Mały wykład z mechaniki kwantowej

Kacper Zalewski

27 sierpnia 2004

Spis treści

Rozdział 1

Stała Plancka $\hbar$

Za dzień narodzenia mechaniki kwantowej jest uważany 14 grudnia roku 1900. Tego dnia, na posiedzeniu Niemieckiego Towarzystwa Fizycznego w Instytucie Fizyki Uniwersytetu Berlińskiego czterdziestodwuletni profesor zwyczajny tego uniwersytetu, Max Planck wygłosił referat pt "O teorii prawa rozkładu energii w widmie normalnym". W ciągu roku akademickiego posiedzenia Niemieckiego Towarzystwa Fizycznego w Berlinie odbywały się co dwa tygodnie. Uczestniczyli w nich fizycy z różnych jednostek naukowych i nie tylko – jednym z założycieli był słynny przemysłowiec Siemens. W porównaniu do naszych konwersatoriów PTF te zebrania były bardziej robocze. W szczególności, w roku 1900 referaty dotyczyły często promieniowania ciała doskonale czarnego, nad którym fizycy berlińscy prowadzili badania zarówno doświadczalne jak i teoretyczne. Referat Plancka też należał do tej serii.

Powszechnie wiadomo, że ciała rozgrzane świecą. Natężenie światła zależy od rodzaju świecącego ciała. Na przykład roztopione szkło świeci znacznie słabiej niż kawałek węgla rozgrzany do tej samej temperatury. Kirchhoff w roku 1859 pokazał metodami termodynamiki, że dla ciała w równowadze z promieniowaniem stosunek natężenia promieniowania emitowanego do natężenia promieniowania absorbowanego przy danej długości fali jest uniwersalną, niezależną od rodzaju ciała i jego powierzchni funkcją długości fali i temperatury

$$\frac{e(\omega)}{a(\omega)} = \rho(\omega, T). \tag{1.1}$$

Zgodnie z obecnymi zwyczajami zamiast długości fali λ wprowadziliśmy tu częstość kołową

$$\omega = \frac{2\pi c}{\lambda}. \tag{1.2}$$

W tym wzorze c oznacza prędkość światła w próżni. Szybko zauważono, że wzór (1.1) upraszcza się, jeśli rozważać ciało, które absorbuje całe padające na nie promieniowanie niezależnie od jego długości fali. Dla takiego ciała, nazwanego ciałem doskonale czarnym, przy odpowiednim doborze normalizacji, $a(\omega) \equiv 1$ i funkcja $\rho(\omega, T)$ jest rozkładem energii w promieniowaniu emitowanym. Planck zaproponował wzór na ten rozkład, który w odróżnieniu od wzorów proponowanych przez jego poprzedników znakomicie zgadza się z doświadczeniem. W dzisiejszych oznaczeniach:

$$\rho(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\frac{\hbar\omega}{k_B T}} - 1}. \tag{1.3}$$

Iloczyn $\rho(\omega, T)d\omega$, daje energię na jednostkę objętości, promieniowania o częstości kołowej z przedziału $d\omega$, znajdującego się w równowadze z ciałem doskonale czarnym o temperaturze T. Ten wzór zawiera dwie stałe fizyczne: stałą Boltzmanna k_B, która występuje także we wzorze Boltzmanna łączącym entropię układu z liczbą dostępnych układowi stanów mikroskopowych, i stałą $\hbar$. Oznaczenie $\hbar$, które zostało wprowadzone znacznie później przez Diraca, jest obecnie powszechnie używane. Planck wprowadził stałą $h = 2\pi\hbar$, która została potem na jego cześć nazwana stałą Plancka. Stała $\hbar$ ma różne nazwy: zredukowana stała Plancka, h-przekreślone, lub po prostu stała Plancka. Stała Boltzmanna występuje w wielu wzorach termodynamicznych. Mimo to Boltzmann nie wyznaczył nigdy jej wartości numerycznej. We wzorach termodynamicznych znanych Boltzmannowi przed rokiem 1900 stała k_B występuje zawsze pomnożona przez stałą Avogadry. Ten iloczyn, oznaczany przez R i znany jako stała gazowa, był dobrze zmierzony, ale ponieważ oszacowania liczby Avogadry były wówczas bardzo niepewne, trudno było przejść od stałej R do stałej k_B. Porównując swój wzór z doświadczalnym rozkładem energii w widmie ciała doskonale czarnego, Planck wyznaczył po raz pierwszy wartość numeryczną stałej Boltzmanna. Znając tę stałą i stałą gazową wyliczył stałą Avogadry. Otrzymał bardzo rozsądny wynik, co stanowiło niezależne od widma ciała doskonale czarnego potwierdzenie wzoru (1.3). Według obecnych danych:

$$\hbar = 1,055 \times 10^{-34}\text{J sek.} \tag{1.4}$$

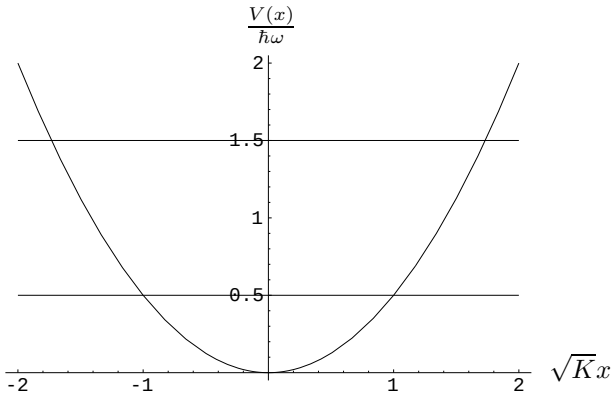
Znane są jeszcze cztery dalsze cyfry za przecinkiem, ale nie podajemy ich tutaj. Stała Plancka ma wymiar działania, czyli iloczynu energii i czasu, lub równoważnie pędu i drogi.

W swoim referacie Planck podał wyprowadzenie wzoru (1.3) w oparciu o zdumiewające współczesnych założenie, że przekazywanie energii między ciałem promieniującym i polem promieniowania odbywa się nie w sposób ciągły, ale "paczkami" o energii

$$E = \hbar\omega. \tag{1.5}$$

Sam Planck, fizyk konserwatywny, głęboko znający i ceniący klasyczną fizykę, nazwał w późniejszych wspomnieniach przyjęcie tego założenia "aktem rozpaczy". Widział, że wzór (1.3) dobrze odtwarza dane i chciał go za wszelką cenę jakoś wyprowadzić. Na podstawie klasycznej elektrodynamiki przyjmowano wtenczas, że źródłem promieniowania o częstości ω jest oscylator harmoniczny o częstości ω. Wzór (1.5) sugerował więc, że oscylatory harmoniczne emitujące promieniowanie nie zmieniają energii w sposób ciągły, ale mają poziomy energetyczne odległe od siebie o wielokrotności kwantu energii $\hbar\omega$. Później, w oparciu o rozumowanie, które ma już wartość tylko historyczną, Planck doszedł do słusznego wniosku, że stan oscylatora o najniższej energii ma energię $\frac{1}{2}\hbar\omega$. Tak więc energia oscylatora harmonicznego może przyjmować tylko wartości

$$E_n = \hbar\omega(n + \frac{1}{2}), \qquad n = 0, 1, 2, \ldots. \tag{1.6}$$



Rysunek 1.1: Energia potencjalna jednowymiarowego oscylatora harmonicznego $V(x) = \frac{1}{2}Kx^2$ w jednostkach $\hbar\omega$. Linie poziome oznaczają, podzielone przez $\hbar\omega$, energie stanu podstawowego i pierwszego stanu wzbudzonego oscylatora.

Te dozwolone wartości energii są nazywane poziomami energetycznymi. Poziom podstawowy ma najniższą z możliwych energię. Potem następują kolejne stany wzbudzone. Częstość kołowa oscylatora harmonicznego ω jest związana z jego stałą siłową K i masą m klasycznym wzorem

$$\omega = \sqrt{\frac{K}{m}}. \tag{1.7}$$

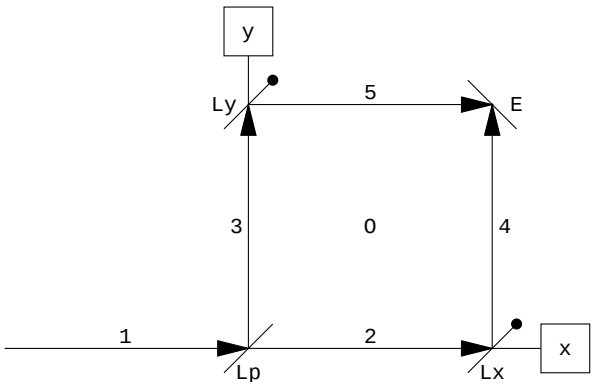
Zależność energii potencjalnej oscylatora harmonicznego od współrzędnej x i dwa pierwsze poziomy energetyczne są pokazane na rysunku 1.

Mówi się, że energia oscylatora jest skwantowana. Kwant energii, czyli różnica energii między sąsiednimi poziomami, zależy od częstości czy długości fali. Dla światła widzialnego, przyjmując $\lambda = 5 \times 10^{-5}$ cm, co odpowiada częstości kołowej $\omega = 4{\times}10^{15}$ sek^{-1}, otrzymujemy $\hbar\omega = 4{\times}10^{-19}$ J. Operowanie wielkościami tak dalekimi od jedności jest niewygodne i dlatego wprowadza się różne inne układy jednostek. Przyjmując za jednostkę energii jeden elektronowolt mamy

$$\hbar = 6,582 \times 10^{-16}\text{eV sek} \tag{1.8}$$

i co za tym idzie dla powyższego przykładu $\hbar\omega = 2,5$ eV. Popularne są też układy jednostek, w których z definicji $\hbar = 1$. Występowanie stałej Plancka jest charakterystyczne dla efektów kwantowych. W przypadku oscylatora harmonicznego różnica energii między sąsiednimi poziomami energetycznymi i różna od zera najniższa możliwa energia oscylatora są proporcjonalne do $\hbar$. W fizyce klasycznej obie te liczby są równe zeru. W roku 1918 Planck otrzymał nagrodę Nobla "w uznaniu wkładu do fizyki teoretycznej przez odkrycie kwantów energii".

W roku 1905 zostało zrobione kolejne wielkie odkrycie związanie ze stałą Plancka. Einstein postawił tezę, że światło składa się z kwantów o energii $\hbar\omega$, które są niepodzielne jak atomy. Powszechnie dziś używana nazwa foton została dla takiego kwantu wprowadzona dopiero w roku 1926 przez G.N. Lewisa. W roku 1921 Einstein otrzymał nagrodę Nobla "za zasługi dla fizyki teoretycznej, a



Rysunek 1.2: Fotony nadlatujące wzdłuż drogi 1 trafiają na lustro półprzepuszczalne Lp. Jeżeli lustra odbijające Lx i Ly są odchylone, to foton zostaje przepuszczony i trafia do detektora x lub odbity i trafia do detektora y. Jeśli lustra odbijające są ustawione jak na rysunku, to fotony mogą wytworzyć prążki interferencyjne na ekranie E. Każdy foton potrzebuje do tego obu dróg $3+5$ i $2+4$. Obserwator w O może spowodować przestawienie luster zanim doleci do nich foton, nawet jeśli decyzję podejmie, kiedy ten foton już opuścił okolice lustra Lp.

w szczególności za odkrycie praw rządzących efektem fotoelektrycznym", a więc za teorię kwantów światła, a nie za teorię względności, która komisji noblowskiej wydawała się jeszcze wtedy nieco podejrzana. Teoria kwantów światła spotkała się ze sprzeciwem większej części ówczesnych czołowych fizyków, a nawet i dziś, kiedy prawie wszyscy się z nią zgadzają, sporna jest interpretacja paradoksów, do których prowadzi. Ograniczymy się tu do przedstawienia cyklu eksperymentów wykonanych przy pomocy układu przedstawionego na rysunku 2.

A1) Rozważmy najpierw następujący eksperyment. Klasyczna wiązka światła leci wzdłuż osi x i trafia na lustro półprzepuszczalne Lp ustawione w początku układu współrzędnych, pod kątem $\pi/4$ do osi x. (rysunek 2). Zgodnie z definicją lustra półprzepuszczalnego połowa natężenia wiązki zostaje przepuszczona i leci dalej wzdłuż osi x. Druga połowa zostaje odbita i leci wzdłuż osi y. Jeżeli gdzieś dalej na osiach x i y ustawimy detektory, powiedzmy w punktach $x = x_0$ i $y = y_0$, to będziemy mogli stwierdzić, że natężenia każdej z wiązek po spotkaniu z lustrem jest równe połowie natężenia wiązki pierwotnej. Żeby nie było problemów z dzieleniem pojedynczych fotonów przyjęliśmy, że wiązka jest klasyczna, to znaczy, że zawiera bardzo dużo fotonów.

A2) Powtórzmy poprzedni eksperyment w sytuacji, kiedy padająca wiązka zawiera tylko jeden foton. Doświadczalnie jest to wykonalne. Ponieważ foton nie może się podzielić albo jeden, albo drugi detektor notuje przybycie fotonu, ale nigdy oba. Już to jest dosyć dziwne. Mamy dokładnie określony stan początkowy i nie potrafimy przewidzieć jaki będzie stan końcowy, ale to jest dopiero początek trudności. Wydaje się oczywiste, że jeśli zadziałał detektor w x_0, to foton przeszedł przez lustro i leciał wzdłuż osi x , a jeśli zadziałał detektor w y_0, to foton się odbił i leciał wzdłuż osi y. Zobaczymy poniżej, że nawet ta sprawa nie jest taka prosta.

B1) Wstawmy teraz na osiach przed detektorami, powiedzmy w punktach

$x_0 - 1$ i $y_0 - 1$, dodatkowe lustra Lx i Ly ustawione pod kątami $\pi/4$ do osi i zorientowane tak, że w przypadku klasycznej wiązki odbite promienie spotykają się w punkcie $(x_0 - 1, y_0 - 1)$. Jak wiadomo, przy takim spotkaniu następuje interferencja i jeżeli wstawimy w obszarze spotkania odpowiedni ekran pokryty emulsją fotograficzną, oznaczony na rysunku E, to powstaną na nim prążki interferencyjne. Można dobrać długość fali i emulsję tak, żeby jeden foton zaczerniał dokładnie jedno ziarno emulsji. Liczne fotony zawarte w klasycznej wiązce trafiają w jedne obszary ekranu częściej, a w inne rzadziej i powstają prążki. Część wiązki, która przeszła przez pierwsze lustro, i część wiązki, która się od niego odbiła, dochodzą do ekranu różnymi drogami. Jeżeli którąkolwiek z tych dróg zablokować, to prążki interferencyjne znikną. Żeby nastąpiła interferencja, potrzebne jest światło idące obu drogami.

B2) Niech teraz w powyższym układzie padająca wiązka składa się z jednego fotonu. Zostanie zaczernione tylko jedno ziarno i trudno stwierdzić, czy była interferencja, czy nie. Możemy jednak powtarzać ten eksperyment z pojedynczymi fotonami co jakiś czas, nie zmieniając emulsji na ekranie. Po dłuższym czasie, kiedy zaczernionych punktów będzie już dużo, stwierdzimy, że powstały prążki interferencyjne takie same jak w klasycznym eksperymencie. Wynikają z tego trzy ważne wnioski:

- Foton interferuje sam ze sobą, bo powstały prążki interferencyjne, a w okolicy ekranu nie było nigdy więcej niż jeden foton naraz. Co więcej, różne fotony nie interferują ze sobą, bo w eksperymencie z klasyczną wiązką, gdzie dużo fotonów nadlatuje jednocześnie, żadna dodatkowa interferencja nie powstaje.

- Skoro do interferencji wystarcza jeden foton, to ten foton musi w jakimś sensie iść obu drogami naraz. Zgadza się to z obserwacją, że zablokowanie którejkolwiek z dwu dróg powoduje, że prążki nie powstają. Widać stąd też, że wprawdzie nie możemy zaobserwować podzielenia się fotonu, ale foton, który nie jest obserwowany, może się zachowywać tak, jakby się podzielił.

- Nie da się przewidzieć, gdzie trafi w ekran pojedynczy foton, tak jak poprzednio nie dawało się przewidzieć, który z liczników go zarejestruje. Jeżeli jednak fotonów jest bardzo dużo, to rozkład punktów trafienia jest przewidywalny. Wynik jest taki, jakby rozkład prawdopodobieństwa punktów trafienia był dobrze określony. Przy rzucaniu monetą to, czy w jednym rzucie wyjdzie orzeł czy reszka, jest nieprzewidywalne, ale przy bardzo wielu rzutach prawo wielkich liczb gwarantuje (dla monety "rzetelnej"), że liczba orłów będzie mniej więcej równa liczbie reszek. Podobnie jest i w omawianym tu eksperymencie. Punkt trafienia jednego fotonu jest nieprzewidywalny, ale rozkład trafień bardzo wielu fotonów musi być mniej więcej taki, jak przewiduje klasyczna teoria.

C) W eksperymentach (A) foton padający na lustro półprzepuszczalne zostaje całkowicie przepuszczony lub całkowicie odbity. Mówimy, że poszedł jedną z dwu możliwych dróg. W eksperymentach (B) foton jakoś idzie obu drogami, choć nie zostaje rozbity na dwa fotony. Powstaje pytanie, skąd foton padający na lustro półprzepuszczalne wie, jak się ma dalej zachowywać. Doświadczenie wskazuje, że w układzie (A) wybiera inną możliwość niż w układzie(B), cho-

ciaż w obu tych układach otoczenie lustra półprzepuszczalnego wygląda dokładnie tak samo. Można by przypuszczać, że wpływa na to ustawienie pozostałej aparatury – konkretnie luster odbijających. Załóżmy więc, że lustra odbijające w punktach $x_0 - 1$ i $y_0 - 1$ są osadzone na zawiasach tak, że można je ustawiać w dwu pozycjach. W pozycji 1 lustro odbija wiązkę jak w eksperymencie (B). W pozycji 2 lustro jest odchylone i światło przechodzi obok. Jeżeli każde z luster ustawimy w położeniu 2, realizujemy układ z eksperymentu (A) i każdy foton idzie jedną drogą. Jeśli każde z luster ustawimy w pozycji 1, to realizujemy układ z eksperymentu (B) i każdy foton idzie obu drogami. Wyobraźmy sobie teraz, że urządzenie eksperymentalne jest bardzo duże. Powiedzmy, że każdy z prostoliniowych odcinków dróg fotonów za lustrem półprzepuszczalnym ma jeden rok świetlny długości. W środku kwadratu utworzonego przez te cztery odcinki siedzi eksperymentator, który w trzy miesiące po tym, jak foton padnie na lustro półprzepuszczalne, decyduje, czy ma być realizowany eksperyment (A) czy eksperyment (B). Łatwo sprawdzić, że jest jeszcze dosyć czasu, żeby nie przekraczając prędkości światła przekazać tę decyzję do obu luster odbijających i odpowiednio je ustawić. W tym wariancie foton decyduje, czy ma iść jedną drogą czy obydwoma, na trzy miesiące zanim powstanie informacja, jaki układ doświadczalny zostanie ustawiony, a mimo to i tak zawsze dobrze zgaduje, co będzie potrzebne.

Niektórzy uważają teorię, która prowadzi do takich paradoksów (poznamy ich więcej w dalszych wykładach), za sprzeczną ze zdrowym rozsądkiem i wyrażają nadzieję, że da się ją zastąpić czymś lepszym, choć na razie nie wiedzą czym. Inni twierdzą, że póki ktoś nie wykaże, że te trudności prowadzą do sprzeczności z doświadczeniem, a nie tylko z czyimś zdrowym rozsądkiem, nie ma się czym przejmować. Prawie wszyscy specjaliści zgadzają się, że do praktycznych zastosowań mechaniki kwantowej jej obecna interpretacja jest zupełnie wystarczająca.

Kolejne wielkie odkrycie zostało dokonane przez de Broglie'a w roku 1924 i uczczone w roku 1929 nagrodą Nobla "za odkrycie falowej natury elektronów". De Broglie założył, że z każdą cząstką swobodną związana jest fala, której wektor falowy jest równoległy do pędu cząstki. Częstość fali jest związana z energią cząstki wzorem (1.5). Wektor falowy fali $\mathbf{k}$ jest związany z pędem cząstki $\mathbf{p}$ wzorem, który obecnie jest znany jako wzór de Broglie'a:

$$\mathbf{p} = \hbar\mathbf{k}. \tag{1.9}$$

Zauważmy, że dla fotonu ten wzór wynika ze wzoru (1.5), bo dla cząstki o masie równej zeru długość wektora pędu pomnożona przez c jest równa energii cząstki, a dla światła długość wektora falowego pomnożona przez c jest równa częstości kołowej ω. Dla cząstki o masie różnej od zera związek między wektorem falowym i długością fali ($k = 2\pi/\lambda$) pozostaje bez zmian, natomiast, jak widać ze wzorów (1.5) i (1.9), związek między częstością kołową i wektorem falowym przybiera postać

$$\omega = \frac{E}{\hbar} = \hbar^{-1}\sqrt{m^2c^4 + \vec{p}^2c^2} = c\sqrt{\vec{k}^2 + \frac{m^2c^2}{\hbar^2}}. \tag{1.10}$$

Tego wzoru nie można bez dodatkowych założeń wyprowadzić ani ze wzoru (1.5), ani z optyki. Z teorii de Broglie'a wynikał zdumiewający dla współczesnych wniosek, że cząstki o masie różnej od zera powinny interferować podobnie

jak fotony. Ten wniosek został wkrótce potwierdzony doświadczalne przez G.P. Thomsona i niezależnie przez C.J. Davissona i L.H. Germera. Davisson i Thomson otrzymali w roku 1937 nagrodę Nobla "za doświadczalne odkrycie dyfrakcji elektronów na kryształach".

Rozdział 2

Pomiary w mechanice kwantowej

W mechanice klasycznej, jeśli podamy położenie i pęd lub prędkość punktu materialnego, to jego stan w danej chwili jest całkowicie określony. Pomiar może zmienić te parametry. Na przykład, mierząc prędkość pocisku przy pomocy wahadła balistycznego, hamujemy bieg pocisku, ale można też, innymi metodami, pomiar przeprowadzić delikatnie, tak że ani pęd, ani położenie pocisku nie zostaną zaburzone. Oczywiście każdy konkretny pomiar wprowadza jakieś zaburzenie, ale przez odpowiedni dobór metody pomiaru możemy to zaburzenie uczynić mniejszym niż jakakolwiek z góry zadana granica.

W mechanice kwantowej sytuacja jest znacznie bardziej skomplikowana. Oprócz ograniczeń wynikających z niedoskonałości zastosowanej metody pomiaru, są jeszcze ograniczenia wynikające z praw Przyrody. Na początek zwracamy uwagę na trzy takie problemy:

- Niektóre pary wielkości nie dadzą się jednocześnie zmierzyć. Na przykład, jakkolwiek pomysłowy byłby eksperymentator, nie wymyśli metody, która pozwoliłaby jednocześnie i z dowolną dokładnością zmierzyć x-ową składową wektora położenia cząstki i x-ową składową jej wektora pędu.

- Pojedynczą wielkość da się w zasadzie zmierzyć z dowolną dokładnością, ale jeżeli mamy kilka układów w tym samym stanie, to naogół dla każdego z nich pomiar da inny wynik. Zwykle nie da się więc przewidzieć wyniku pomiaru znając (dokładnie!) stan układu.

- Nawet jeżeli układ jest w takim stanie, że potrafimy dokładnie przewidzieć wynik pomiaru jakiejś wielkości, to zwykle nie można gwarantować, że stan układu po pomiarze będzie taki sam jak przed pomiarem. W szczególności ponowny pomiar wielkości, która przedtem była dobrze określona, może już dać zupełnie inny wynik.

Dla człowieka wychowanego na klasycznej fizyce, każde z tych stwierdzeń jest wielkim odkryciem lub wielkim błędem. Oba te stanowiska bywały bronione przez największych fizyków. Obecnie jest prawie powszechnie przyjęte, że do celów praktycznych dzisiejsza mechanika kwantowa jest dobra. Niektórzy sądzą jednak, że jest jeszcze do odkrycia jakaś nowa interpretacja, dzięki której

ta teoria stanie się łatwiejsza do przyjęcia, a może i poszerzy zakres stosowalności. Obecna sytuacja stwarza pułapkę, która wciąga niektórych amatorów. Dowiadują się z wiarygodnych źródeł, że oczekiwana jest nowa teoria na miejsce mechaniki kwantowej, a kiedy nie szczędząc wolnego czasu zbudują taką teorię, nikt nie traktuje ich poważnie. Dlatego dobrze jest wiedzieć, że po to żeby taka nowa teoria mogła być potraktowana poważnie, musi zawierać następujące dwa elementy: po pierwsze dowód, że wszystkie dobre wyniki mechaniki kwantowej wynikają też z proponowanej teorii, po drugie precyzyjne stwierdzenie, które z podstawowych założeń mechaniki kwantowej zostały usunięte, co wprowadzono na ich miejsce i dlaczego to ma być ulepszenie.

Jednym z ciekawych epizodów dyskusji wokół podstaw mechaniki kwantowej była praca Einsteina Podolsky'ego i Rosena z roku 1935[1]. W tej pracy autorzy zaproponowali sposób jednoczesnego pomiaru położenia i pędu cząstki wbrew pierwszemu z powyższych punktów. Rozważali następujący eksperyment myślowy. Niech cząstka A rozpada się na cząstki B i C. W układzie spoczynkowym cząstki A, na mocy zasady zachowania pędu, która obowiązuje także w mechanice kwantowej, cząstki B i C lecą z równymi co do wielkości pędami w przeciwnych kierunkach. Kiedy oddalą się od siebie tak znacznie, że wszelkie ich wzajemne oddziaływania możemy zaniedbać, mierzymy pęd cząstki B i stąd wyliczamy pęd cząstki C. Ten pęd zachowuje się póki cząstka C nie oddziałuje z niczym. Mierzymy teraz położenie cząstki C, co podobnie jak pomiar samego pędu jest dozwolone przez mechanikę kwantową. W ten sposób wyznaczyliśmy w danej chwili i pęd, i położenie cząstki C. Jak się wkrótce wyjaśniło, kluczowe jest tu założenie, że kiedy cząstki B i C są bardzo daleko od siebie, to pomiar wykonany na jednej z nich nie zaburza stanu drugiej. To założenie, znane obecnie jako założenie lokalnego realizmu, jest sprzeczne z mechaniką kwantową i, jak już dziś wiemy, z doświadczeniem. Obecnie mówi się o paradoksie EPR, od pierwszych liter nazwisk autorów, polegającym na tym, że wbrew intuicji stany nawet bardzo od siebie odległych cząstek mogą być skorelowane.

Drugi z powyższych punktów budził nie mniejsze emocje. Kiedyś filozofowie uczyli, że różne skutki dowodzą różnych przyczyn. Mechanika kwantowa dostarcza kontrprzykładów na to twierdzenie. Wyobraźmy sobie zbiór atomów radu. Każdy z tych atomów jest dokładnie w tym samym stanie i ewoluuje dokładnie w takich samych warunkach. Mimo to może się zdarzyć, że atom A rozpadnie się po godzinie, a atom B po tysiącu lat. Według mechaniki kwantowej nie ma w tym żadnej sprzeczności i nie da się znaleźć żadnej przyczyny różnych losów atomów A i B. Ten wynik jest sprzeczny z naszym codziennym doświadczeniem. Jest też sprzeczny z programem fizyki, w który wierzono od czasów Newtona. Według tego programu celem fizyki jest znajdowanie równań, które na podstawie warunków początkowych pozwolą jednoznacznie przewidzieć dalsze losy układu. Długo nie wątpiono, że takie równania istnieją. Żeby ratować ten program zaproponowano, że atomy radu są wprawdzie identyczne we wszystkich parametrach, które potrafimy zmierzyć, ale istnieją jeszcze inne "ukryte"parametry, których nie potrafimy zmierzyć, ale gdybyśmy potrafili, to moglibyśmy przewidzieć indywidualny czas życia każdego z atomów. Na razie nikomu nie udało się zbudować teorii z ukrytymi parametrami, która byłaby lepsza od standardowej mechaniki kwantowej. Hipotezy o ukrytych parametrach nie da się obalić, jeśli się precyzyjniej nie sformułuje jej założeń. John

[1] A. Einstein, B. Podolsky i N. Rosen, *Physical Review* **47**(1935)777.

Bell podał założenia, które wydają się bardzo naturalne dla teorii z ukrytymi parametrami. W oparciu o te założenia wyprowadził pewne nierówności, znane obecnie jako nierówności Bella, które w niektórych sytuacjach są sprzeczne z mechaniką kwantową. Dziś już wiadomo, że te nierówności są także sprzeczne z doświadczeniem. To wyklucza szeroką klasę modeli z ukrytymi parametrami, ale można szukać dalej. W tych wykładach będę się trzymał strony praktycznej mechaniki kwantowej unikając subtelności interpretacyjnych. Bardzo dobry i przystępny opis kilku popularnych obecnie interpretacji mechaniki kwantowej można znaleźć w książce Daviesa i Browna "Duch w atomie"[2].

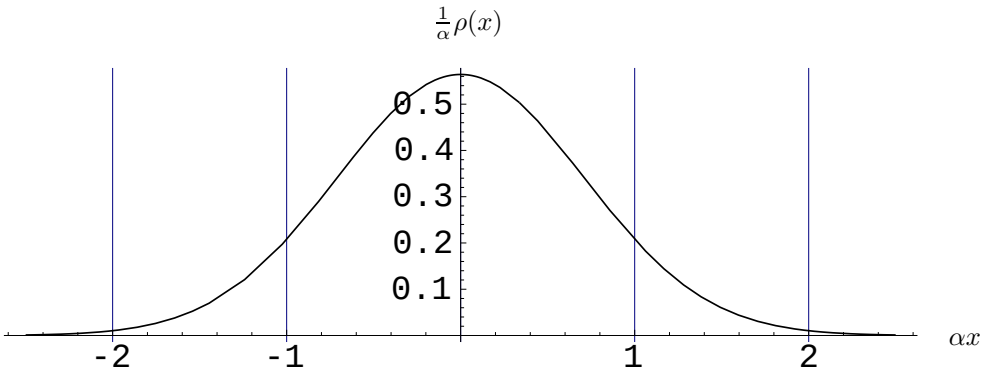
To, które wielkości można jednocześnie zmierzyć, jest jednym z ważnych przewidywań teorii. Wrócimy do niego później. Zastanówmy się teraz, jak można by stwierdzić doświadczalnie, że układ znajduje się w takim, a nie innym stanie. Żeby uniknąć problemów z zaburzaniem układu przez pomiar, należy do pomiarów użyć nie jednego układu, ale zbioru układów, z których każdy jest w tym samym stanie. Taki zbiór nazywa się zespołem i zakładamy, że potrafimy go przygotować. Na każdym z elementów zespołu dokonujemy pomiaru tylko raz. Musimy się liczyć z tym, że nawet przy bardzo dobrze przeprowadzonych pomiarach różne pomiary dadzą różne wyniki. Jak więc uzyskać informacje o stanie układu? Okazuje się, że przy zwiększaniu liczby pomiarów na układzie w danym stanie, wymaga to oczywiście dostatecznie licznego zespołu, częstości poszczególnych wyników (stosunki liczby pomiarów, które dały ten wynik do całkowitej liczby pomiarów) dążą do określonych granic – rozkład prawdopodobieństwa wyników pomiaru jest dobrze określony i mierzalny. Zauważmy, że chociaż nie da się jednocześnie zmierzyć położenia i pędu cząstki, to można wyznaczyć osobno rozkład prawdopodobieństwa dla pędów i rozkład prawdopodobieństw dla położeń w danym stanie układu. Mierząc dostatecznie dużo takich rozkładów prawdopodobieństwa, można określić stan układu. Ile? to zależy od stanu i od rodzaju układu.

Jako prosty przykład rozważmy stan podstawowy (o najniższej energii) jednowymiarowego oscylatora harmonicznego. Oznaczymy ten stan $|E_0\rangle$. Taki zapis stanów został wprowadzony przez Diraca. Zauważmy na razie, że między jego dwa nawiasy można wygodnie wpisywać dowolną ilość informacji o stanie. Na przykład, dla atomu wodoru często rozważa się stany $|n, l, m, s_z\rangle$, gdzie n, l, m, s_z określają odpowiednio energię, kręt, rzut krętu na oś z i rzut spinu na oś z. Stan $|E_0\rangle$ jest określony przez jedną liczbę rzeczywistą – energię, chociaż określenie stanu klasycznego jednowymiarowego oscylatora harmonicznego wymagało by podania dwu liczb, na przykład położenia i pędu. Precyzyjny pomiar energii w tym stanie daje z prawdopodobieństwem jeden (na pewno) wartość E_0. Prawdziwe jest i twierdzenie odwrotne. Jeżeli pomiar energii oscylatora daje zawsze wynik E_0, to oscylator jest w stanie $|E_0\rangle$. Pomiary pędu dają rozkład Gaussa

$$\rho(p) = Ne^{-\frac{p^2}{\hbar\sqrt{mK}}}, \tag{2.1}$$

gdzie m oznacza masę cząstki, K stałą siłową oscylatora, a N jest stałą normalizacyjną zależna od liczby wykonanych pomiarów. Często używa się też rozkładów znormalizowanych do jedności, to znaczy ze współczynnikiem normalizacyjnym N dobranym tak, żeby sumowanie po wszystkich możliwych wynikach

[2] P. Davies i T.R. Brown *Duch w atomie* Wydawnictwo CIS Warszawa (1996)



Rysunek 2.1: Znormalizowany do jedności rozkład prawdopodobieństwa (2.2) dla współrzędnej oscylatora harmonicznego. Parametr $\alpha = \sqrt[4]{\frac{mK}{\hbar^2}}$. Klasycznie dostępny jest tylko obszar $\alpha|x| < 1$. Prawdopodobieństwo znalezienia cząstki poza tym obszarem wynosi około 15,7%. Prawdopodobieństwo znalezienia jej poza obszarem $\alpha|x| < 2$ wynosi niecałe 0,5%

pomiaru – w obecnym przypadku całkowanie po pędzie p od minus nieskończoności do plus nieskończoności – dawało wynik jeden. Zauważmy, że w fizyce klasycznej wartość bezwzględna pędu nie mogłaby przekraczać wartości $\sqrt{2mE_0}$, przy której energia kinetyczna staje się równa energii całkowitej oscylatora. W mechanice kwantowej, natomiast, pomiar (poprawny!) może dać dowolnie dużą wartość pędu, choć pędy większe od dozwolonych klasycznie zdarzają się rzadko. Pomiary położenia też dają rozkład Gaussa

$$\rho(x) = Ne^{-\frac{\sqrt{mK}x^2}{\hbar}} \tag{2.2}$$

Ten rozkład jest pokazany na rysunku 1, gdzie parametr $\alpha = \sqrt[4]{\frac{mK}{\hbar^2}}$. Można było dla prostoty położyć $\alpha = 1$, ale przy wyborze zmiennych jak na rysunku, wykres jest poprawny dla każdej dodatniej wartości parametru α. Znów dozwolone, choć mało prawdopodobne, są wartości bezwzględne x większe od największej wartości dopuszczalnej klasycznie $\sqrt{\frac{2E_0}{K}}$, czyli $1/\alpha$, przy której energia potencjalna staje się równa energii całkowitej, a energia kinetyczna osiąga wartość zero i nie może się już dalej zmniejszać. Znając rozkłady, możemy wyliczać wartości średnie. Dla wielkości mierzalnych x i p otrzymujemy zero, co w notacji Diraca zapisujemy $\langle x\rangle = 0$ i $\langle p\rangle = 0$. Ciekawsze są wyniki dla średnich kwadratów, równych w tych przypadkach wariancjom

$$\langle x^2\rangle = \frac{\hbar}{2\sqrt{mK}}; \quad \langle p^2\rangle = \frac{\hbar\sqrt{mK}}{2}. \tag{2.3}$$

Dobierając iloczyn mK coraz większy lub coraz mniejszy, możemy dowolnie zmniejszać wariancję odpowiednio położenia lub pędu, aż uznamy, że wartość jest z dostateczną dla nas dokładnością jednoznacznie określona, ale nie da się jednocześnie zmniejszyć tych wariancji, bo ich iloczyn jest równy $\frac{\hbar^2}{4}$ niezależnie od wartości mK. To potwierdza, że jednoczesne ustalenie wartości pędu i współrzędnej nie jest możliwe.

Rozdział 3

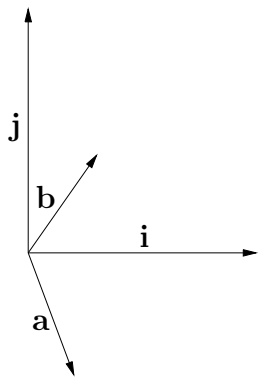
Wektory stanu – zasada superpozycji

Dirac zauważył, że wiele rozumowań w mechanice kwantowej bardzo się upraszcza, jeżeli obiekty określające stan, takie jak $|E_0\rangle$ wprowadzony w poprzednim rozdziale, interpretować jako wektory. W dalszym ciągu będziemy je nazywali wektorami stanu.

