P\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} \qquad 5.3.9$$

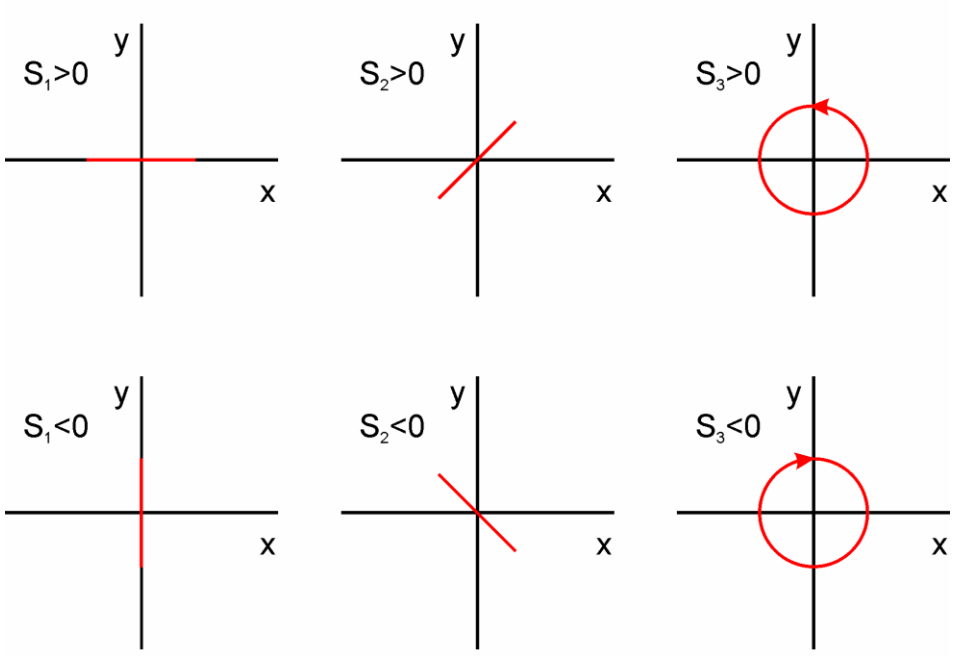
Widać, że dla światła całkowicie spolaryzowanego P=1, a dla światła całkowicie niespolaryzowanego mamy P=0. Pośrednie wartości oznaczają światło częściowo spolaryzowanym. O świetle częściowo spolaryzowanym pisałem już wcześniej (1.1.4), ale ograniczając się do polaryzacji liniowej. Wzór (5.3.8) jest zgodny z (1.1.4), tyle że zapisany z użyciem parametrów Stokesa i jest w mocy dla dowolnych stanów polaryzacji. Wartość P=0 oznacza, że w wiązce świetlnej dominuje pewien stan polaryzacji.

Wracam do omówienia znaczenia fizycznego parametrów Stokesa. Pierwszy parametr S_1 określa różnicę pomiędzy średnim natężeniem składowej wiązki świetlnej spolaryzowanej liniowo wzdłuż osi x-ów, a średnim natężeniem składowej wiązki świetlnej spolaryzowanej liniowo wzdłuż osi y-ów. Możemy to rozumieć tak. Ustawiam w bieg wiązki analizator o osi zgodnej z osią x-ów i mierzymy natężenie światła. Mówiąc kolokwialnie ile tego światła spolaryzowanego liniowo wzdłuż osi x-ów będzie, tyle go przejdzie przez analizator. Następnie robimy to samo dla osi y-ów i odejmujemy obie liczby, co daje nam wartości parametru S_1. Widać przy tym, że dla światła spolaryzowanego liniowo wzdłuż osi x-ów $S_1=I$, a dla polaryzacji liniowej zgodnej z osią y-ów mamy $S_1=-I$.

Drugi parametr Stokesa S_2 określa podobną różnicę, ale przy osiach układu współrzędnych obróconych o $\pi/4$. Teraz analizator jest obrócony raz

o $+\pi/4$ a drugim razem o $-\pi/4$. Parametr S_2 jest różnicą natężeń mierzonych w tych dwóch przypadkach i jest równy I, dla światła spolaryzowanego liniowo zgodnie z obróconą osią x-ów, oraz $-I$ dla światła spolaryzowanego liniowo zgodnie z obróconą osią y-ów.

Trzeci parametr S_3 określa różnicę w natężeniu składowej kołowej polaryzacji prawoskrętnej i lewoskrętnej. Tu musimy zastosować specjalny analizator, tzw. analizator kołowy. Jest to analizator który przepuszcza bezstratnie wiązki światła o stanach polaryzacji R lub L. Można go uzyskać w układzie ćwierćfalówka i polaryzator liniowy. Rysunek (5.3.2) ilustruje kwestie znaków kolejnych parametrów Stokesa, dla charakterystycznych dla nich stanów polaryzacji. Tabela (5.3.1) przedstawia wartości parametrów Stokesa dla wybranych stanów polaryzacji.



Rysunek 5.3.2. Znaki współczynników Stokesa dla różnych czystych stanów polaryzacji powiązanych z charakterem każdego ze współczynników (oprócz współczynnika S_0). Czerwone linie oznaczają stan polaryzacji.

To co się mocno podkreśla przy formalizmie Stokesa to fakt, że wartość współczynników wektora Stokesa mogą być wyznaczone bezpośrednio z pomiarów. Istnieje kilka ekonomicznych (tzn. minimalizujących liczbę pomiarów) schematów pomiarowych dających wartość współczynników Stokesa, których nie będę tu przytaczał[8].