Żeby przypomnieć terminologię zaczniemy od szczególnie łatwych do wyobrażenia sobie wektorów – wektorów na płaszczyźnie. Nie jest to mechanika kwantowa, ale może ułatwić zrozumienie podanych dalej w tym rozdziale rozważań z mechaniki kwantowej. Rozważamy wektory zaczepione w początku układu, tak jak wektory narysowane na Rys. 1. Te wektory mają następujące dwie własności: jeżeli $\mathbf{a}$ jest wektorem i c liczbą, to iloczyn $c\mathbf{a}$ jest wektorem. Jeżeli $\mathbf{a}$ i $\mathbf{b}$ są wektorami, to $\mathbf{a} + \mathbf{b}$ też jest wektorem. Równoważnym, a krótszym, sformułowaniem tych własności jest stwierdzenie, że wektory na płaszczyźnie stanowią przestrzeń wektorową. Stwierdzenie, że stanowią przestrzeń wektorową nad ciałem liczb rzeczywistych, oznacza dodatkowo, że liczby c są rzeczywiste. Wektor $c_1\mathbf{a} + c_2\mathbf{b}$, gdzie c_1, c_2 są liczbami, nazywamy kombinacją liniową wektorów $\mathbf{a}$ i $\mathbf{b}$.

W przestrzeni wektorowej możemy wprowadzić bazę, to znaczy zbiór wektorów taki, że każdy wektor z rozważanej przestrzeni da się przedstawić jako kombinacja liniowa wektorów z tego zbioru i że żaden podzbiór tego zbioru nie jest bazą. Najbardziej znana baza dla wektorów na płaszczyźnie składa się z dwu wektorów, jednego leżącego na osi x i drugiego leżącego na osi y. Ale istnieje nieskończenie wiele innych baz. Każde dwa wektory nie leżące na jednej prostej stanowią bazę. Liczba wektorów bazowych nie zależy od wyboru bazy i nazywa się wymiarem przestrzeni wektorowej. Wymiar przestrzeni wektorów na płaszczyźnie wynosi oczywiście dwa. O wektorach mówi się, że stanowią układ zupełny, jeżeli każdy wektor z rozważanej przestrzeni wektorowej da się przedstawić jako ich kombinacja liniowa. Baza jest więc układem zupełnym zawierającym minimalna liczbę wektorów.

Wektor można przedstawić jako uporządkowany zbiór liczb. Na przykład wektor $\mathbf{a}$ można przedstawić jako parę liczb (a_x, a_y) gdzie a_i oznacza rzut wektora $\mathbf{a}$ na oś i. Elementy tego zbioru liczb, w naszym przykładzie a_x i a_y, nazywamy współrzędnymi wektora. Porządek jest ważny, bo wektor (a_1, a_2) jest



Rysunek 3.1: Każde dwa wektory z czterech przedstawionych na rysunku stanowią bazę dla wektorów na płaszczyźnie. Wektor **b** ma inne współrzędne (inną reprezentację) w bazie (**i**, **j**), a inne w bazie (**a**,**j**)

.

naogół różny od wektora (a_2, a_1). Uporządkowany zbiór współrzędnych wektora (liczb) nazywamy reprezentacją wektora. Każdy wektor może być reprezentowany na nieskończenie wiele różnych sposobów w zależności od tego, jak wybieramy wektory bazowe, na które rzutujemy. Do definicji rzutów potrzebne jest pojęcie iloczynu skalarnego. Iloczyn skalarny wektorów **a** i **b** definiujemy wzorem

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y \tag{3.1}$$

Wektor nazywamy jednostkowym, jeśli kwadrat jego długości, $\mathbf{a}^2 \equiv \mathbf{a} \cdot \mathbf{a} = 1$. Jeżeli $\mathbf{a} \cdot \mathbf{b} = 0$, to mówimy, że wektory **a** i **b** są wzajemnie ortogonalne. Bazę złożoną z wektorów jednostkowych i parami ortogonalnych nazywamy bazą ortonormalną.

Dla przestrzeni wektorowych nad ciałem liczb zespolonych wzór (3.1) trzeba uogólnić, bo w przeciwnym razie, na przykład, wektor o współrzędnych $(i, 0)$ miałby ujemny kwadrat długości. W tym celu dla każdego wektora **a** wprowadzamy wektor dualny $\mathbf{a}_D$. Dla wektorów na płaszczyźnie można przyjąć, że współrzędnymi wektora $\mathbf{a}_D$ są (a_x^*, a_y^*), gdzie gwiazdka oznacza sprzężenie zespolone. Zamiast iloczynu skalarnego (3.1) będziemy stosować iloczyn

$$\mathbf{a}_D \cdot \mathbf{b} = a_x^* b_x + a_y^* b_y. \tag{3.2}$$

Definicje długości wektora i ortogonalności dwu wektorów pozostają bez zmian. Ta definicja iloczynu skalarnego redukuje się do definicji (3.1), jeśli wektor **a** ma współrzędne rzeczywiste Wynika z niej też, że długość każdego wektora jest liczbą rzeczywistą i nieujemną, przy czym długość wektora zero oznacza, że wektor jest zerowy, czyli że każda z jego współrzędnych jest równa zeru. Naogół iloczyn skalarny nie jest już jednak przemienny

$$\mathbf{a}_D \cdot \mathbf{b} = (\mathbf{b}_D \cdot \mathbf{a})^* \tag{3.3}$$

Zestawimy teraz własności iloczynu skalarnego, które są oczywiste, jeśli się używa definicji (3.2), ale które w ogólniejszym przypadku są przyjmowane jako

aksjomaty. Wprowadzamy oznaczenia Diraca, w których iloczyn skalarny (3.2) jest pisany jako $\langle \mathbf{a}|\mathbf{b}\rangle$. Zapiszmy jego właściwości.

1. Kwadrat długości wektora $\mathbf{a}$, $\langle \mathbf{a}|\mathbf{a}\rangle$, jest liczbą rzeczywistą nieujemną, równą zeru wtedy i tylko wtedy kiedy wektor $\mathbf{a}$ jest wektorem zerowym

2. Iloczyn skalarny wektorów $\mathbf{a}$ i $\mathbf{b}$ jest liniowy to znaczy, że dla dowolnej liczby c i dowolnych wektorów $\mathbf{a}$, $\mathbf{b}$, $\mathbf{b}_1$ i $\mathbf{b}_2$

$$\langle \mathbf{b}_1 + \mathbf{b}_2|\mathbf{a}\rangle = \langle \mathbf{b}_1|\mathbf{a}\rangle + \langle \mathbf{b}_2|\mathbf{a}\rangle. \tag{3.4}$$

$$\langle \mathbf{b}|c\mathbf{a}\rangle = c\langle \mathbf{b}|\mathbf{a}\rangle \tag{3.5}$$

3. Iloczyn skalarny $\langle \mathbf{a}|\mathbf{b}\rangle$ jest liczbą, naogół zespoloną. Zmiana porządku czynników jest równoważna ze sprzężeniem zespolonym:

$$\langle \mathbf{b}|\mathbf{a}\rangle = \langle \mathbf{a}|\mathbf{b}\rangle^*. \tag{3.6}$$

Wracamy teraz do mechaniki kwantowej. Należy zwracać uwagę zarówno na podobieństwa, jak i na różnice z omówionym powyżej prostym przykładem. Zbiór wektorów stanu danego układu uzupełniony wektorem zerowym stanowi przestrzeń wektorową nad ciałem liczb zespolonych. To, oczywiście, nie jest twierdzenie z matematyki, tylko ważny fakt z fizyki. Wektor zerowy nie odpowiada żadnemu stanowi układu. Jeżeli dwa wektory różnią się tylko o stały czynnik liczbowy, to odpowiadają temu samemu stanowi układu. Tak więc każdemu różnemu od wektora zerowego wektorowi z tej przestrzeni przypisany jest jednoznacznie stan układu, natomiast każdemu stanowi układu odpowiada nieskończony zbiór wektorów $c|\psi\rangle$, gdzie wektor $|\psi\rangle$ określa stan układu, a c może być dowolną liczbą zespoloną różną od zera. Podkreślamy, że wektory stanów różnych układów, na przykład jednocząstkowe wektory stanu dwu elektronów, nie należą do tej samej przestrzeni wektorowej. Wektory stanu pary elektronów uzupełnione wektorem zerowym natomiast, stanowią przestrzeń wektorową.

Twierdzeniu, że kombinacja liniowa dwu wektorów należących do danej przestrzeni wektorowej jest wektorem należącym do tej samej przestrzeni, odpowiada jedno z podstawowych praw mechaniki kwantowej znane jako zasada superpozycji.

Zasada superpozycji: Jeżeli $|\phi\rangle$ i $|\psi\rangle$ są wektorami stanu tego samego układu, a liczby (naogół zespolone) c_1 i c_2 nie są jednocześnie równe zeru, to $c_1|\phi\rangle + c_2|\psi\rangle$ jest też wektorem stanu tego układu.

Z punktu widzenia klasycznej mechaniki ta zasada jest zupełnie niezrozumiała. Punkt materialny może mieć położenie $\vec{x}_1$ i pęd $\vec{p}_1$ lub położenie $\vec{x}_2$ i pęd $\vec{p}_2$, ale co znaczy kombinacja liniowa takich stanów? Łatwiej jest znaleźć analogię w optyce. Rozważmy monochromatyczną wiązkę światła, spolaryzowaną liniowo i propagującą wzdłuż osi z. Płaszczyzna polaryzacji jest jednoznacznie określona przez wektor polaryzacji $\vec{\varepsilon}_1$. Jest to wektor jednostkowy w płaszczyźnie (x, y). Niech teraz ta wiązka interferuje z inną wiązką o tej samej długości fali i tej samej fazie, też propagującą wzdłuż osi z i mającą polaryzację liniową określoną przez wektor polaryzacji $\vec{\varepsilon}_2$. Wiązka wypadkowa będzie spolaryzowana liniowo i jej wektor polaryzacji będzie kombinacją liniową wektorów $\vec{\varepsilon}_1$ i $\vec{\varepsilon}_2$. Współczynniki c_1, c_2 tej kombinacji liniowej będą zależały od stosunku amplitud interferujących

wiązek. Weźmy teraz przykład z mechaniki kwantowej. Oscylator harmoniczny może być w stanie podstawowym $|E_0\rangle$, lub w pierwszym stanie wzbudzonym $|E_1\rangle$. Wynika stąd, że jeżeli liczby c_1, c_2 nie są równocześnie równe zeru, to oscylator może być także w stanie

$$|\psi\rangle = c_1|E_0\rangle + c_2|E_1\rangle. \tag{3.7}$$

Przypadki, kiedy $c_1 = 0$ lub $c_2 = 0$ są mało interesujące. Wtedy oscylator jest odpowiednio w stanie $|E_1\rangle$ lub $|E_0\rangle$. Jeżeli natomiast żadna z liczb c_1, c_2 nie jest równa zeru, to stan $|\psi\rangle$ jest różny od każdego ze stanów składowych. W szczególności dokładny pomiar energii będzie czasami dawał wynik E_0, a czasami wynik E_1. Odpowiedź na pytanie, jak często otrzymamy który wynik, wymaga znajomości dodatkowego prawa fizyki i będzie omówiona w następnym rozdziale.

Definicja iloczynu skalarnego jest zbudowana podobnie jak dla dwuwymiarowych wektorów, ale z jedną istotną zmianą. Wektory $\mathbf{a}$ i $\mathbf{a}^*$ należały do tej samej przestrzeni w tym sensie, że ich suma była dobrze określona. Dla wektorów stanu z mechaniki kwantowej wektory stanu $|\psi\rangle$ i wektory dualne do nich, które oznacza się $\langle\psi|$, należą do różnych przestrzeni i dodawanie ich do siebie nie ma sensu, podobnie jak nie ma sensu dodawanie sekund do kilogramów.

Jak zobaczymy dalej, przestrzeń stanów danego układu w mechanice kwantowej ma często wymiar nieskończony. W takim razie obliczanie iloczynów skalarnych wymaga sumowania szeregów nieskończonych lub liczenia całek. Może się okazać, że suma czy całka jest rozbieżna i wtedy iloczyn skalarny nie istnieje. W dalszym ciągu będziemy rozróżniać wektory, których kwadrat jest liczba skończoną, i wektory uogólnione, dla których sytuacja jest bardziej skomplikowana. Jedne i drugie są używane w mechanice kwantowej. Okazuje się, że wektory stanu odpowiadające stanom związanym, jak stan oscylatora harmonicznego czy stan elektronu w atomie wodoru, są wektorami. Natomiast inne ważne stany, jak stan cząstki o ściśle określonym pędzie czy położeniu, są wektorami uogólnionymi.

Rozdział 4

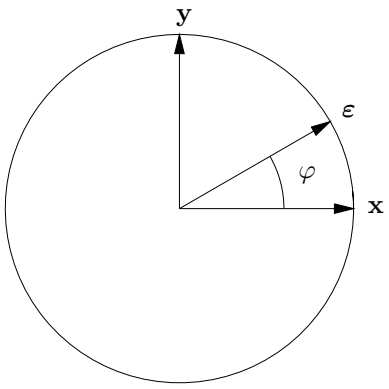
Interpretacja probabilistyczna

Omówimy teraz bardzo ważne zagadnienie interpretacji fizycznej współczynników liczbowych w kombinacjach liniowych wektorów stanu. Zaczniemy od przykładu z fizyki klasycznej. Rozważmy równoległą wiązkę światła monochromatycznego spolaryzowaną liniowo i biegnącą równolegle do osi z. Stan polaryzacyjny wiązki jest określony przez jednostkowy wektor polaryzacji $\boldsymbol{\varepsilon}$ leżący w płaszczyźnie (**x**,**y**) prostopadłej do kierunku biegu wiązki. Orientację wektora polaryzacji można ustalić przy pomocy pryzmatu Nicola. Kiedy światło spolaryzowane liniowo pada na taki pryzmat, przechodzi w całości jeśli wektor polaryzacji jest równoległy lub antyrównoległy do osi pryzmatu, którą ustawiamy równolegle do osi **x**. Jeśli wektor polaryzacji jest prostopadły do tej osi, natężenie wiązki przepuszczonej spada do zera. Jeżeli kąt między osią **x** i wektorem $\boldsymbol{\varepsilon}$ wynosi φ, i natężenie padającej wiązki jest I, to przechodzi wiązka o natężeniu $I\cos^2\varphi$. Ten fakt można interpretować następująco. Natężenie I wiązki jest proporcjonalne do kwadratu długości wektora $\boldsymbol{\varepsilon}$. Wiązkę interpretujemy jako superpozycję dwu wiązek: wiązki z wektorem polaryzacji o długości ε cosφ skierowanym równolegle do osi **x** i wiązki o wektorze polaryzacji o długości ε sinφ skierowanym prostopadle do osi **x** (Rys. 4.1). Oznaczając jednostkowe wektory polaryzacji skierowane odpowiednio wzdłuż osi **x** i **y** przez $\boldsymbol{\varepsilon}_x$ i $\boldsymbol{\varepsilon}_y$ otrzymujemy

$$\boldsymbol{\varepsilon} = \cos\varphi\boldsymbol{\varepsilon}_x + \sin\varphi\boldsymbol{\varepsilon}_y. \tag{4.1}$$

Ten rozkład jest zgodny ze zwykłymi regułami rozkładu wektora na składowe. Pierwsza z wiązek w całości przechodzi, a druga jest w całości zatrzymana. Stąd natężenie przepuszczonej wiązki jest zredukowane o czynnik $\cos^2\varphi$.

Sprawa się komplikuje, jeśli rozpatrujemy światło jako zbiór fotonów. Natężenie można dobrać tak małe, że będzie padał jeden foton dziennie i wobec niemożliwości rozbicia fotonu, nie jest możliwe przepuszczenie części fotonu i zatrzymanie reszty. Postulujemy, że w takim wypadku każdy foton ma prawdopodobieństwo $\cos^2\varphi$ przejścia przez pryzmat. Na mocy twierdzenia wielkich liczb, jeśli fotonów pada bardzo dużo, to stosunek liczby tych, które przejdą do całkowitej liczby padających fotonów, będzie bardzo bliski $\cos^2\varphi$. Ponieważ każdy foton niesie taką samą energię, tak samo podzieli się natężenie padającej wiązki. Otrzymujemy więc wynik zgodny z doświadczeniami dla wiązek zawierających wiele fotonów, jeśli założymy że prawdopodobieństwo przejścia fotonu przez



Rysunek 4.1: Pryzmat Nicola jest ustawiony tak, że przepuszcza światło spolaryzowane liniowo, jeśli wektor polaryzacji jest równoległy do wektora **x** i całkowicie je zatrzymuje, jeśli wektor polaryzacji jest równoległy do wektora **y**. Jeśli wektor polaryzacji $\boldsymbol{\varepsilon}$ tworzy kąt φ z wektorem **x**, to każdy foton ma prawdopodobieństwo przejścia $\cos^2\varphi$ i co za tym idzie prawdopodobieństwo $\sin^2\varphi$ zostania zatrzymanym.

pryzmat Nicola, to znaczy zachowania się tak jakby foton miał polaryzację $\boldsymbol{\varepsilon}_x$, i prawdopodobieństwo zatrzymania fotonu, to znaczy zachowania takiego jakby foton miał polaryzację $\boldsymbol{\varepsilon}_y$, dane są odpowiednio przez kwadraty współczynników we wzorze (4.1): $\cos^2\varphi$ i $\sin^2\varphi$. Można te prawdopodobieństwa wyrazić przez iloczyny skalarne odpowiednich wektorów. Oznaczmy przez $P_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_x)$ prawdopodobieństwo, że foton z polaryzacją $\boldsymbol{\varepsilon}$ zachowa się jak foton z polaryzacją $\boldsymbol{\varepsilon}_x$, czyli przejdzie przez pryzmat. Oznaczmy przez $P_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_y)$ prawdopodobieństwo, że foton z polaryzacją $\boldsymbol{\varepsilon}$ zachowa się jak foton z polaryzacją $\boldsymbol{\varepsilon}_y$, czyli nie przejdzie przez pryzmat. Wtedy

$$P_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_x) = (\boldsymbol{\varepsilon}_x \cdot \boldsymbol{\varepsilon})^2; \qquad P_{\boldsymbol{\varepsilon}}(\boldsymbol{\varepsilon}_y) = (\boldsymbol{\varepsilon}_y \cdot \boldsymbol{\varepsilon})^2. \tag{4.2}$$

Powyższe wzory dopuszczają jeszcze jedną interpretację, i tę właśnie stosuje się zwykle w mechanice kwantowej. Kiedy foton pada na pryzmat Nicola, następuje pomiar jego polaryzacji. Jeżeli wektor polaryzacji fotonu $\boldsymbol{\varepsilon}$ nie jest równoległy, antyrównoległy czy prostopadły do osi **x**, to wynik takiego pomiaru nie da się przewidzieć, chociaż stan polaryzacyjny padającego fotonu jest dokładnie znany. Można tylko obliczyć prawdopodobieństwa wyników ze wzorów (4.2). Jeżeli foton przejdzie przez pryzmat Nicola, to uznajemy, że był w stanie polaryzacyjnym $\boldsymbol{\varepsilon}_x$ (a nie tylko, że zachował się tak jakby był w tym stanie). Jeżeli foton zostanie zatrzymany wnioskujemy, że był w stanie polaryzacyjnym $\boldsymbol{\varepsilon}_y$. Można więc wyliczyć, lub wyznaczyć z pomiarów, prawdopodobieństwo, że foton, który ma wektor polaryzacji $\boldsymbol{\varepsilon}$, okaże w pomiarze, że ma wektor polaryzacji $\boldsymbol{\varepsilon}_x$, lub że ma wektor polaryzacji $\boldsymbol{\varepsilon}_y$. Ten wynik może się wydawać naturalną konsekwencją wzoru (4.1), ale jest tu trudność. W jaki sposób foton w stanie $\boldsymbol{\varepsilon}$, który jest superpozycją fotonów w stanach $\boldsymbol{\varepsilon}_x$ i $\boldsymbol{\varepsilon}_y$, przechodzi w foton w jednym z tych stanów i już na pewno nie w tym drugim? Ta trudność będzie lepiej widoczna w paradoksie z kotem Schrödingera, który przedyskutujemy w dalszym ciągu rozdziału.

Ogólnie, w mechanice kwantowej postulujemy, że **jeżeli stany $|\psi\rangle$ i $|\phi\rangle$ tego**

samego układu są znormalizowane do jedności, co znaczy że $\langle\psi|\psi\rangle = \langle\phi|\phi\rangle = 1$, to prawdopodobieństwo wyniku pomiaru wskazującego, że układ jest w stanie, $|\phi\rangle$, jeśli wiadomo, że układ jest w stanie $|\psi\rangle$, wynosi

$$P_\psi(\phi) = |\langle\phi|\psi\rangle|^2. \tag{4.3}$$

Liczbę, naogół zespoloną, $\langle\phi|\psi\rangle$ nazywamy amplitudą prawdopodobieństwa. Rozumując jak w fizyce klasycznej, można by się spodziewać, że jeśli wiemy że układ jest w stanie $|\psi\rangle$, to tym samym wiemy, że nie jest w żadnym innym stanie i w szczególności prawdopodobieństwo, że układ jest w stanie $|\phi\rangle \neq |\psi\rangle$ powinno być równe zero. W mechanice kwantowej jednak, pomiar zaburza układ i może dać wynik $|\phi\rangle$, także jeśli przed pomiarem układ był na pewno w stanie $|\psi\rangle$.

Rozpatrzmy kilka przykładów. Jeżeli jednowymiarowy oscylator harmoniczny jest w stanie $|E_0\rangle$, w którym pomiar energii na pewno daje wynik E_0, to prawdopodobieństwo wyniku E_1, wskazującego że układ jest w stanie E_1, jest zero. Stąd

$$\langle E_1|E_0\rangle = 0. \tag{4.4}$$

Ogólniej, dwa wektory stanu odpowiadające jednoznacznie określonym, różnym wartościom jakiejś wielkości mierzalnej są zawsze ortogonalne. Jeżeli takie wektory stanowią bazę w przestrzeni wektorów stanu danego układu, to normalizując je do jedności otrzymujemy bazę ortonormalną. W dalszym ciągu będziemy prawie zawsze używać baz ortonormalnych, bo to upraszcza wzory.

Rozważmy teraz stan jednowymiarowego oscylatora harmonicznego

$$|\psi\rangle = c_0|E_0\rangle + c_1|E_1\rangle. \tag{4.5}$$

Na mocy zasady superpozycji jednowymiarowy oscylator harmoniczny może być w takim stanie. Zakładamy, że każdy z wektorów stanu $|\psi\rangle, |E_0\rangle, |E_1\rangle$ jest znormalizowany do jedności. Amplituda prawdopodobieństwa, że pomiar energii da wynik E_0, wynosi

$$\langle E_0|\psi\rangle = c_0\langle E_0|E_0\rangle + c_1\langle E_0|E_1\rangle = c_0. \tag{4.6}$$

Pierwszy iloczyn skalarny w środkowym wyrażeniu jest równy jeden z założenia o normalizacji, a drugi zero jak widać z poprzedniego przykładu. Rozumując podobnie dla amplitudy prawdopodobieństwa, że wynik pomiaru będzie E_1, otrzymujemy

$$P_\psi(E_0) = |c_0|^2; \qquad P_\psi(E_1) = |c_1|^2. \tag{4.7}$$

Ponieważ wyniki E_0 i E_1 nawzajem się wykluczają i są jedynymi możliwymi wynikami pomiarów w stanie $|\psi\rangle$, to z teorii prawdopodobieństwa wiadomo, że powinno zachodzić $|c_0|^2+|c_1|^2 = 1$. Łatwo sprawdzić, że to wynika z normalizacji wektora $|\psi\rangle$. Wektorem dualnym do wektora $|\psi\rangle$ jest

$$\langle\psi| = c_0^*\langle E_0| + c_1^*\langle E_1|. \tag{4.8}$$

stąd

$$1 = \langle\psi|\psi\rangle = |c_0|^2 + |c_1|^2, \tag{4.9}$$

gdzie skorzystaliśmy z ortogonalności i normalizacji wektorów $|E_0\rangle$ $|E_1\rangle$.

Wyniki uzyskane tu dla kombinacji liniowej dwu wektorów stanu przenoszą się na kombinacje liniowe większej liczby, skończonej lub nieskończonej, wzajemnie ortogonalnych, znormalizowanych do jedności wektorów stanu. Zawsze prawdopodobieństwo otrzymania wyniku pomiaru odpowiadającego i-temu wektorowi stanu jest równe kwadratowi modułu jego współczynnika $|c_i|^2$ i zawsze suma wszystkich tych prawdopodobieństw jest równa jeden (ćwiczenie).

Sprawa się komplikuje, jeśli wynik pomiaru może przebiegać ciągłe widmo wartości. Na przykład położenie x cząstki w jednym wymiarze może przyjmować dowolną wartość od minus do plus nieskończoności. Kombinację liniową wektorów $|x\rangle$ trzeba teraz pisać jako całkę

$$|\psi\rangle = \int dx\ \psi(x)|x\rangle, \tag{4.10}$$

Prawdopodobieństwo uzyskania w wyniku pomiaru danej wartości położenia x jest naogół równe zeru. Rozkład prawdopodobieństwa należy więc określić nie przez podanie prawdopodobieństw wszystkich możliwych wyników, jak w poprzednim przypadku, ale przez podanie gęstości prawdopodobieństwa $\rho(x)$ związanej z prawdopodobieństwem wzorem

$$P_\psi(x_0 \le x \le x_1) = \int_{x_0}^{x_1} dx'\ \rho(x'). \tag{4.11}$$

Odpowiednikiem wzorów (4.7) jest w tym przypadku wzór

$$\rho(x) = |\psi(x)|^2. \tag{4.12}$$

Obiekty takie jak $|x\rangle$, które nazwaliśmy tu wektorami, są ściślej mówiąc wektorami uogólnionymi. Wektora uogólnionego nie można znormalizować do jedności, bo jego kwadrat jest nieokreślony. Jakieś określenie normalizacji jest jednak konieczne, bo inaczej nie byłaby określona normalizacja funkcji $\psi(x)$, a wiadomo z teorii prawdopodobieństwa, że powinno być

$$\int_{-\infty}^{+\infty} dx\ \rho(x) = 1. \tag{4.13}$$

Poprawna interpretacja wektorów uogólnionych wymaga zastosowania dość zaawansowanej matematyki, ale to się opłaca, bo wiele rachunków bardzo się skraca i upraszcza, jeśli dopuścimy wektory uogólnione. Wektory uogólnione omówimy w rozdziale szóstym.

Trudności pojęciowe związane z interpretacją probabilistyczną mechaniki kwantowej dobrze ilustruje eksperyment myślowy znany jako kot Schrödingera. Wyobraźmy sobie kota zamkniętego w szczelnym pokoju. W pokoju, poza zasięgiem kota znajduje się urządzenie, które ma 50% szans na otworzenie przełącznika p. Urządzenie ma być kwantowe, takie że po jego zadziałaniu stan przełącznika jest

$$|\psi\rangle = \frac{1}{\sqrt{2}}|p+\rangle + \frac{1}{\sqrt{2}}|p-\rangle, \tag{4.14}$$

gdzie $|p+\rangle$ oznacza stan z otwartym przełącznikiem, a $|p-\rangle$ stan z zamkniętym przełącznikiem. Jeżeli przełącznik został otwarty, to uruchamia ciężarek, który spada na butelkę z cyjanowodorem, który jak wiadomo jest silnie trującym gazem. Butelka zostaje stłuczona, cyjanowodór dociera do kota i kot zdycha. Jeśli przełącznik nie został otwarty, nic się nie dzieje i kot pozostaje w dobrym zdrowiu. Takie urządzenie uruchamiające przełącznik łatwo jest zaprojektować. Można na przykład wykorzystać układ opisany w rozdziale pierwszym. Pojedynczy foton pada na lustro półprzepuszczalne, jeśli przejdzie jest rejestrowany przez detektor, który następnie otwiera przełącznik, jeśli się odbije, nic ciekawego się nie dzieje. Jakiś czas po zadziałaniu urządzenia mamy więc stan, który wyewoluował ze stanu przełącznika

$$|\psi\rangle = \frac{1}{\sqrt{2}}|p+, \text{zdechły kot}\rangle + \frac{1}{\sqrt{2}}|p-, \text{żywy kot}\rangle. \tag{4.15}$$

W tym zapisie pominęliśmy mniej interesujące cechy układu jak to, czy butelka jest stłuczona, czy nie. Powyższy stan, zupełnie poprawny z punktu widzenia mechaniki kwantowej, jest czymś bardzo dziwnym. Nikt nigdy nie widział superpozycji żywego i zdechłego kota. Przy czym nie chodzi tu o zdychającego kota, ale o kota, który z prawdopodobieństwem 50% jest całkiem zdechły i z prawdopodobieństwem 50% jest żywy, wesoły i tryskający energią. Ale to jest dopiero początek trudności. Kiedy ktoś zajrzy do pokoju po zakończeniu eksperymentu z kotem (przez okienko, żeby się nie zatruć) , zobaczy kota żywego (na pewno), lub zdechłego (na pewno). Wygląda, jakby wzrok zaglądającego mógł dobić lub uratować tego kwantowego kota miedzy życiem i śmiercią. Istnieje jeszcze dziwniejsza wersja tego eksperymentu myślowego znana jako "przyjaciel Wignera". Przypuśćmy, że zaglądający sam znajduje się w zamkniętym domku zawierającym pokój z kotem i że obiecał podnieść prawą rękę do góry, jeśli zobaczy kota żywego, a opuścić ja, jeśli zobaczy kota martwego. Póki nikt nie obserwuje domku, można twierdzić, że stan wewnątrz jest opisany przez wektor

$$|\psi\rangle = \frac{1}{\sqrt{2}}|p+, \text{zdechły kot, ręka w dół}\rangle + \frac{1}{\sqrt{2}}|p-, \text{żywy kot, ręka do góry}\rangle. \tag{4.16}$$

Dopiero kiedy przyjaciel zaglądającego spojrzy przez okno w domku i zobaczy jak zaglądający trzyma rękę, nastąpi pomiar i kot zostanie dobity, lub uratowany. Schrödinger uznał, że w tym przykładzie jest tyle sprzeczności, że mechanika kwantowa powinna zostać gruntownie przerobiona. Dopiero teoria bez takich sprzeczności mogłaby być do przyjęcia. Inny pogląd, prezentowany na przykład przez Feynmana w jego słynnych wykładach, jest następujący. Celem fizyki jest przewidywanie wyników doświadczeń. W doświadczeniu z kotem teoria przewiduje, że jeśli przeprowadzimy ten eksperyment bardzo wiele razy, to mniej więcej w połowie przypadków kot przeżyje, a w połowie zdechnie. Dopiero gdyby to przewidywanie okazało się sprzeczne z doświadczeniem, czego nikt się nie spodziewa, byłoby się nad czym zastanawiać. Jeszcze inne poglądy w tej sprawie można znaleźć w cytowanej już książce Davisa i Browna.

Szczególnie dyskutowane jest pytanie, jak się to dzieje, że stan taki jak kot żywy+zdechły przechodzi w procesie pomiaru w kota całkiem żywego lub całkiem zdechłego. Wymyślono dla tego procesu nazwę: redukcja pakietu, ale to niewiele pomaga. Ten proces zachodzi w czasie, ale czy da się opisać równaniami ruchu mechaniki kwantowej? Jedni odpowiadają nie, inni odpowiadają

tak, ale tylko jeśli się uwzględni szczególne własności układów makroskopowych o bardzo wielu stopniach swobody, jak każdy klasyczny aparat pomiarowy, czy w naszym przykładzie kot i każdy z obserwatorów. Co to są za własności i jak je uwzględnić, niezbyt wiadomo. Jeżeli mechanika kwantowa nie wystarcza, to powstaje pytanie w którym miejscu się załamuje i dla czego. Próbowano odpowiedzi, że do redukcji pakietu potrzeby jest świadomy obserwator i mechanikę kwantową trzeba uzupełnić nieznanymi jeszcze prawami Przyrody dotyczącymi świadomości. Ale ta odpowiedź nasuwa następne pytania. Czy to musi być świadomość człowieka, czy wystarczy dinozaur, czy jakiś pierwotny stwór jednokomórkowy? Od tego zależy, od jak dawna można było redukować pakiety. Co się działo przedtem, czy wszystkie możliwości współistniały jak w przykładzie z kotem? Wszechświat jest zapewne też układem opisywanym przez teorię kwantową, ale skąd tu wziąć zewnętrznego obserwatora. Odpowiedź Berkeleya, że tym zewnętrznym obserwatorem jest pan Bóg nie jest obecnie popularna wśród fizyków. Jak widać można tu stawiać wiele pytań, ale nie wiadomo nie tylko jakie są odpowiedzi, ale nawet, czy same pytania maja sens.

Interpretację probabilistyczną mechaniki kwantowej zaproponował w roku 1926 Max Born, który otrzymał w roku 1954 nagrodę Nobla "za podstawowe badania w dziedzinie mechaniki kwantowej, w szczególności za jego statystyczną interpretacje funkcji falowej". Historia tego odkrycia dobrze ilustruje trudności, które napotykają badacze dochodząc do wyników. Potem, na wykładach, te wyniki są często przedstawiane jako coś niemal oczywistego. W mającej niecałe cztery strony, ale bardzo ciekawej, pracy o rozpraszaniu[1] Born uzyskał wynik, który w naszej notacji można by zapisać jako $\rho(x) = \psi(x)$. Na dole strony jest notka dodana w korekcie: staranniejsze rozważanie wskazuje, że prawdopodobieństwo jest proporcjonalne do kwadratu wielkości (znów w naszej notacji) $\psi(x)$. Po paru miesiącach stało się jasne, głównie pod wpływem zależnego od czasu równania Schrödingera, w którym występuje jednostka urojona i, że funkcje falowe naogół przyjmują wartości zespolone i już jako coś naturalnego zaczęto interpretować we wzorze na prawdopodobieństwo kwadrat amplitudy jako kwadrat modułu amplitudy.

[1] M. Born, Zeischrift für Physik **37**(1926)863

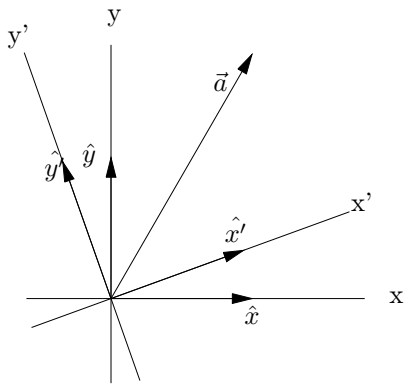
Rozdział 5

Reprezentacje wektorów stanu – funkcje falowe

Jeżeli w przestrzeni stanów danego układu wybierzemy bazę, to każdy wektor z tej przestrzeni można jednoznacznie określić przez podanie jego składowych względem tej wybranej bazy. Liczba wektorów bazowych jest z definicji równa wymiarowi przestrzeni, w której ta baza jest określona, i wymiarowi reprezentacji w tej bazie. Operowanie uporządkowanymi zbiorami liczb, czyli reprezentacjami wektorów, często okazuje się dogodniejsze niż operowanie abstrakcyjnymi wektorami.W tym rozdziale omówimy niektóre własności takich reprezentacji.

Należy rozróżnić dwa rodzaje baz. W jednym, zbiór wektorów bazowych można ponumerować wskaźnikiem $k = 1, 2, \ldots$ Jeżeli liczba wektorów bazowych jest skończona i wynosi n, to baza, ma wymiar n. Jeżeli liczba wektorów bazowych jest nieskończona, to mówimy, że baza jest przeliczalna. W tym przypadku wymiar jest nieskończony. W obu tych przypadkach mówimy, że baza jest dyskretna. Dla danego wektora i danej bazy dyskretnej, reprezentantem jest skończony, lub nieskończony, ciąg składowych wektora, czyli liczb c_k, naogół zespolonych.

Inny rodzaj stanowią bazy, gdzie wektory bazowe można określić przy pomocy wskaźnika α, który zmienia się w sposób ciągły. W tym przypadku mówimy o bazie i reprezentacji ciągłej. Symbol α może oznaczać jedną liczbę lub skończony zbiór liczb, na przykład wektor. Reprezentacje ciągłe są oczywiście nieskończenie wielo wymiarowe. W tych reprezentacjach zamiast listy współczynników c_k podaje się funkcje, na przykład $\psi^{(\alpha)}(\alpha)$, gdzie ψ oznacza wektor stanu, który opisujemy; górny wskaźnik α, który się zwykle pomija, przypomina, że używamy bazy α, a argument α jest konkretną wartością wskaźnika, któremu odpowiada współczynnik — liczba $\psi(\alpha)$. Ze względów historycznych takie funkcje są nazywane funkcjami falowymi. Jak zobaczymy w tym rozdziale i w następnym, stosowanie baz ciągłych wymaga pokonania pewnych trudności matematycznych. Na przykład okazuje się, że wektory bazowe nie są takimi wektorami, jakie dotąd rozważaliśmy, bo ich iloczyny skalarne naogół nie są liczbami. Mogłyby więc być określane jako wektory uogólnione. W dalszym ciągu będziemy używali terminu wektory uogólnione, tylko kiedy będzie nam zależało na podkreśleniu, że dany obiekt nie jest wektorem z dobrze określonymi



Rysunek 5.1: Bazy $(\hat{x}, \hat{y})$ oraz $(\hat{x'}, \hat{y'})$ i wektor $\vec{a}$.

iloczynami skalarnymi, ale może być wektorem stanu jakiegoś układu. W innych przypadkach będziemy, niezbyt ściśle, nazywali wektorami zarówno wektory zwykłe jak i uogólnione. W praktyce, stosowanie baz ciągłych jest naturalne i powszechnie stosowane. Na przykład stany cząstki o danym położeniu $|\vec{x}\rangle$, jak również stany cząstki o określonym pędzie $|\vec{p}\rangle$ tworzą bazy ciągłe.

Istnieją i bardziej skomplikowane przypadki, kiedy wektory bazowe mają jednocześnie wskaźniki dyskretne i ciągłe, albo kiedy w jednej bazie niektóre wektory bazowe mają wskaźniki dyskretne, a niektóre wskaźniki ciągłe, ale uogólnienie na takie przypadki jest już łatwe i nie będziemy go osobno dyskutować. Omówimy teraz reprezentacje dyskretne i reprezentacje ciągłe, podkreślając podobieństwa i różnice. Najpierw jednak, jako wprowadzenie, przypomnimy odpowiednie własności zwykłych, znanych z geometrii wektorów na płaszczyźnie.

Każdy wektor $\vec{a}$ na płaszczyźnie można zapisać w postaci

$$\vec{a} = c_x\hat{x} + c_y\hat{y}, \tag{5.1}$$

gdzie $\hat{x}$ i $\hat{y}$ oznaczają wektory jednostkowe równoległe odpowiednio do osi x i do osi y. Para liczb rzeczywistych (c_x, c_y) określa jednoznacznie wektor $\vec{a}$. Mówimy, że ta para liczb stanowi reprezentacje wektora $\vec{a}$. Mówi się także, że liczby (c_x, c_y) przedstawiają wektor $\vec{a}$ w reprezentacji $(\hat{x}, \hat{y})$. Te dwa znaczenia słowa reprezentacja mogą prowadzić do nieporozumień, więc widząc to słowo należy się zawsze zastanowić, w którym z nich zostało użyte. Dla ułatwienia, tam gdzie ta niejednoznaczność mogła by być szczególnie kłopotliwa, będziemy nazywali parę liczb (c_x, c_y) reprezentantem wektora $\vec{a}$ w reprezentacji $(\hat{x}, \hat{y})$. Para wektorów $\hat{x}$ i $\hat{y}$ stanowi dla wektorów na płaszczyźnie bazę. Jest to baza ortonormalna, bo długość każdego z wektorów bazowych jest równa jeden ($\hat{x}^2 = \hat{y}^2 = 1$) i wektory bazowe są do siebie ortogonalne ($\hat{x} \cdot \hat{y} = 0$). Wymiar przestrzeni jest zdefiniowany jako równy liczbie jej wektorów bazowych. Z tego, że baza składa się z dwu wektorów wynika więc, że wektory na płaszczyźnie stanowią przestrzeń dwuwymiarową.

Gdybyśmy wybrali inną bazę, otrzymalibyśmy inną reprezentacje wektora $\vec{a}$. Na przykład, wybierając jako wektory bazowe wektory $\hat{x'}, \hat{y'}$ (rysunek 1), otrzymalibyśmy inną reprezentację $(c_{x'}, c_{y'})$ tego samego wektora $\vec{a}$. Jak łatwo

sprawdzić przy pomocy rysunku,

$$\begin{aligned} c_{x'} &= V_{x'x}c_x + V_{x'y}c_y, \\ c_{y'} &= V_{y'x}c_x + V_{y'y}c_y, \end{aligned} \tag{5.2}$$

gdzie elementy macierzowe V_{ij} są równe kosinusom kątów między odpowiednimi osiami, lub równoważnie iloczynom skalarnym odpowiednich wektorów bazowych: $V_{x'x} = \hat{x'} \cdot \hat{x}$ itd. Powyższy wzór można przepisać w postaci macierzowej

$$\begin{pmatrix} c_{x'} \\ c_{y'} \end{pmatrix} = \begin{pmatrix} V_{x'x} & V_{x'y} \\ V_{y'x} & V_{y'y} \end{pmatrix} \begin{pmatrix} c_x \\ c_y \end{pmatrix}, \tag{5.3}$$

czyli oznaczając macierze pojedynczymi literami, jako $C' = VC$. Macierz V jest ortogonalna, to znaczy że jej odwrotnością jest jej macierz transponowana: $V^T V = VV^T = 1$. Dla macierzy transponowanej z definicji $(V^T)_{ij} = V_{ji}$.

Przez operator będziemy tu rozumieć obiekt, który działając na wektor z jakiejś przestrzeni daje ten sam lub inny wektor z tej samej przestrzeni. W szczególności operator $\hat{x}\hat{x} + \hat{y}\hat{y}$, którego działanie na wektory jest określone wzorami

$$\begin{aligned} (\hat{x}\hat{x} + \hat{y}\hat{y})\vec{a} &= \hat{x}(\hat{x} \cdot \vec{a}) + \hat{y}(\hat{y} \cdot \vec{a}), \\ \vec{a}(\hat{x}\hat{x} + \hat{y}\hat{y}) &= (\vec{a} \cdot \hat{x})\hat{x} + (\vec{a} \cdot \hat{y})\hat{y}, \end{aligned} \tag{5.4}$$

jest operatorem jednostkowym, bo działając prawostronnie, czy lewostronnie na dowolny wektor $\vec{a}$ daje ten sam wektor $\vec{a}$.

W mechanice kwantowej sytuacja jest podobna, ale pod kilkoma względami bardziej skomplikowana. Zacznijmy od wybrania w przestrzeni stanów jakiejś ortonormalnej bazy. Dla bazy dyskretnej ortonormalność oznacza, że

$$\langle \alpha_i | \alpha_j \rangle = \delta_{ij}, \tag{5.5}$$

gdzie wskaźniki całkowite i, j numerują wektory bazowe, a delta Kroneckera po prawej stronie jest z definicji równa jeden, jeśli $i = j$ i zero jeśli $i \neq j$. Napisanie odpowiedniego wzoru dla bazy ciągłej jest znacznie trudniejsze i zostanie omówione dopiero w następnym rozdziale.

Z definicji bazy, każdy wektor stanu układu da się zapisać jako (skończona lub nieskończona) kombinacja liniowa wektorów bazowych. Dla bazy dyskretnej

$$|\psi\rangle = \sum_i c_i^{(\alpha)} |\alpha_i\rangle, \tag{5.6}$$

gdzie sumowanie jest po wszystkich wartościach wskaźnika i. Dla bazy ciągłej

$$|\psi\rangle = \int \psi(\alpha) |\alpha\rangle d\alpha. \tag{5.7}$$

To jest całka określona, obszarem całkowania jest cały zakres zmienności wskaźnika α. W praktyce całkuje się zwykle po przestrzeni n-wymiarowej, gdzie n jest liczbą składowych wskaźnika α, lub po jakimś obszarze w tej przestrzeni. W

dalszym ciągu będziemy często stosowali ten skrócony sposób zapisywanie całek określonych.

Współczynniki $c_i^{(\alpha)}$ we wzorze (5.6) są naogół liczbami zespolonymi. Mnożąc tę równość obustronnie przez wektor dualny $\langle\alpha_k|$ i korzystając z warunku ortonormalności otrzymujemy wzór na współczynniki

$$c_k^{(\alpha)} = \langle\alpha_k|\psi\rangle. \tag{5.8}$$

Dla przypadku ciągłego powinno wyjść

$$\psi(\alpha) = \langle\alpha|\psi\rangle, \tag{5.9}$$

ale nie mając odpowiednika wzorów (5.5) nie potrafimy tego udowodnić.

Podstawiając wynik (5.8) do wzoru (5.6) otrzymujemy

$$|\psi\rangle = \sum_i |\alpha_i\rangle\langle\alpha_i|\psi\rangle. \tag{5.10}$$

Widać stąd, że operator

$$\sum_i |\alpha_i\rangle\langle\alpha_i| = \hat{1} \tag{5.11}$$

działając na dowolny wektor stanu $|\psi\rangle$ daje znów ten sam wektor. Jest to więc operator jednostkowy. Łatwo sprawdzić (ćwiczenie), że ten operator działając na dowolny wektor dualny też działa jak operator jednostkowy. Zobaczymy w dalszym ciągu, że znajomość tego operatora często się przydaje. Podobnie dla bazy ciągłej, podstawiając wzór (5.9) do wzoru (5.7), otrzymujemy wzór

$$|\psi\rangle = \int |\alpha\rangle\langle\alpha|\psi\rangle d\alpha \tag{5.12}$$

i stąd

$$\int |\alpha\rangle\langle\alpha| d\alpha = \hat{1}, \tag{5.13}$$

który znaczy, że działając operatorem po lewej stronie tej równości na dowolny wektor stanu $|\psi\rangle$ odtwarzamy ten sam wektor.

Jako zastosowanie wzoru (5.11) znajdziemy wzór łączący współczynniki $c_k^{(\alpha)}$ ze współczynnikami $c_i^{(\beta)} = \langle\beta_i|\psi\rangle$. Mnożąc przez operator jednostkowy wektor $|\psi\rangle$ we wzorze (5.8), co nie psuje równości, otrzymujemy

$$c_k^{(\alpha)} = \sum_i \langle\alpha_k|\beta_i\rangle\langle\beta_i|\psi\rangle. \tag{5.14}$$

Podobnie dla baz ciągłych, mnożąc obie strony równości (5.7) przez wektor bazowy $\langle\beta|$ z bazy β otrzymujemy wzór na zmianę funkcji falowej ze zmianą reprezentacji

$$\psi^{(\beta)}(\beta) = \int \langle\beta|\alpha\rangle\psi^{(\alpha)}(\alpha) d\alpha \tag{5.15}$$

Jeżeli zapiszemy wzór (5.14) w postaci macierzowej $C^{(\alpha)} = VC^{(\beta)}$, to na elementy macierzy V otrzymujemy wzór

$$V_{ki} = \langle \alpha_k | \beta_i \rangle. \tag{5.16}$$

Pokażemy teraz, że macierz V jest unitarna, to znaczy, że spełnia warunki $V^\dagger V = 1$ i $VV^\dagger = 1$, gdzie $V^\dagger$ oznacza macierz hermitowsko sprzężoną do macierzy V, to znaczy taką, że $(V^\dagger)_{ik} = V^*_{ki}$. Gwiazdka, jak zwykle, oznacza sprzężenie zespolone. Dla macierzy skończonych wystarczy udowodnić jeden z tych warunków. Drugi już z niego wynika. Dla macierzy nieskończonych jednak, może się zdarzyć, że jeden z warunków jest spełniony, a drugi nie. Trzeba więc sprawdzić oba. W naszym przypadku dowody obu warunków są bardzo podobne, więc udowodnimy tylko pierwszy, pozostawiając dowód drugiego jako ćwiczenie. Należy wykazać, że

$$(V^\dagger V)_{ik} = \sum_j (V^\dagger)_{ij} V_{jk} = \delta_{ik}, \tag{5.17}$$

gdzie pierwsza równość wynika z definicji iloczynu macierzy, a druga z definicji macierzy jednostkowej. Podstawiając wzór na elementy macierzy V_{ik} i wykorzystując definicję sprzężenia hermitowskiego otrzymujemy wyrażenie na iloczyn macierzy

$$\sum_j \langle \alpha_j | \beta_i \rangle^* \langle \alpha_j | \beta_k \rangle = \sum_j \langle \beta_i | \alpha_j \rangle \langle \alpha_j | \beta_k \rangle = \langle \beta_i | \beta_k \rangle = \delta_{ik}, \tag{5.18}$$

co było do udowodnienia. Pierwsza równość wynika z własności iloczynu skalarnego, druga z definicji operatora jednostkowego, a trzecia z ortonormalności bazy $|\beta_j\rangle$. Analogiczny wynik można wyprowadzić i dla wektorów uogólnionych, ale dowód znów wymaga znajomości odpowiednika wzoru (5.5).

Zastąpienie abstrakcyjnego wektora ciągiem zwykłych liczb jest często wygodne, ale należy rozumieć, że nie każdy ciąg liczb $c_i^{(\alpha)}$ reprezentuje wektor. Ponieważ zbiór wartości wskaźnika jest dyskretny, wchodzi w grę tylko wektor zwykły, a nie uogólniony. Wektor $|\psi\rangle$ musi mieć dobrze określony kwadrat długości $\langle \psi | \psi \rangle$. Mnożąc rozwinięcie (5.6) przez wektor do niego dualny

$$\langle \psi | = \sum_i c_i^{(\alpha)*} \langle \alpha_i | \tag{5.19}$$

i wykorzystując ortonormalność bazy $|\alpha_i\rangle$ otrzymujemy

$$\langle \psi | \psi \rangle = \sum_i |c_i^{(\alpha)}|^2. \tag{5.20}$$

Żeby ciąg liczb $c_i^{(\alpha)}$ przedstawiał wektor, szereg po prawej stronie musi być zbieżny. Podobnie dla przypadku ciągłego, zakładając wzór (5.13), otrzymujemy

$$\langle \psi | \psi \rangle = \int \langle \psi | \alpha \rangle \langle \alpha | \psi \rangle d\alpha = \int |\psi(\alpha)|^2 d\alpha. \tag{5.21}$$

Po to żeby funkcja falowa $\psi(\alpha)$ odpowiadała zwykłemu wektorowi stanu, a nie na przykład wektorowi uogólnionemu, całka z kwadratu jej modułu po całym obszarze zmienności wskaźnika α musi być zbieżna.

Jeżeli normalizacja wybrana jest tak, że suma szeregu (5.20) jest równa jeden, mamy prostą interpretację probabilistyczną współczynników

$$P_\psi(\alpha_i) = |c_i^{(\alpha_i)}|^2. \tag{5.22}$$

W przypadku ciągłym prawdopodobieństwo, że układ jest w stanie $|\alpha\rangle$ jest naogół równe zero. Inaczej nie dało by się spełnić warunku, że suma prawdopodobieństw wszystkich możliwych stanów $|\alpha\rangle$ jest równa jeden. Jeżeli jednak przyjmiemy, że funkcja falowa jest znormalizowana według wzoru

$$\int |\psi(\alpha)|^2 d\alpha = 1, \tag{5.23}$$

to można interpretować kwadrat modułu funkcji falowej jako gęstość prawdopodobieństwa. To znaczy, że jeśli V_α jest obszarem w przestrzeni α, to prawdopodobieństwo, że układ znajduje się w jakimś stanie $|\alpha\rangle$ należącym do V_α wynosi

$$P(V_\alpha) = \int_{V_\alpha} |\psi(\alpha)|^2 d\alpha. \tag{5.24}$$

W szczególności, jeśli parametr α jest jedną liczbą rzeczywistą, to

$$P(\alpha_0 < \alpha < \alpha_1) = \int_{\alpha_0}^{\alpha_1} |\psi(\alpha)|^2 d\alpha. \tag{5.25}$$

Na przykład, dla jednej cząstki, funkcja falowa $\psi(\vec{x}) \equiv \psi^{(x)}(\vec{x})$ daje rozkład prawdopodobieństwa znalezienia cząstki w zwykłej przestrzeni, a funkcja falowa $\phi(\vec{p}) \equiv \psi^{(p)}(\vec{p})$ daje rozkład prawdopodobieństwa znalezienia cząstki w przestrzeni pędów. Znalezienie rozkładu prawdopodobieństwa dla pędów $|\phi(\vec{p})|^2$, jeśli znana jest funkcja falowa $\psi(\vec{x})$, też jest możliwe, ale trzeba najpierw obliczyć funkcję falową w reprezentacji pędów $\phi(\vec{p})$, stosując wzór (5.15).

Rozdział 6

Funkcje falowe i wektory uogólnione

W poprzednim rozdziale pozostawiliśmy nierozwiązany problem, jak określić dla baz ciągłych relacje odpowiadające relacjom ortonormalności (5.5) dla baz dyskretnych. Określenie powinno być takie, żeby wynikała z niego relacja $\psi(\alpha) = \langle\alpha|\psi\rangle$, lub równoważnie (korzystamy ze wzoru na operator jednostkowy (5.13)) wzór

$$\psi(\alpha) = \int \delta(\alpha' - \alpha)\psi(\alpha')d\alpha', \tag{6.1}$$

gdzie wprowadziliśmy słynną deltę Diraca

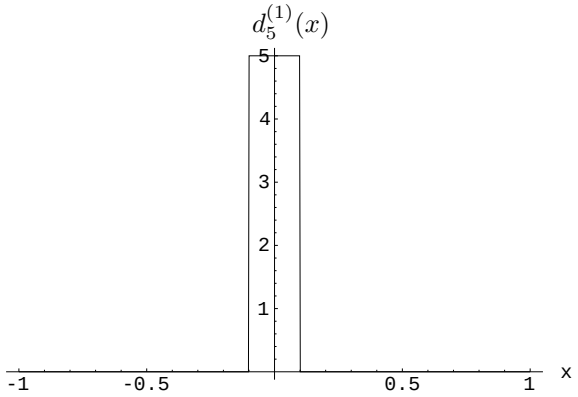
$$\delta(\alpha' - \alpha) = \langle\alpha|\alpha'\rangle. \tag{6.2}$$

Trudność polega na tym, że jak udowodniono nie ma takiej funkcji dwu zmiennych α i α', która podstawiona za $\langle\alpha|\alpha'\rangle$ spełniałaby równość (6.1) dla dowolnej funkcji $\psi(\alpha)$. Wymaganie, żeby równość zachodziła dla dowolnej funkcji $\psi(\alpha)$ jest za silne, ale przy rozsądnych założeniach, na przykład, że funkcja $\psi(\alpha)$ jest ciągła i całkowalna z kwadratem modułu, twierdzenie o niemożliwości pozostaje prawdziwe. Kiedy Dirac w latach dwudziestych wprowadził swoją deltę, którą zresztą niepoprawnie nazywał funkcją, powstała ciekawa sytuacja. Matematycy uważali, że rachunki z użyciem delty Diraca są błędne i pracowicie wykazywali, że te same wyniki można uzyskać inaczej. Fizycy stosowali w rachunkach deltę Diraca, bo dzięki temu rachunki stawały się znacznie krótsze i prostsze, a wyniki były takie same jak otrzymane żmudnymi metodami dopuszczanymi przez matematyków. Dopiero około roku 1950 wyjaśniło się że delta Diraca istnieje, ale jest dystrybucją, a nie funkcją.

Zauważmy najpierw, że n-wymiarową deltę Diraca można wyrazić przez delty jednowymiarowe

$$\delta^n(\alpha) = \prod_{i=1}^{n} \delta^1(\alpha_i) \tag{6.3}$$

gdzie α_i oznacza i-tą składowa wskaźnika α, a jednowymiarowa delta spełnia dla każdej ciągłej w otoczeniu zera funkcji $f(x)$ równanie



Rysunek 6.1: Funkcja $d_5^{(1)}(x)$ ze wzoru (6.5).

$$f(0) = \int_{-\infty}^{+\infty} dx f(x)\delta^1(x). \tag{6.4}$$

Problem sprowadza się więc do zdefiniowania dystrybucji $\delta^1(x)$ zależnej od jednej zmiennej rzeczywistej x. W dalszym ciągu będziemy zwykle opuszczali wskaźnik 1, bo z kontekstu będzie widać, kiedy mówimy o delcie jednowymiarowej, a kiedy o wielowymiarowej.

Dystrybucje można wprowadzić albo przez wzory całkowe takie jak równość (6.4), albo przez ciągi zwykłych funkcji tak, jak liczby niewymierne można określić jako granice ciągów liczb wymiernych. To drugie podejście wydaje się lepiej dostosowane do sytuacji, które spotykamy w fizyce. Dalsze rozważanie dotyczy delty jednowymiarowej. Coś trzeba też założyć o dopuszczalnych funkcjach $f(x)$. Przyjmujemy, że są ciągłe i że ich moduły (wartości bezwzględne) są ograniczone .

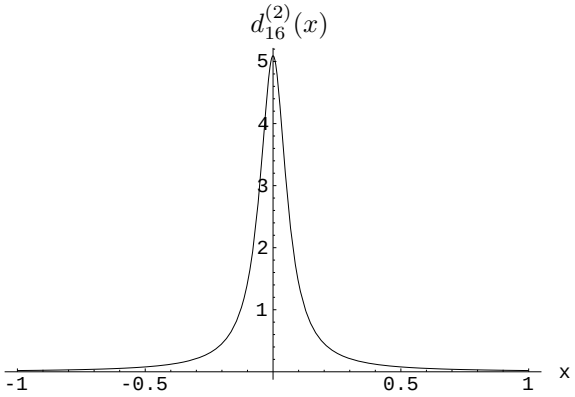
Rozważmy ciąg funkcji $d_n^{(1)}$ zdefiniowanych dla $n = 1, 2, \ldots$ następująco

$$d_n^{(1)}(x) = \begin{cases} n & \text{dla } |x| < \frac{1}{2n} \\ 0 & \text{dla } |x| \geq \frac{1}{2n} \end{cases} \tag{6.5}$$

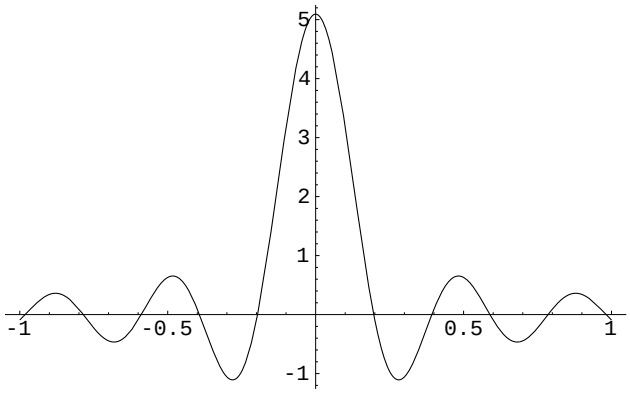
Dla $n = 5$, ta funkcja jest narysowana na rysunku 1. Ze wzrostem n słupek na wykresie staje się coraz wyższy i coraz węższy, ale pole pod nim jest zawsze równe jeden. Granica tego ciągu przy $n \to \infty$ nie istnieje jako funkcja, bo ciąg $d_n^{(1)}(0)$ jest rozbieżny do nieskończoności. Jeżeli jednak podstawimy funkcję $d_n^{(1)}(x)$ w miejsce delty Diraca w całce po prawej stronie wzoru (6.4), to granica całki przy $n \to \infty$ istnieje i jest równa $f(0)$ (ćwiczenie). W tym sensie, którego nie będziemy tu dalej uściślać, granica ciągu $d_n^{(1)}(x)$ istnieje, ale nie jest funkcją tylko dystrybucją, którą nazywamy deltą Diraca. Ta sama dystrybucja może być przedstawiona jako granica ciągu funkcji na wiele sposobów. Na przykład zamiast ciągu funkcji $d_n^{(1)}(x)$ można wybrać ciąg funkcji

$$d_n^{(2)}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{ikx - |k|/n} = \frac{1}{\pi} \frac{n}{n^2x^2 + 1}. \tag{6.6}$$

Ta funkcja, pokazana dla $n = 16$ na rysunku 2, jest jakościowo podobna do



Rysunek 6.2: Funkcja $d_{16}^{(2)}(x)$ ze wzoru (6.6).



Rysunek 6.3: Funkcja $d_{16}^{(3)}(x)$ ze wzoru (6.8).

funkcji $d_n^{(1)}(x)$. Przy wzrastającym n, maksimum przy $x = 0$ staje się coraz wyższe i coraz węższe, ale tak, że pole pod krzywą pozostaje równe jeden. Funkcja (6.6) podstawiona do wzoru (6.4) w miejsce delty Diraca, daje w granicy $n \to \infty$ poprawny wynik $f(0)$. Przechodząc z n do nieskończoności w funkcji podcałkowej we wzorze (6.6), otrzymuje się często przez fizyków używany wzór

$$\int_{-\infty}^{+\infty} dk e^{ikx} = 2\pi\delta^1(x). \tag{6.7}$$

Z punktu widzenia klasycznej definicji całki niewłaściwej ten wzór jest bez sensu, ale rozumiany dystrybucyjnie jest dobry i jak dalej zobaczymy bardzo pożyteczny. Jako kolejny przykład ciągu funkcji zbieżnego do delty Diraca rozważmy blisko związany z poprzednim ciąg

$$d_n^{(3)}(x) = \frac{1}{2\pi}\int_{-n}^{n} dk e^{ikx} = \frac{\sin nx}{\pi x}. \tag{6.8}$$

Ta funkcja dla n=16 jest pokazana na rysunku 3. Jak poprzednio w otoczeniu zera pojawia się maksimum, które ze wzrostem n staje się coraz wyższe i coraz węższe, ale dla $x \neq 0$ funkcja nie dąży do zera jak w poprzednich przypadkach.

Obszar x znacząco różnych od zera, który nie powinien dawać wkładu do całki we wzorze (6.4), jest wyeliminowany nie przez to, że funkcja $d_n^{(3)}(x)$ jest tam mała, jak to miało miejsce w poprzednich przykładach, ale przez to, że bardzo szybko oscyluje wokół zera. To też działa. Z tego ostatniego przykładu widać także, że spotykane niekiedy w starych podręcznikach sformułowanie, że delta Diraca ma wartość zero dla różnych od zera wartości argumentu jest błędne. Tak było w pierwszych dwu przykładach, ale w trzecim granica funkcji $d_n^{(3)}(x)$ przy ustalonym x i n dążącym do nieskończoności nie istnieje. Ponieważ dystrybucja delta jest określona jako wspólna granica wszystkich ciągów funkcyjnych, które podstawione do wzoru (6.4) dają w granicy równość prawdziwą, ten jeden kontrprzykład wystarcza, żeby stwierdzić, że wartość dystrybucji dla danej wartości argumentu jest naogół wielkością nieokreśloną. Natomiast, jak łatwo się przekonać choćby na podstawie rysunków 1-3, prawdziwe jest twierdzenie, że całka z iloczynu delty Diraca i dowolnej funkcji po obszarze, w którym argument delty nie przyjmuje wartości zero, znika. Dlatego twierdzenie, że $\delta(x \neq 0) = 0$, choć ściśle biorąc fałszywe, często daje dobry wynik.

Wracamy do ogólnej dyskusji przypadku n-wymiarowego. Ze wzoru (6.2) widać, że dla stanu $|\alpha\rangle$ długość zdefiniowana jako $\sqrt{\langle\alpha|\alpha\rangle}$ jest nieokreślona. Dlatego obiekty $|\alpha\rangle$ nie są zwykłymi wektorami, tylko wektorami uogólnionymi. Rozważmy, co z tego wynika. Przedstawienie zwykłego wektora stanu $|\psi\rangle$ w reprezentacji wektorów uogólnionych zostało opisane w poprzednim rozdziale i nie prowadzi do trudności. Twierdzenia podane tam bez dowodu staja się łatwe do udowodnienia przy zastosowaniu delty Diraca. Zastosowanie funkcji falowych do opisu stanów układów jest tak powszechne, że początkujący czasem nawet nie zdają sobie sprawy, że to nie jest jedyna możliwość. Używanie wektorów uogólnionych jako wektorów stanu układów fizycznych ma istotną niedogodność. Ponieważ długość takiego wektora jest nieokreślona, nie da się odpowiadającego mu rozkładu prawdopodobieństwa znormalizować do jedności. Z tego powodu, w niektórych problemach unika się wprowadzania stanów układu odpowiadających wektorom uogólnionym. Przykład takiej sytuacji spotkamy w rozdziale 29. Można jednak obliczać prawdopodobieństwa względne. Jeśli funkcja falowa $\psi(\alpha)$ odpowiada uogólnionemu wektorowi stanu $|\psi\rangle$, to stosunek prawdopodobieństwa, że α zawiera się w obszarze V_1 do prawdopodobieństwa, że zawiera się o obszarze V_2 wynosi

$$\frac{P(V_1)}{P(V_2)} = \frac{\int_{V_1} |\psi(\alpha)|^2 d\alpha}{\int_{V_2} |\psi(\alpha)|^2 d\alpha}. \tag{6.9}$$

W praktyce opis stanów za pomocą wektorów uogólnionych jest stosowany bardzo często. Ważny przykład przedyskutujemy w rozdziale 9.

To, czy stan układu będzie opisany przez wektor zwykły czy uogólniony, można często łatwo rozpoznać. Cząstka związana w polu jakiegoś potencjału prawie na pewno znajduje się niedaleko centrum tego potencjału, więc gęstość prawdopodobieństwa da się znormalizować i stan jest opisywany przez zwykły wektor stanu. Cząstka swobodna, która z tym samym prawdopodobieństwem może się znaleźć w każdym punkcie przestrzeni, ma gęstość prawdopodobieństwa stałą, a więc nie dającą się znormalizować, i wektor stanu jest wektorem uogólnionym.

W dalszych rozdziałach będziemy większość wzorów pisali dla baz dyskretnych. Przejście do bazy ciągłej jest zwykle łatwe i sprowadza się do zastąpienia

sumowań całkowaniami i delt Kroneckera deltami Diraca.

Rozdział 7

Kwantowanie

Przez kwantowanie rozumiemy przechodzenie od teorii klasycznej do teorii kwantowej. Nie jest znana żadna ogólna metoda kwantowania, za pomocą której można by skwantować dowolną teorię klasyczną. Nawet nie jest jasne, które teorie klasyczne należy kwantować. Trwają, na przykład, dyskusje czy i jak kwantować ogólną teorię względności. Pierwszą metodę kwantowania mechaniki którą do dziś dnia uważamy za dobrą, wymyślił Heisenberg, w oparciu o wcześniejsze idee Bohra, w lecie roku 1925. Na początku roku 1926 Schrödinger, zainspirowany wynikami de Broglie'a, podał swoją metodę kwantowania mechaniki, która wydawała się zupełnie inna od metody Heisenberga. Nawet nazwy były różne: mówiło się o mechanice kwantowej Heisenberga i o mechanice falowej Schrödingera. Wkrótce Dirac podał jeszcze inną metodę kwantowania, znacznie ogólniejszą, z której metody Heisenberga i Schrödingera wynikały jako szczególne przypadki. Te wyniki zostały szybko docenione. Heisenberg otrzymał nagrodę Nobla w roku 1932 "za stworzenie mechaniki kwantowej, której zastosowanie doprowadziło miedzy innymi do odkrycia alotropowych postaci wodoru", a nagrodę w roku 1933 przyznano za "odkrycie nowych produktywnych postaci teorii atomowej" i podzielono po połowie między Diraca i Schrödingera.

Metoda, którą tu omówimy, została wybrana ze względu na jej prostotę. W miarę potrzeby będziemy ją później uogólniać. Podstawowy pomysł polega na tym, że każdej wielkości mierzalnej przypisujemy operator liniowy i hermitowski. Niech wielkości mierzalnej A odpowiada operator $\hat{A}$. Tu i w dalszym ciągu daszek nad symbolem wskazuje, że symbol oznacza operator i że, co za tym idzie, można nim działać na wektor, zwykły lub dualny, i w wyniku otrzymuje się znów wektor, odpowiednio zwykły lub dualny. Rozwiązując problem własny tego operatora

$$\hat{A}|\alpha_n\rangle = a_n|\alpha_n\rangle, \tag{7.1}$$

otrzymujemy wszystkie możliwe wyniki dokładnego pomiaru wielkości mierzalnej A – to są wartości własne a_n – i odpowiadające im wektory stanu $|\alpha_n\rangle$. Stwierdzenie, że operator jest hermitowski, oznacza że jego wektory własne stanowią układ zupełny i że wszystkie jego wartości własne są rzeczywiste. Zauważmy, że hermitowskość operatora $\hat{A}$ jest bardzo naturalnym założeniem. Możliwe wyniki pomiaru a_n muszą być liczbami rzeczywistymi. Dla dowolnego stanu $|\psi\rangle$ powinny być określone amplitudy prawdopodobieństwa wyników pomiaru $a_1, a_2, \ldots$, czyli przedstawienie tego stanu jako kombinacji liniowej stanów

$|\alpha_1\rangle, |\alpha_2\rangle, \ldots$ (Rozdział 4), a to właśnie znaczy, że wektory własne stanowią układ zupełny.

Zapis (7.1) jest uproszczony. Jak zaraz zobaczymy na przykładzie, jednej wartości własnej a_n czy równoważnie jednej wartości wskaźnika n, może odpowiadać wiele różnych wektorów własnych. Dlatego wektor $|\alpha_n\rangle$ należy naogół rozumieć jako skrócony zapis wektora $|\alpha_n, \beta\rangle$, gdzie β zawiera dodatkowe informacje potrzebne do jednoznacznego określenia wektora odpowiadającego danej wartości własnej a_n. Wektory $|\alpha_n, \beta\rangle$, dla danego n, stanowią przestrzeń wektorową (ćwiczenie). W tej przestrzeni można wybrać bazę ortonormalną. Zespół wektorów bazowych wszystkich tych przestrzeni (jedna baza dla każdego n) stanowi bazę ortonormalną w przestrzeni wektorów stanu układu. Każdy inny wektor własny operatora $\hat{A}$ może być przedstawiony jako kombinacja liniowa tych wektorów bazowych. Wprowadzoną powyżej bazę będziemy nazywali bazą ortonormalną w przestrzeni wektorów własnych operatora $\hat{A}$. Z definicji

$$\langle \alpha'_n, \beta' | \alpha_n, \beta \rangle = \delta_{n,n'} \delta_{\beta,\beta'}. \tag{7.2}$$

Jeżeli wskaźniki n lub β przebiegają ciągły zbiór wartości, to odpowiednio należy deltę Kroneckera zastąpić deltą Diraca. Jeżeli wskaźnik β składa się z kilku liczb, to odpowiednią deltę Kroneckera czy Diraca należy zastąpić iloczynem delt.

Dla przypomnienia niektórych pojęć związanych z problemami własnymi, rozważmy szczególnie prosty przykład, kiedy operator $\hat{A}$ jest operatorem jednostkowym:

$$\hat{1}|\psi\rangle = \lambda|\psi\rangle. \tag{7.3}$$

To równanie ma dla dowolnego λ rozwiązanie $|\psi\rangle = 0$, ale takie, zerowe, rozwiązania nie liczą się jako rozwiązania problemu własnego. Dla $|\psi\rangle \neq 0$ rozwiązania istnieją tylko dla $\lambda = 1$. Mówimy, że jedyną wartością własną jest $\lambda = 1$. Dla $\lambda = 1$ każdy niezerowy wektor $|\psi\rangle$ jest rozwiązaniem, czyli wektorem własnym. Jest więc oczywiste, że te wektory stanowią układ zupełny i co za tym idzie, że operator jednostkowy jest hermitowski. Wektory własne różniące się od siebie tylko o stały czynnik, odpowiadają tej samej wartości własnej. Nie liczymy ich jako różnych rozwiązań. Jeżeli danej wartości własnej odpowiada $k > 1$ liniowo niezależnych wektorów własnych, mówimy że ta wartość własna jest k-krotnie zdegenerowana. W tym przykładzie wartość własna $\lambda = 1$ jest nieskończenie wielokrotnie zdegenerowana. Dwa dowolnie wybrane, liniowo niezależne wektory własne naogół nie są ortogonalne, ale da się wybrać w przestrzeni wektorów własnych operatora bazę ortogonalną.

Wracamy do mechaniki kwantowej. Narzucają się dwa pytania:

- Czy dla każdej wielkości mierzalnej A istnieje operator $\hat{A}$ spełniający równość (7.1)?
- Jak wyznaczyć ten operator?

Żeby odpowiedzieć na pierwsze pytanie rozważmy operator

$$\hat{A} = \sum_{i,\beta} |\alpha_i, \beta\rangle a_i \langle \alpha_i, \beta|, \tag{7.4}$$

gdzie sumowanie przebiega po wszystkich wektorach z bazy ortonormalnej przestrzeni wektorów własnych operatora $\hat{A}$. Wartościami własnymi tego operatora są liczby a_i, którym odpowiadają wektory własne $|\alpha_i, \beta\rangle$, bo

$$\hat{A}|\alpha_k,\beta\rangle = \sum_{i,\beta'} |\alpha_i,\beta'\rangle a_i \langle\alpha_i,\beta'|\alpha_k,\beta\rangle = a_k|\alpha_k,\beta\rangle. \tag{7.5}$$

Widać stąd, że operator $\hat{A}$ istnieje, jakikolwiek byłby zbiór (rzeczywistych!) możliwych wyników pomiaru wielkości mierzalnej A i zbiór (zupełny!) odpowiadających im wektorów własnych.