W literaturze parametry S_0-S_3 określane są jako współrzędne wektora Stokesa. Jednak matematycznie rzecz biorąc możliwe stany polaryzacji nie rozpinają przestrzeni wektorowej. Oczywiście istnieją wektory czterowymiarowe. Ale ich przestrzeń jest bogatsza od przestrzeni utworzonej

[8] Zobacz na przykład F. Ratajczy, Optyka ośrodków anizotropowych, PWN, 1994

przez możliwe stany polaryzacji. Pomyślmy o przykładzie światła niespolaryzowanego. Wtedy wszystkie parametry Stokesa oprócz zerowego są równe zeru; $S_0(1,0,0,0)$. W przestrzeni wektorowej istnieje wektor o współrzędnych (-1,0,0,0), co oznaczałoby ujemną wartość natężenie światła. Zatem temu wektorowi nie odpowiada, żadne stan światła. Jest jednak poważniejszy problem. Gdy mamy wektor prędkości **v** o współrzędnych (v_x,v_y)

Stan polaryzacji	S_0	S_1	S_2	S_3
Światło niespolaryzowane	1	0	0	0
Liniowa wzdłuż osi *x* (H)	1	1	0	0
Liniowa +45°	1	0	1	0
Liniowa -45°	1	0	-1	0
Liniowa wzdłuż osi (V)	1	-1	0	0
Kołowa prawo skrętna	1	0	0	1
Kołowa lewo skrętna	1	0	0	-1

Tabela 5.3.1. Wartość współrzędnych wektora Stokesa dla wybranych stanów polaryzacji światła, przy natężeniu światła równym jeden.

w danym układzie współrzędnych, to zawsze możemy zmienić układ współrzędnych tak, ze ten sam wektor będzie miał współrzędne $(0,v)$. W przypadku wektora Stokesa nie możemy wyzerować zerowego parametru, gdyż z definicji jest on przyporządkowany natężeniu światła. Przyznasz, że zmiana układu współrzędnych nie może zmienić natężenia światła. To przypisanie parametrom Stokesa ściśle określonych ról fizycznych wykreśla całość z rodziny wektorów. Będę jednak dalej mówił „wektor Stokesa", ale z powodów historycznych i literaturowych. Nie chcę po prostu wprowadzać nazw niezgodnych z konwencją literaturową.
Podsumujmy tą część definicją

Definicja 5.3.1: Sfera Poincarego

Sferę, która reprezentuje możliwe wektory Stokesa dla światła całkowicie spolaryzowanego o danym natężeniu nazywamy sferą Poincarego.

5.4. Macierze Muellera

Liniowe zmiany stanu polaryzacji reprezentowanego przez wektor Stokesa opisujemy za pomocą macierzy 4x4. Macierze te nazywamy macierzami Muellera[9]. Oczywiście nie dlatego, że są to od strony matematycznej jakieś szczególne macierze. Nazwa macierze Muellera oznacza tylko, że intepretujemy

[9] W oryginalnej pisowni są to macierze Müllera. Ale w polskiej (i anglojęzycznej pisowni mamy macierze Muellera)

je jako reprezentujące działanie liniowego elementu polaryzacyjnego w formalizmie wektorów Stokesa. Wypiszę przykładowe macierze Müllera.

Jeżeli element nie zmienia stanu polaryzacji, to opisuje go oczywiście macierz jednostkowa. Nietrywialne przykłady, tradycyjnie, zacznę od idealnego polaryzatora liniowego. Niech kąt ψ, będzie kątem między osią polaryzatora a osią x.

$$\frac{1}{2}\begin{bmatrix} 1 & c_2 & s_2 & 0 \\ c_2 & c_2^2 & c_2 s_2 & 0 \\ s_2 & c_2 s_2 & s_2^2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad 5.4.1$$

Gdzie współczynniki c_2 i s_2 są zdefiniowane wzorami (5.2.20). W szczególnych przypadkach mamy

$$\psi = 0 \qquad \psi = \pm\pi/4 \qquad \psi = \pi/2$$

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \frac{1}{2}\begin{bmatrix} 1 & 0 & \pm 1 & 0 \\ 0 & 0 & 0 & 0 \\ \pm 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad 5.4.1a$$

Dla ćwierćfalówki macierz Muellera, ma postać

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_2^2 & c_2 s_2 & -s_2 \\ 0 & c_2 s_2 & s_2^2 & c_2 \\ 0 & s_2 & -c_2 & 0 \end{bmatrix} \qquad 5.4.2$$

W szczególnych przypadkach mamy

$$\psi = 0 \qquad \psi = \pm\pi/4 \qquad \psi = \pi/2$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \pm 1 \\ 0 & 0 & 1 & 0 \\ 0 & \pm 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad 5.4.3a$$

Dla półfalówki macierz Muellera, dla kąta ψ orientacji szybkiej osi, ma postać

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_4 & s_4 & 0 \\ 0 & s_4 & -c_4 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \qquad 5.4.4$$

$$c_4 = \cos(4\psi) \qquad 5.4.4a$$

$$s_4 = \sin\left(4\psi\right) \qquad 5.4.4b$$

W szczególnych przypadkach mamy

$\psi = 0$	$\psi = \pm\pi/4$	$\psi = \pi/2$	
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	5.4.4c

Dla dowolnej falówki wprowadzającej przesuniecie fazowe Δ między szybką a wolną osią i o szybkiej osi zorientowanej pod kątem ψ do osi *x*, ma postać

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c_2^2 + s_2^2\cos\left(\Delta\right) & c_2 s_2\left(1-\cos\left(\Delta\right)\right) & -s_2\sin\left(\Delta\right) \\ 0 & c_2 s_2\left(1-\cos\left(\Delta\right)\right) & c_2^2 + s_2^2\cos\left(\Delta\right) & c_2\sin\left(\Delta\right) \\ 0 & s_2\sin\left(\Delta\right) & -c_2\sin\left(\Delta\right) & \cos\left(\Delta\right) \end{bmatrix} \qquad 5.4.5$$

W dalszym toku wykładu będę używał sfery Poincarego do ilustrowania zjawisk związanych z polaryzacją.

6. Interferencja fal spolaryzowanych ♣

Aby zamknąć to krótkie wprowadzenie do teorii światła spolaryzowanego opowiem pokrótce o wpływie polaryzacji światła na zjawisko interferencji i dyfrakcji. Zacznę od interferencji.

Żeby obliczyć interferencję fal musimy znać ich fazę. Tak przynajmniej to wyglądało do tej pory (§TX 1). Jak jest faza fali elektromagnetycznej o polaryzacji – na przykład kołowej? Nie mamy tam jednego wyrażenia typu

$$u = ae^{i\delta} \tag{6.1}$$

Gdzie φ jest fazą fali, a dwa takie wyrażenia, osobno dla osi x i y (zobacz kolumnę Jonesa (4.1.2)). Dla polaryzacji liniowej, można wybrać taki układ współrzędnych, że jedna ze współrzędnych wektora Jonesa jest równa zeru i wtedy nie mamy problemu z jednoznacznym określeniem fazy fali. Wygląda na to, że do tej pory opisywaliśmy interferencję fal liniowo spolaryzowanych. A co w innych przypadkach?

Przypomnę wzór na natężenie światła dwóch interferujących fal monochromatycznych (TX 1.20)

$$i_{12} = a_1^2 + a_2^2 + 2a_1a_2\cos\left(\underbrace{\delta_1 - \delta_2}_{\Delta}\right) \tag{6.2}$$

Człon interferencyjny zawiera cosinus różnicy faz $\Delta\varphi$ tych fal. Słowem interferencja fal daje nam obraz tego jak faza jednej fali zmienia się względem fazy drugiej fali. Gdy fazy obu fal zmieniają się tak samo to ich różnica jest stała, i nie widzimy prążków interferencyjnych tylko jednolicie oświetlony ekran (przy założeniu, że częstości fal są takie same). Różnicę faz $\Delta\varphi$ mogę wyrazić w następujący sposób

$$\Delta\varphi = \arg\left(\left\langle a_1e^{i\delta_1}\middle|a_2e^{i\delta_2}\right\rangle\right) = \mathrm{Arg}\left(e^{i(\delta_1-\delta_2)}\right) \tag{6.3}$$

Funkcja arg wybiera argument liczby zespolonej (def. DD 1.1.2). Zatem dla fal spolaryzowanych liniowo i zgodnie, różnicę ich faz możemy obliczyć z iloczynu skalarnego odpowiednich wektorów Jonesa

$$\Delta\varphi = a\mathrm{rg}\left(\left\langle \mathbf{E_1}\middle|\mathbf{E_2}\right\rangle\right) \tag{6.4}$$

Zobaczmy co się dzieje w ogólnym przypadku.

Napiszę wzór na natężenie sumy dwóch fal danych wzorami (5.1.2)

$$\left\langle \mathbf{E_1}+\mathbf{E_2}\middle|\mathbf{E_1}+\mathbf{E_2}\right\rangle = \left\langle \mathbf{E_1}\middle|\mathbf{E_1}\right\rangle + \left\langle \mathbf{E_2}\middle|\mathbf{E_2}\right\rangle + \left\langle \mathbf{E_1}\middle|\mathbf{E_2}\right\rangle + \left\langle \mathbf{E_2}\middle|\mathbf{E_1}\right\rangle = i_1 + i_2 + \left\langle \mathbf{E_1}\middle|\mathbf{E_2}\right\rangle + \left\langle \mathbf{E_2}\middle|\mathbf{E_1}\right\rangle \tag{6.5}$$

Otrzymaliśmy sumę natężeń tych fal plus człony mieszane, czyli podobną strukturę wyrażenia jak dla przypadku skalarnego (przy pominięciu efektu polaryzacji). Rozpiszę we współrzędnych człony mieszane

$$\left\langle \mathbf{E_1}\middle|\mathbf{E_2}\right\rangle = E_{1x}E_{2x}^* + E_{1y}E_{2y}^* \tag{6.6a}$$

$$\left\langle \mathbf{E_2} \middle| \mathbf{E_1} \right\rangle = E_{1x}^{*} E_{2x} + E_{1y}^{*} E_{2y} \qquad 6.6b$$

Korzystając z postaci (5.1.5) możemy zapisać

$$E_{1x} E_{2x}^{*} + E_{1y} E_{2y}^{*} = E_{0;1x} E_{0;2x} + E_{0;1y} E_{0;2y} e^{i(\Delta_1 - \Delta_2)} \qquad 6.6c$$

$$E_{1x}^{*} E_{2x} + E_{1y}^{*} E_{2y} = E_{0;1x} E_{0;2x} + E_{0;1y} E_{0;2y} e^{-i(\Delta_1 - \Delta_2)} \qquad 6.6d$$

Iloczyn skalarny (5.3.5) jest zatem równy

$$\left\langle \mathbf{E_1} + \mathbf{E_2} \middle| \mathbf{E_1} + \mathbf{E_2} \right\rangle = i_1 + i_2 + 2E_{0;1x} E_{0;2x} + E_{0;1y} E_{0;2y} \left(e^{i(\Delta_1 - \Delta_2)} + e^{-i(\Delta_1 - \Delta_2)} \right) \qquad 6.7$$

lub po rozpisaniu części ekspotencjalnej

$$\left\langle \mathbf{E_1} + \mathbf{E_2} \middle| \mathbf{E_1} + \mathbf{E_2} \right\rangle = i_1 + i_2 + 2E_{0;1x} E_{0;2x} + 2E_{0;1y} E_{0;2y} \cos\left(\Delta_1 - \Delta_2\right) \qquad 6.8$$

Człon interferencyjny jest teraz liczbą zależną od cosinusa z Δ_1-Δ_2. Przypomina to wyrażenie na różnicę faz w (6.3). Nie możemy jednak napisać

$$\arg\left(\left\langle \mathbf{E_1} \middle| \mathbf{E_2} \right\rangle\right) = \Delta_1 - \Delta_2 \qquad 6.9$$

Bo to nie jest prawda, gdyż przeszkadza nam człon $E_{0;1x}E_{0;2x}$ we wzorze (5.3.7c) Mamy co najwyżej zależność

$$\arg\left(E_{0;1y} E_{0;2y} e^{i(\Delta_1 - \Delta_2)} \right) = \Delta_1 - \Delta_2 \qquad 6.9a$$