Powyższa konstrukcja wymaga znajomości wszystkich możliwych wyników pomiaru wielkości A i odpowiadających im wektorów stanu. Nie jest więc użyteczna, jeśli tych wielkości dopiero szukamy. Bardzo użyteczna, choć nie całkiem ogólna, jest natomiast następująca metoda zaproponowana przez Schrödingera dla reprezentacji położeń. Ta metoda stosuje się, choć nie zawsze, do wielkości mierzalnych $O(\vec{p},\vec{x})$, które w fizyce klasycznej są funkcjami pędów i położeń. Przyjmujemy, że

$$\langle\vec{x}|\hat{O}|\psi\rangle = \hat{O}(\vec{x})\psi(\vec{x}). \tag{7.6}$$

Zauważmy, że wstawiając w elemencie macierzowym po lewej stronie operator jednostkowy zbudowany z wektorów $|\vec{x'}\rangle$ i korzystając z definicji funkcji falowej otrzymujemy bez dalszych założeń

$$\langle\vec{x}|\hat{O}|\psi\rangle = \int dx' \langle\vec{x}|\hat{O}|\vec{x'}\rangle\psi(x'). \tag{7.7}$$

Jeżeli elementy macierzowe operatora $\hat{O}$ dadzą się zapisać w postaci

$$\langle\vec{x}|\hat{O}|\vec{x'}\rangle = \delta(\vec{x}-\vec{x'})\hat{O}(\vec{x}), \tag{7.8}$$

to wynika stąd wzór (7.6). Operatory spełniające ten warunek nazywa się lokalnymi. Kwantowanie, które tu opisujemy, prowadzi do operatorów lokalnych, pozostaje więc znaleźć operator $\hat{O}(\vec{x})$. Służy do tego następująca recepta[1].

- Tam gdzie we wzorze klasycznym występuje składowa wektora położenia cząstki x_i wstawić liczbę x_i. Tę liczbę możemy interpretować jako operator, którego działanie na funkcje falową polega na pomnożeniu tej funkcji przez x_i.
- Tam gdzie we wzorze klasycznym występuje składowa wektora pędu cząstki p_i wstawić operator $-i\hbar\frac{\partial}{\partial x_i}$. Przyjęte jest oznaczenie $\vec{\nabla}$ na wektor nabla, którego składowymi są pochodne $\frac{\partial}{\partial x_i}$, więc pisze się często

$$\hat{\vec{p}} = -i\hbar\vec{\nabla} \tag{7.9}$$

Zauważmy, że powyższe podstawienia nie miałyby sensu, gdyby je stosować do operatora $\hat{O}$ po lewej stronie wzoru (7.6). Widać to choćby stąd, że wektory $|\psi\rangle$, na które ten operator działa, nie zależą naogół od położenia $\vec{x}$. Zobaczmy teraz na przykładach, jak się stosuje ten przepis.

Przykład 1. Żeby się dowiedzieć jakie wartości może przyjmować x-owa składowa pędu cząstki i co stąd wynika dla funkcji falowych, należy rozwiązać problem własny

[1] Ogólniej, można wybrać dowolne współrzędne q_i – niekoniecznie kartezjańskie – i stosować opisaną poniżej metodę do nich i do sprzężonych z nimi kanonicznie pędów p_i. Przykłady zastosowania tej ogólniejszej recepty można znaleźć w rozdziałach 11, 16 i 31.

$$-i\hbar\frac{\partial\psi_{p_x}(x)}{\partial x_i} = p_x\psi_{p_x}(x). \tag{7.10}$$

Funkcja falowa $\psi_{p_x}(x)$ może zależeć od innych jeszcze zmiennych, ale pochodna cząstkowa po zmiennej x nie działa na nie. Zbiór liczb p_x, dla których rozwiązanie powyższego problemu własnego istnieje, jest zbiorem możliwych wyników pomiaru x-owej składowej wektora pędu. Funkcja falowa $\psi_{p_x}(x)$ zależy od wartości własnej p_x, ale może też zależeć od innych wskaźników. Przedyskutujemy ten przykład dokładniej w rozdziale dziewiątym.

Przykład 2. Klasyczny hamiltonian (energia) jednowymiarowego oscylatora harmonicznego dany jest wzorem

$$H = \frac{p_x^2}{2m} + \frac{K}{2}x^2, \tag{7.11}$$

gdzie m i K oznaczają odpowiednio masę cząstki i stałą siłową oscylatora. Otrzymujemy stąd operator

$$\hat{H}(\vec{x}) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{K}{2}x^2 \tag{7.12}$$

i problem własny

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi_n(x)}{\partial x^2} + \frac{K}{2}x^2\psi_n(x) = E_n\psi_n(x). \tag{7.13}$$

Z interpretacji fizycznej wynika, że mamy to do czynienia ze stanem związanym, bo cząstka nie powinna się znajdować bardzo daleko punktu $x = 0$. Przyjmujemy więc warunek na rozwiązanie, że funkcje $\psi_n(x)$ mają dążyć na tyle szybko do zera przy x dążącym zarówno do minus jak i do plus nieskończoności, żeby całka $\int|\psi_n(x)|^2dx$ była zbieżna[2] . To jest ważny punkt. Oprócz znalezienia równania trzeba jeszcze nałożyć odpowiednie warunki na dopuszczalne rozwiązania. Wybór tych warunków jest podyktowany przez sytuacje fizyczną, którą chcemy opisać.

Rozwiązanie powyższego problemu własnego jest dość trudnym problemem z matematyki, podamy więc tylko wynik końcowy. Rozwiązania istnieją tylko dla

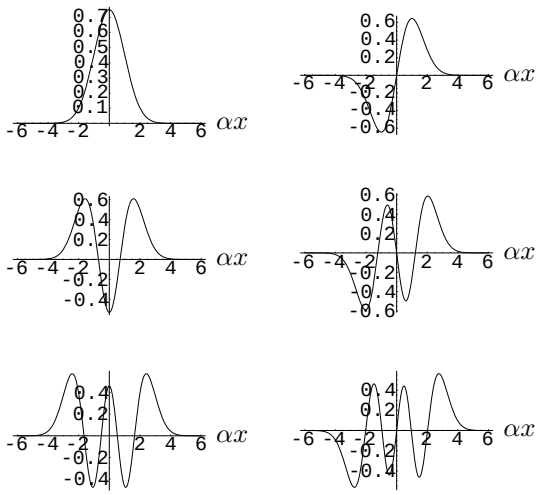
$$E_n = \hbar\omega(n + \frac{1}{2}), \qquad n = 0, 1, \ldots, \tag{7.14}$$

gdzie $\omega = \sqrt{\frac{K}{m}}$. Odtwarza się więc wynik Plancka opisany w rozdziale pierwszym. Każdej wartości własnej odpowiada jedna tylko funkcja własna (funkcji różniących się o stały czynnik nie uważamy za różne):

$$\psi_n(x) = \sqrt{\frac{\alpha}{\sqrt{\pi}2^n n!}}H_n(\alpha x)e^{-\frac{1}{2}\alpha^2x^2}. \tag{7.15}$$

W tym wzorze $\alpha = \sqrt[4]{\frac{mK}{\hbar^2}}$, a $H_n(x)$ oznaczają wielomiany Hermite'a, które można znaleźć w tablicach, albo za pomocą programów komputerowych. Stały

[2]Ta całka mogłaby być zbieżna także, jeśli funkcja $\psi(x)$ nie dąży do zera przy $x \to \pm\infty$. Dla stanów związanych odrzucamy takie rozwiązania, jeśli się pojawią, jako niefizyczne



Rysunek 7.1: Funkcje falowe jednowymiarowego oscylatora harmonicznego (7.15) (podzielone przez $\sqrt{\alpha}$) dla $n = 0, 1, \ldots, 5$. Funkcja $\psi_n(x)$ przecina oś αx w n miejscach

współczynnik jest dobrany tak, żeby funkcje ψ_n były znormalizowane do jedności. Zauważmy, że funkcje $\frac{1}{\sqrt{\alpha}}\psi_n(\alpha x)$ przy danym n zależą tylko od bezwymiarowego parametru αx. Ich wykresy dla $n = 0, 1, \ldots, 5$ są podane na rysunku 1.

Problem własny rozpatrywany w tym przykładzie jest szczególnym przypadkiem równania Schrödingera niezależnego od czasu, które będzie omawiane w rozdziale dziesiątym. W mechanice kwantowej, podobnie jak w mechanice klasycznej, tylko bardzo nieliczne problemy da się dokładnie rozwiązać analitycznie. Te rozwiązania są jednak bardzo cenne jako podstawa różnych prostych modeli i jako punkt wyjścia dla metod przybliżonych.

Przedstawimy teraz ważne ograniczenie opisanej w tym rozdziale metody kwantowania mechaniki. Przypuśćmy, że chcemy znaleźć operator odpowiadający wielkości klasycznej – iloczynowi $p_x x$. Stosując naszą receptę moglibyśmy otrzymać

$$\widehat{p_x x}\psi(x) = -i\hbar\frac{\partial}{\partial x}\left(x\psi(x)\right) = -i\hbar\left(\psi(x) + x\frac{\partial\psi(x)}{\partial x}\right). \tag{7.16}$$

Pisząc to samo wyrażenie klasyczne w postaci xp_x otrzymalibyśmy

$$\widehat{xp_x}\psi(x) = -i\hbar x\frac{\partial\psi(x)}{\partial x}, \tag{7.17}$$

a więc inny wynik. Opisana metoda kwantowania jest w tym przypadku niejednoznaczna i nie nadaje się do zastosowania. Komutator dwu operatorów, powiedzmy $\hat{A}$ i $\hat{B}$, jest zdefiniowany wzorem

$$\left[\hat{A}, \hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{7.18}$$

Jeżeli dla dwu operatorów ich komutator jest równy zeru, mówimy że te operatory komutują ze sobą. Dla operatorów $\hat{x}$ i $\hat{p}_x$ znaleźliśmy

$$[\hat{x}, \hat{p_x}]\,\psi(x) = i\hbar\psi(x). \tag{7.19}$$

Ponieważ ta równość zachodzi dla każdej funkcji $\psi(x)$, możemy ją zapisać jako równość operatorów[3]

$$[\hat{x}, \hat{p}_x] = i\hbar. \tag{7.20}$$

Operatory $\hat{p}_x$ i $\hat{x}$ nie komutują między sobą. Łatwo sprawdzić, że każdy operator komutuje sam ze sobą i że komutują ze sobą zarówno operatory odpowiadające różnym składowym wektora położenia jak i operatory odpowiadające różnym składowym wektora pędu. Metoda kwantowania opisana w tym rozdziale stosuje się tylko wtedy, kiedy kwantowana wielkość nie zawiera iloczynów wielkości, którym odpowiadają nie komutujące ze sobą operatory. Stosuje się więc w szczególności do hamiltonianów będących sumą energii kinetycznej, zależnej tylko od pędów, i energii potencjalnej, zależnej tylko od położeń cząstek. To jest bardzo szeroka klasa problemów, obejmująca między innymi fizykę atomów cząsteczek i jąder atomowych.

Zwracamy uwagę, że podając receptę na kwantowanie nie określaliśmy liczby wymiarów przestrzeni . Tę samą metodę kwantowania można więc stosować do trzech składowych pędu jednej cząstki, jak i do trzydziestu składowych pędów dziesięciu cząstek.

[3]Dwa operatory, które działając na każdy wektor dają ten sam wynik, są z definicji identyczne

Rozdział 8

Wartości średnie

Jeżeli wielkość mierzalna A może przyjmować wartości $a_1, a_2, \ldots$ z prawdopodobieństwami odpowiednio $P_1, P_2, \ldots$ to z definicji wartość średnia A:

$$\bar{A} = \sum_i P_i a_i. \tag{8.1}$$

Żeby posłużyć się tym wzorem, trzeba znać wszystkie możliwe wyniki pomiaru a_i i wszystkie prawdopodobieństwa P_i oraz jeszcze zwykle wykonać nieskończone sumowanie lub całkowanie. Dlatego ważne jest, że w mechanice kwantowej istnieje jeszcze inny, zwykle znacznie prostszy, sposób obliczania wartości średnich.

Twierdzenie: Niech $\bar{A}$ oznacza wartość średnią wielkości mierzalnej A w stanie $|\psi\rangle$, a $\hat{A}$ odpowiadający tej wielkości operator kwantowo-mechaniczny. Wtedy

$$\bar{A} = \langle\psi|\hat{A}|\psi\rangle. \tag{8.2}$$

Wstawiając operatory jednostkowe i korzystając z definicji funkcji falowej można ten wzór przepisać w postaci

$$\bar{A} = \int dx \int dx' \langle\psi|\vec{x}\rangle\langle\vec{x}|\hat{A}|\vec{x'}\rangle\langle\vec{x'}|\psi\rangle = \int dx \int dx' \psi^*(x)\langle\vec{x}|\hat{A}|\vec{x'}\rangle\psi(\vec{x'}). \tag{8.3}$$

Całkowania dx i dx' oznaczają, jak zwykle, całkowania po wszystkich składowych wektora położenia i, w razie potrzeby, jeszcze sumowania po zmiennych dyskretnych. Zwykle będziemy mieli do czynienia z operatorami "lokalnymi", których elementy macierzowe dadzą się zapisać w postaci

$$\langle\vec{x}|\hat{A}|\vec{x'}\rangle = \delta(\vec{x} - \vec{x'})\hat{A}(\vec{x}). \tag{8.4}$$

Dla takich operatorów całkowanie po dx' można wykonać dzięki delcie Diraca i otrzymujemy prostszy wzór

$$\bar{A} = \int dx \psi^*(x)\hat{A}(x)\psi(x). \tag{8.5}$$

Przykład: dla stanu podstawowego jednowymiarowego oscylatora harmonicznego funkcja falowa $\psi_0(x) = \sqrt{\frac{\alpha}{\sqrt{\pi}}} e^{-\frac{1}{2}\alpha^2 x^2}$, gdzie $\alpha^2 = \frac{\sqrt{mK}}{\hbar}$. Stąd

$$
\begin{aligned}
\bar{x} &= \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} dx x e^{-\alpha^2 x^2} = 0 \\
\bar{p}_x &= \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} dx e^{-\frac{1}{2}\alpha^2 x^2} \left(-i\hbar \frac{\partial}{\partial x} \right) e^{-\frac{1}{2}\alpha^2 x^2} = 0 \\
\bar{x}^2 &= \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} dx x^2 e^{-\alpha^2 x^2} = \frac{1}{2\alpha^2} \\
\bar{p}_x^2 &= \frac{\alpha}{\sqrt{\pi}} \int_{-\infty}^{+\infty} dx e^{-\frac{1}{2}\alpha^2 x^2} \left(-\hbar^2 \frac{\partial^2}{\partial x^2} \right) e^{-\frac{1}{2}\alpha^2 x^2} = \frac{\hbar^2 \alpha^2}{2}
\end{aligned} \tag{8.6}
$$

zgodnie z wynikami cytowanymi w rozdziale drugim.

Dowód twierdzenia (8.2) jest bardzo prosty. Niech wektory $|\alpha_i\rangle$ będą wektorami własnymi operatora $\hat{A}$ odpowiadającymi odpowiednio wartościom własnym a_i i niech znormalizowany do jedności wektor stanu $|\psi\rangle$ ma rozwinięcie

$$
|\psi\rangle = \sum_i c_i |\alpha_i\rangle. \tag{8.7}
$$

Na mocy interpretacji probabilistycznej mechaniki kwantowej i definicji wartości średniej

$$
\bar{A} = \sum_i |c_i|^2 a_i. \tag{8.8}
$$

Z drugiej strony, $c_i = \langle \alpha_i | \psi \rangle$, $c_i^* = \langle \psi | \alpha_i \rangle$ i $\langle \alpha_j | \hat{A} | \alpha_i \rangle = a_i \langle \alpha_j | \alpha_i \rangle = a_i \delta_{ij}$. Tak więc

$$
\bar{A} = \sum_{i,j} \langle \psi | \alpha_i \rangle \langle \alpha_i | \hat{A} | \alpha_j \rangle \langle \alpha_j | \psi \rangle. \tag{8.9}
$$

Człony $i = j$ z tej sumy dają prawą stronę równości (8.8), a człony $i \neq j$ są równe zeru. Usuwając dwa czynniki $\hat{1}$ z tego wzoru otrzymujemy tezę (8.2).

Podamy teraz bez dowodu ważne twierdzenie o wartościach średnich w mechanice kwantowej. Niech $\sigma_\psi^2(O)$ oznacza wariancję wielkości O w stanie $|\psi\rangle$ czyli $\langle \psi | \hat{O}^2 | \psi \rangle - \langle \psi | \hat{O} | \psi \rangle^2$. Jeżeli operatory $\hat{A}$ i $\hat{B}$ są hermitowskie oraz $|\psi\rangle$, $\hat{A}|\psi\rangle$ i $\hat{B}|\psi\rangle$ są wektorami (nie wektorami uogólnionymi!), to zachodzi nierówność

$$
\sigma_\psi^2(A) \sigma_\psi^2(B) \geq \frac{1}{4} |\langle \psi | \left[\hat{A}, \hat{B} \right] | \psi \rangle|^2, \tag{8.10}
$$

gdzie $[\cdots]$ oznacza komutator. To jest jedna z postaci słynnej zasady nieokreśloności Heisenberga. W szczególności, podstawiając wzór z poprzedniego rozdziału $[\hat{x}, \hat{p}_x] = i\hbar$, otrzymujemy najsłynniejszy przykład tej zasady

$$
\sigma_\psi^2(x) \sigma_\psi^2(p_x) \geq \frac{1}{4} \hbar^2. \tag{8.11}
$$

Analogiczne wzory stosują się do pozostałych składowych wektorów położenia i pędu. Można natomiast jednocześnie dokładnie zmierzyć, na przykład, x i p_y.

Przykład Zasada nieokreśloności Heisenberga pozwala prosto zrozumieć, dlaczego elektron w atomie nie spada na jądro. Wyjaśnimy to na prostszym przykładzie jednowymiarowego oscylatora harmonicznego o masie m i stałej

siłowej K, z $\langle x\rangle = 0$ i $\langle p_x\rangle = 0$. Dla takiego oscylatora, $\langle x^2\rangle = \sigma^2(x)$ i $\langle p_x^2\rangle = \sigma^2(p_x)$. Wobec tego energia

$$E = \frac{\sigma^2(p_x)}{2m} + \frac{1}{2}K\sigma^2(x) \geq \frac{\hbar^2}{8m\sigma^2(x)} + \frac{1}{2}K\sigma^2(x) \equiv E_{min}(\sigma^2(x)), \qquad (8.12)$$

gdzie nierówność wynika z zasady Heisenberga. Ograniczenie od dołu E_{min} robi się bardzo duże zarówno przy spadaniu cząstki na punkt przyciągania ($\sigma^2(x) \to 0$) jak i przy próbie unieruchomienia cząstki ($\sigma^2(p_x) \to 0$ czyli $\sigma^2(x) \to \infty$). Minimum ograniczenia E_{min} jest osiągane dla $\sigma^2(x) = \frac{\hbar^2}{2\sqrt{mK}}$ i wynosi $\frac{1}{2}\hbar\omega$, co (przypadkowo!) zgadza się dokładnie z energią stanu podstawowego oscylatora.

W mechanice klasycznej czasem określa się stan cząstki dokładnie, na przykład podając jej pęd i położenie, a czasem podaje się tylko rozkład prawdopodobieństwa dla pędów i położeń cząstki. Na przykład w fizyce statystycznej, przy odpowiednich założeniach, dowodzi się, że gęstość prawdopodobieństwa znalezienia cząstki o pędzie $\vec{p}$ w punkcie $\vec{x}$ wynosi

$$\rho(\vec{p}, \vec{x}) = \frac{1}{Z}e^{-\frac{E(\vec{p},\vec{x})}{k_B T}}. \qquad (8.13)$$

To jest słynny rozkład Boltzmanna: $E(\vec{p}, \vec{x})$ oznacza energię cząstki o pędzie $\vec{p}$ i położeniu $\vec{x}$, $k_B T$ jest iloczynem stałej Boltzmanna i temperatury bezwzględnej, a $\frac{1}{Z}$ jest współczynnikiem normalizacyjnym.

Podobne rozróżnienie istnieje i w mechanice kwantowej. Stan opisany przez wektor stanu czy funkcję falową nazywa się **stanem czystym**. Poza tym rozpatruje się stany mieszane określone jak następuje. W przestrzeni wektorów stanu układu wybieramy bazę $|\psi_1\rangle, |\psi_2\rangle, \ldots$ Gdybyśmy powiedzieli, że układ znajduje się w stanie $|\psi_{10}\rangle$, czy w stanie $\sum_i c_i|\psi_i\rangle$, to określilibyśmy stan czysty w którym znajduje się układ. Zamiast tego mówimy, że układ znajduje się z prawdopodobieństwem p_1 w stanie $|\psi_1\rangle$, z prawdopodobieństwem p_2 w stanie $|\psi_2\rangle$ itd. Ważne jest, że między tymi możliwościami nie ma interferencji. Tak więc stan czysty $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ naogół zmienia się, jeśli wektorowi $|\psi_2\rangle$ zmienimy fazę, a pozornie podobny stan mieszany z prawdopodobieństwami stanów $|\psi_1\rangle, |\psi_2\rangle$ odpowiednio $|c_1|^2$ i $|c_2|^2$ nie. W ten sposób podajemy określenie stanu układu analogiczne do podania rozkładu prawdopodobieństwa dla pędów i położeń w mechanice klasycznej. Tak określony stan w mechanice kwantowej nazywamy **stanem mieszanym**, w odróżnieniu od stanów czystych.

Zrozumienie różnicy między stanami czystymi i mieszanymi może ułatwić wyobrażenie sobie odpowiednich zespołów. Dla stanu czystego $|\psi\rangle$ wszystkie elementy zespołu są w tym samym stanie $|\psi\rangle$. Dla stanu mieszanego, niektóre elementy są w stanie $|\psi_1\rangle$, inne w $|\psi_2\rangle$ itd. przy czym w stanie $|\psi_1\rangle$ jest ułamek p_1 całkowitej liczby elementów, w stanie $|\psi_2\rangle$ ułamek p_2 itd.

Stan mieszany można określić krócej przez podanie **operatora gęstości**

$$\hat{\rho} = \sum_i |\psi_i\rangle p_i \langle\psi_i|. \qquad (8.14)$$

Rozumując jak dla operatorów przypisanych wielkościom mierzalnym, łatwo sprawdzić, że prawdopodobieństwa p_i są wartościami własnymi tego operatora, a stany $|\psi_i\rangle$ są odpowiadającymi im wektorami własnymi. Operator gęstości zawiera więc pełną informację o stanie mieszanym.

Zamiast operatora gęstości można też używać jego macierzowych reprezentacji. W dowolnej bazie $|\alpha_i\rangle$ elementy macierzy gęstości zdefiniowane są wzorem analogicznym do wzorów określające reprezentacje macierzowe operatorów odpowiadających wielkościom mierzalnym

$$\rho_{ij}^{(\alpha)} = \langle \alpha_i | \hat{\rho} | \alpha_j \rangle. \tag{8.15}$$

Przykład: Rozkład Boltzmanna opisuje stan mieszany. Wybierając jako wektory bazowe $|\psi_n\rangle$ wektory stanów o określonej energii $|E_i, \beta\rangle$, gdzie β oznacza pozostałe parametry potrzebne do jednoznacznego określenia stanu o danej energii, mamy

$$p(E_i, \beta) = \frac{1}{Z} e^{-\frac{E_i}{k_B T}}. \tag{8.16}$$

Macierz gęstości w reprezentacji stanów $|E_i, \beta\rangle$ ma postać

$$\langle E_j, \beta' | \hat{\rho} | E_i, \beta \rangle = \frac{1}{Z} e^{-\frac{E_i}{k_B T}} \delta_{i,j} \delta_{\beta,\beta'}. \tag{8.17}$$

Łatwo sprawdzić (ćwiczenie), że są to elementy macierzowe operatora

$$\hat{\rho} = \frac{1}{Z} e^{-\frac{\hat{H}}{k_B T}}, \tag{8.18}$$

gdzie $\hat{H}$ jest operatorem energii, dla którego $\hat{H}|E_i, \beta\rangle = E_i |E_i, \beta\rangle$.

Twierdzenie: W stanie mieszanym wartości średnie operatorów odpowiadających wielkościom mierzalnym dane są wzorem:

$$\bar{A} = Tr\left(\hat{\rho}\hat{A}\right). \tag{8.19}$$

Ślad (Tr) operatora jest zdefiniowany jako suma wszystkich elementów diagonalnych macierzy reprezentującej ten operator. Dowodzi się, że wartość śladu nie zależy od wyboru reprezentacji.

Dla dowodu twierdzenia wystarczy wykazać, że w jakiejś bazie $|\alpha_i\rangle$ zachodzi

$$\bar{A} = \sum_{i,j} \langle \alpha_i | \hat{\rho} | \alpha_j \rangle \langle \alpha_j | \hat{A} | \alpha_i \rangle, \tag{8.20}$$

Wybierzmy jako bazę $|\alpha_i\rangle$, bazę złożoną z wektorów własnych operatora gęstości. W tej bazie

$$\langle \alpha_i | \hat{\rho} | \alpha_j \rangle = p_i \delta_{ij}, \tag{8.21}$$

więc wzór poprzedni przyjmuje postać

$$\bar{A} = \sum_i p_i \langle \alpha_i | \hat{A} | \alpha_i \rangle. \tag{8.22}$$

Zgodnie z udowodnionym powyżej twierdzeniem o średniej dla stanów czystych, ten wzór jest prawdziwy, bo prawa strona jest średnią po stanach czystych i ze średniej w danym stanie czystym i.

Rozdział 9

Pęd cząstki w mechanice kwantowej

W tym rozdziale omówimy kwantowanie pędu w mechanice kwantowej. Zaczniemy od przestrzeni jednowymiarowej. Pęd cząstki (punktu materialnego) ma więc jedną tylko składową, którą w klasycznej teorii oznaczymy p. Należy określić według mechaniki kwantowej, jakie wartości może dać dokładny pomiar p i jakie wektory stanu opisują układ, dla którego wynik pomiary pędu da na pewno wynik p. Zamiast wektorów stanu można, równoważnie, podać jakąś ich reprezentację.

Jak widzieliśmy w rozdziale 7 (przykład 1) możliwe wyniki pomiaru p i odpowiadające im funkcje falowe w reprezentacji położeń $\psi_p(x)$ występują w problemie własnym

$$-i\hbar\frac{\partial\psi_p(x)}{\partial x} = p\psi_p(x). \tag{9.1}$$

Pochodna cząstkowa oznacza, że różniczkowanie po x przebiega przy ustalonych wartościach wszystkich innych zmiennych, od których mogła by jeszcze zależeć funkcja falowa $\psi_p(x)$. Na rozwiązanie nakładamy warunek, że ma być wektorem dającym się znormalizować do jedności lub wektorem uogólnionym dającym się znormalizować do delty Diraca. Ta druga możliwość pojawia się, bo, jak zobaczymy, stan cząstki o określonym pędzie nie jest stanem związanym.

Rozwiązanie równania (9.1), jak łatwo sprawdzić przez podstawienie, istnieje dla każdej liczby rzeczywistej p i ma postać

$$\psi_p(x) = Ce^{i\frac{px}{\hbar}}, \tag{9.2}$$

gdzie C jest dowolną stałą różną od zera. Na podstawie równania, współczynnik C nie zależy od x. Jeśli x jest jedyną zmienną w problemie, to C jest stałą liczbową. W ogólnym przypadku C może zależeć od dowolnych zmiennych niezależnych od x. Dla zespolonych wartości parametru p funkcja $\psi_p(x)$ też spełnia równanie, ale rośnie wykładniczo przy $x \to -\infty$, lub przy $x \to +\infty$, więc nie da się znormalizować. Potwierdza to wybór warunku nałożonego na rozwiązanie – ten warunek wyklucza zespolone wartości parametru p, które ani nigdy nie są otrzymywane jako wynik dokładnego pomiaru pędu, ani nie mogą

być wartościami własnymi operatora hermitowskiego. Ponieważ kwadrat modułu funkcji $\psi_p(x)$ jest równy stałej $|C|^2$, nie da się tej funkcji znormalizować przez warunek $\int_{-\infty}^{+\infty}|\psi(x)|^2dx = 1$. Nie ma w tym zresztą nic dziwnego, bo nie założyliśmy, że na cząstkę działa jakaś siła, która by ją zmuszała do trzymania się w niewielkiej odległości od wybranego punktu w przestrzeni. Wektor stanu $|p\rangle$ jest więc wektorem uogólnionym i należy go znormalizować do delty Diraca. Należy tu zwrócić uwagę na różnicę w stosunku do cząstki o ustalonej energii w potencjale oscylatora harmonicznego, której funkcje falowe można było znormalizować do delty Kroneckera, i wobec tego odpowiadające im wektory stanu były zwykłymi (nie uogólnionymi) wektorami .

Warunek normalizacji przyjmujemy jak w rozdziale szóstym. Powinno być

$$\langle\psi_p|\psi_{p'}\rangle = \delta(p-p'). \tag{9.3}$$

Wstawiając po lewej stronie operator jednostkowy utworzony z wektorów $|x\rangle$ i korzystając z definicji funkcji falowej i ze wzoru (9.2) otrzymujemy

$$\langle\psi_p|\psi_{p'}\rangle = \int dx\langle\psi_p|x\rangle\langle x|\psi_{p'}\rangle = \int dx\psi_p^*(x)\psi_{p'}(x) = |C|^2\hbar\int d\left(\frac{x}{\hbar}\right)e^{\frac{i}{\hbar}(p'-p)x}. \tag{9.4}$$

Dzięki pomnożeniu i podzieleniu ostatniego wyrażenia przez $\hbar$ możemy bezpośrednio zastosować wzór (6.7), co daje

$$\langle\psi_p|\psi_{p'}\rangle = 2\pi\hbar|C|^2\delta(p-p'). \tag{9.5}$$

Przyrównując do jedynki współczynnik przy delcie Diraca wyliczamy stałą $|C|$ i otrzymujemy

$$\psi_p(x) = \frac{1}{\sqrt{2\pi\hbar}}e^{i\frac{px}{\hbar}}. \tag{9.6}$$

Warunek normalizacyjny daje tylko wartość bezwzględną współczynnika C. Dla prostoty przyjęliśmy konwencję, że ten współczynnik jest rzeczywisty i dodatni. Jego stała faza nie wpływa na przewidywania fizyczne.

Podsumujmy wyniki. Wynikiem dokładnego pomiaru pędu w przypadku jednowymiarowym może być dowolna liczba rzeczywista. Dla danej wartości pędu p, funkcja falowa w reprezentacji położeń opisuje falę płaską z wektorem falowym k związanym z pędem p wzorem de Broglie'a $p = \hbar k$.

Uogólnienie powyższych wyników na układ d-wymiarowy jest bardzo proste. Zamiast jednego równania własnego mamy układ d równań, które możemy zapisać jako jedno równanie wektorowe:

$$-i\hbar\vec{\nabla}\psi_{\vec{p}}(\vec{x}) = \vec{p}\psi_{\vec{p}}(\vec{x}). \tag{9.7}$$

Każde z tych równań określa zależność funkcji falowej $\psi_{\vec{p}}(\vec{x})$ od jednej ze składowych wektora położenia $\vec{x}$. Na przykład równanie pierwsze z pochodną cząstkowa po x_1 daje

$$\psi_{\vec{p}}(\vec{x}) = f(x_2,\ldots,x_d)e^{i\frac{p_1x_1}{\hbar}}, \tag{9.8}$$

gdzie czynnik f jest dowolną funkcją zmiennych $x_2,\ldots,x_d$. Udowodnić to można przez podstawienie do równania. Uwzględniając wszystkie równania i

pamiętając wzór na iloczyn skalarny dwu wektorów w d wymiarach ($\vec{x} \cdot \vec{p} = x_1 p_1 + \cdots + x_d p_d$) otrzymujemy

$$\psi_{\vec{p}}(\vec{x}) = C e^{i \frac{\vec{p} \cdot \vec{x}}{\hbar}}. \tag{9.9}$$

Normalizację wprowadzamy podobnie jak w przypadku jednowymiarowym. Przyjmujemy

$$\langle \psi_{\vec{p}} | \psi_{\vec{p'}} \rangle = \delta^d(\vec{p} - \vec{p'}), \tag{9.10}$$

i otrzymujemy

$$\psi_{\vec{p}}(\vec{x}) = \frac{1}{(2\pi\hbar)^{d/2}} e^{i \frac{\vec{p} \cdot \vec{x}}{\hbar}}. \tag{9.11}$$

W stanie opisywanym przez funkcję falową $\psi_{\vec{p}}(\vec{x})$ można zmierzyć wszystkie składowe pędu i dla każdej z nich wiadomo na pewno, jaki będzie wynik pomiaru. Ponieważ stany te stanowią układ zupełny (każdą funkcje można przedstawić jako kombinacje liniową funkcji $\psi_{\vec{p}}(\vec{x})$), mówimy ogólnie, że składowe pędu są jednocześnie mierzalne. W tej definicji zupełność układu jest ważna. Zdarza się, że jakieś wielkości mierzalne są jednocześnie dobrze określone dla zbioru stanów, który nie jest układem zupełnym. To nie wystarczy, żeby uważać te wielkości za jednocześnie mierzalne. Wrócimy do tego zagadnienia w rozdziale 11. Pozostałe uwagi są jak dla układu jednowymiarowego. Dowolna liczba rzeczywista może być wynikiem pomiaru dowolnej składowej wektora pędu cząstki. Stanowi z dobrze określonym wektorem pędu odpowiada funkcja falowa opisująca falę płaską, przy czym wektor falowy tej fali jest związany z pędem cząstki (lub cząstek!) wzorem de Broglie'a $\vec{p} = \hbar \vec{k}$.

Znając funkcje falowe $\psi_{\vec{p}}(\vec{x}) = \langle \vec{x} | \vec{p} \rangle$, łatwo jest powiązać funkcje falowe układu cząstek w reprezentacji położeń z odpowiadającymi im funkcjami w reprezentacji pędów. Oznaczmy funkcje falowe danego układu, w danym stanie, w reprezentacji położeń i w reprezentacji pędów odpowiednio przez

$$\psi^(\vec{x}) = \langle \vec{x} | \psi \rangle, \tag{9.12}$$

$$\phi(\vec{p}) = \langle \vec{p} | \psi \rangle. \tag{9.13}$$

Wstawiając do pierwszego wzoru jedynkę zbudowaną z wektorów $|\vec{p}\rangle$ i do drugiego jedynkę zbudowaną z wektorów $|\vec{x}\rangle$ oraz pamiętając, że zmiana kolejności czynników w iloczynie skalarnym jest równoważna ze sprzężeniem zespolonym, otrzymujemy

$$\psi(\vec{x}) = \frac{1}{(2\pi\hbar)^{d/2}} \int d^d p e^{i \frac{\vec{p} \cdot \vec{x}}{\hbar}} \phi(\vec{p}), \tag{9.14}$$

$$\phi(\vec{p}) = \frac{1}{(2\pi\hbar)^{d/2}} \int d^d x e^{-i \frac{\vec{p} \cdot \vec{x}}{\hbar}} \psi(\vec{x}). \tag{9.15}$$

Takie wzory pozwalają powiązać, na przykład dla stanu podstawowego atomu wodoru, rozkład gęstości elektronu w przestrzeni $|\psi(\vec{x})|^2$ z jego rozkładem pędu $|\phi(\vec{p})|^2$.

Według terminologii matematycznej funkcja $\phi(\vec{p})$ jest transformatą Fouriera funkcji $\psi(\vec{x})$. Z punktu widzenia fizyki mamy tu przykład zastosowania zasady

superpozycji i interpretacji probabilistycznej mechaniki kwantowej. Funkcja falowa $\psi(\vec{x})$ została przedstawiona jako superpozycja fal płaskich, a kwadrat modułu współczynnika $\phi(\vec{p})$ daje gęstość prawdopodobieństwa otrzymania wyniku $\vec{p}$ przy dokładnym pomiarze pędu w stanie $|\psi\rangle$.

Ta uwaga pozwala na prostą interpretację zasady nieokreśloności Heisenberga dla współrzędnej i odpowiadającego jej pędu. Stan o określonym pędzie $\vec{p}$ odpowiada fali płaskiej z określonym wektorem falowym $\vec{k} = \vec{p}/\hbar$ i amplitudą stałą w całej przestrzeni. Chcąc wytworzyć pakiet falowy zlokalizowany gdzieś w przestrzeni, musimy wykorzystać interferencję – konstruktywną w obszarze koncentracji i destruktywna poza tym obszarem – fal o różnych wektorach falowych. Im lepsza jest wymagana lokalizacja w przestrzeni, tym szerszy zakres długości fal czy wektorów falowych, czy pędów musi być użyty. Wzór Heisenberga odpowiada znanemu od dawna matematykom twierdzeniu o całkach Fouriera.

Niekiedy zamiast normalizacji do delty Diraca stosuje się tak zwaną normalizację w pudle. Zacznijmy od przypadku jednowymiarowego. Intuicyjnie wydaje się rozsądne, że jeśli założymy, że funkcje falowe są okresowe z okresem L, dużo większym od rozmiarów badanego układu, to to nieznacznie wpłynie na wyniki. Założenie o okresowości prowadzi do warunku na wartości własne

$$\frac{pL}{\hbar} = 2n\pi, \tag{9.16}$$

gdzie n jest liczbą całkowitą, czyli do wartości własnych

$$p_n = \frac{2\pi\hbar}{L}n; \qquad n = 0, \pm 1, \pm 2, \ldots \tag{9.17}$$

Otrzymaliśmy rozkład dyskretny, ale w praktyce dla $L \to \infty$ nie wiele różniący się od ciągłego. Wybieramy teraz jako warunek normalizacyjny założenie, że całka z kwadratu funkcji falowej po odcinku o długości L jest równa jeden. To daje $|C|^2 L = 1$, a zatem znormalizowana funkcja falowa ma postać

$$\psi_p(x) = \frac{1}{\sqrt{L}} e^{i\frac{px}{\hbar}}. \tag{9.18}$$

Analogicznie w trzech wymiarach, zastępując odcinek L objętością V, otrzymujemy

$$\psi_{\vec{p}}(\vec{x}) = \frac{1}{\sqrt{V}} e^{i\frac{\vec{p}\cdot\vec{x}}{\hbar}}. \tag{9.19}$$

W praktyce, jeśli da się zrobić przejście graniczne, odpowiednio $L \to \infty$ albo $V \to \infty$, normalizacja w pudle daje poprawne wyniki. Tę metodę normalizacji zastosujemy w rozdziale 29, gdzie zobaczymy jej zalety.

Wybór warunku brzegowego prowadzącego do okresowości funkcji falowej nie jest jedynym możliwym. Możliwości jest nieskończenie wiele. Można by, na przykład, przyjąć tak zwany warunek anty-okresowości

$$\frac{pL}{\hbar} = (2n+1)\pi, \quad \text{czyli} \quad p_n = \frac{\pi\hbar}{L}(2n+1). \tag{9.20}$$

Jak łatwo sprawdzić, ta zmiana nie wpływa na wyniki w granicy $L \to \infty$. To jest szczególny przypadek twierdzenia Weyla, stosującego się też w przypadku trójwymiarowym, o niezależności wyników dla $L \to \infty$ czy $L_x \to \infty$, $L_y \to \infty$, $L_z \to \infty$ od warunków na brzegach przedziału czy pudła.

Rozdział 10

Równanie Schrödingera niezależne od czasu

W tym rozdziale przedyskutujemy problem znajdowania możliwych wyników dokładnego pomiaru energii układu i znajdowania funkcji falowych stanów, dla których wynik dokładnego pomiaru energii da się z pewnością przewidzieć. Możliwe wyniki pomiaru energii są nazywane poziomami energetycznymi układu. Stan o najniższej energii, jeśli istnieje dla danego układu, jest nazywany stanem podstawowym. Pozostałe stany to są stany wzbudzone. Danej wartości energii może odpowiadać dokładnie jeden stan – wtedy mówimy, że ten poziom energetyczny jest niezdegenerowany – lub nieskończenie wiele różnych stanów. Jak zostanie pokazane w dalszym ciągu tego rozdziału, stany odpowiadające danej wartości energii stanowią przestrzeń wektorową. Jeżeli ta przestrzeń ma wymiar $k > 1$, skończony lub nieskończony, to mówimy, że odpowiednia wartość własna jest k krotnie zdegenerowana.

Weźmy jako przykład stany elektronu w polu kulombowskim w przybliżeniu, które daje nierelatywistyczna mechanika kwantowa bez uwzględnienia spinu elektronu. Ten układ może służyć jako uproszczony model atomu wodoru czy jonu z jednym elektronem, takiego jak He^+, Li^{++} itd. Niektóre poprawki do tego modelu zostaną przedyskutowane w dalszych rozdziałach. Już teraz, dla uproszczenia sformułowań, będziemy nazywali źródło pola kulombowskiego jądrem. Jako poziom zerowy energii przyjmujemy energię spoczynkową elektronu.

Oznaczmy ładunek jądra przez $-Ze$, gdzie liczba (ujemna) e oznacza ładunek elektronu, a masę elektronu przez m. Wtedy, jak można pokazać, poziomy energetyczne o ujemnej energii dane są wzorem

$$E_n = -\frac{me^4Z^2}{2\hbar^2 n^2}; \qquad n = 1, 2, \ldots \tag{10.1}$$

To, że te wartości energii są ujemne oznacza, że dla tych stanów energia elektronu w polu jądra jest niższa niż energia swobodnego elektronu, a zatem że elektron jest związany. Energia stanu podstawowego, którą otrzymujemy podstawiając $n = 1$ do powyższego wzoru, może być interpretowana jako różnica energii miedzy elektronem w stanie podstawowym i nieruchomym elektronem nie oddziałującym z jądrem. Energia stanu podstawowego ze zmienionym znakiem

daje więc minimalną pracę potrzebną na wyrwanie elektronu z atomu czy jonu i przeniesienie go na tak dużą odległość, że energia oddziaływania z jądrem staje się do zaniedbania. W zwykłej terminologii to jest energia jonizacji. Oprócz wartości (10.1), energia może jeszcze przyjmować dowolne wartości nieujemne. Zbiór możliwych wartości energii nazywamy widmem energii. W omawianym przykładzie to widmo składa się z części dyskretnej $E < 0$ i z części ciągłej przy $E \geq 0$. Stany odpowiadające energiom ujemnym są, jak pokazaliśmy, stanami związanymi i ich wektory powinny dać się znormalizować do jedności. Stany odpowiadające energiom nieujemnym są stanami, w których elektron nie jest zlokalizowany w żadnej skończonej części przestrzeni i powinny być wektorami uogólnionymi. Takie stany są też nazywane stanami rozproszeniowymi.

Żeby określić jednoznacznie wektor stanu w rozpatrywanym przykładzie, można podać trzy liczby: liczbę naturalną n, która określa energię według wzoru (10.1), nieujemną, mniejszą od n liczbę całkowitą l i liczbę całkowitą m nie większą co do wartości bezwzględnej od liczby l. Tak więc stan można jednoznacznie określić jako $|n, l, m\rangle$. Takie liczby określające stan, jak n, l, m w rozważanym tu przykładzie, nazywa się niekiedy liczbami kwantowymi. Dla $n = 1$ możliwe są tylko $l = 0$ i $m = 0$ – poziom jest niezdegenerowany. Dla każdego wyższego n występuje natomiast degeneracja. Na przykład dla $n = 2$ możliwe są $l = 0$ i $l = 1$. Dla $l = 0$ musi być $m = 0$, natomiast dla $l = 1$ możliwe są $m = -1, 0, 1$. Tak więc mamy cztery różne wektory stanu odpowiadające tej samej energii E_2: $|2, 0, 0\rangle, |2, 1, -1\rangle, |2, 1, 0\rangle$ i $|2, 1, 1\rangle$. Liczba n określa energię układu według wzoru (10.1). Jak pokażemy w rozdziale szesnastym, liczby l i m określają odpowiednio kwadrat krętu elektronu i rzut tego krętu na oś z. Z interpretacji probabilistycznej mechaniki kwantowej wynika więc (patrz rozdział siódmy), że dwa stany $|n, l, m\rangle$ różniące się choć jedną z liczb kwantowych n, l, m są do siebie ortogonalne. Stąd dalej wynika, że poziom energetyczny $n = 2$ jest czterokrotnie zdegenerowany. Schemat poziomów energetycznych atomu wodoru jest narysowany w rozdziale dwunastym.

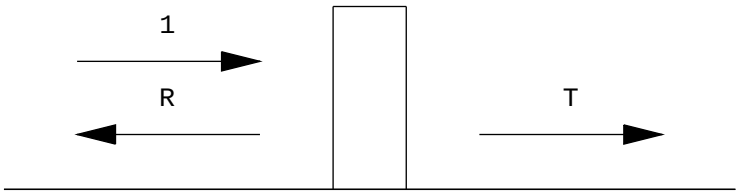
Jak wynika z podanych w rozdziale siódmym zasad kwantowania, żeby uzyskiwać wyniki takie, jak omówione powyżej dla modelu atomu wodoru, należy rozwiązać problem własny dla operatora energii

$$\hat{H}|\psi\rangle = E|\psi\rangle. \tag{10.2}$$

To równanie, jedno z najsłynniejszych w całej mechanice kwantowej, jest znane jako niezależne od czasu równanie Schrödingera. Równanie Schrödingera zależne od czasu, które można uważać za uogólnienie równania (10.2), jest też bardzo ważne i zostanie omówione w rozdziale dwudziestym pierwszym. Z równania (10.2) można w znaleźć wszystkie możliwe wartości energii, jako wartości własne E, jak również odpowiadające im wektory stanu $|\psi\rangle$. Równanie jest liniowe, więc jeżeli wektory $|\psi_1\rangle$ i $|\psi_2\rangle$ są jego rozwiązaniami odpowiadającymi tej samej wartości własnej E, to dowolna kombinacja liniowa $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ też jest rozwiązaniem odpowiadającym tej samej wartości własnej. To dowodzi, że jak zapowiedzieliśmy wcześniej, wektory stanu odpowiadające danej wartości własnej stanowią przestrzeń wektorową.

W praktyce, zwykle przechodzi się do reprezentacji położeń, gdzie równanie Schrödingera, przy założeniu że operatora Hamiltona jest lokalny (rozdział siódmy), przyjmuje postać

$$\hat{H}(x)\psi(x) = E\psi(x), \tag{10.3}$$



Rysunek 10.1: Rozpraszanie na jednowymiarowej, prostokątnej barierze potencjału. Amplituda fali dla cząstki nadlatującej z lewej strony jest umownie przyjęta równa jedności. Amplitudy fal odbitej i przepuszczonej są odpowiednio równe R i T.

Argument x oznacza zbiór wszystkich zmiennych występujących w problemie, na przykład wektor $\vec{x}$, a hamiltonian $\hat{H}(x)$ jest kwantowym odpowiednikiem klasycznego hamiltonianu zbudowanym tak, jak to zostało opisane w rozdziale siódmym. To równanie też jest nazywane równaniem Schrödingera. Na przykład, dla elektronu w polu kulombowskim z przykładu podanego powyżej w tym rozdziale, równanie Schrödingera niezależne od czasu przybiera postać[1]

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\psi(\vec{x}) - \frac{Ze^2}{r}\psi(x) = E\psi(x). \tag{10.4}$$

Założyliśmy tu, że jądro znajduje się w początku układu współrzędnych i odległość elektronu od jądra oznaczyliśmy r. Symbol nabla kwadrat $\vec{\nabla}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ oznacza się niekiedy $\triangle$ i nazywa laplasjanem. Rozwiązanie tego równania jest dosyć trudne. Wartości własne podaliśmy w tym rozdziale. Do odpowiadających im funkcji własnych wrócimy w rozdziale szesnastym.

Żeby znaleźć interesujące nas wartości własne i funkcje własne trzeba wiedzieć, jakie warunki mają spełniać rozwiązania. Niektóre z tych warunków wynikają z samego równania. Na przykład, jeżeli potencjał $V(x)$ nie ma osobliwości, funkcja ψ powinna być co najmniej dwa razy różniczkowalna. Inne warunki wynikają z opisu sytuacji fizycznej. Na przykład, jeśli funkcja falowa ma opisywać stan związany, to cząstka powinna być nieźle zlokalizowana i nakłada się zwykle warunek normalizacji

$$\int |\psi(x)|^2 dx = 1. \tag{10.5}$$

Zauważmy, że jeśli funkcja $\psi(x)$ jest rozwiązaniem równania Schrödingera, to jest nim także funkcja $c\psi(x)$, dla dowolnej różnej od zera stałej c. Istotą powyższego warunku jest więc to, że całka z kwadratu modułu funkcji falowej po całej dostępnej przestrzeni jest zbieżna. Wybór normalizacji do jedynki jest zrobiony tylko dla wygody i czasem wygodniej jest przyjąć inną konwencję. Przykład takiej sytuacji spotkamy w rozdziale czternastym.

Dla stanów nie odpowiadających układom związanym, czyli dla stanów rozproszeniowych, oprócz warunku normalizowalności do delty Diraca nakładamy jeszcze czasem dalsze warunki. Rozważmy na przykład jednowymiarowy układ przedstawiony na rysunku 10.1. Cząstka nadlatuje z lewej strony i trafia na barierę

[1] Zgodnie ze zwyczajem, powszechnie przyjmowanym w wykładach mechaniki kwantowej, używamy układu jednostek Gaussa, a nie zalecanego przez różne władze układu MKSA.

potencjału. Klasycznie, jeśli energia cząstki jest niższa od wysokości bariery, cząstka odbija się i wraca skąd przyszła. Jeśli energia cząstki jest większa od wysokości bariery, cząstka przelatuje przez barierę i zmierza ku dużym x-om. Według mechaniki kwantowej, cząstka przy każdej energii ma pewne prawdopodobieństwo odbicia się od bariery i pewne prawdopodobieństwo przeniknięcia przez barierę. Przenikanie przez barierę cząstek o energii niższej niż wysokość bariery było tak zaskakujące, że otrzymało osobna nazwę: efekt tunelowy albo tunelowanie. Ten efekt często występuje w przyrodzie. Na przykład rozpady α jąder atomowych są możliwe tylko dzięki efektowi tunelowemu. W omawianym przez nas przykładzie przyjmujemy następujące warunki brzegowe:

$$\psi(x) \to e^{ikx} + Re^{-ikx}, \qquad x \to -\infty, \tag{10.6}$$

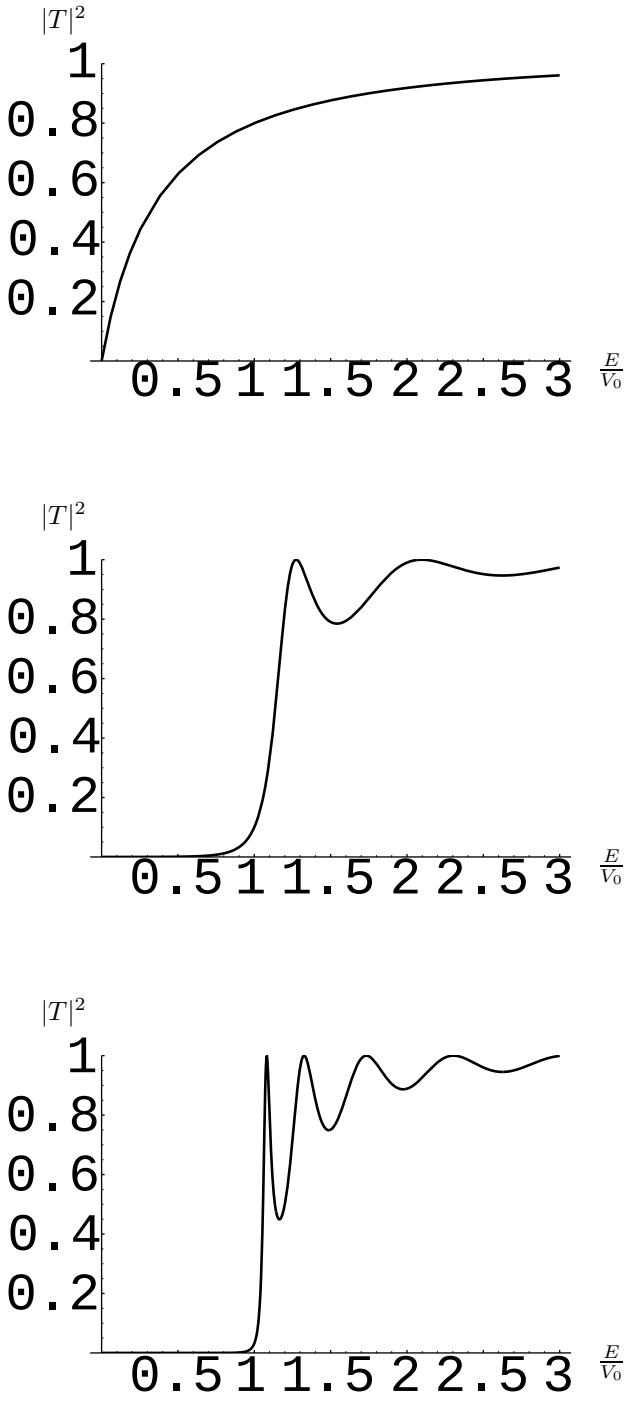
$$\psi(x) \to Te^{ikx}, \qquad x \to +\infty. \tag{10.7}$$

Te warunki mają prosty sens fizyczny. W obszarze dużych co do wartości bezwzględnej ujemnych x-ów, mamy superpozycję fali padającej odpowiadającej cząstce o pędzie $p = \hbar k > 0$ i fali odbitej o pędzie przeciwnym. W obszarze dużych dodatnich x-ów mamy tylko falę przepuszczoną o pędzie takim jak fala padająca. Stosunek amplitudy fali odbitej do amplitudy fali padającej wynosi R, a stosunek amplitudy fali przepuszczonej do amplitudy fali padającej wynosi T. Wynika stąd, że prawdopodobieństwo odbicia cząstki od bariery wynosi $|R|^2$, a prawdopodobieństwo przeniknięcia cząstki przez barierę wynosi $|T|^2$. Rozwiązując równanie Schrödingera niezależne od czasu z powyższymi warunkami brzegowymi i z warunkiem, że funkcja falowa i jej pierwsza pochodna po x mają być ciągłe, stwierdza się, że dla danej bariery i przy danym k rozwiązanie istnieje tylko dla jednej, ściśle określonej, pary liczb (R, T). W ten sposób wyznacza się prawdopodobieństwo przejścia cząstki przez barierę i prawdopodobieństwo odbicia. Oczywiście wychodzi $|R|^2 + |T|^2 = 1$.

Wyliczona zależność prawdopodobieństwa przejścia $|T|^2$ od energii jest pokazana na rysunku 10.2. W tym problemie występują trzy parametry wymiarowe: wysokość bariery V_0, szerokość bariery a oraz energia E cząstki padającej. Są jeszcze dwie stałe wymiarowe: stała Plancka i masa cząstki m. Wielkości bezwymiarowe, jak prawdopodobieństwo przejścia, zależą tylko od dwu bezwymiarowych kombinacji tych parametrów. Wybraliśmy stosunek energii cząstki do wysokości bariery $\frac{E}{V_0}$ i bezwymiarowy parametr utworzony z masy cząstki i parametrów bariery $2\hbar^{-2}mV_0a^2$. Klasycznie, $|T|^2$ jest równe jeden dla $E > V_0$ i zero dla $E < V_0$. Z górnej części rysunku widać, że jeśli bariera jest niska lub wąska, to punkt $E = V_0$ niczym się nie wyróżnia – prawdopodobieństwo przejścia rośnie od progu, i nie osiąga wartości jeden nawet dla $E \approx 3V_0$. Ze wzrostem szerokości lub wysokości bariery, prawdopodobieństwo tunelowania, to znaczy transmisji, przy $E < V_0$, maleje, ale pojawiają się silne oscylacje dla $E > V_0$ – cząstka przy niektórych energiach ma duże prawdopodobieństwo, że zostanie odbita, chociaż rozumując klasycznie bariera jest na to za niska.

Bardziej skomplikowany jest przypadek trójwymiarowy, gdzie zakładając, że cząstka pada z pędem $\vec{k}$ równoległym do osi z otrzymuje się warunek brzegowy dla $r \to \infty$:

$$\psi(\vec{x}) \to e^{ikz} + \frac{f(\theta, \phi)}{r} e^{ikr}. \tag{10.8}$$



Rysunek 10.2: Prawdopodobieństwo przejścia przez barierę $|T|^2$ dla cząstki o masie m, padającej na prostokątną barierę o wysokości V_0 i szerokości a. Bezwymiarowy parametr $\frac{2mV_0a^2}{\hbar^2}$ dla kolejnych wykresów licząc od góry wynosi 1, 6 i 11.

Pierwszy człon po prawej stronie oznacza falę padającą, drugi falę rozproszoną. Amplituda fali rozproszonej zależy od kątów sferycznych θ i ϕ kierunku, w którym poleciała cząstka rozproszona. Rozwiązując równanie Schrödingera niezależne od czasu, wyznacza się jedyną funkcję $f(\theta, \phi)$, dla której rozwiązanie istnieje. Kwadrat modułu tej funkcji daje różniczkowy przekrój czynny (rozdział 28) na rozpraszanie elastyczne pod kątami θ, ϕ.

Jako przykład rozwiązania równania Schrödingera niezależnego od czasu, rozważmy przypadek jednowymiarowy, kiedy potencjał $V(x)$ jest równy zero dla $0 < x < 1$ i nieskończoność poza tym odcinkiem. Cząstka nie może się znaleźć w obszarze, gdzie potencjał jest nieskończony, więc możemy przyjąć, że tam funkcja falowa $\psi(x) \equiv 0$ i ograniczyć się do szukania funkcji falowej w obszarze $0 < x < 1$. Równanie Schrödingera ma tam postać

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = E\psi(x). \tag{10.9}$$

Żeby funkcja falowa była ciągła w punktach $x = 0$ i $x = 1$ nakładamy warunki brzegowe

$$\psi(0) = \psi(1) = 0. \tag{10.10}$$

Ciągłość funkcji falowej jedni uważają za wniosek z interpretacji fizycznej, inni wyprowadzają z samego równania jak następuje. Zastępując nieskończony potencjał poza studnią przez stały potencjał $V(x) = V_0$, można udowodnić, że funkcja falowa i jej pierwsza pochodna są ciągłe. Przechodząc z V_0 do nieskończoności traci się ciągłość pochodnej, ale funkcja falowa pozostaje ciągła[2].

Ogólne rozwiązanie powyższego równania Schrödingera ma postać

$$\psi(x) = c_1 \sin(kx) + c_2 \cos(kx), \tag{10.11}$$

gdzie $k = \hbar^{-1}\sqrt{2mE}$. Z warunku w $x = 0$ otrzymujemy $c_2 = 0$. Warunek w $x = 1$ powoduje, że nie dla każdej wartości parametru k, a więc i nie dla każdej wartości energii E, otrzymujemy dopuszczalne rozwiązanie. Rozważany tu problem jest na tyle prosty, że można go rozwiązać analitycznie, ale pouczające jest następujące, znacznie ogólniejsze rozumowanie.

Istnieje metoda numerycznego rozwiązywania równań różniczkowych znana jako metoda wstrzeliwania się. Wprowadzamy mały parametr o wymiarze długości Δx. W naszym przypadku dla dowolnego $x < 1 - \Delta x$ zachodzi w przybliżeniu

$$\begin{aligned} \psi(x + \Delta x) &= \psi(x) + \psi'(x)\Delta x + \frac{1}{2}\psi''(x)(\Delta x)^2, \\ \psi'(x + \Delta x) &= \psi'(x) + \psi''(x)\Delta x, \\ \psi''(x + \Delta x) &= -\frac{2mE}{\hbar^2}\psi(x + \Delta x). \end{aligned} \tag{10.12}$$

Pierwsza równość ma błąd rzędu $(\Delta x)^3$, co jest do zaniedbania, jeśli "krok"Δx został wybrany dostatecznie mały. Druga równość ma błąd rzędu $(\Delta x)^2$, ale pochodna $\psi'(x)$ w najważniejszym dla nas wzorze na funkcję falową $\psi(x)$ występuje

[2]Tu też jest potrzebne założenie fizyczne, choć bardzo naturalne, że granica rozwiązania przy $V_0 \to \infty$ jest rozwiązaniem dla $V_0 = \infty$

pomnożona przez Δx, więc błąd na funkcje falową spowodowany niedokładnym wyliczeniem pochodnej jest znów rzędu $(\Delta x)^3$. Trzecia równość wynika z równania Schrödingera i jest dokładna. W punkcie $x = 0$ mamy $\psi(0) = 0$ i stąd $\psi''(0) = 0$. Pierwszej pochodnej w punkcie $x = 0$ nie znamy, ale ponieważ normalizacja funkcji $\psi(x)$ jest dowolna, możemy wybrać ją tak, żeby było $\psi'(0) = 1$[3]. Żeby zacząć rozwiązywać równanie, potrzebna nam jest jeszcze wartość własna E, którą można zgadnąć lepiej lub gorzej. Znając $\psi(0), \psi'(0), \psi''(0)$ i E możemy obliczyć, z podanych wyżej równości, funkcję falową i jej pierwsze dwie pochodne dla $x = \Delta x$. Powtarzając tę procedurę obliczamy w końcu wartość funkcji $\psi(1)$. Jeżeli zgadliśmy wartość E trochę mniejszą od E_0, to otrzymamy $\psi(1) > 0$. Jeśli zgadliśmy wartość E trochę większą od E_0, to otrzymamy $\psi(1) < 0$. Tylko jeśli zgadliśmy $E = E_0$, otrzymamy zgodną z warunkami zadania wartość $\psi(1) = 0$. Te trzy przypadki są przedstawione na rysunku 3. Oczywiście trzeba by tu jeszcze uwzględnić błąd spowodowany zastosowanym przybliżeniem, ale to jest problem z metod numerycznych, którego dyskusję pominiemy. Powyższa dyskusja pozwala zrozumieć, co w ogólnym przypadku jednowymiarowego równania Schrödingera wyróżnia wartości własne. Podobne rozumowanie stosuje się, jeśli szukamy rozwiązania dla $-\infty < x < \infty$. Na przykład, na rysunku 7.1 są pokazane funkcje falowe jednowymiarowego oscylatora harmonicznego. Te funkcje odpowiadają poprawnym wartościom własnym i dlatego dążą do osi x-ów przy $x \to \pm\infty$. Gdybyśmy próbowali konstruować rozwiązania używając zgadniętych wartości własnych, stwierdzilibyśmy, że naogół ze wzrostem $|x|$ otrzymane funkcje uciekają do plus lub minus nieskończoności.

Wracamy do rozwiązania analitycznego naszego problemu. Jak łatwo sprawdzić podstawiając rozwiązanie do równania, $E = \hbar^2 k^2/2m$. Po to, żeby funkcji $\psi(x)$ była równa zero dla $x = 1$, potrzeba i wystarcza, żeby k było całkowitą wielokrotnością π. Stąd na wartości własne energii otrzymujemy

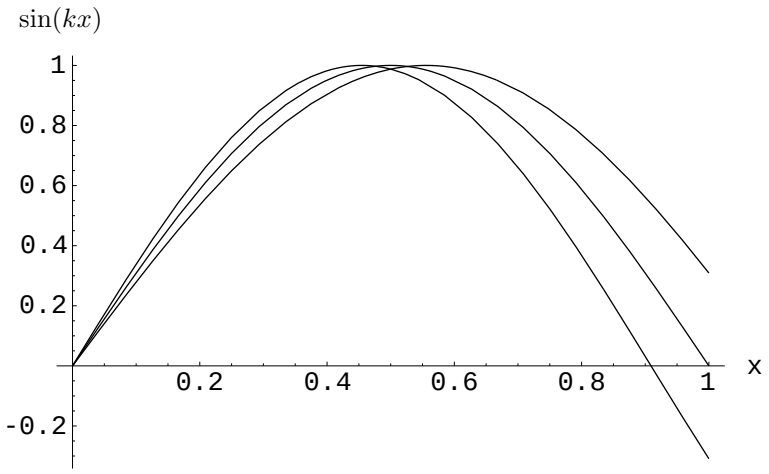
$$E_n = \frac{\hbar^2 \pi^2 n^2}{2m}; \qquad n = 1, 2, \ldots. \tag{10.13}$$

Wartość współczynnika c_1 jest kwestią umowy, ale jeśli przyjmiemy, że chcemy funkcje falowe znormalizować do jedności i że stałe współczynniki normalizacyjne powinny być rzeczywiste i dodatnie, to otrzymamy

$$\psi_n(x) = \sqrt{2}\sin(n\pi x); \qquad n = 1, 2, \ldots. \tag{10.14}$$

Funkcje $\psi_0(x)$ i $\psi_{-n}(x)$ też spełniają równanie i warunki brzegowe, ale $\psi_0(x) \equiv 0$, więc nie opisuje żadnego stanu, a $\psi_{-n}(x) \equiv -\psi_n(x)$, więc opisuje ten sam stan $|n\rangle$.

[3]Przypadku $\psi'(0) = 0$ nie rozważamy, bo prowadzi do rozwiązania $\psi(x) \equiv 0$, które jest nie do przyjęcia.



Rysunek 10.3: Funkcje $\sin(kx)$ dla $k = 0.9\pi,\ \pi,\ 1.1\pi$.

Rozdział 11

Wielkości jednocześnie mierzalne

Według mechaniki kwantowej niektóre pary wielkości mogą być jednocześnie ustalone i zmierzone, a inne nie. Na przykład, można jednocześnie ustalić dwie składowe wektora położenia cząstki czy dwie składowe wektora jej pędu, ale nie da się jednocześnie ustalić x-owej składowej wektora położenia i x-owej składowej wektora pędu. W tym rozdziale podamy ogólne twierdzenie, które dla dwu dowolnych wielkości mierzalnych A i B pozwala stwierdzić, czy te wielkości dadzą się jednocześnie ustalić.

Mówimy, że wielkości mierzalne A i B dadzą się jednocześnie ustalić, jeżeli istnieje baza wspólnych wektorów własnych odpowiadających im operatorów. Te wektory bazowe z definicji spełniają równania:

$$\hat{A}|a_i, b_j, \alpha\rangle = a_i|a_i, b_j, \alpha\rangle, \tag{11.1}$$

$$\hat{B}|a_i, b_j, \alpha\rangle = b_j|a_i, b_j, \alpha\rangle. \tag{11.2}$$

Parametr α oznacza zbiór parametrów potrzebny, poza parametrami a_i i b_j, do jednoznacznego określenia wektora bazowego. Zgodnie z zasadami podanymi w rozdziale piątym, w stanie $|a_i, b_j, \alpha\rangle$ pomiar wielkości A da na pewno wynik a_i i pomiar wielkości B da na pewno wynik b_j. Przyjmujemy, że w tym stanie wielkości A i B są jednocześnie ustalone i dadzą się jednocześnie zmierzyć. Może się zdarzyć, że wspólne wektory własne operatorów $\hat{A}$ i $\hat{B}$ istnieją, ale nie stanowią układu zupełnego. W takim razie nie mówimy, że wielkości A i B są jednocześnie mierzalne.

Łatwo jest sprawdzić, że jeżeli wielkości A i B są jednocześnie mierzalne, to odpowiadające im operatory komutują. Rzeczywiście

$$\hat{A}\hat{B}|a_i, b_j, \alpha\rangle = b_j\hat{A}|a_i, b_j, \alpha\rangle = a_i b_j|a_i, b_j, \alpha\rangle, \tag{11.3}$$

$$\hat{B}\hat{A}|a_i, b_j, \alpha\rangle = a_i\hat{B}|a_i, b_j, \alpha\rangle = a_i b_j|a_i, b_j, \alpha\rangle, \tag{11.4}$$

więc operator $\hat{A}\hat{B}$ jest równy operatorowi $\hat{B}\hat{A}$ w działaniu na wektory z rozpatrywanej bazy. Ponieważ dowolny wektor da się przedstawić jako kombinacja

liniowa tych wektorów bazowych i operatory $\hat{A}$ i $\hat{B}$ sa liniowe, mamy dla dowolnego wektora $|\psi\rangle$ wynik $(\hat{A}\hat{B} - \hat{B}\hat{A})|\psi\rangle = 0$, co daje dla operatorów

$$[\hat{A}, \hat{B}] = 0. \tag{11.5}$$

Zamiast mówić, że komutator dwu operatorów jest równy zeru, można też mówić, że te dwa operatory komutują ze sobą. Prawdziwe, chociaż nieco trudniejsze do udowodnienia, jest też twierdzenie odwrotne, które podajemy tu bez dowodu: jeżeli operatory $\hat{A}$ i $\hat{B}$ komutują, to ich wspólne wektory własne stanowią układ zupełny.

Otrzymane wyniki można krótko podsumować jako:
Twierdzenie: po to żeby wielkości mierzalne A i B były jednocześnie mierzalne, potrzeba i wystarcza, żeby odpowiadające im operatory komutowały ze sobą. Podamy teraz trzy przykłady zastosowania tego twierdzenia.

Przykład 1: Cząstka swobodna w przestrzeni trójwymiarowej

Trzy składowe pędu cząstki p_x, p_y, p_z są jednocześnie mierzalne. Żeby to wykazać, trzeba udowodnić znikanie trzech komutatorów, ale ponieważ wszystkie te komutatory liczy się tak samo, ograniczymy się do komutatora $[\hat{p}_x, \hat{p}_y]$. W reprezentacji położeń mamy

$$\hat{p}_x\hat{p}_y\psi(\vec{x}) = -\hbar^2\frac{\partial^2\psi(\vec{x})}{\partial x\partial y}, \tag{11.6}$$

$$\hat{p}_y\hat{p}_x\psi(\vec{x}) = -\hbar^2\frac{\partial^2\psi(\vec{x})}{\partial y\partial x}. \tag{11.7}$$

Pochodne po prawych stronach tych równań są sobie równe, co ponieważ funkcja $\psi(\vec{x})$ jest dowolna, prowadzi do wniosku, że komutator jest równy zeru. Łatwo też sprawdzić, że hamiltonian cząstki swobodnej

$$H(\vec{x}) = \frac{\vec{p}^2}{2m} \tag{11.8}$$

komutuje z operatorami składowych pędu cząstki. Istnieje więc układ zupełny wspólnych funkcji własnych operatorów $\hat{H}, \hat{p}_x, \hat{p}_y, \hat{p}_z$. W rozdziale dziewiątym stwierdziliśmy jednak, że wymaganie, żeby jakaś funkcja $\psi(\vec{x})$ była wspólną funkcją własną operatorów wszystkich składowych wektora pędu, określa ją z dokładnością do stałego czynnika, którego wartość bezwzględną można wyznaczyć z warunku normalizacji. Na mocy twierdzenia takie funkcje stanowią także układ zupełny funkcji własnych hamiltonianu. Nie ma więc potrzeby rozwiązywania równania Schrödingera dla cząstki swobodnej. Wystarczyło znaleźć wspólne funkcje własne operatorów składowych wektora pędu, co jak widać z dyskusji w rozdziale dziewiątym sprowadza się do rozwiązania trzech identycznych równań różniczkowych zwyczajnych pierwszego rzędu. Jest to zadanie znacznie prostsze niż rozwiązywanie równania Schrödingera, które jest cząstkowym równaniem różniczkowym trzech zmiennych i drugiego rzędu. Znając funkcje własne, łatwo jest już znaleźć wartości własne hamiltonianu

$$\hat{H}(\vec{x})\psi_{p_x,p_y,p_z}(\vec{x}) = \frac{\vec{p}^2}{2m}\psi_{p_x,p_y,p_z}(\vec{x}). \tag{11.9}$$

Ten przykład jest typowy. Jeśli znamy jakiś operator prostszy od hamiltonianu i komutujący z hamiltonianem, to zwykle można przy jego pomocy uprościć rozwiązywanie równania Schrödingera niezależnego od czasu. Zauważmy,

że znaleźliśmy układ zupełny funkcji własnych hamiltonianu, a nie wszystkie funkcje własne hamiltonianu. Na przykład funkcja $\psi_{p_x,p_y,p_z}(\vec{x})+\psi_{-p_x,-p_y,-p_z}(\vec{x})$ jest funkcją własną hamiltonianu, choć nie jest funkcją własną żadnego z operatorów $\hat{p}_x, \hat{p}_y, \hat{p}_z$.

Przykład 2: Trójwymiarowy oscylator harmoniczny.