6.1. Fale spolaryzowane liniowo

Jeżeli dwie fale są spolaryzowane liniowo wzajemnie ortogonalnie to człony mieszane w (6.5) są równe zeru. Nakładające się fale sumują się natężeniowo. Jeżeli polaryzacja tych fal jest współliniowa to sprowadza się to do przypadku interferencji fal skalarnych, który był analizowany w (§TX 1). Aby to zobaczyć, załóżmy, że fale są spolaryzowane wzdłuż osi *x*-ów. (zawsze możemy tak wybrać układ współrzędnych by było to prawdą). Wtedy *y*-owe współrzędne wektora pola elektrycznego są równe zeru. Korzystając z wzoru na składowe pola postaci (5.1.2), jako wynik interferencji otrzymujemy wyrażenie typu (6.2). Powiedzmy zatem, że kierunki polaryzacji interferujących fal tworzą kąt $\alpha \in [0, \pi)$. Zorientuję układ współrzędnych tak, że polaryzacja jednej z fal jest zgodna z kierunkiem osi *x*. Wtedy wektory Jonesa obu fal wyrażą się przez współrzędne

$$\mathbf{E_1} = E_{0;1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad 6.1.1a$$

$$\mathbf{E_2} = E_{0;2} \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \qquad 6.1.1b$$

Obliczę kąt fazowy, części interferencyjnej, ze wzoru (6.9)

$$\arg\left(\langle \mathbf{E_1} | \mathbf{E_2} \rangle\right) = \arg\left(E_{0;1}E_{0;2}\cos(\alpha)\right) = 0 \qquad 6.1.2$$

Zobaczmy jak wygląda wzór opisujący natężenie interferujących fal. Zgodnie z (6.5) ma on postać

$$i = i_1 + i_2 + 2E_{0;1}E_{0;2}\cos(\alpha) \qquad 6.1.3$$

Wyrażenie

$$E_{0;2}\cos(\alpha) \qquad 6.1.3a$$

Jest rzutem wektora natężenia pola elektrycznego $\mathbf{E_2}$ na wektor $\mathbf{E_1}$ (możemy oczywiście interpretować ten wzór jako rzut $\mathbf{E_1}$ na **E2)**. Rzut ten wybiera z wektora $\mathbf{E_2}$ część zgodną z kierunkiem wektora $\mathbf{E_1}$. Ta część zgodna zachowuje się tak jak fala spolaryzowana liniowo zgodne z falą $\mathbf{E_1}$. Część prostopadła do $\mathbf{E_1}$ nie wnosi żadnego wkładu do członu interferencyjnego (mieszanego).

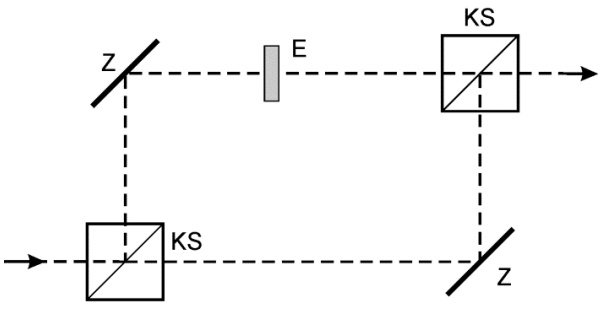
Sprawa wymaga jeszcze nieco uwagi. Skupmy się na interferometrze w układzie Macha-Zendera (rys. 6.1.1). Powiedzmy, że propagują się w nim dwie fale liniowo spolaryzowane, przy czym na jednej ze ścieżek jest element polaryzacyjny, który może zmienić kierunek polaryzacji biegnącej tam fali. Jeżeli fale są zgodnie spolaryzowane to efekt interferencji zależy tylko od względne różnicy faz tych fal. Możemy więc na wyjściu napisać

$$\mathbf{E_1} = E_{0;1}\begin{pmatrix}1\\0\end{pmatrix} \qquad 6.1.4a$$

$$\mathbf{E_2} = E_{0;2}\begin{pmatrix}e^{ik\Delta z}\\0\end{pmatrix} \qquad 6.1.4b$$

Wielkość Δz opisuje różnicę dróg geometrycznych między gałęziami interferometru (przyjmujemy tu, że interferometr pracuje w powietrzu, dla którego współczynnik załamania n=1). Wyrażenie na natężenie interferujących fal przyjmie postać

$$i = i_1 + i_2 + 2\mathbf{E_1}\cdot\mathbf{E_2} = i_1 + i_2 + 2\sqrt{i_1 i_2}\cos(k\Delta z) \qquad 6.1.5$$



Rysunek 6.1.1. Schemat układu do badania interferencji spolary-zowanych fal świetlnych. Z – zwierciadła, KS – kostki światłodzielące, E - element polaryzacyjny do zmiany stanu polaryzacji. Do układu wchodzi światło o określonej polaryzacji. Zakładamy, że zwierciadła i kostki nie zmieniają stanu polaryzacji światła.

Czyli tak jak w przypadku, gdy nie uwzględniamy polaryzacji. Różnica faz pomiędzy falami wynosi teraz $k\Delta z$ Jeżeli obrócimy polaryzację jednej z fal o kąt α, to otrzymamy

$$i = i_1 + i_2 + 2\mathbf{E_1}\cdot\mathbf{E_2} = i_1 + i_2 + 2E_{0;1x}E_{0;2x} + 2\sqrt{i_1 i_2}\cos(k\Delta z)\cos(\alpha) \qquad 6.1.6$$

Składają się tu dwa efekty, jeden związany z różnicą dróg optycznych dla obu ramion interferometru, a drugi z nierównoległością kierunków polaryzacji. W tym drugim przypadku czynnik z cosinusem wkłada do członu interferencyjnego tą składową wektora polaryzacji jednej fali, która jest zgodna z kierunkiem polaryzacji drugiej fali. I dla tej części określamy różnicę faz fali, ponieważ dla części ortogonalnych nie jest to możliwe. Różnica faz fali, przynajmniej rozumiana klasycznie, jest różnicą przesunięć fazowych wynikających z różnic dróg propagacji tych fal. Podsumowując, korzystając z (6.1.2) możemy obliczyć tą składową jednego wektora Jonesa, która jest zgodna z kierunkiem drugiego wektora Jonesa. Ta zgodna składowa interferuje tak jak interferują dwie fale o tej samej polaryzacji.

Do tej pory zakładaliśmy milczącą, że interferują dwie poosiowe fale płaskie, tyle, że o różnych stanach polaryzacji liniowej. Gdy w układzie interferometru Macha-Zendera pochylimy jedno zwierciadło, to doprowadzimy do interferencji dwóch pochylonych fal płaskich. Wtedy różnica faz wynikającą z różnicy dróg optycznych będzie się zmieniała przy zmianie punktu obserwacji (rys. (TX 1.1.1)).

6.2. Fale spolaryzowane kołowo

Rozważmy interferencję fal spolaryzowanych kołowo. Gdy polaryzacje są ortogonalne (prawo i lewo skrętna), to wtedy człon interferencyjny jest równy zeru. Niech zatem będą to polaryzacje zgodne co do skrętności. Kolumny Jonesa tych fal mają postać (5.1.9a)

$$\mathbf{E_1} = E_{0;1}\begin{bmatrix}1\\ i\end{bmatrix} \qquad 6.2.1a$$

$$\mathbf{E_2} = E_{0;2}\begin{bmatrix}1\\ i\end{bmatrix}e^{i\psi} \qquad 6.2.1b$$

Czynnik $e^{i\psi}$ uwzględnia możliwość przesunięcia fazowego między obiema falami na przykład na skutek nierównych dróg optycznych w interferometrze Macha-Zendera (rys. 6.1.1). Obliczymy iloczyn skalarny tych fal

$$\langle \mathbf{E_1} | \mathbf{E_2} \rangle = E_{0;1}E_{0;2}\left(e^{-i\psi} + e^{-i\psi}\right) = 2E_{0;1}E_{0;2}e^{-i\psi} \qquad 6.2.2$$

$$\arg\langle \mathbf{E_1} | \mathbf{E_2} \rangle = \arg\left(2E_{0;1}E_{0;2}e^{-i\psi}\right) = -\psi \qquad 6.2.3$$

W tym przypadku możemy dobrze określić kąt ψ między oboma wektorami natężenia pola elektrycznego. Jest on równy różnicy faz ψ wynikającej z różnicy dróg optycznych między dwoma wiązkami, tak jak to ma miejsce w przypadku, gdy zaniedbujemy polaryzację. Jest to wynik, który łatwo uogólnić na dowolne ale takie same stany polaryzacji fal.

Fakt 6.2.1:

Jeżeli interferują dwie fale o jednakowej polaryzacji, to faza między tymi falami jest dobrze określona i równa jest fazie jaką dla tego przypadku wyliczylibyśmy przy zaniedbaniu efektów polaryzacyjnych.

6.3. Fala spolaryzowana liniowo i kołowo

Przyjrzę się interferencji fali liniowo spolaryzowanej, zgodnie z osią *x*, z falą kołowa spolaryzowaną. Odpowiednie kolumny Jonesa mają postać

$$\mathbf{E_1} = E_{0;1}\begin{bmatrix}1\\0\end{bmatrix} \qquad 6.3.1a$$

$$\mathbf{E_2} = E_{0;2}\begin{bmatrix}1\\ \pm i\end{bmatrix} = E_{0;2}\begin{bmatrix}1\\ e^{\pm i\frac{\pi}{2}}\end{bmatrix} \qquad 6.3.1b$$

Warto w tym miejscu zastanowić się nad bazą, w której reprezentujemy wektory Jonesa. Aby obliczyć iloczyn skalarny współrzędne wektorów muszą być dane w tej samej bazie. Nasze wzory napisane są w bazie V, L (def. 1.1.5 i 1.1.6), to znaczy wektory bazy są zgodne z kierunkami V i L polaryzacji liniowej. Dla polaryzacji kołowej, wzór (6.1.9b) możemy zapisać w postaci

$$E_{0;2}\begin{bmatrix}1\\ \pm i\end{bmatrix} = E_{0;2}\left(|H\rangle \pm i|V\rangle\right) = E_{0;2}\begin{bmatrix}1\\0\end{bmatrix} + E_{0;2}e^{\pm i\frac{\pi}{2}}\begin{bmatrix}0\\1\end{bmatrix} \qquad 6.3.2b$$

Podobnie możemy przedstawić pierwszą falę (6.3.19a)

$$E_{0;1}\begin{bmatrix}1\\0\end{bmatrix} = E_{0;1}\left(|H\rangle \pm i|V\rangle\right) = E_{0;1}\begin{bmatrix}1\\0\end{bmatrix} + 0\begin{bmatrix}0\\1\end{bmatrix} \qquad 6.3.2a$$

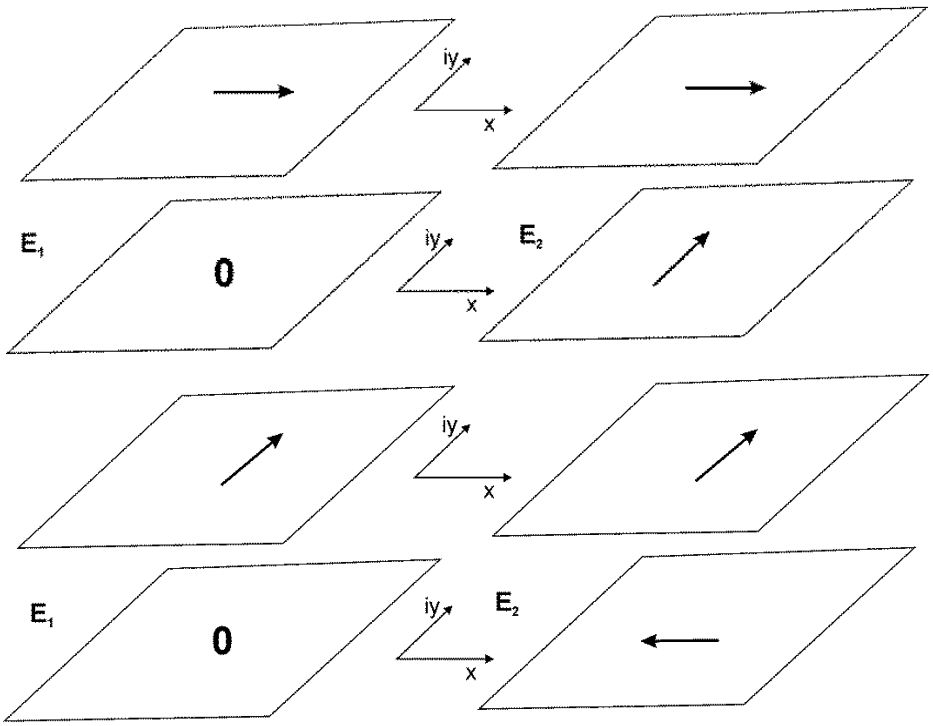
Obliczymy iloczyn skalarny dla fal (6.3.21)

$$\langle \mathbf{E_1}|\mathbf{E_2}\rangle = E_{0;1}E_{0;2} \Rightarrow \arg\langle \mathbf{E_1}|\mathbf{E_2}\rangle = 0 \qquad 6.3.3$$