Hamiltonian trójwymiarowego oscylatora harmonicznego można zapisać w postaci

$$\hat{H} = \hat{H}_x + \hat{H}_y + \hat{H}_z, \tag{11.10}$$

gdzie

$$\hat{H}_u = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial u^2} + \frac{1}{2}Ku^2; \qquad u = x, y, z. \tag{11.11}$$

Operatory $\hat{H}_x, \hat{H}_y, \hat{H}_z, \hat{H}$ komutują między sobą, więc istnieje układ zupełny ich wspólnych funkcji własnych. Dla każdego z operatorów $\hat{H}_u$, funkcjami własnymi są funkcje falowe jednowymiarowego oscylatora harmonicznego podane w rozdziale siódmym. Interpretacja fizyczna jest następująca. Drgania trójwymiarowego oscylatora harmonicznego można sobie wyobrażać jako superpozycję drgań w kierunku osi x, drgań w kierunku osi y i drgań w kierunku osi z. Te drgania składowe są od siebie niezależne i energia oscylatora jest sumą energii tych trzech jednowymiarowych oscylacji. Dalsze rozumowanie jest takie, jak dla omówionego w rozdziale dziewiątym pędu cząstki w przestrzeni trójwymiarowej. Wspólne funkcje własne operatorów $\hat{H}_x, \hat{H}_y, \hat{H}_z$ są dane wzorem

$$\psi_{n_x,n_y,n_z}(\vec{x}) = \psi_{n_x}(x)\psi_{n_y}(y)\psi_{n_z}(z), \tag{11.12}$$

gdzie liczby n_x, n_y, n_z przybierają niezależnie od siebie wartości całkowite nieujemne, a funkcje $\psi_{n_u}(u); \quad u = x, y, z \quad$ są funkcjami własnymi jednowymiarowego oscylatora harmonicznego, podanymi w rozdziale dziesiątym. Na mocy twierdzenia, te same funkcje (11.12) są też funkcjami własnymi hamiltonianu $\hat{H}$ trójwymiarowego oscylatora harmonicznego. Każdy z hamiltonianów $\hat{H}_u$ ma w stanie (11.12) ustaloną energię $\hbar\omega(n_u+\frac{1}{2})$. Wynika z tego, że energia oscylatora trójwymiarowego też jest ustalona i wynosi

$$E_{n_x+n_y+n_z} = \hbar\omega\left(n_x + n_y + n_z + \frac{3}{2}\right). \tag{11.13}$$

Zauważmy, że wszystkie poziomy energetyczne poza podstawowym są zdegenerowane, bo każdą liczbę całkowitą większą od zera można na różne sposoby przedstawiać jako sumę trzech liczb całkowitych nieujemnych. Dla danej wartości liczby kwantowej $n_x + n_y + n_z$, funkcje (11.12) stanowią bazę w przestrzeni wektorowej funkcji własnych odpowiadających energii (11.13). Jeżeli jednak $n_x + n_y + n_z > 0$, to ta przestrzeń jest więcej niż jednowymiarowa i można w niej wybrać bazę na nieskończenie wiele sposobów. Baza z funkcji (11.12) jest poprawna, ale warto pamiętać, że jakaś inna baza może być wygodniejsza.

Przykład 3: Problem dwu ciał.

Rozważmy w fizyce klasycznej dwie cząstki o masach m_1 i m_2, położeniach $\vec{x}_1$ i $\vec{x}_2$ oraz pędach $\vec{p}_1$ i $\vec{p}_2$. Zakładamy, że na te cząstki nie działają żadne siły zewnętrzne i że energia ich wzajemnego oddziaływania zależy tylko od ich wzajemnego położenia $\vec{x}_1 - \vec{x}_2$. Klasyczny hamiltonian ma postać

$$H = \frac{\vec{p}_1^2}{2m_1} + \frac{\vec{p}_2^2}{2m_2} + V(\vec{x}_1 - \vec{x}_2). \tag{11.14}$$

Równanie Schrödingera odpowiadające temu hamiltonianowi jest cząstkowym równaniem różniczkowym w sześciu zmiennych. Pokażemy, jak wykorzystując twierdzenie o wielkościach jednocześnie mierzalnych, można zastąpić to równanie równaniem w trzech zmiennych.

Z klasycznej mechaniki wiadomo, że jeśli wprowadzić nowe masy, współrzędne i sprzężone z niemi pędy według wzorów:

$$\begin{aligned} M &= m_1 + m_2, \\ \mu &= \frac{m_1 m_2}{M}, \\ \vec{X} &= \frac{m_1\vec{x}_1 + m_2\vec{x}_2}{M}, \\ \vec{x} &= \vec{x}_1 - \vec{x}_2, \\ \vec{P} &= \vec{p}_1 + \vec{p}_2, \\ \vec{p} &= \frac{m_2\vec{p}_1 - m_1\vec{p}_2}{M}, \end{aligned} \tag{11.15}$$

to hamiltonian można przepisać w postaci

$$H = \frac{\vec{P}^2}{2M} + \frac{\vec{p}^2}{2\mu} + V(\vec{x}). \tag{11.16}$$

Nowe zmienne mają prostą interpretacje fizyczną. M i $\vec{P}$ oznaczają całkowitą masę i całkowity pęd układy dwu ciał, a $\vec{x}_1 - \vec{x}_2$ jest położeniem cząstki 1, jeśli położenie cząstki 2 przyjąć za początek układu współrzędnych. $\vec{X}$ jest położeniem środka masy układu. Pęd $\vec{p}$ ma najprostszą interpretację w układzie środka masy, gdzie $\vec{p}_1 = -\vec{p}_2 = \vec{p}$. Jeżeli stosunek mas cząstek jest bardzo różny od jedności, to masa zredukowana μ jest niewiele mniejsza od masy lżejszej z cząstek. Na przykład, dla atomu wodoru masa zredukowana układu elektron-proton jest mniejsza od masy elektronu o około 0.05%. Jeżeli masy są równe, jak w przypadku układu e^+e^-, to masa zredukowana jest równa połowie masy cząstki. We wzorze (11.16) pierwszy człon po prawej stronie daje energię swobodnego ruchu układu jako całości. Następne dwa człony dają energię kinetyczną i energię potencjalną ruchu względnego.

Kwantując hamiltonian (11.16) według reguł z rozdziału siódmego, co stanowi uogólnienie od pary (współrzędna cząstki, pęd cząstki) na parę (współrzędna układu, pęd kanonicznie sprzężony z tą współrzędną), otrzymujemy

$$\hat{H} = \hat{H}_c + \hat{H}_w, \tag{11.17}$$

gdzie hamiltonian ruchu całości

$$\hat{H}_c = -\frac{\hbar^2}{2M}\vec{\nabla}_X^2 \tag{11.18}$$

i hamiltonian ruchu względnego

$$\hat{H}_w = -\frac{\hbar^2}{2\mu}\vec{\nabla}_x^2 + V(\vec{x}). \tag{11.19}$$

Wskaźniki przy nablach przypominają, że we wzorze na $\hat{H}_c$ różniczkowania są po składowych wektora $\vec{X}$, a we wzorze na $\hat{H}_w$ po składowych wektora $\vec{x}$. Operatory $\hat{H}_c$ i $\hat{H}_w$ komutują ze sobą. Istnieje więc układ zupełny wspólnych funkcji własnych operatorów $\hat{H}, \hat{H}_c, \hat{H}_w$. Hamiltonian $\hat{H}_c$ jest hamiltonianem cząstki swobodnej o masie M. Wobec tego wspólne funkcje własne można zapisać w postaci

$$\Psi_{\vec{P}}(\vec{x}, \vec{X}) = \psi(\vec{x}) e^{i\frac{\vec{P}\cdot\vec{X}}{\hbar}}. \tag{11.20}$$

Jeżeli nie interesuje nas ruch całości, to możemy wybrać stany odpowiadające wartościom własnym $\vec{P} = 0$ i wspólne funkcje własne sprowadzają się do funkcji $\psi(\vec{x})$. Wtedy wkład energii ruchu całości do pełnej energii układu jest równy zeru i energia E sprowadza się do energii w układzie środka masy. Warunek, że wspólne funkcje własne są też funkcjami własnymi hamiltonianu $\hat{H}_w$, daje równanie

$$-\frac{\hbar^2}{2\mu}\vec{\nabla}_x^2\psi(\vec{x}) + V(\vec{x})\psi(\vec{x}) = E\psi(\vec{x}). \tag{11.21}$$

To równanie różni się od równania Schrödingera dla jednej cząstki w polu sił o potencjale $V(\vec{x})$ tylko zamianą masy cząstki m na masę zredukowaną pary cząstek μ. Na przykład, w rozdziale dziesiątym był dyskutowany model atomu wodoru i jonów wodoropodobnych, w którym elektron poruszał się w zewnętrznym polu kulombowskim $\frac{-Ze}{r}$, które mogło być interpretowane jako pole kulombowskie pochodzące od jądra tak ciężkiego, że jego ruchy można zaniedbać. Z dyskusji podanej w bieżącym rozdziale wynika, że uwzględnienie realistycznej, skończonej masy jądra sprowadza się do zastąpienia we wzorach masy elektronu przez masę zredukowaną układu elektron-jądro. Jest to poprawka nieduża, ale przy dokładnościach uzyskiwanych w spektroskopii atomowej bardzo ważna.

Rozdział 12

Cząstki identyczne

W fizyce klasycznej znany był następujący paradoks Gibbsa. Rozważmy naczynie rozdzielone przegrodą na dwie części o takiej samej objętości. W jednej z tych części znajduje się N cząsteczek tlenu, a w drugiej N cząsteczek azotu. Kiedy usuwamy przegrodę, następuje nieodwracalny proces mieszania się gazów i jak łatwo obliczyć entropia wzrasta o $2k_B N \ln 2$, gdzie k_B oznacza stałą Boltzmanna. Gdyby po obu stronach przegrody znajdował się tlen, to po usunięciu przegrody makroskopowo nic by się nie stało, więc i nie byłoby wzrostu entropii. Gibbs i jego następcy zastanawiali się, jak interpolować miedzy tymi dwoma sytuacjami. Co się dzieje, kiedy cząsteczki po obu stronach przegrody stają się coraz podobniejsze do siebie? Czy zmiana przyrostu entropii, kiedy różnica miedzy cząsteczkami dąży do zera, zmienia się nagle w jakimś punkcie czy przebiega stopniowo?

Mechanika kwantowa dała zaskakujące rozwiązanie tego paradoksu: jest zasadnicza różnica między cząstkami czy układami cząstek, jak cząsteczki czy jądra, identycznymi a nieidentycznymi, choćby bardzo podobnymi do siebie. Nie ma więc powodu spodziewać się ciągłego przejścia, kiedy cząsteczki zmieniają się z bardzo podobnych do siebie w identyczne. Już samo stwierdzenie, że cząstki identyczne istnieją jest ciekawe. Filozofowie uczyli kiedyś, że jeśli dwa obiekty są pod każdym względem identyczne, to są tym samym obiektem. Mechanika kwantowa dostarcza licznych kontrprzykładów na to twierdzenie. Na przykład, każde dwa elektrony są identyczne. To nie tylko znaczy, że mają dokładnie takie same masy, ładunki i inne parametry. To także znaczy, że nie ma sensu stwierdzenie, że ten elektron jest w Krakowie, a tamten w Żywcu. W pewnym sensie, który omówimy w dalszym ciągu tego rozdziału, każdy z tych elektronów jest i tu, i tam, ale w taki sposób, że poprawne jest stwierdzenie, że jeden elektron znajduje się w Krakowie, a drugi w Żywcu. Nie ma sensu tylko pytanie, który z nich jest gdzie.

Rozważmy funkcję falową N cząstek identycznych. Argumentami tej funkcji będą wielkości $x_k^{(i)}$. W tym zapisie x_k oznacza zbiór liczb potrzebny do pełnego określenia stanu jednej cząstki. Dla dodatniego mezonu π mogłyby to być trzy składowe wektora położenia, lub trzy składowe wektora pędu. Dla elektronu, oprócz tych trzech liczb, trzeba by jeszcze dodać informacje o orientacji spinu, na przykład rzut spinu na dowolnie wybraną, ale ustaloną, oś z. Dla nukleonu, oprócz danych jak dla elektronu, potrzebna jest jeszcze informacja o ładunku, pozwalająca stwierdzić, czy mówimy o protonie, czy o neutronie.

Górny wskaźnik określa, która z N cząstek znajduje się w stanie x_k. Zauważmy, że wobec nierozróżnialności cząstek ta informacja, której użyjemy do budowy funkcji falowej, nie da się uzyskać doświadczalnie. Nałoży to warunki na możliwe funkcje falowe. Zgodnie z przyjętymi oznaczeniami funkcję falową N cząstek identycznych możemy zapisać w postaci $\Psi(x_1^{(1)}, \ldots, x_N^{(N)})$.

Uwzględnimy teraz fakt, że zamiana stanami dwu cząstek identycznych nie prowadzi do innego stanu N-cząstkowego. Zgodnie z interpretacją funkcji falowej to znaczy, że na skutek takiej zamiany funkcja falowa może co najwyżej zostać pomnożona przez stały czynnik. Przypuśćmy, że zamieniliśmy stanami cząstki i-tą i k-tą, wtedy

$$\Psi(x_1^{(1)}, \ldots, x_i^{(i)}, \ldots x_k^{(k)}, \ldots, x_N^{(N)}) = \epsilon \Psi(x_1^{(1)}, \ldots, x_k^{(i)}, \ldots, x_i^{(k)}, \ldots, x_N^{(N)}), \tag{12.1}$$

gdzie, jeśli umówimy się, że funkcje Ψ są znormalizowane do jedności, współczynnik ϵ jest stałą, równą co do wartości bezwzględnej jeden. Wzór (12.1) jest tylko zapisem faktu, że zamiana cząstek identycznych jest niewykrywalna doświadczalnie, czyli nie zmienia stanu układu. Metodami kwantowej teorii pola udowodniono jednak następujące twierdzenie. W przestrzeni trójwymiarowej, współczynnik ϵ zależy tylko od rodzaju cząstek identycznych opisywanych przez funkcję falową i może przyjmować tylko wartości $\epsilon = \pm 1$. Współczynnik ϵ nie zależy więc od tego, ile identycznych cząstek jest w układzie, które z nich zamieniamy, czy w jakich te cząstki są stanach. Twierdzenie uogólnia się na cztery i więcej wymiarów, natomiast nie jest prawdziwe w przestrzeni dwuwymiarowej, gdzie w związku z tym zachodzą różne ciekawe efekty, których tu nie będziemy omawiać.

Wykorzystując powyższe twierdzenie możemy przepisać wzór (12.1) w postaci

$$\Psi(x_1^{(1)}, \ldots, x_i^{(i)}, \ldots x_k^{(k)}, \ldots, x_N^{(N)}) = \pm \Psi(x_1^{(1)}, \ldots, x_k^{(i)}, \ldots, x_i^{(k)}, \ldots, x_N^{(N)}). \tag{12.2}$$

Cząstki, dla których stosuje się znak plus, nazywamy bozonami, a cząstki, dla których stosuje się znak minus, fermionami. Jak zobaczymy w dalszym ciągu rozdziału, własności fermionów istotnie różnią się od własności bozonów. Dlatego dobrze jest wiedzieć, które cząstki są bozonami, a które fermionami. Przydają się tu dwie reguły. Każda cząstka ma spin. Jak pokażemy w rozdziale siedemnastym, w przestrzeni trójwymiarowej, dla cząstek o masie różnej od zera, spin cząstki może przyjmować tylko wartości całkowite $s = 0, 1, 2, \ldots$ lub połówkowe $s = \frac{1}{2}, \frac{3}{2}, \ldots$ Znów metodami kwantowej teorii pola udowodniono, że wszystkie cząstki o spinie całkowitym są bozonami, a wszystkie cząstki o spinie połówkowym są fermionami. Ten wynik stosuje się także do cząstek o masie zero. Istnieją obszerne tablice cząstek, w których między innymi podane są ich spiny, więc ta reguła jest łatwa do stosowania. Są jednak typy cząstek (czy quasicząstek), których się nie umieszcza w takich tablicach. Przykładami są fonony, rotony, magnony itd. używane w fizyce fazy skondensowanej. Wtedy przydaje się następna reguła. Bozonami są cząstki, które mogą powstawać i znikać pojedynczo, jak foton absorbowany, czy emitowany przez atom. Fermionami są cząstki, które mogą powstawać i znikać tylko parami, jak pary e^+e^- w procesie anihilacji. Wśród najbardziej rozpowszechnionych cząstek, elektrony, nukleony i intensywnie teraz badane neutrina są fermionami. Natomiast cząstki pośredniczące w oddziaływaniach silnych (gluony), elektromagnetycznych (fotony), słabych (bozony pośredniczące $W^{\pm}$ i Z^0) i grawitacyjnych (grawitony) są bo-

zonami. Jądra atomowe, jeśli odległości między niemi są dużo większe od ich rozmiarów, często można traktować jako cząstki. W takim razie są fermionami, jeśli składają się z nieparzystej liczby nukleonów i bozonami, jeśli składają się z parzystej liczby nukleonów. To powoduje, na przykład, że widmo rotacyjne cząsteczki $^{16}O - ^{16}O$ jest zupełnie inne od widma rotacyjnego cząsteczki $^{16}O - ^{17}O$.

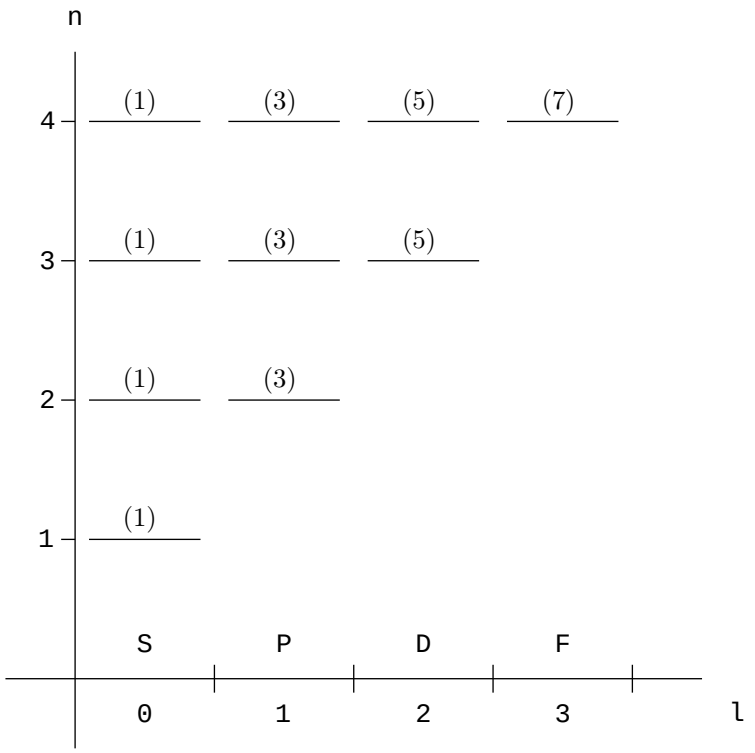
Dla fermionów wzór (12.2) oznacza, że funkcja falowa jest antysymetryczna względem zamiany cząstek, czyli zmienia znak przy takiej zamianie. Szczególnie ciekawy jest przypadek, kiedy dwa lub więcej ze stanów cząstek są identyczne. Przypuśćmy, że we wzorze (12.2) zachodzi $x_i = x_k$. Wtedy funkcje Ψ po obu stronach równości są identyczne, a zarazem na mocy równości muszą mieć przeciwne znaki, co może być prawdą tylko jeśli

$$\Psi(x_1^{(1)}, \ldots, x_k^{(i)}, \ldots x_k^{(k)}, \ldots, x_N^{(N)}) = 0. \tag{12.3}$$

Ten wzór wyraża zakaz Pauli'ego. Dwa identyczne fermiony nie mogą znajdować się w tym samym stanie. Zakaz Pauli'ego ma podstawowe znaczenie przy opisie układów złożonych z wielu identycznych fermionów, jak atomy, cząsteczki czy jądra atomowe. W roku 1945 Pali dostał nagrodę Nobla "za odkrycie zasady wykluczania zwanej także zakazem Pauli'ego". Należy jednak pamiętać, że zakaz Pauli'ego jest tylko jednym z wniosków z bardziej ogólnego twierdzenia o antysymetrii funkcji falowej wielu identycznych fermionów. To uogólnienie wprowadził Dirac.

Jako przykład zastosowania zakazu Pauli'ego omówimy układ okresowy pierwiastków, czyli tak zwaną tablicę Mendelejewa. Dokładna teoria jest bardzo skomplikowana, więc ograniczymy się do prostego modelu. W dokładnym hamiltonianie występuje energia oddziaływania między elektronami i energia oddziaływania elektronów z jądrem. Jako uproszczenie założymy, że elektrony nie oddziałują między sobą, ale poruszają się w pewnym średnim potencjale pochodzącym od jądra i od chmury elektronowej. Założymy też, że w pierwszym przybliżeniu ten potencjał jest podobny do potencjału kulombowskiego w atomie wodoru i jonach wodoropodobnych. Takie przybliżenia często się stosuje w fizyce. Mówi się, że prawdziwy hamiltonian zastępujemy hamiltonianem efektywnym, który powinien być wybrany tak, żeby dawał wyniki możliwie bliskie do prawdziwych w zakresie, który nas interesuje. Schemat poziomów energetycznych elektronu w polu kulombowskim jest pokazany na rysunku 12.1. Energia zależy tylko od liczby kwantowej n. Dla danej pary liczb kwantowych $n, l < n$, poziom energetyczny jest zdegenerowany $2l+1$ razy. Stany, dla których $l = 0, 1, 2, 3, \ldots$, nazywa się tradycyjnie stanami $S, P, D, F, \ldots$. Liczbę n dopisujemy przed literą definiującą wartość l. Na przykład stan $l = 2$ z $n = 4$ oznacza się jako stan $4D$. Musimy jeszcze uwzględnić spin elektronu. Zakładamy, że każdy ze stanów przedstawionych na rysunku 12.1 występuje w dwu wersjach: ze spinem skierowanym do góry i ze spinem skierowanym w dół.

Pierwsze dwa elektrony można umieścić w stanach $n = 1$. Na rysunku jest tylko jeden taki stan, ale ta liczba się podwaja ze względu na spin elektronu. Trzeci elektron musi już obsadzić stan z $n = 2$ o znacznie wyższej energii. Zgodnie z tym, w pierwszym wierszu tablicy Mendelejewa mamy dwa atomy: wodór z jednym elektronem i hel z dwoma. Dla helu wypełniona jest powłoka $1S$, czyli elektrony zajmują wszystkie miejsca na poziomie $n = 1$. Dla $n = 2$ jest osiem miejsc (dwa razy cztery) i rzeczywiście w drugim wierszu tablicy Mendelejewa



Rysunek 12.1: Stany atomu wodoru dla $n = 1, 2, 3, 4$ i $l = 0, 1, 2, 3$. Energia zależy tylko od liczby kwantowej n. Dla danych n, l kreska oznacza $2l + 1$ stanów, co przypomina liczba w nawiasie nad kreską. Nad osią l podane są spektroskopowe oznaczenia stanów z $l = 0, 1, 2, 3$.

jest osiem atomów od litu do neonu. Dla atomu neonu wypełniona jest druga powłoka elektronowa, która zawiera dwa elektrony $2S$ i sześć elektronów $2P$. Zauważmy, że cała ta konstrukcja opiera się na zakazie Pauli'ego – w jednym stanie wolno nam umieścić co najwyżej jeden elektron. W trzecim wierszu pojawia się pierwsza trudność. Spodziewamy się osiemnastu atomów, a jest ich osiem tak, jak w wierszu drugim. Ten wynik nie jest zaskakujący. Degeneracja $(2l+1)$-krotna stanów przy danych n i l wynika z symetrii sferycznej atomu i powinna być zachowana także dla hamiltonianu efektywnego. Natomiast niezależność energii przy danym n od liczby kwantowej l, jest szczególną cechą potencjału kulombowskiego i nie ma powodu oczekiwać, że będzie się stosować dokładnie także dla hamiltonianu efektywnego. Dane wskazują więc, że energia stanów $3D$ jest na tyle wysoka, że powłoka zamyka się bez nich i kończący ten wiersz atom argonu ma w zewnętrznej powłoce elektrony $3S$ i $3P$. W następnym wierszu jest osiemnaście atomów i zewnętrzna powłoka atomu kryptonu, który kończy ten wiersz tabeli, zawiera dwa elektrony $4S$, sześć elektronów $4P$ i dziesięć elektronów $3D$. Energie poziomów $4D$ i $4F$ przesunęły się do góry. W następnym, piątym wierszu dochodzą elektrony $5S$, $5P$ i $4D$. W szóstym wierszu obserwujemy nowe ciekawe zjawisko. Dochodzą elektrony $6S$, $6P$,$5D$ i $4F$. Elektrony $4F$ pojawiają się tak późno, bo ich energie są znacznie wyższe niż byłyby, gdyby potencjał był kulombowski. Niemniej, odpowiadają one liczbie kwantowej $n=4$ i w związku z tym znajdują się wyraźnie bliżej jądra niż pozostałe elektrony, które doszły w tym wierszu. Własności chemiczne atomu zależą głównie od jego zewnętrznych elektronów. Dlatego 15 pierwiastków różniących się między sobą tylko liczbą elektronów $4F$, $n_f = 0, \ldots, 14$, ma prawie identyczne własności chemiczne. To są pierwiastki ziem rzadkich, których rozdzielenie sprawiło kiedyś chemikom dużo kłopotu. Sprawę dodatkowo komplikuje fakt, że poziomy energetyczne $4F$ i $6S$ są bardzo bliskie siebie. W następnym, jak dotąd ostatnim, wierszu tablicy Mendelejewa rzecz się powtarza i odpowiednikiem pierwiastków ziem rzadkich są aktynowce, ale to już są bardzo ciężkie pierwiastki promieniotwórcze i związane z niemi problemy chemiczne są zupełnie inne. Zauważmy, że gdyby nie zakaz Pauli'ego, wszystkie elektrony umieszczały by się na poziomie $1S$ i Świat – taki jak go znamy – nie mógłby istnieć.

W badaniach nad układami wielu identycznych cząstek ważną rolę odgrywa sposób liczenia stanów układu. Mówi się o trzech statystykach: statystyce Boltzmanna dla cząstek rozróżnialnych, statystyce Fermi'ego-Diraca dla fermionów i statystyce Bosego-Einsteina dla bozonów. W klasycznej fizyce nie ma pojęcia cząstek identycznych, w takim sensie jak tu omawiamy, i zawsze stosuje się statystyka Boltzmanna. Rozważmy najpierw dwie cząstki rozróżnialne w dwu wzajemnie ortogonalnych stanach 1 i 2, na przykład w stanie podstawowym i w pierwszym stanie wzbudzonym oscylatora harmonicznego. Jeżeli funkcję falową cząstki w pierwszym z tych stanów oznaczymy $\psi_1(x)$ i cząstki w drugim $\psi_2(x)$, to stan tej pary cząstek opisuje funkcja falowa $\psi_1(x^{(1)})\psi_2(x^{(2)})$. Ten stan, w którym cząstka pierwsza jest w stanie pierwszym, a cząstka druga w stanie drugim, jest różny od stanu $\psi_1(x^{(2)})\psi_2(x^{(1)})$, w którym cząstka pierwsza jest w stanie drugim, a cząstka druga w stanie pierwszym. To jest sposób liczenia stanów, do którego jesteśmy przyzwyczajeni w fizyce klasycznej. Żadna z tych funkcji falowych pary cząstek nie może jednak opisywać pary cząstek identycznych, bo żadna nie spełnia warunku 12.2. Spełniają natomiast ten warunek, odpowiednio dla bozonów i dla fermionów, funkcje

$$\Psi_{BE}(x^{(1)}, x^{(2)}) = \frac{1}{\sqrt{2}}(\psi_1(x^{(2)})\psi_2(x^{(1)} + \psi_1(x^{(1)})\psi_2(x^{(2)}), \tag{12.4}$$

$$\Psi_{FD}(x^{(1)}, x^{(2)}) = \frac{1}{\sqrt{2}}(\psi_1(x^{(2)})\psi_2(x^{(1)} - \psi_1(x^{(1)})\psi_2(x^{(2)}). \tag{12.5}$$

Z tych funkcji nie da się już jednak odczytać, która cząstka jest w stanie 1, a która w stanie 2. Stosując probabilistyczną interpretację mechaniki kwantowej możemy tylko stwierdzić, że każda z dwu cząstek znajduje się z równym prawdopodobieństwem w każdym z tych dwu stanów. Wobec tego, trzeba się przyzwyczaić, że mając dane dwie cząstki identyczne i dwa stany możemy utworzyć tylko jeden stan dwucząstkowy, a nie dwa jak w fizyce klasycznej. Przykład z czterema stanami i czterema cząstkami identycznymi przedstawiony jest na rysunku 12.2. Na rysunku każda z kratek oznacza stan jednocząstkowy i każda kropka cząstkę. W części (b) cztery cząstki znajdują się w jednym stanie. Według statystyk Boltzmanna i Bosego-Einsteina odpowiada tej sytuacji dokładnie jeden stan czterocząstkowy. Według statystyki Fermiego-Diraca takiego stanu czterocząstkowego nie ma ze względu na zakaz Pauli'ego. W części (a) w każdym ze stanów jednocząstkowych znajduje się dokładnie jedna cząstka. Według statystyk "kwantowych" Fermiego-Diraca i Bosego-Einsteina, istnieje dokładnie jeden stan czterocząstkowy odpowiadający tej sytuacji. Według statystyki Boltzmanna jest takich stanów $4! = 24$, bo na tyle sposobów można rozmieścić cztery cząstki rozróżnialne w czterech stanach tak, żeby w każdym stanie była dokładnie jedna cząstka. Jeżeli każdy stan czterocząstkowy jest jednakowo prawdopodobny, jak się to często zakłada w fizyce statystycznej, to stosunek prawdopodobieństwa stanu (b) do prawdopodobieństwa stanu (a) wynosi zero dla statystyki Fermiego-Diraca, jeden dla statystyki Bosego-Einsteina i $\frac{1}{24}$ dla statystyki Boltzmanna. Jeżeli te stany nie są dokładnie równoprawdopodobne, to i tak w porównaniu do statystyki klasycznej prawdopodobieństwo fluktuacji polegającej na tym, że wszystkie cząstki spotkają się w jednym stanie jednocząstkowym, zostaje znacznie zwiększone dla bozonów i zredukowane do zera dla fermionów. O konsekwencjach zakazu Pauli'ego dla fermionów mówiliśmy już. Zwróćmy teraz uwagę na bozony.

W bardzo niskich temperaturach, ciekły hel ^4He, którego atomy są bozonami, wykazuje nadciekłość, to znaczy płynie bez lepkości. Można to jakościowo zrozumieć w oparciu o powyższą dyskusję. Atomy płynącego helu mają w przybliżeniu wszystkie ten sam pęd. Działanie lepkości polegałoby na zmianie pędu niektórych atomów, ale dla bozonów prawdopodobieństwo tego, że wszystkie atomy są w tym samym stanie jest znacznie większe niż w fizyce klasycznej, stąd działanie lepkości jest znacznie słabsze niż można by się klasycznie spodziewać. Podobnie można tłumaczyć (jakościowo) brak oporu elektrycznego w nadprzewodnikach. Wprawdzie elektrony są fermionami, ale w nadprzewodniku są skorelowane w dwuelektronowe "pary Coopera", które są (w przybliżeniu) bozonami.

Innym przykładem ciekawych własności bozonów jest kondensacja Einsteina. Rozważmy gaz doskonały złożony z identycznych bozonów. Licząc stany według statystyki Bosego-Einsteina i robiąc zwykłe przybliżenia polegające na zastępowaniu sum przez całki, otrzymuje się dla liczby cząstek gazu o energii niższej niż E wzór

$$N(E) = AV \int_0^E dx \frac{\sqrt{x}}{fe^x - 1}, \tag{12.6}$$

gdzie energia jest liczona w jednostkach $k_B T$, A jest stałą, V oznacza objętość, a lotność f należy dobrać tak, żeby całkując po całym zakresie energii od zera do nieskończoności uzyskać całkowitą liczbę cząstek w układzie. Jeżeli f jest bardzo duże, jedynkę w mianowniku można zaniedbać i otrzymujemy wzór znany z klasycznej fizyki. Im mniejsza wartość lotności f, tym większa wychodzi liczba cząstek w układzie, ale lotność f nie może być mniejsza od jedynki, bo wtedy całka staje się rozbieżna. Wykresy funkcji $N(E)$ dla $f = 1.1$, $f = 1.01$ i $f = 1$ są pokazane na rysunku 12.3 Widać, że kiedy f osiąga minimalną wartość jeden, otrzymuje się pewną skończoną graniczną liczbę cząstek w układzie i nie jest jasne, jak opisywać układ, kiedy liczba cząstek jest większa od tej granicy. Einstein zauważył jednak, że w jednym tylko stanie o energii zero (cząstki nieruchome) jest miejsce na

$$n_0 = \frac{1}{f - 1} \tag{12.7}$$

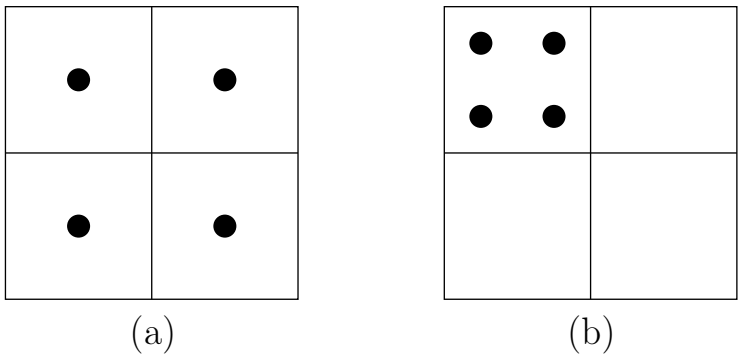
cząstek, a więc można ich tam pomieścić dowolnie dużo, kiedy f zmierza do jedności. Wkład od tego stanu jest zaniedbywany, kiedy się przybliża sumy przez całki. Jak będzie więc wyglądał poprawny rozkład energii cząstek? Póki lotność f nie jest bardzo bliska jedności, wkład od stanu zerowego jest zaniedbywalny i można używać krzywych jak na rysunku 12.3. Na przykład dla $f = 1.01$, na poziomie zerowym znajduje się (średnio!) sto cząstek, co jest całkowicie zaniedbywalną częścią liczby 6×10^{23} cząstek w molu gazu. Jeżeli jednak liczba cząstek w układzie jest znacząco większa niż maksymalna liczba, którą można otrzymać ze wzoru całkowego, to cały nadmiar siedzi w stanie o energii zerowej (i sąsiednich, bo rozkład energii cząstki swobodnej jest ciągły) i może być dowolnie duży. Wtedy cząstki dzielą się na dwie grupy: część ma rozkład odpowiadający rysunkowi 12.3 dla $f = 1$, a reszta, zwana kondensatem, ma energię równą zeru, lub prawie. Przechodzenie cząstek do kondensatu jest znane jako kondensacja Einsteina. Należy zwrócić uwagę na analogię z kondensacją pary wodnej w ciekłą wodę przy skraplaniu. Dla skraplania jednak, kondensacja zachodzi w zwykłej przestrzeni, a dla kondensacji Einsteina w przestrzeni pędów. Kondensacja Einsteina jest ważnym modelem przejścia fazowego.

Prostsze przykłady kondensacji Einsteina otrzymujemy w przypadkach, kiedy widmo energii jest dyskretne. Przypuśćmy, na przykład, że gaz omawiany powyżej znajduje się w polu sił o potencjale $V(\vec{x}) = \frac{1}{2} K \vec{x}^2$. To jest potencjał trójwymiarowego oscylatora harmonicznego, więc poziomy energii dane są wzorem (11.12). Średnia liczba cząstek w danym stanie i o energii E_i dana jest znanym z fizyki statystycznej wzorem

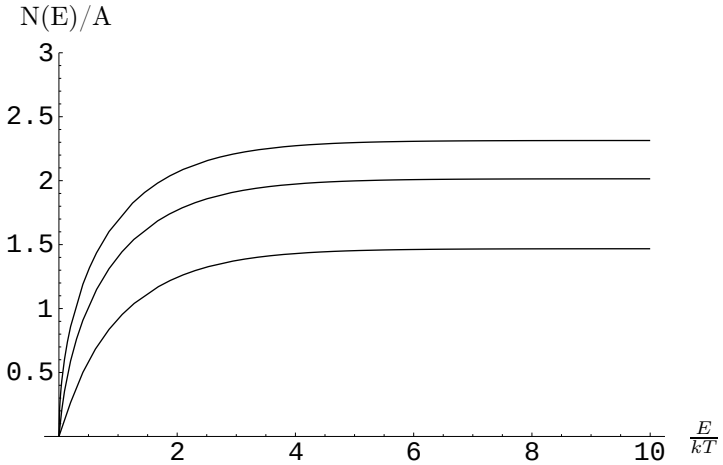
$$\langle n_i \rangle = \frac{1}{fe^{\frac{E_i}{k_B T}} - 1}, \tag{12.8}$$

gdzie lotność f należy wyznaczyć ze wzoru na całkowitą liczbę cząstek N:

$$\sum_i \langle n_i \rangle = N. \tag{12.9}$$



Rysunek 12.2: Dwa rozkłady czterech cząstek w czterech stanach



Rysunek 12.3: Wykres funkcji $N(E)$ zdefiniowanej w tekście. Krzywe licząc od dołu odpowiadają wartościom lotności:$f = 1.1$, $f = 1.01$ i $f = 1$.

Kiedy przy stałej temperaturze lotność maleje, w każdym ze stanów (średnio) przybywa cząstek. Ale lotność nie może być mniejsza niż e^{-E_0/k_BT}, bo wyszłaby ujemna liczba cząstek w stanie podstawowym. Przy lotności dążącej z góry do tej granicy, średnia liczba cząstek w stanie podstawowym dąży do nieskończoności. Średnie liczby cząstek w pozostałych stanach, natomiast, dążą do granic skończonych i to tak, że liczba $N - \langle n_0 \rangle$ też dąży do skończonej granicy. Otrzymywanie i badanie różnych kondensatów Einsteina jest obecnie ważnym działem optyki.