Rysunek (6.3.1) wyjaśnia dlaczego argument iloczynu skalarnego jest równy zeru. Widać, że wobec zerowania się drugiej składowej pierwszej fali iloczyn skalarny określa zgodność faz pierwszych składowych obu wektorów. Gdy obie fale mają jednostkowe natężenie, to ich kolumny Jonesa mają postać

$$\mathbf{E_1} = \begin{bmatrix}1\\0\end{bmatrix} \qquad 6.3.4a$$

$$\mathbf{E_2} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\ e^{\pm i\frac{\pi}{2}}\end{bmatrix} \qquad 6.3.4b$$



Rysunek 6.1.1. Góra – wektor Jonesa ma dwie zespolone współrzędne, które same można traktować jako dwu wymiarowe wektory. Dlatego do wyrysowania stanu $\mathbf{E_1}$ i $\mathbf{E_2}$ potrzebowałem dwóch płaszczyzn. W iloczynie skalarnym mnożymy odpowiednie współrzędne, po sprzężeniu drugiego składnika iloczynu (4.1.15). Górne płaszczyzny pokazują pierwsze współrzędne wektorów $\mathbf{E_1}$ i $\mathbf{E_2}$. Jako wektory na płaszczyźnie zespolonej obie współrzędne mają taki sam kąt. Dolna płaszczyzna pokazuje drugie współrzędne wektorów $\mathbf{E_1}$ i $\mathbf{E_2}$. Jedna ze współrzędnych jest równa zeru, co oznacza, że fazy nie są w tym przypadku określone. O wzajemnej fazie decyduje tu pierwsza para, która jest fazowo zgodna. Dolny rysunek pokazuje sytuację w chwili późniejszej. Wszystkie fazory ulegają obróceniu o ten sam kąt, bo obie fale mają tą samą częstość i wzajemne kąty nie ulegają zmianie.

Wzór na natężenie interferujących fal (6.1.5) (przy założeniu, że obie mają jednostkową amplitudę) przyjmuje postać

$$i = i_1 + i_2 + \left\langle \mathbf{E_1} \middle| \mathbf{E_2} \right\rangle + \left\langle \mathbf{E_2} \middle| \mathbf{E_1} \right\rangle \qquad 6.3.5a$$

$$\left\langle \mathbf{E_1} \middle| \mathbf{E_2} \right\rangle = 1 \cdot \frac{1}{\sqrt{2}} + 0 \cdot \frac{1}{\sqrt{2}} e^{\mp i\frac{\pi}{2}} = \frac{1}{\sqrt{2}} \qquad 6.3.5b$$

$$\left\langle \mathbf{E_2} \middle| \mathbf{E_1} \right\rangle = \frac{1}{\sqrt{2}} \cdot 1 + \frac{1}{\sqrt{2}} e^{\pm i\frac{\pi}{2}} \cdot 0 = \frac{1}{\sqrt{2}} \qquad 6.3.5c$$

$$i = 1 + 1 + \frac{2}{\sqrt{2}} = 2 + \sqrt{2} \qquad 6.3.6$$

Jest to mniej niż gdyby obie fale były zgodnie i liniowo spolaryzowane, wtedy i=4. Zmniejszenie wartości maksymalnej wynika z tego, że jedna ze składowych fali o polaryzacji liniowej mnoży się przez zerową składową fali liniowo spolaryzowanej i nie daje wkładu do członu interferencyjnego.

Należy zwrócić uwagę, że natężenie fali (6.1.2) (czyli również (6.3.6)) zostało policzone dla ustalonej chwili czasu i w danym punkcie. Ponieważ częstości obu fal są takie same, a człon interferencyjny w (6.3.24) jest od czasu niezależny, to zmiana chwili czasu niczego w wyrażeniu na natężenie światła nie zmienia. Natomiast zmiana punktu obserwacji może zmienić obliczoną wartość natężenia. Jedna z fal może być pochylona względem drugiej co oznacza, że w innym punkcie ta pochylona fala zostanie przemnożona przez dodatkowy czynnik fazowy, co spowoduje zmianę wartości natężenia pola.

Otrzymany wynik pokazuje, że określenie kąta między wektorami Jonesa zwykle się nie udaje. Choć argument iloczynu skalarnego między falą liniowo i kołowo spolaryzowaną jest równy zeru (6.3.5), to nie oznacza to, że możemy uznać iż fale te są w tej samej fazie. Wektor fali kołowo spolaryzowanej kręci się a liniowo spolaryzowanej zachowuje kierunek (ale zmienia się co pół okresu zwrot). Kąt między tymi wektorami zmienia się w czasie, ale nie widać tego we wzorze (6.3.5). Zadziałała tu magia zapisu zespolonego, który jak wiemy (§TX 1.4), przy opisie interferencji, daje wynik równy średniej wartości po okresie (z dokładnością do stałej mnożącej). Spróbujmy przekonać się, że tak rzeczywiście jest.

W tym celu zapiszę wektory natężenia pola elektrycznego w postaci (1.2).

$$\begin{cases} E_{x1} = 1\cos(\omega t) \\ E_{y1} = 0\cos(\omega t + \Delta_1) \end{cases} \qquad 6.3.7a$$

$$\begin{cases} E_{x2} = \dfrac{1}{\sqrt{2}}\cos(\omega t) \\ E_{y2} = \dfrac{1}{\sqrt{2}}\cos\left(\omega t \pm \dfrac{\pi}{2}\right) \end{cases} \qquad 6.3.7b$$

Obliczę średnią po czasie z iloczynu skalarnego sumy wektorów $\mathbf{E_1}$ i $\mathbf{E_2}$.