Rozdział 13

Metoda zaburzeń

W mechanice kwantowej, podobnie jak w mechanice klasycznej, tylko bardzo nieliczne problemy da się dokładnie rozwiązać analitycznie. Zwykle trzeba stosować metody przybliżone. Jedną z najczęściej używanych metod przybliżonych jest metoda zaburzeń, nazywana także metodą perturbacji. Tę metodę stosuje się wtedy, kiedy hamiltonian układu da się zapisać w postaci

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1, \tag{13.1}$$

gdzie dla "hamiltonianu niezaburzonego" $\hat{H}_0$ znamy dokładne wartości własne ε_n i odpowiadające im wektory własne $|n\rangle$ spełniające równania

$$\hat{H}_0|n\rangle = \varepsilon_n|n\rangle, \tag{13.2}$$

λ jest parametrem liczbowym i "zaburzenie" $\lambda\hat{H}_1$ jest małe w tym sensie, że powoduje tylko nieznaczne zmiany rozwiązań niezaburzonych. Jeśli jakaś wartość własna ε_n jest zdegenerowana, to odpowiada jej nieskończenie wiele różnych wektorów własnych stanowiących przestrzeń wektorową (rozdział dziesiąty). W takim przypadku do zbioru wektorów $|n\rangle$ włączamy tylko jakąś bazę ortonormalną tej przestrzeni.

Na przykład, zewnętrzne pole elektryczne powoduje przesunięcia i rozszczepienia poziomów energetycznych atomów (efekt Starka). Dla atomu wodoru w jednorodnym polu elektrycznym o natężeniu $\mathcal{E}$ woltów na centymetr, skierowanym wzdłuż osi z, hamiltonian ma postać

$$\hat{H} = -\frac{\hbar^2}{2m}\triangle - \frac{e^2}{r} + e\mathcal{E}z. \tag{13.3}$$

Równania Schrödingera z tym hamiltonianem nie da się rozwiązać analitycznie, ale jeżeli za hamiltonian niezaburzony wybierzemy hamiltonian atomu wodoru bez pola zewnętrznego

$$\hat{H}_0 = -\frac{\hbar^2}{2m}\triangle - \frac{e^2}{r} \tag{13.4}$$

i jeśli natężenie pola $\mathcal{E}$ jest dostatecznie małe, to można próbować zastosować metodę zaburzeń. Żeby nabrać wyczucia, co tu znaczy, że natężenie pola jest

małe, załóżmy, że $\mathcal{E} = 10^4$V/cm. Średnica atomu wodoru w stanie podstawowym wynosi około 10^{-8} cm, więc wywołane zewnętrznym polem zmiany energii przy przesunięciach elektronu wewnątrz atomu są rzędu 10^{-4} elektronowolta. Energia przejścia między stanem podstawowym i pierwszym stanem wzbudzonym atomu wodoru wynosi kilka elektronowoltów, więc wkład zewnętrznego pola do energii, przy tym natężeniu pola, jest rzeczywiście niewielki. Na tej podstawie można mieć nadzieję, że metoda zaburzeń da dobre wyniki, ale jak zobaczymy z przykładów w dalszej części tego rozdziału, pewności nie ma.

Wartości własne i wektory własne pełnego hamiltonianu $\hat{H}$ zależą naogół od parametru λ. Podstawowym założeniem metody zaburzeń jest, że te zależności są analityczne w otoczeniu punktu $\lambda = 0$, to znaczy że można napisać tak zwane rozwinięcia perturbacyjne

$$E_n(\lambda) = \sum_{k=0}^{\infty} E_n^{(k)} \lambda^k, \tag{13.5}$$

$$\psi_n(x,\lambda) = \sum_{k=0}^{\infty} \psi_n^{(k)}(x) \lambda^k. \tag{13.6}$$

Szeregi po prawych stronach tych wzorów powinny być zbieżne, i to szybko, bo zwykle nie potrafimy wyliczyć więcej niż parę pierwszych wyrazów. Dla $\lambda = 0$ mamy problem niezaburzony, więc na podstawie równania (13.2) możemy napisać

$$E_n^{(0)} = \varepsilon_n, \tag{13.7}$$

$$\psi_n^{(0)}(x) = \langle x|n\rangle. \tag{13.8}$$

Metody obliczania dalszych współczynników zostaną podane w następnych dwu rozdziałach. Tu przedyskutujemy trzy przykłady ilustrujące stosowalność metody zaburzeń w różnych sytuacjach.

Przykład 1: Jednowymiarowy oscylator harmoniczny zaburzony stałą siłą.

Przyjmijmy

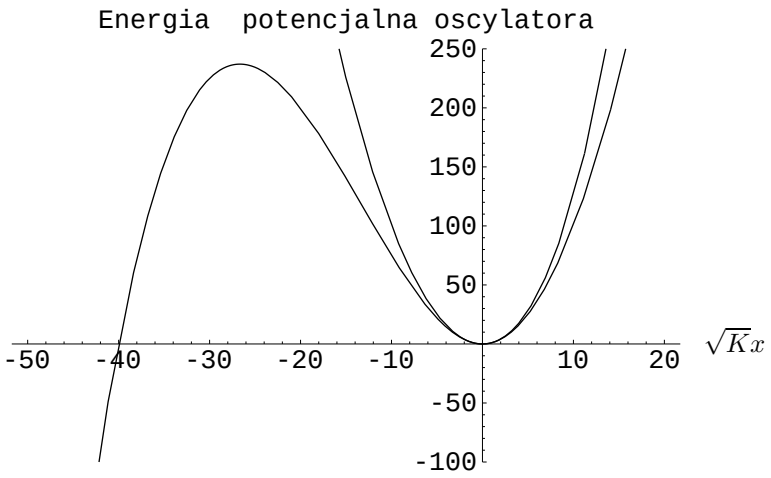
$$H_0 = \frac{p^2}{2m} + \frac{1}{2}Kx^2, \tag{13.9}$$

$$\lambda H_1 = \lambda K x. \tag{13.10}$$

Fizycznie jest to jednowymiarowy oscylator harmoniczny zaburzony stałą siłą. Pełny hamiltonian możemy po prostych przekształceniach zapisać w postaci

$$H = \frac{p^2}{2m} + \frac{1}{2}K(x+\lambda)^2 - \frac{K\lambda^2}{2}. \tag{13.11}$$

Zaburzenie nie zmieniło więc masy, kształtu potencjału ani stałej siłowej oscylatora, natomiast przesunęło minimum potencjału do punktu $x = -\lambda$ i spowodowało dodanie do energii potencjalnej przyczynku $-\frac{K\lambda^2}{2}$. Jeżeli dla problemu niezaburzonego wartości własne energii były ε_n, to dla problemu zaburzonego



Rysunek 13.1: Energia potencjalna oscylatora harmonicznego niezaburzonego (linia symetryczna względem osi $x = 0$) i tego samego oscylatora zaburzonego członem anharmonicznym $0.025Kx^3$ (linia przesunięta w lewo, dla $x > 0$ w górę, a dla $x < 0$ w dół).

$$E_n = \varepsilon_n - \frac{K}{2}\lambda^2, \tag{13.12}$$

bo samo przesunięcie potencjału bez zmiany kształtu nie wpływa na położenie poziomów energetycznych. Widać, że w tym przypadku rozwinięcie perturbacyjne istnieje dla każdego z poziomów energetycznych i jest zbieżne dla każdej wartości parametru λ. Dla każdego n różne od zera są tylko dwa współczynniki: $E_n^{(0)} = \varepsilon_n$ i $E_n^{(2)} = -\frac{K}{2}$. Jeżeli dla problemu niezaburzonego funkcje falowe $\langle x|n\rangle$ odpowiadające wartościom własnym ε_n oznaczymy $\phi_n(x)$, to odpowiadające im funkcje własne dla problemu zaburzonego maja postać $\phi_n(x + \lambda)$, bo dodanie stałej do energii potencjalnej układu nie wpływa na funkcje falowe. Korzystając ze znanych wzorów (7.15) na funkcje własne jednowymiarowego oscylatora harmonicznego, można sprawdzić, że rozwinięcia perturbacyjne dla funkcji falowych też istnieją dla każdego n i są zbieżne dla każdej wartości parametru λ.

Przykład 2: Jednowymiarowy oscylator anharmoniczny.

Rozważmy teraz pozornie bardzo podobny problem, gdzie hamiltonian niezaburzony jest znów hamiltonianem oscylatora harmonicznego, ale zaburzenie ma postać

$$\lambda\hat{H}_1 = \lambda K x^3; \qquad \lambda > 0. \tag{13.13}$$

Zależność od x energii potencjalnej, zaburzonej i niezaburzonej, jest pokazana na rysunku 1. Dla małych wartości $|x|$ energia potencjalna w problemie zaburzonym niewiele się różni od energii potencjalnej bez zaburzenia, bo $|x^3| \ll x^2$. Jeżeli parametr λ jest mały, to ten obszar, gdzie zaburzenie nieznacznie tylko zmienia energie potencjalną, może być całkiem szeroki, ale w końcu, dla dostatecznie dużych wartości $|x|$, zaburzenie zaczyna być większe co do modułu od niezaburzonej energii potencjalnej. W szczególności, dla $x \to -\infty$ potencjał

niezaburzony zmierza do plus nieskończoności, a potencjał zaburzony do minus nieskończoności. Wynika z tego, że układ zaburzony, w przeciwieństwie do układu niezaburzonego, nie ma stanu podstawowego o najniższej energii. Jakkolwiek niską energię miałby jakiś stan zaburzony, zawsze można uzyskać stan o energii jeszcze niższej, przesuwając cząstkę dalej na lewo. Dla tego szeregi perturbacyjne nie mogą być zbieżne do poprawnych wartości własnych energii i odpowiadających im funkcji własnych – w tym przykładzie szeregi perturbacyjne są rozbieżne dla wszystkich n i dla wszystkich wartości λ! Mogło by się wydawać, że w takim razie metoda zaburzeń nic nie daje, ale to nie musi tak być. Jeżeli obszar, w którym wpływ zaburzenia jest niewielki, jest dostatecznie duży, to okazuje się, że istnieje stan "rezonansowy" o energii bliskiej energii niezaburzonego stanu podstawowego, a czasem także stany rezonansowe o energiach bliskich energiom niezbyt wysokich niezaburzonych stanów wzbudzonych. Te stany nie są stanami o najniższej energii, nie są też stanami trwałymi. Po jakimś czasie cząstka znajdująca się początkowo w takim stanie ucieknie do obszaru dużych ujemnych x-ów. Ale czas życia takich stanów rezonansowych może być całkiem długi i można ich energie wyznaczać doświadczalnie. Wrócimy do tego zagadnienia w rozdziale trzydziestym. Okazuje się, że niekiedy szeregi perturbacyjne, chociaż są rozbieżne, mogą być wykorzystane do oszacowań energii i funkcji falowych takich stanów rezonansowych. Dzieje się tak wtedy, kiedy te szeregi są "zbieżne asymptotycznie". Kiedy sumujemy wyraz po wyrazie szereg asymptotycznie zbieżny, początkowo jest tak, jak dla szeregów zbieżnych – wartości bezwzględne kolejnych wyrazów maleją, choć nie zawsze monotonicznie, sumy częściowe stopniowo zbliżają się do dokładnego wyniku. Ale tak jest tylko z początku. Potem wyrazy znowu rosną i wartości sum częściowych oddalają się od wyników dokładnych. Dobrze to oddaje nazwa rosyjska takich szeregów: szeregi zbieżne z początku. Warto zapamiętać, że sukcesy metody zaburzeń w praktyce nie zawsze wiążą się z prawdziwą zbieżnością szeregów perturbacyjnych. Na przykład atom wodoru umieszczony w jednorodnym polu elektrycznym staje się nietrwały – po jakimś czasie następuje jonizacja. Wobec tego poziomy energetyczne, które znajdujemy badając efekt Starka, są poziomami rezonansowymi i szeregi perturbacyjne nie mogą być zbieżne. Mimo to, energie wyliczone perturbacyjnie, dla niezbyt silnych pól, dobrze zgadzają się z doświadczeniem, bo szeregi są asympotyczne i w tym przypadku pozwalają podejść bardzo blisko do dokładnych wartości energii stanów rezonansowych. Podobnie w elektrodynamice kwantowej, szeregi perturbacyjne są rozbieżne, a mimo to perturbacyjnie, używając grafów Feynmana, uzyskuje się niezwykle dokładne wyniki.

Przykład 3: Stany zdegenerowane — przykład klasyczny

Przedyskutujemy teraz przypadek, kiedy metodę zaburzeń stosujemy do stanu zdegenerowanego. Ten przypadek jest bardzo ważny w praktyce. Na przykład, jeśli chcielibyśmy stosować metodę zaburzeń do atomu wodoru w jednorodnym polu elektrycznym, tak jak to opisaliśmy na początku tego rozdziału, musielibyśmy uwzględnić fakt, że wszystkie stany niezaburzone oprócz stanu podstawowego są zdegenerowane. Dyskusję teorii kwantowej odłożymy do rozdziału piętnastego, a tu wyjaśnimy podstawową trudność na przykładzie z fizyki klasycznej. Igła kompasu w polu magnetycznym Ziemi wskazuje na północ. Jeżeli to pole magnetyczne uznamy za zaburzenie, to dla układu niezaburzonego – kompas bez pola magnetycznego – wszystkie orientacje igły odpowiadają tej samej energii, mamy więc degeneracyję. Rozważmy teraz stan igły jako

funkcję natężenia pola magnetycznego Ziemi. Oczywiście w rzeczywistości to natężenie ma jakąś jedną dobrze określoną wartość, ale w metodzie zaburzeń tak się robi – rozważa się zależność zaburzenia od wielkości członu zaburzającego (parametru λ) i dopiero na końcu podstawia się interesująca nas wielkość tego zaburzenia. Jakkolwiek słabe byłoby pole magnetyczne Ziemi, igła wskazuje na północ. Jeżeli szeregi perturbacyjne mają być zbieżne, to granica rozwiązania zaburzonego przy zaburzeniu dążącym do zera, powinna być równa rozwiązaniu niezaburzonemu. Przypuśćmy, że jako stan niezaburzony wybraliśmy stan z igłą wskazującą na wschód. Z punktu widzenia układu niezaburzonego ten wybór jest równie dobry jak każdy inny. Przy tym wyborze jednak, metoda zaburzeń się nie stosuje. Stan niezaburzony nie przechodzi, przy włączeniu bardzo słabego pola magnetycznego Ziemi, w bliski mu stan zaburzony – po włączeniu dowolnie słabego pola, następuje skok igły o 90^o. Jeżeli natomiast jako stan niezaburzony wybierzemy stan z igłą wskazującą na północ, to metoda zaburzeń stosuje się bardzo dobrze. Po włączeniu pola magnetycznego orientacja igły magnetycznej w ogóle się nie zmienia, a energia układu rośnie liniowo z natężeniem przyłożonego pola. Widzimy więc, że po to żeby stosować metodę zaburzeń do stanów zdegenerowanych, trzeba umieć odpowiednio wybrać stany niezaburzone. Mówi się, że trzeba wybrać stany niezaburzone "dopasowane do zaburzenia". To jest ważny problem fizyczny, który omówimy w rozdziale 15.

Przykład 4: Stany zdegenerowane — przykład kwantowy

Zaburzenia stanów zdegenerowanych są tak ważne, że zilustrujemy je jeszcze jednym przykładem. Rozważmy problem własny

$$\begin{pmatrix} 1 & \lambda \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix} = E_n \begin{pmatrix} a_n \\ b_n \end{pmatrix}. \tag{13.14}$$

To jest przybliżone równanie Schrödingera dla nieruchomej cząstki o spinie $\frac{1}{2}$ w polu magnetycznym skierowanym wzdłuż osi x (rozdział 20). Pole o zerowym natężeniu odpowiada problemowi niezaburzonemu. W problemie niezaburzonym $\hat{H}_0 = \hat{1}$, więc $\varepsilon_n = 1$ i każdy wektor jest wektorem własnym hamiltonianu. Wartość własna jest dwukrotnie zdegenerowana. Jako bazę w przestrzeni rozwiązań można wybrać na przykład wektory

$$\phi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \phi_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{13.15}$$

Fizycznie, te stany odpowiadają rzutom spinu $\pm\frac{1}{2}\hbar$ na oś z (rozdział 19). Dokładne rozwiązanie problemu własnego daje

$$E_0 = 1+\lambda, \quad \psi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \qquad E_1 = 1-\lambda, \quad \psi_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{13.16}$$

Warunkiem stosowalności metody zaburzeń jest żeby w granicy $\lambda \to 0$ rozwiązanie zaburzone przechodziło w rozwiązanie niezaburzone. Dla rozwiązań niezaburzonych (13.15) ten warunek nie jest spełniony, więc metoda zaburzeń się nie stosuje. Możemy natomiast wybrać funkcje niezaburzone "dopasowane do zaburzenia"

$$\chi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \chi_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{13.17}$$

Przy tym wyborze metoda zaburzeń już w pierwszym przybliżeniu daje dokładne rozwiązanie. Fizycznie, funkcje χ_i odpowiadają stanom z rzutami spinu $\pm\frac{1}{2}\hbar$ na kierunek pola. Widać tu bliską analogię z przykładem klasycznym omówionym poprzednio.

Rozdział 14

Metoda Rayleigha-Schrödingera

Metoda Rayleigha-Schrödingera jest najprostszą wersją metody zaburzeń. Stosuje się do stanów niezdegenerowanych, więc nie ma problemu szukania stanów niezaburzonych dostosowanych do zaburzenia, i nie stosuje sztuczek przyspieszających zbieżność szeregów perturbacyjnych. Takie sztuczki, stosowane w innych metodach zaburzeń, w pewnych wypadkach pozwalają uzyskać lepsze wynik przy mniejszym wkładzie pracy, ale nie będziemy ich tu omawiać. Problem polega na znalezieniu rozwinięć perturbacyjnych dla rozwiązań równania (patrz poprzedni rozdział)

$$(\hat{H}_0 + \lambda\hat{V})|\psi_n\rangle = E_n|\psi_n\rangle. \tag{14.1}$$

Z założenia, znane są rozwiązania równania niezaburzonego, które można zapisać w dwu równoważnych postaciach

$$\hat{H}_0|n\rangle = \varepsilon_n|n\rangle, \tag{14.2}$$
$$\langle n|\hat{H}_0 = \varepsilon_n\langle n|. \tag{14.3}$$

Drugie z tych równań powstaje z pierwszego przez sprzężenie hermitowski (przejście do wektorów dualnych). Zakładamy, że układ wektorów $|n\rangle$ stanowi bazę ortonormalną w przestrzeni możliwych stanów układu. Mnożąc równanie (14.1) lewostronnie przez wektor dualny $\langle k|$ i wykorzystując równania niezaburzone, otrzymujemy układ równań (po jednym dla każdego k) równoważny z równaniem (14.1)

$$\varepsilon_k\langle k|\psi_n\rangle + \lambda\langle k|\hat{V}|\psi_n\rangle = E_n\langle k|\psi_n\rangle. \tag{14.4}$$

Żeby skrócić wzory, wprowadza się zwykle oznaczenie

$$V_{kn} \equiv \langle k|\hat{V}|n\rangle = \int dx\psi_k^*(x)\hat{V}(x)\psi_n(x). \tag{14.5}$$

Drugą z tych równości otrzymuje się przez wstawienie jedynek zbudowanych z wektorów własnych operatora położenia $|x\rangle$ przy założeniu, że operator $\hat{V}$ jest

lokalny (rozdział siódmy). Na mocy tej definicji $V_{kn} = V_{nk}^*$. Przyjmiemy teraz założenie, że

$$\langle n|\psi_n\rangle = 1, \tag{14.6}$$

lub równoważnie

$$|\psi_n\rangle = |n\rangle + \sum_{k\neq n} c_{nk}|k\rangle. \tag{14.7}$$

Ponieważ równanie (14.1) nie określa normalizacji wektora $|\psi_n\rangle$, ten warunek nie jest założeniem fizycznym, tylko konwencją, która okaże się wygodna w dalszych rachunkach. Ściśle mówiąc, przyjęliśmy tu założenie, że współczynnik przy wektorze $|n\rangle$ w rozwinięciu wektora $|\psi_n\rangle$ nie jest równy zero. Ponieważ jednak w warunkach stosowalności metody zaburzeń wektor $|\psi_n\rangle$ ma być bliski odpowiadającemu mu wektorowi niezaburzonemu $|n\rangle$, to założenie jest zawsze spełnione.

Rozważmy najpierw równanie (14.4) dla $k = n$:

$$\varepsilon_n + \lambda\langle n|\hat{V}|\psi_n\rangle = E_n. \tag{14.8}$$

Podstawiając za $|\psi_n\rangle$ i E_n ich rozwinięcia perturbacyjne i przyrównując do zera współczynniki przy kolejnych potęgach parametru λ, otrzymujemy związki dla współczynników rozwinięć perturbacyjnych. Dla członów niezależnych od λ równanie jest prawdziwe, jeśli $E_n^{(0)} = \varepsilon_n$, co już wiemy, że zachodzi. Dla członów rzędu λ otrzymujemy

$$E_n^{(1)} = V_{nn}. \tag{14.9}$$

Na przykład, dla stanu podstawowego atomu wodoru zburzonego polem elektrycznym o natężeniu $\mathcal{E}$, otrzymujemy w pierwszym przybliżeniu rachunku zaburzeń

$$E_0 = \varepsilon_0 + e\mathcal{E}\int d^3x|\psi_0(\vec{x})|^2 z + O(\lambda^2), \tag{14.10}$$

gdzie symbol $O(\lambda^2)$ oznacza człon rzędu λ^2, czyli w tym przypadku rzędu $(e\mathcal{E})^2$. Jeżeli w całce zrobimy podstawienie $z \to -z$, to funkcja $|\psi_0(\vec{x})|^2$, która jest sferycznie symetryczna, nie ulegnie zmianie, a czynnik z i z nim cała funkcja podcałkowa i cała całka zmieni znak. Ponieważ zmiana zmiennych w całce nie może zmienić jej wartości, wynika stąd, że całka jest równa zeru – przesunięcie poziomu podstawowego pod wpływem pola elektrycznego w pierwszym rzędzie rachunku zaburzeń znika. Tak więc, to przesunięcie będzie proporcjonalne do kwadratu natężenia pola elektrycznego[1]. Mamy tu do czynienia z tak zwanym kwadratowym efektem Starka.

Ogólnie, wyliczenie pierwszej poprawki perturbacyjnej do energii jest dosyć łatwe – wymaga policzenia jednej całki. Przyrównując do zera współczynnik przy członach rzędu λ^2, otrzymujemy relację

$$E_n^{(2)} = \langle n|\hat{V}(x)|\psi_n^{(1)}\rangle. \tag{14.11}$$

[1]Zakładając, że $E_0^{(2)} \neq 0$, co w tym przypadku jest prawdą.

Żeby policzyć tę poprawkę, trzeba znać pierwszą poprawkę perturbacyjną do funkcji falowej.

Zauważmy, że wobec przyjęcia normalizacji (14.6), współczynniki $\psi_n^{(k)} = 0$ dla $k = 1, 2, \ldots$ i nie musimy ich wyliczać. Żeby wyliczyć poprawki do funkcji falowej wykorzystujemy równania (14.4) dla $k \neq n$. Podstawiając rozwinięcia perturbacyjne dla E_n i $|\psi_n\rangle$, i przyrównując do zera współczynniki przy kolejnych potęgach parametru λ, otrzymujemy kolejne związki dla współczynników rozwinięć perturbacyjnych. Dla członów niezależnych od λ otrzymujemy równość $\varepsilon_k\langle k|n\rangle - \varepsilon_n\langle k|n\rangle = 0$, która jest prawdziwa, bo iloczyny skalarne ortogonalnych wektorów bazowych są równe zeru. Dla członów proporcjonalnych do pierwszej potęgi parametru λ otrzymujemy, podstawiając znane energie i wektory własne niezaburzone,

$$\varepsilon_k\langle k|\psi_n^{(1)}\rangle + V_{kn} = \varepsilon_n\langle k|\psi_n^{(1)}\rangle. \tag{14.12}$$

To pozwala wyliczyć

$$\langle k|\psi_n^{(1)}\rangle = \frac{V_{kn}}{\varepsilon_n - \varepsilon_k} \qquad \text{dla} \qquad k \neq n. \tag{14.13}$$

Ten wzór traci sens, jeśli $\varepsilon_k = \varepsilon_n$, więc założenie o braku degeneracji jest tu naprawdę potrzebne. Pełny wektor $|\psi_n^{(1)}\rangle$ możemy teraz znaleźć ze wzoru:

$$|\psi_n^{(1)}\rangle = \sum_{k\neq n} |k\rangle \frac{V_{kn}}{\varepsilon_n - \varepsilon_k}. \tag{14.14}$$

Podstawiając pierwszą poprawkę dla funkcji falowej do wzoru na drugą poprawkę dla energii, otrzymujemy

$$E_n^{(2)} = \sum_{k\neq n} \frac{|V_{kn}|^2}{\varepsilon_n - \varepsilon_k}. \tag{14.15}$$

Dla stanu podstawowego, powiedzmy $n = 0$, energia niezaburzona ε_0 jest niższa od wszystkich innych energii niezaburzonych. Wynika z tego, że wszystkie mianowniki w sumie po prawej stronie poprzedniego wzoru są ujemne i, ponieważ żaden z liczników nie jest ujemny, otrzymujemy ogólny wynik

$$E_0^{(2)} \leq 0. \tag{14.16}$$

Na przykładzie stanu podstawowego atomu wodoru zaburzonego jednorodnym polem elektrycznym łatwo jest zrozumieć powyższe wyniki. W przybliżeniu, w którym rozkład elektronu elektronu wokół jądra opisujemy sferycznie symetryczną funkcja falową stanu podstawowego, energia oddziaływania z jednorodnym polem elektrycznym jest równa zero. To jest pierwsza poprawka perturbacyjna do energii. Pierwsza poprawka perturbacyjna do funkcji falowej czy wektora stanu polega na przesunięciu chmury elektronowej względem pola tak, że powstaje dipol elektryczny z momentem równoległym do pola elektrycznego. Ten moment dipolowy jest rzędu λ, czyli w tym konkretnym przypadku proporcjonalny do $\mathcal{E}$. Jego oddziaływania z polem elektrycznym, które też jest rzędu λ, powoduje obniżenie energii rzędu λ^2. Dla małych wartości λ jest to bardzo małe przesunięcie. Na przykład, dla stanu podstawowego atomu wodoru zaburzonego polem elektrycznym o natężeniu 10^4 V/cm, energia obniża się o

około 10^{-8} elektronowolta. Zobaczymy w następnym rozdziale, jak dla stanów zdegenerowanych powstaje znacznie większy, liniowy efekt Starka.

Rachunki wygodnie jest prowadzić przy normalizacji (14.6), ale mając już wynik, można wrócić do zwykłej normalizacji do jedności. Taka zmiana normalizacji jest nazywana renormalizacją. W tym przypadku jest to renormalizacja funkcji falowej. W kwantowej teorii pola renormalizuje się różne wielkości, więc to dookreślenie jest użyteczne. Zapisujemy wektor stanu w postaci

$$|\psi_n\rangle = Z|\psi_n\rangle_{pert}, \tag{14.17}$$

gdzie wektor $|\ldots\rangle_{pert}$ jest obliczony w normalizacji (14.6), a stała rzeczywista Z ma być tak dobrana, żeby wektor po prawej stronie był znormalizowany do jedności. Mamy

$$|\psi_n\rangle = Z(|n\rangle + \lambda \sum_{k\neq n} |k\rangle \frac{V_{kn}}{\varepsilon_n - \varepsilon_k}), \tag{14.18}$$

$$\langle\psi_n| = Z(\langle n| + \lambda \sum_{k\neq n} \langle k| \frac{V^*_{kn}}{\varepsilon_n - \varepsilon_k}). \tag{14.19}$$

Mnożąc przez siebie te wektory, otrzymujemy po lewej stronie $\langle\psi_n|\psi_n\rangle = 1$, a po lewej, uwzględniając ortogonalność wektorów własnych problemu niezaburzonego i pomijając człony rzędu w λ wyższego niż pierwszy, Z^2. W tym przybliżeniu $Z = 1$, to znaczy, że otrzymane wektory własne już są znormalizowane do jedności. Łatwo sprawdzić, że w drugim przybliżeniu metody zaburzeń stała Z jest już różna od jedności.

Rozdział 15

Metoda zaburzeń dla stanów zdegenerowanch

W tym rozdziale zastosujemy metodę zaburzeń do zdegenerowanych poziomów energii. Zakładamy, że dla badanego układu N stanów niezaburzonych ma dokładnie tę samą energię:

$$\varepsilon_1 = \varepsilon_2 = \cdots = \varepsilon_N. \tag{15.1}$$

Poza tym zwykle układ ma jeszcze wiele stanów z innymi energiami. Jak zwracaliśmy uwagę w poprzednich dwu rozdziałach, do zdegenerowanego poziomu energii nie można stosować metody Rayleigha-Schrödingera i trzeba najpierw, z pośród nieskończenie wielu wektorów stanu, stanowiących przestrzeń wektorową odpowiadającą tej energii, wybrać odpowiednią bazę – bazę dostosowaną do zaburzenia. Przypuśćmy, że znamy jakąś bazę w tej przestrzeni wektorowej: $|1\rangle, \ldots, |N\rangle$. Dalej będziemy nazywali tę bazę pierwotną. Nie zakładamy, że baza pierwotna jest dostosowana do zaburzenia. Wektory z bazy dostosowanej do zaburzenia, które będziemy oznaczali $|\chi_1\rangle, \ldots, |\chi_N\rangle$, są z definicji bazy kombinacjami liniowymi wektorów bazowych $|j\rangle$ z bazy pierwotnej:

$$|\chi_k\rangle = \sum_{j=1}^{N} c_{kj}|j\rangle. \tag{15.2}$$

Równanie Schrödingera

$$(\hat{H}_0 + \lambda\hat{V})|\psi_n\rangle = E_n|\psi_n\rangle \tag{15.3}$$

daje jak w poprzednim rozdziale

$$\varepsilon_k\langle\chi_k|\psi_n\rangle + \lambda\langle\chi_k|\hat{V}|\psi_n\rangle = E_n\langle\chi_k|\psi_n\rangle. \tag{15.4}$$

Zakładając, że po wybraniu bazy $|\chi_k\rangle$ można stosować metodę zaburzeń, podstawiając rozwinięcia perturbacyjne dla energii E_n oraz dla wektora stanu $|\psi_n\rangle$, i przyrównując do zera współczynniki przy kolejnych potęgach parametru λ, otrzymujemy, jak w poprzednim rozdziale, relacje dla współczynników rozwinięć perturbacyjnych. Dla członów niezależnych od λ otrzymujemy $\varepsilon_k\langle\chi_k|\chi_n\rangle = \varepsilon_n\langle\chi_k|\chi_n\rangle$, co jest prawdą podobnie jak w metodzie Rayleigha-Schrödingera,

bo baza $|\chi_k\rangle$ jest ortonormalna. W pierwszym rzędzie po λ otrzymujemy dla $\varepsilon_k = \varepsilon_n$:

$$\langle\chi_k|\hat{V}|\chi_n\rangle = 0, \quad \text{dla} \quad k \neq n, \tag{15.5}$$

$$\langle\chi_n|\hat{V}|\chi_n\rangle = E_n^{(1)}. \tag{15.6}$$

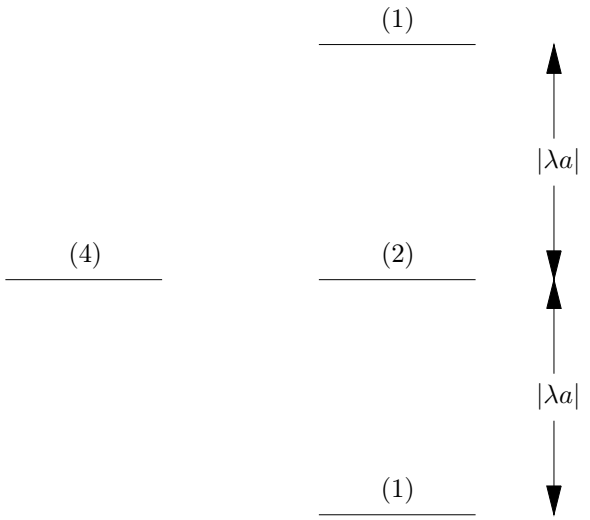
Drugie z tych równań jest jak w metodzie Rayleigha-Schrödingera, ale pierwsze jest nowe i jak zaraz zobaczymy często pozwala znaleźć stany dopasowane do zaburzenia $|\chi_k\rangle$.

Zachowujemy dla elementów macierzowych operatora $\hat{V}$, w reprezentacji stanów z pierwotnej bazy, oznaczenie z poprzedniego rozdziału $V_{kn} = \langle k|\hat{V}|n\rangle$. Macierz V w tej reprezentacji naogół nie jest diagonalna. Ze wzorów (15.5) wynika natomiast, że ta sama macierz jest diagonalna w reprezentacji stanów $|\chi_k\rangle$. Problem: jak mając daną macierz znaleźć reprezentację, w której ta macierz jest diagonalna i jak znaleźć elementy diagonalne w tej reprezentacji, jest dobrze znany z algebry. Współczynniki pierwszej poprawki do energii $E_n^{(1)}$ są wartościami własnymi macierzy V, a wektory bazowe dopasowane do zaburzenia $|\chi_k\rangle$ są jej wektorami własnymi. Jeżeli wszystkie wartości własne są różne, to wektory własne są jednoznacznie określone (jak zwykle z dokładnością do stałego czynnika, którego moduł można wyznaczyć z warunku normalizacji). W problemie fizycznym mówimy wtedy, że zaburzenie całkowicie usuwa czy zdejmuje degenerację – w pierwszym przybliżeniu rachunku zaburzeń poziomy są już niezdegenerowane. Poziom energetyczny, który był N-krotnie zdegenerowany w problemie niezaburzonym, rozszczepia się pod wpływem zaburzenia, w tym przypadku, na N poziomów o różnych energiach. Jeżeli wszystkie wartości własne są równe, to dają pierwsza poprawkę perturbacyjną do energii – wszystkie N poziomów przesuwa się jednakowo o iloczyn parametru λ i otrzymanej wartości własnej – ale nie dają żadnej informacji o stanach niezaburzonych dopasowanych do zaburzenia. W takim wypadku, jeśli chcemy się czegoś dowiedzieć o tych stanach, musimy badać wyższe przybliżenia metody zaburzeń, co zresztą też nie zawsze jest skuteczne. Mówimy wtenczas, że zaburzenie w pierwszym przybliżeniu metody zaburzeń nie zdejmuje degeneracji. Może się wreszcie zdarzyć, że zaburzenie w pierwszym przybliżeniu metody zaburzeń zdejmuje degenerację częściowo. Niektóre wartości własne są równe, inne są różne, i otrzymujemy częściową informacje o stanach dopasowanych do zaburzenia.

Jako przykład przedyskutujemy w pierwszym przybliżeniu rachunku zaburzeń rozszczepienie pierwszego poziomu wzbudzonego atomu wodoru przez pole elektryczne (efekt Starka). Wybieramy oś z układu współrzędnych równoległą do kierunku pola. Operator zaburzenia dla tego problemu został wprowadzony w rozdziale trzynastym

$$\lambda\hat{V}(\vec{x}) = e\mathcal{E}z. \tag{15.7}$$

Możemy przyjąć, że $\lambda = e\mathcal{E}$. Pierwszy poziom wzbudzony atomu wodoru jest czterokrotnie zdegenerowany. Jako wektory bazowe pierwotne możemy wybrać wektory stanu $|n,l,m\rangle = |2,0,0\rangle, |2,1,1\rangle, |2,1,0\rangle, |2,1,-1\rangle$ (rozdział dziesiąty). Macierz V, której wartości własne i wektory własne będą nam potrzebne, jest więc macierzą o czterech wierszach i czterech kolumnach. Taka macierz ma 16 elementów, ale ze względu na relację $V_{jk}^* = V_{kj}$ wystarczy obliczyć dziesięć z nich.



Rysunek 15.1: Rozszczepienie czterokrotnie zdegenerowanego pierwszego stanu wzbudzonego atomu wodoru (n=2) pod wpływem pola elektrycznego (linowy efekt Starka). Na osi pionowej (nie narysowanej) odkładana jest energia stanu. Liczby w nawiasach podają degeneracje poziomów energetycznych. Definicja parametru a podana jest w tekście.