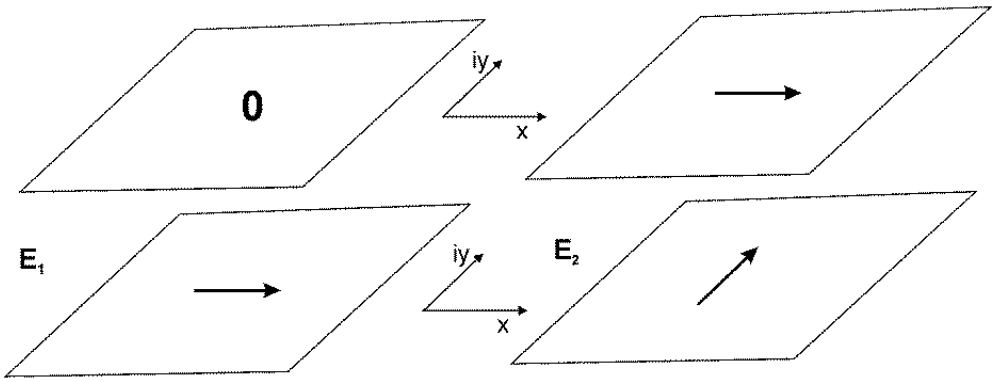
$$\left\langle \mathbf{E_1} + \mathbf{E_2} \middle| \mathbf{E_1} + \mathbf{E_2} \right\rangle = \left(E_1^2 + E_2^2 + \frac{2}{\sqrt{2}} \right)\cos^2(\omega t) \qquad 6.3.8$$

Średnia po czasie po jednym okresie z tego wyrażenia jest równa średniej z cosinusa do kwadratu co jak wiemy jest równe ½ (TX 1.4.5). Zatem z dokładnością do czynnika ½ otrzymaliśmy to samo co z rachunku na wielkościach zespolonych, podobnie jak to było w przypadku skalarnym (§TX 1.4).

Można się z powyższego spodziewać, że dla fali o liniowej polaryzacji danej wzdłuż osi *y*

$$\mathbf{E_1} = E_{0;1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad 6.3.9$$

i dla fali kołowej (5.3.12b) względny kąt dla członu interferencyjnego wynosi $\pm\pi/2$ (rys. 6.3.2)



Rysunek 6.3.2. Gdy wektor pierwsze fali dany jest przez (6.3.9), to pierwsza współrzędna wektora $\mathbf{E_1}$ jest równa zeru. W efekcie o kącie fazowym członu interferencyjnego decyduje dolna para.

I rzeczywiście

$$\langle \mathbf{E_1} | \mathbf{E_2} \rangle = E_{0;1} E_{0;2} e^{\pm i\frac{\pi}{2}} \Rightarrow \arg\langle \mathbf{E_1} | \mathbf{E_2} \rangle = \pm\frac{\pi}{2} \qquad 6.3.10$$

Natężenie w obrazie interferencyjnym (6.3.5) wynosi

$$i = i_1 + i_2 \qquad 6.3.11$$

Jest zatem sumą natężeń fal. Ponieważ fazory dla jednej współrzędnej są ortogonalne a dla drugiej nie są określone (bo jedna ze współrzędnych jest równa zeru) to decyduje ta współrzędna, która jest określona i czynnik interferencyjny zeruje się. Powtórzenie rachunku na średnią po okresie pokazuje, że z dokładnością do czynnika ½ średnia po okresie jest równa (5.3.31).

W ogólnym przypadku interferencję dwóch fal dowolnie spolaryzowanych możemy zapisać w postaci

$$\mathbf{E_1} = \begin{cases} E_{x1} = E_{01x}\cos(\omega t) \\ E_{y1} = E_{01y}\cos(\omega t + \Delta_1) \end{cases} \qquad 6.3.12a$$

$$\mathbf{E_2} = \begin{cases} E_{x2} = E_{02x}\cos(\omega t)e^{i\psi} \\ E_{y2} = E_{02y}\cos(\omega t + \Delta_2)e^{i\psi} \end{cases} \qquad 6.3.12b$$

$$\langle\langle \mathbf{E_1}+\mathbf{E_2} | \mathbf{E_1}+\mathbf{E_2} \rangle\rangle_T = \langle\langle \mathbf{E_1} | \mathbf{E_1} \rangle\rangle_T + \langle\langle \mathbf{E_2} | \mathbf{E_2} \rangle\rangle_T + \langle\langle \mathbf{E_1} | \mathbf{E_2} \rangle\rangle_T + \langle\langle \mathbf{E_2} | \mathbf{E_1} \rangle\rangle_T \qquad 6.3.12c$$

Obliczając średnie po czasie mamy

$$\langle\langle \mathbf{E_1} | \mathbf{E_1} \rangle\rangle_T = \left(E_{01x}^2 + E_{01y}^2\right)\langle\cos^2(\omega t)\rangle_T = \frac{1}{2}i_1 \qquad 6.3.13a$$

Podobnie

$$\langle\langle \mathbf{E_2} | \mathbf{E_2} \rangle\rangle_T = \left(E_{02x}^2 + E_{02y}^2\right)\langle\cos^2(\omega t)\rangle_T = \frac{1}{2}i_2 \qquad 6.3.13b$$

Człon interferencyjny daje wkład

$$\left\langle\left\langle \mathbf{E_1}\middle|\mathbf{E_2}\right\rangle\right\rangle_T =$$

$$E_{01x}E_{02x}e^{-i\psi}\underbrace{\left\langle\cos^2(\omega t)\right\rangle_T}_{=\frac{1}{2}} + E_{01y}E_{02y}e^{-i\psi}\left\langle\cos(\omega t+\Delta_1)\cos(\omega t+\Delta_2)\right\rangle_T \qquad 6.3.13c$$

Oraz

$$\left\langle\left\langle \mathbf{E}_2\middle|\mathbf{E}_1\right\rangle\right\rangle_T =$$

$$E_{01x}E_{02x}e^{i\psi}\underbrace{\left\langle\cos^2(\omega t)\right\rangle_T}_{=\frac{1}{2}} + E_{01y}E_{02y}e^{i\psi}\left\langle\cos(\omega t+\Delta_1)\cos(\omega t+\Delta_2)\right\rangle_T \qquad 6.3.13d$$

Dalej mamy

$$\left\langle\cos(\omega t+\Delta_1)\cos(\omega t+\Delta_2)\right\rangle_T = \qquad 6.3.13e$$

$$\left\langle\begin{matrix}\left(\cos(\omega t)\cos(\Delta_1)-\sin(\omega t)\sin(\Delta_1)\right)\\ \left(\cos(\omega t)\cos(\Delta_2)-\sin(\omega t)\sin(\Delta_2)\right)\end{matrix}\right\rangle_T$$

Składowe tego wyrażenia są równe

$$\left\langle\cos^2(\omega t)\cos(\Delta_1)\cos(\Delta_2)\right\rangle_T = \frac{1}{2}\cos(\Delta_1)\cos(\Delta_2) \qquad 6.3.13f$$

$$\left\langle\cos(\omega t)\sin(\omega t)\cos(\Delta_1)\sin(\Delta_2)\right\rangle_T = 0 \qquad 6.3.13g$$

$$-\left\langle\cos(\omega t)\sin(\omega t)\sin(\Delta_1)\cos(\Delta_2)\right\rangle_T = 0 \qquad 6.3.13h$$

$$\left\langle\sin^2(\omega t)\sin(\Delta_1)\sin(\Delta_2)\right\rangle_T = \frac{1}{2}\sin(\Delta_1)\sin(\Delta_2) \qquad 6.3.13i$$

Zbierając wszystkie składowe mamy

$$\left\langle\left\langle \mathbf{E_1}+\mathbf{E_2}\middle|\mathbf{E_1}+\mathbf{E_2}\right\rangle\right\rangle_T = \frac{1}{2}\begin{pmatrix} i_1+i_2+2E_{01x}E_{02x}\cos(\psi)+\\ 2E_{01y}E_{02y}\cos(\Delta_1-\Delta_2)\cos(\psi)\end{pmatrix} \qquad 6.3.14$$

Oblicz wartość (6.3.13c) w zapisie zespolonym (nie ma wtedy uśredniania po czasie) a otrzymasz to samo, tylko bez czynnika ½.

6.4. Podsumowanie

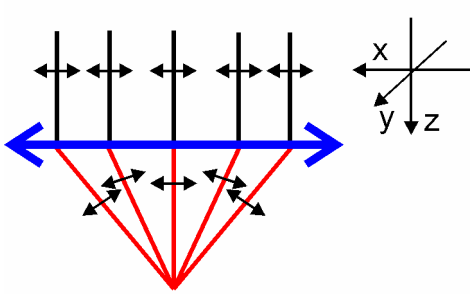
Polaryzacja mocno komplikuje kwestie interferencji. Wynik interferencji zależy od dwóch czynników: Wzajemnego stanu polaryzacji i różnicy dróg optycznych. Ogólny wzór opisujący wpływ czynnika polaryzacyjnego dany jest przez iloczyn skalarny sumy wektorów Jonesa dla pola elektrycznego $\mathbf{E_1}$ i $\mathbf{E_2}$.

Nawet jeżeli ustawimy na wejściu interferometru z rysunku (6.1.1) dwie fale o jednakowej polaryzacji, to obecność elementów odbijających może zmienić ich stan (zobacz §1.2.1). Problem ten dotyczy również interferometrów w innych konfiguracjach. W efekcie na wyjściu otrzymamy dwa różne stany polaryzacji i wynik interferencji przestanie być zgodny z oczekiwanym. Wyjściem z sytuacji jest zastosowanie specjalnych elementów

światłodzielących, które zachowują stan polaryzacji, lub ustawnie polaryzatora na wyjściu interferometru. Polaryzator ustawi obie wiązki w tym samym stanie polaryzacji. Zwykle ustawia się go tak aby na jego wyjściu natężenie obu wiązek było takie samo, wtedy kontrast prążków jest najlepszy.

6.5. Dyfrakcja

Obliczając obraz dyfrakcyjny w punkcie P sumowaliśmy fazory, które reprezentowały amplitudy i fazy fal wtórnych, dochodzących do danego punktu P (§TXI 2.1). Następnie sumę fazorów podnosiliśmy do kwadratu, co uwzględniało w danym punkcie obserwacji P efekt interferencji fal wtórnych. Polaryzacja komplikuje ten obraz, co ilustruje rysunek (6.5.1)



Rysunek 6.5.1. Na soczewkę pada fala płaska spolaryzowana liniowo. Po przejściu przez soczewkę kierunki promieni zmieniają się (tylko centralny nie zmienia kierunku). Zmienia się również kierunek polaryzacji fal i obraz interferencji komplikuje się.

Widać z niego, że po przejściu przez soczewkę promienie reprezentujące padającą poosiowo falę płaską zmieniają swój kierunek. Zmienia się również orientacja polaryzacji (polaryzacja dalej jest prostopadła do kierunku rozchodzenia się światła, czyli do kierunku promienia). Oznacza to, że w płaszczyźnie obserwacji musimy wybrać sobie układ współrzędnych *x*-*y*-*z* i dla każdego promienia dochodzącego z soczewki znaleźć składowe wektora pola elektrycznego w kierunku tych osi. Składowe zgodne z osią *x* (*y*, *z*) będą ze sobą interferowały tak jak w przypadku, gdy nie uwzględnialiśmy polaryzacji. W wyniku otrzymamy amplitudę fali dla składowych polaryzacji o kierunku osi *x* i *y* i *z* (odpowiednio E_x, E_y, E_z). Natężenie światła jest równe

$$i = E_x^2 + E_y^2 + E_z^2 \qquad 6.5.1$$

Dodajemy kwadraty składowych amplitud, bo dla składowych ortogonalnych czynnik interferencyjny jest równy zeru. W przypadku obiektywów o małych aperturach numerycznych (zwykle przyjmuje się NA<0.4), (def. TXI 4.3.1), co oznacza małe kąty ugięcia promieni na obiektywie wpływ polaryzacji jest na tyle mały, że zwykle można go pominąć.