Okazuje się, że dziewięć z tych dziesięciu elementów jest równych zero, co można prosto pokazać wykorzystując symetrie tak, jak w dyskusji pierwszej poprawki perturbacyjnej do energii stanu podstawowego atomu wodoru w poprzednim rozdziale, albo wypisując elementy macierzowe za pomocą funkcji falowych podanych dalej w tym rozdziale, przechodząc do współrzędnych sferycznych i wykonując całkowania po kątach. Jedyny niezerowy element oznaczymy a. Mamy

$$a = \langle 2,0,0|\hat{V}|2,1,0\rangle. \tag{15.8}$$

Pierwsze poprawki do energii wyliczamy jako pierwiastki "równania sekularnego"

$$\begin{vmatrix} -x & a & 0 & 0 \\ a^* & -x & 0 & 0 \\ 0 & 0 & -x & 0 \\ 0 & 0 & 0 & -x \end{vmatrix} = x^2(x^2 - |a|^2) = 0. \tag{15.9}$$

To równanie ma pierwiastek pojedynczy $x_1 = -|a|$, pierwiastek podwójny $x_2 = x_3 = 0$ i jeszcze jeden pierwiastek pojedynczy $x_4 = |a|$. To znaczy, że czterokrotnie zdegenerowany poziom ε_2 rozszczepi się na trzy poziomy, przesunięte w stosunku do energii niezaburzonej ε_2 odpowiednio o: $\lambda E^{(1)}_{2,1} = -|\lambda a|$, $\lambda E^{(1)}_{2,2} = \lambda E^{(1)}_{2,3} = 0$ (podwójnie zdegenerowany) i $\lambda E^{(1)}_2 = +|\lambda a|$ (rysunek 15.1). Degeneracja nie została całkowicie zdjęta, więc stany niezaburzone dopasowane do zaburzenia nie zostaną jednoznacznie określone.

Łatwe całkowanie z wykorzystaniem funkcji falowych atomu wodoru podanych poniżej daje $|\lambda a| = 3|e\mathcal{E}|a_0$, gdzie "promień Bohra" $a_0 \approx 0.53 \times 10^{-8}$

cm. Podstawiając natężenie pola 10^4 woltów/cm otrzymujemy przesunięcia poziomów o około 1.5×10^{-4} elektronowolta, a zatem małe w porównaniu z energiami wzbudzeń atomu wodoru, ale duże w porównaniu z kwadratowym efektem Starka omawianym w poprzednim rozdziale. Ciekawszy od tego niewielkiego przesunięcia jest fakt rozszczepienia poziomu. Z tego powodu linie odpowiadające przejściom z tego poziomu i na ten poziom też mogą się rozszczepić, chociaż nie muszą, bo czasem niektóre przejścia mogą być zakazane. Mogą się też zdarzać przejścia między rozszczepionymi poziomami tego samego multipletu.

Żeby znaleźć wektory stanu dopasowane do zaburzenia musimy znaleźć niezerowe rozwiązania równań

$$\begin{pmatrix} -x_k & a & 0 & 0 \\ a^* & -x_k & 0 & 0 \\ 0 & 0 & -x_k & 0 \\ 0 & 0 & 0 & -x_k \end{pmatrix} \begin{pmatrix} c_{1k} \\ c_{2k} \\ c_{3k} \\ c_{4k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}; \qquad k = 1, 2, 3, 4, \quad (15.10)$$

albo tych samych równań w krótszej notacji macierzowej

$$V\vec{c}_k = x_k\vec{c}_k; \qquad k = 1, 2, 3, 4. \quad (15.11)$$

Czasem wygodnie jest zapisywać wektory $\vec{c}_k$ w postaci transponowanej, bo wtedy mieszczą się w jednej linijce tekstu. Ogólnie $\vec{c}_k^T = (c_{1k}, c_{2k}, c_{3k}, c_{4k})$. Wektorowi $\vec{c}_k$ odpowiada wektor stanu

$$|\chi_k\rangle = c_{1k}|200\rangle + c_{2k}|2, 1, 0\rangle + c_{3k}|2, 1, 1\rangle + c_{4k}|2, 1, -1\rangle. \quad (15.12)$$

Mnożąc tę równość lewostronnie przez wektor dualny $\langle x|$, otrzymujemy rozwinięcie niezaburzonej funkcji falowej dopasowanej do zaburzenia $\chi_k(x)$ na niezaburzone funkcje pierwotne $\psi_{2,l,m}(x)$. Możemy więc te same rozwiązania przedstawiać za pomocą wektora $\vec{c}_k$, wektora transponowanego względem niego $\vec{c}_k^T$, niezaburzonego wektora stanu dopasowanego do zaburzenia $|\chi_k\rangle$, lub niezaburzonej funkcji falowej dopasowanej do zaburzenia $\chi_k(x)$. Poniżej używany funkcji falowych.

Dla $k = 1$ otrzymujemy

$$\chi_1(x) = \frac{1}{\sqrt{2}} \left(\psi_{2,0,0}(x) + \psi_{2,1,0}(x)\right). \quad (15.13)$$

Znak między dwoma składnikami w nawiasie zależy od znaku parametru a, który przyjęliśmy ujemny. Par rozwiązań odpowiadających wartości własnej zero, dla $k = 2$ i $k = 3$, jest nieskończenie wiele, ale można je wszystkie zapisać w postaci

$$\chi_2(x) = \cos\theta\psi_{2,1,1}(x) - \sin\theta\psi_{2,1,-1}(x), \quad (15.14)$$
$$\chi_3(x) = \sin\theta\psi_{2,1,1}(x) + \cos\theta\psi_{2,1,-1}(x), \quad (15.15)$$

gdzie kąt θ jest dowolny. Widać, że tych funkcji falowych dopasowanych do zaburzenia nie da się jednoznacznie określić. To jest zresztą jeden z tych przypadków, kiedy przechodzenie do wyższych rzędów rachunku zaburzeń też nie pozwala jednoznacznie określić tych funkcji. Dla $k = 4$ otrzymujemy

$$\chi_4(x) = \frac{1}{\sqrt{2}}\left(\psi_{2,0,0}(x) - \psi_{2,1,0}(x)\right) \tag{15.16}$$

Znając funkcje falowe dla pierwszego poziomu wzbudzonego atomu wodoru:

$$\langle \vec{x}|2,0,0\rangle = \frac{1}{\sqrt{8\pi a_0^3}} e^{-\frac{r}{2a_0}}(1 - \frac{r}{2a_0}), \tag{15.17}$$

$$\langle \vec{x}|2,1,0\rangle = \frac{1}{\sqrt{32\pi a_0^5}} r e^{-\frac{r}{2a_0}} \cos\theta, \tag{15.18}$$

$$\langle \vec{x}|2,1,\pm 1\rangle = \mp\frac{1}{8\sqrt{\pi a_0^5}} r e^{-\frac{r}{2a_0}} \sin\theta e^{\pm i\varphi}, \tag{15.19}$$

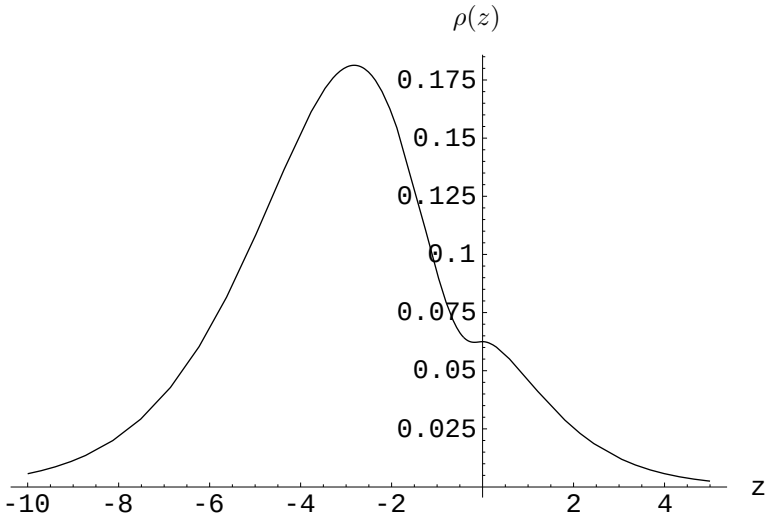
łatwo jest znaleźć interpretację fizyczną funkcji niezaburzonych dopasowanych do zaburzenia. Dla stanu $|\chi_1\rangle$, rozkład składowej z wektora położenia elektronu jest pokazany na rysunku 2. Widać, że ładunek ujemny został przesunięty w kierunku przeciwnym do kierunku pola. Wskutek tego powstał moment dipolowy równoległy do pola i energia obniżyła się. Ważne jest, że to przesunięcie ładunku, a zatem i wytworzony moment dipolowy, nie zależy od natężenia pola. Stworzenie superpozycji dwu stanów o tej samej energii nie wymaga dodatkowej energii. Energia sprzężenia niezależnego od pola momentu dipolowego z polem jest proporcjonalna do pierwszej potęgi natężenia pola i dla tego mamy tu do czynienia z liniowym efektem Starka, a nie z kwadratowym, jak w przypadku niezdegenerowanego stanu podstawowego atomu wodoru, który dyskutowaliśmy w poprzednim rozdziale. Dla stanu $|\chi_4\rangle$ przesunięcie elektronu jest w przeciwnym kierunku. Wobec tego powstaje dipol, taki sam, ale zorientowany antyrównolegle do pola i mamy wzrost energii, a nie spadek. W stanach $|\chi_2\rangle$ i $|\chi_3\rangle$ nie ma przesunięcia ładunku i w związku z tym nie ma zmiany energii.

Możliwość tworzenia ze zdegenerowanych, lub nieznacznie różniących się energią, stanów jednocząstkowych, stanów o nie zwiększonej energii, gdzie chmura elektronowa jest wysunięta w jakimś kierunku daleko do jądra, ułatwia tworzenie wiązań chemicznych. To zjawisko jest znane w chemii jako hybrydyzacja. Na przykład w atomie węgla ze stanu $2S$ i jednego stanu $2P$ można utworzyć stany takie jak stany $|\chi_1\rangle$ i $|\chi_4\rangle$ dla pierwszego stanu wzbudzonego atomu wodoru. Może to prowadzić do powstania cząsteczki, w której atom węgla jest związany z dwoma innymi atomami tak, że wszystkie trzy atomy leżą na jednej prostej. Tak jest zbudowana cząsteczka acetylenu C_2H_2. Ze stanu $2S$ i dwu stanów $2P$ można zbudować trzy kombinacje liniowe, w których elektrony są wysunięte w trzech kierunkach leżących w jednej płaszczyźnie i tworzących ze sobą kąty 120^0 (znak Mercedesa). Tak są zhybrydyzowane atomy węgla na przykład w cząsteczce benzenu C_6H_6 i w graficie. Wreszcie ze stanu $2S$ i trzech stanów $2P$ można zbudować cztery kombinacje liniowe, w których elektrony są wysunięte ku wierzchołkom czworościanu regularnego, w którego środku masy znajduje się atom węgla. Tak jest zbudowana na przykład cząsteczka metanu CH_4 i diament.

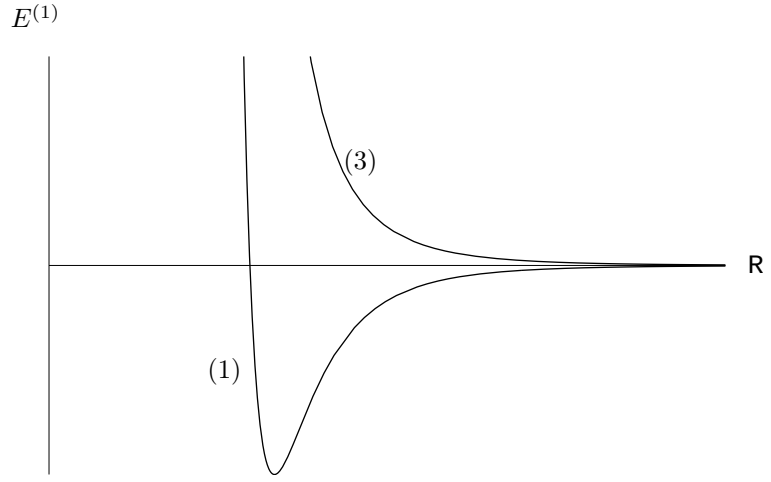
Jako kolejny przykład rozpatrzymy rozszczepienie poziomu podstawowego pary atomów wodoru znajdujących się w odległości R od siebie. Jako zaburzenie wydzielamy wzajemne oddziaływanie tych dwu atomów. W stanie niezaburzonym mamy więc dwa atomy wodoru, każdy w swoim stanie podstawowym. Uwzględniamy spiny, więc stan podstawowy każdego z atomów jest dwukrotnie

zwyrodniały a stan podstawowy całego układu czterokrotnie. Jeżeli odległość miedzy atomami jest bardzo duża, to ich wzajemne oddziaływanie jest zaniedbywalne. Kiedy atomy zbliżają się do siebie, poziom niezaburzony rozszczepia się na poziom niezdegenerowany (singlet) i poziom trzykrotnie zdegenerowany (triplet) rysunek 3. Pokażemy w rozdziale dziewiętnastym, jak ten fakt można zrozumieć bez żadnych rachunków. W stanach tripletowych energia jest większa niż suma energii dwu izolowanych atomów i nic szczególnie ciekawego się nie dzieje. W stanie singletowym natomiast, energia jest niższa niż energia dwu izolowanych atomów, co stanowi warunek konieczny, chociaż w mechanice kwantowej nie dostateczny, powstania trwałej cząsteczki. Przy bardzo małych odległościach R, dominuje odpychanie elektrostatyczne miedzy protonami i energia szybko rośnie. Po drodze występuje więc gdzieś minimum energii. Klasycznie, odległość między protonami dająca minimum energii jest wielkością cząsteczki H_2 liczoną od protonu do protonu i wartość energii w tym punkcie daje minus energię dysocjacji cząsteczki wodoru. Kwantowo, sytuacja jest nieco bardziej skomplikowana. Ze względu na zasadę nieokreśloności protony nie mogą utrzymywać stałej odległości R, ale wykonują drgania, co z kolei zmniejsza energie potrzebną do rozerwania cząsteczki. Poza tym, przesunięcie energii wyliczone w pierwszym przybliżeniu metody zaburzeń nie jest dokładne. Mimo to, wyniki uzyskane w pierwszym przybliżeniu metody zaburzeń dla stanów zdegenerowanych są na tyle rozsądne, że właśnie taki rachunek został kiedyś opublikowany jako pierwszy dowód, że mechanika kwantowa dobrze tłumaczy budowę cząsteczki wodoru.

Rozważmy teraz kryształ sodu zawierający jeden mol, czyli około 6×10^{23} atomów. Każdy taki atom na zewnątrz zamkniętej powłoki elektronowej ma jeden elektron. Jeżeli zaniedbamy oddziaływanie między atomami, to energie wszystkich tych elektronów zewnętrznych są równe, ale rozumując jak dla dwu atomów wodoru widzimy, że degeneracja tego poziomu jest olbrzymia. Kiedy włączymy oddziaływanie, poziom się rozszczepi, ale na takie mnóstwo poziomów, że z jednego poziomu powstanie pasmo wypełniające praktycznie w sposób ciągły pewien przedział energii. Takie pasma są jednym z podstawowych pojęć w fizyce ciała stałego.



Rysunek 15.2: Rozkład gęstości prawdopodobieństwa znalezienia elektronu (wycałkowany po x i y) w stanie odpowiadającym funkcji falowej $\frac{1}{\sqrt{2}}\left(\langle\vec{x}|2,0,0\rangle + \langle\vec{x}|2,1,0\rangle\right)$. Analogiczny rozkład dla stanu odpowiadającego funkcji falowej $\frac{1}{\sqrt{2}}\left(\langle\vec{x}|2,0,0\rangle - \langle\vec{x}|2,1,0\rangle\right)$ można otrzymać przez odbicie rozkładu na rysunku względem osi $z = 0$.



Rysunek 15.3: Jakościowe przedstawienie zależności rozszczepienia, w pierwszym przybliżeniu rachunku zaburzeń, czterokrotnie zdegenerowanego stanu podstawowego pary atomów wodoru (uwzględniamy spiny), od odległości R między atomami. Liczby w nawiasach oznaczają degeneracje stanów po rozszczepieniu

Rozdział 16

Kręt cząstki o spinie zero

Przez cząstkę będziemy tu rozumieli odpowiednik cząstki punktowej z klasycznej mechaniki. Stan klasycznej cząstki punktowej jest jednoznacznie określony przez podanie jej pędu i położenia. W najprostszym uogólnieniu kwantowym, stan cząstki jest określony przez podanie pędu albo położenia. Takie cząstki omawiamy w bieżącym rozdziale. Okazuje się jednak, że chcąc opisać dane doświadczalne trzeba wprowadzić ogólniejsze cząstki, też punktowe, które oprócz pędu czy położenia mają jeszcze spin. Teoria cząstek ze spinem jest omawiana w następnych czterech rozdziałach. Cząstki punktowe nie występują w przyrodzie, ale są wygodną idealizacją. W nierelatywistycznej mechanice kwantowej elektrony, nukleony i zwykle jądra atomowe można traktować jako cząstki punktowe. Atomu wodoru zwykle nie możemy traktować jako cząstki punktowej. Nie chodzi tu o to, że wiemy, że atom składa się z protonu i elektronu i że ma skończony promień. Nukleon też składa się z kwarków i gluonów i ma skończony promień. Atom wodoru jednak, nawet przy określonym pędzie i spinie, ma wiele stanów o różnych energiach. Ma więc dodatkowe, "wewnętrzne" stopnie swobody. Jądro atomowe jest traktowane jako cząstka punktowa tylko w sytuacjach, kiedy prawdopodobieństwo wzbudzenia go jest do zaniedbania.

W klasycznej mechanice dużą rolę odgrywa kręt, zdefiniowany dla cząstki punktowej wzorem $\vec{L} = \vec{r} \times \vec{p}$ lub, rozpisując na współrzędne,

$$\begin{aligned} L_x &= yp_z - zp_y, \\ L_y &= zp_x - xp_z, \\ L_z &= xp_y - yp_x. \end{aligned} \tag{16.1}$$

Ponieważ w tych wzorach nie występują iloczyny wielkości, którym odpowiadają niekomutujące ze sobą operatory, kwantowanie można przeprowadzić przez zwykłe podstawienie $\vec{p} = -i\hbar\vec{\bigtriangledown}$:

$$\begin{aligned} \hat{L}_x &= -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \\ \hat{L}_y &= -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \\ \hat{L}_z &= -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \end{aligned} \tag{16.2}$$

Znajdziemy teraz wartości własne i funkcje własne operatora $\hat{L}_z$. Równanie własne ma postać

$$\hat{L}_z\psi_m(\vec{x}) = \hbar m\psi_m(\vec{(x)}). \tag{16.3}$$

Zgodnie ze zwyczajem oznaczyliśmy tu wartość własną przez $\hbar m$. Rachunki wygodniej jest prowadzić we współrzędnych cylindrycznych, w których

$$x = r\cos\varphi; \qquad y = r\sin\varphi. \tag{16.4}$$

Zgodnie z regułą różniczkowania funkcji złożonych

$$\frac{\partial}{\partial\varphi} = \frac{\partial x}{\partial\varphi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\varphi}\frac{\partial}{\partial y}, \tag{16.5}$$

ale na mocy poprzednich wzorów

$$\begin{aligned} \frac{\partial x}{\partial\varphi} &= -r\sin\varphi = -y, \\ \frac{\partial y}{\partial\varphi} &= r\cos\varphi = x. \end{aligned} \tag{16.6}$$

Wobec tego[1]

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\varphi} \tag{16.7}$$

i równanie własne dla L_z przyjmuje prostszą formę

$$-i\frac{\partial\psi_m(r,\varphi,z)}{\partial\varphi} = m\psi_m(r,\varphi,z). \tag{16.8}$$

Ogólne rozwiązanie tego równania ma postać

$$\psi_m(r,\varphi,z) = Ce^{im\varphi}, \tag{16.9}$$

gdzie C jest dowolną funkcją zmiennych r, z. Równie dobrze można było prowadzić rozumowanie używając współrzędnych sferycznych i wtedy C byłoby dowolną funkcją zmiennych r, θ. Pozostaje nałożyć warunki, które musi spełniać rozwiązanie, żeby było fizycznie do przyjęcia. Kiedy kąt φ wzrasta o 2π przy ustalonych wartościach pozostałych zmiennych, każdy punkt w płaszczyźnie (x, z) zatacza koło wokół osi z i wraca tam skąd wyszedł. Żeby funkcja falowa nie miała nieciągłości, musi zachodzić

$$e^{2\pi im} = 1, \tag{16.10}$$

a to znaczy, że m musi być liczbą całkowitą. Udowodniliśmy więc dwa ważne twierdzenia:

- Wynik dokładnego pomiaru z-owej składowej krętu L_z może dać tylko wynik $\hbar m$, gdzie m jest liczbą całkowitą.

[1]Ten wzór można otrzymać prościej zauważając, że L_z jest pędem kanonicznie sprzężonym do współrzędnej φ i stosując uogólnioną regułę kwantowania z rozdziału siódmego.

- Funkcja falowa $\psi_m(\vec{x})$ jest proporcjonalna do funkcji $e^{im\varphi}$, przy czym współczynnik proporcjonalności nie zależy od zmiennej φ.

Analogiczne rozumowanie można oczywiście przeprowadzić dla każdej ze składowych wektora $\vec{L}$.

Wyliczymy teraz komutatory operatorów odpowiadających składowym krętu, żeby sprawdzić, czy da się jednocześnie ustalić wartość więcej niż jednej składowej. Ponieważ rachunki dla wszystkich trzech par składowych są takie same, wystarczy wyliczyć jeden komutator, na przykład

$$\left[\hat{L}_x, \hat{L}_y\right] = -\hbar^2 \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) + \hbar^2 \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right). \tag{16.11}$$

Działając obu stronami tej równości na dowolną funkcję f, otrzymujemy po redukcjach wyrażenia po prawej stronie

$$\left[\hat{L}_x, \hat{L}_y\right] f = i\hbar \hat{L}_z f. \tag{16.12}$$

Ponieważ funkcja f jest dowolna, z tej równości wynika też równość między operatorami: $[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z$. Analogicznie oblicza się pozostałe dwa komutatory i w końcu otrzymujemy trzy reguły komutacji przechodzące w siebie przez cykliczne przestawienia wskaźników:

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \tag{16.13}$$

$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \tag{16.14}$$

$$[\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y. \tag{16.15}$$

Można też zapisać te trzy równości jako jedną równość wektorową

$$\hat{\vec{L}} \times \hat{\vec{L}} = i\hbar\hat{\vec{L}}. \tag{16.16}$$

Z tego, że komutatory nie są równe zero, wyciągamy wniosek, że żadne dwie z trzech składowych krętu nie są jednocześnie mierzalne[2]. Sytuacja jest tu inna niż dla składowych wektora pędu czy wektora położenia.

Kwadratowi długości wektora krętu odpowiada operator

$$\hat{\vec{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \tag{16.17}$$

Za pomocą reguł komutacji (16.13) łatwo jest sprawdzić, że:

$$[\hat{\vec{L}}^2, \hat{L}_x] = [\hat{\vec{L}}^2, \hat{L}_y] = [\hat{\vec{L}}^2, \hat{L}_z] = 0. \tag{16.18}$$

Można więc jednocześnie określić kwadrat krętu i jedną ze współrzędnych wektora krętu. W dalszym ciągu będziemy najczęściej używać wspólnych funkcji własnych operatorów $\hat{\vec{L}}^2$ i $\hat{L}_z$.

Problem własny dla operatora kwadratu krętu zapiszemy w postaci:

[2]Można jednocześnie zmierzyć L_x i L_y w stanach, w których $L_z = 0$, ale takie stany nie stanowią układu zupełnego.

$$\hat{\vec{L}}^2\psi_{l,m}(\vec{x}) = \hbar^2 l(l+1)\psi_{l,m}(\vec{x}). \tag{16.19}$$

Operator $\hat{\vec{L}}^2$ można przedstawić jako operator zbudowany ze zmiennych sferycznych θ i φ, ale wzór jest dość skomplikowany, więc nie będziemy go tu podawać. Rozwiązania problemu własnego istnieją dla l całkowitych, nieujemnych. Dla każdej wartości l możliwe są wartości m całkowite, nie większe co do wartości bezwzględnej od l. Funkcje własne mają postać

$$\psi_{l,m}(\vec{x}) = f(r)Y_l^m(\theta,\varphi), \tag{16.20}$$

gdzie $f(r)$ jest dowolną funkcją odległości r punktu od początku układu współrzędnych (względem którego jest mierzony kręt), a funkcje Y_l^m są znane jako "funkcje kuliste" i znajduje się je w odpowiednich tablicach lub za pomocą programów komputerowych. Normalizacja funkcji kulistych wybrana jest tak, że kwadrat modułu takiej funkcji, zcałkowany po pełnym kącie bryłowym ($d\Omega = d\varphi d\cos\vartheta$), daje jedynkę. Z faktu, że są to zarazem funkcje własne z-owej składowej krętu, wynika że muszą być postaci: funkcja od θ razy $e^{im\varphi}$ Poniżej podajemy dla przykładu funkcje kuliste dla $l \leq 2$:

$$\begin{aligned}
Y_0^0(\theta,\varphi) &= \frac{1}{\sqrt{4\pi}}; \\
Y_1^0(\theta,\varphi) &= \sqrt{\frac{3}{4\pi}}\cos\theta; \\
Y_1^{\pm 1}(\theta,\varphi) &= \mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\pm i\varphi}; \\
Y_2^0(\theta,\varphi) &= \sqrt{\frac{5}{4\pi}}(\frac{3}{2}\cos^2\theta - \frac{1}{2}); \\
Y_2^{\pm 1}(\theta,\varphi) &= \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{\pm i\varphi}; \\
Y_2^{\pm 2}(\theta,\varphi) &= \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{\pm 2i\varphi}.
\end{aligned} \tag{16.21}$$

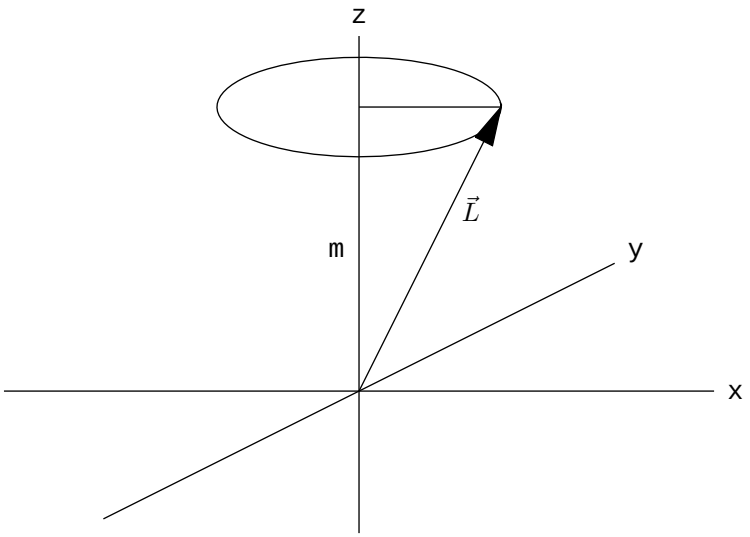
(16.22)

Jak widać, czynniki zależne od kąta θ są niezbyt skomplikowanymi wyrażeniami zbudowanymi z sinusów i kosinusów tego kąta.

Niektóre z podanych tu faktów prosto ilustruje klasyczny "model wektorowy" przedstawiony na rysunku 1. Wektor krętu precesuje wokół osi z tworząc z nią stały kąt. W ten sposób długość wektora krętu i jego rzut na oś z są jednoznacznie określone, natomiast rzuty na osie x i y nie mają stałych wartości. W rozdziale dwudziestym podamy dalsze zastosowania tego modelu.

Ważny praktycznie jest przypadek, kiedy hamiltonian układu – dla prostoty przyjmujemy, że układem jest jedna cząstka – komutuje ze wszystkimi składowymi operatora krętu i co za tym idzie, także z operatorem kwadratu krętu. Załóżmy, że hamiltonian komutuje z operatorem $\hat{L}_z$. To znaczy, że dla dowolnej funkcji $f(\vec{x})$:

$$-i\hbar\hat{H}\frac{\partial f(\vec{x})}{\partial\varphi} = -i\hbar\frac{\partial(\hat{H}f(\vec{x}))}{\partial\varphi} = -i\hbar\frac{\partial\hat{H}}{\partial\varphi}f(\vec{x}) - i\hbar\hat{H}\frac{\partial f(\vec{x})}{\partial\varphi}, \tag{16.23}$$



Rysunek 16.1: **Model wektorowy:** Wektor krętu $\vec{L}$ procesuje wokół osi z. Długość wektora i jego rzut na oś z są określone, a rzuty na osie x i y zmieniają się

gdzie druga równość wynika z zastosowania wzoru (Leibnitza) na pochodną iloczynu. Porównując pierwsze wyrażenie z ostatnim widzimy, że założenie o komutacji hamiltonianu z operatorem z-owej składowej krętu jest równoważne z założeniem, że hamiltonian nie zależy od kąta φ, lub równoważnie: nie zmienia się przy obrotach wokół osi z. Założenie, że hamiltonian komutuje z operatorami odpowiadającymi wszystkim trzem składowym operatora krętu oznacza, że hamiltonian nie zmienia się przy obrotach wokół żadnej z osi x, y, z; ale z takich obrotów można złożyć dowolny obrót, więc otrzymujemy następujące twierdzenie: stwierdzenie, że hamiltonian komutuje z operatorami odpowiadającymi wszystkim trzem składowym operatora krętu jest równoważne ze stwierdzeniem, że hamiltonian jest sferycznie symetryczny.

Dla hamiltonianów sferycznie symetrycznych można znaleźć układ zupełny wspólnych funkcji własnych operatorów $\hat{H}, \hat{\vec{L}}^2, \hat{L}_z$. Takie funkcje własne muszą mieć postać

$$\psi_{n,l,m}(\vec{x}) = f_{nl}(r) Y_l^m(\theta, \varphi). \tag{16.24}$$

Na przykład, funkcje falowe dla pierwszego poziomu wzbudzonego atomu wodoru, które były podane w poprzednim rozdziale, są tego kształtu. Podstawiając taką funkcję do równania Schrödingera, które jest w tym przypadku cząstkowym równaniem różniczkowym trzech zmiennych, otrzymujemy zwyczajne równanie różniczkowe na funkcję jednej zmiennej $f_{nl}(r)$, co jest znacznym uproszczeniem zadania. Okazuje się, że jeszcze wygodniej jest wprowadzić funkcje

$$\chi_{nl}(r) = r f_{nl}(r). \tag{16.25}$$

Dla takich funkcji, z równania Schrödingera otrzymujemy po przekształceniach (trzeba znać wzór na laplasjan we współrzędnych sferycznych) równanie:

$$-\frac{\hbar^2}{2m}\frac{d^2\chi_{nl}(r)}{dr^2}+\left[V(r)+\frac{\hbar^2 l(l+1)}{2mr^2}\right]\chi_{nl}(r)=E\chi_{nl}(r). \tag{16.26}$$

Oprócz zwykłych warunków dla funkcji falowych funkcja $\chi_{nl}(r)$ musi spełniać warunek

$$\chi_{nl}(0)=0, \tag{16.27}$$

bo inaczej funkcja $f_{nl}(r)$, a zatem i funkcja $\psi_{nlm}(\vec{x})$, miałaby niedopuszczalną osobliwość w początku układu współrzędnych. Zauważmy, że równanie na funkcje χ, a zatem i ta funkcja, nie zależy od liczby kwantowej m. Ten fakt ma prostą interpretacje fizyczną. Rzut m zależy od wyboru kierunku osi z w przestrzeni, ale dla sferycznie symetrycznego hamiltonianu wszystkie takie wybory są równoważne. Równanie zależy natomiast od liczby kwantowej l. Dodatkowy, zależny od l człon, który dodaje się do potencjału, można interpretować jako odpowiednik klasycznego potencjału sił odśrodkowych. Na rysunku 2 porównane są efektywne potencjały

$$V_{eff}=V(r)+\frac{\hbar^2 l(l+1)}{2mr^2} \tag{16.28}$$

dla stanów $l=0$ i $l=1$ atomu wodoru. Rysunek ułatwia zrozumienie, dla czego dla stanów z $l=0$ w punkcie $r=0$ jest maksimum gęstości elektronów, a dla $l>0$ gęstość elektronów w tym punkcie jest równa zero. Skala trudności rozwiązywania równania Schrödingera dla cząstki poruszającej się w trzech wymiarach w polu sił o potencjale sferycznie symetrycznym jest podobna jak dla cząstki poruszającej się w przestrzeni jednowymiarowej. Za pomocą komputera problem jest zwykle łatwy, ale znaleźć rozwiązanie analityczne udaje się tylko wyjątkowo.

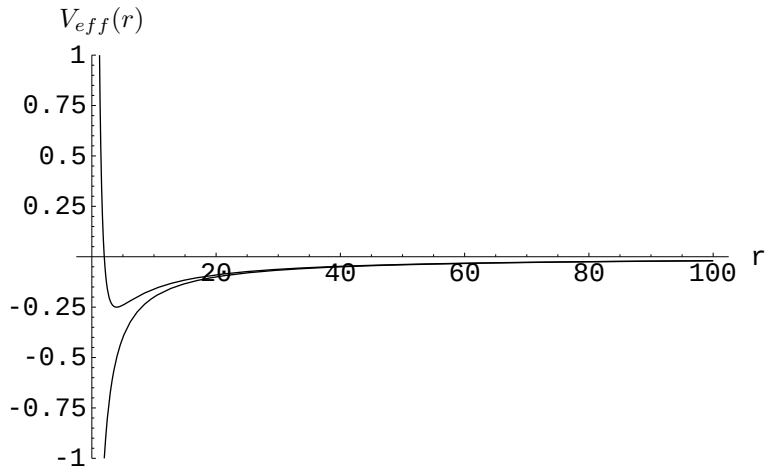
Podajemy tu dla przykładu rozwiązanie analityczne dla elektronu w polu kulombowskim pochodzącym od ładunku $Z|e|$, punktowego jądra znajdującego się w początku układu współrzędnych. Rozważamy tylko stany związane, to znaczy stany odpowiadające ujemnym wartościom własnym energii. Ładunek elektronu i masę zredukowaną układu elektron-jądro oznaczamy odpowiednio: e i μ. Części radialne funkcji falowych falowe mają postać :

$$f_{nl}(r)=\sqrt{\left(\frac{2Z}{na_0}\right)^3\frac{(n-l-1)!}{2n[(n+l)!]^3}}e^{-\frac{\rho}{2}}\rho^l L_{n+l}^{2l+1}(\rho). \tag{16.29}$$

Stała $a_0=\frac{\hbar^2}{\mu e^2}\approx 0.53\times 10^{-8}$cm jest bardzo bliska promieniowi Bohra, $\rho=\frac{2Zr}{na_0}$, a $L_q^p(r)$ są stowarzyszonymi wielomianami Laguerre'a, które podobnie jak wielomiany Hermite'a z problemu oscylatora harmonicznego, znajduje się w tablicach lub za pomocą komputera. Liczba kwantowa n może przyjmować dowolne wartości całkowite dodatnie, a liczba kwantowa l dowolne wartości całkowite nieujemne i mniejsze od n. Wartości własne energii dane są wzorem różniącym się od cytowanego już w rozdziale 10 tylko podstawieniem μ za m :

$$E_{nlm}=-\frac{\mu e^4 Z^2}{2\hbar^2 n^2}. \tag{16.30}$$

Niezależność wartości własnej energii od rzutu krętu, czyli od liczby kwantowej m, wynika z podanej w tym rozdziale teorii. Niezależność od kwadratu krętu,



Rysunek 16.2: Efektywna energia potencjalna dla stanów z l=0 (dolna krzywa) i dla stanów z $l = 1$ (górna krzywa) dla elektronu w polu kulombowskim. Jednostką odległości jest promień Bohra, a jednostką energii jeden Rydberg ($1\text{Ry} = \frac{e^2}{2a_0} \approx 13,6$ elektronowolta).

czyli od liczby kwantowej l, jest zjawiskiem wyjątkowym, zachodzącym akurat dla potencjału kulombowskiego.

Rozdział 17

Przestrzenie niezmiennicze i nieredukowalne względem grupy obrotów

Według teorii rozwiniętej w poprzednim rozdziale, jeżeli pęd cząstki jest równy zero, to równy zero jest także jej kręt. Ten fakt jest oczywisty w fizyce klasycznej, gdzie jeżeli $\vec{p} = 0$, to też $\vec{L} = \vec{r} \times \vec{p} = 0$; ale według wzorów podanych w poprzednim rozdziale, zachodzi i w fizyce kwantowej: z podanych tam wzorów na składowe wektora krętu widać, że jeśli dla każdej ze składowych operatora pędu zachodzi $\hat{p}_i \psi(\vec{x}) = 0$, to także dla każdej składowej krętu $\hat{L}_i \psi(\vec{x}) = 0$. Można ten wynik sformułować inaczej: Jeżeli wszystkie trzy operatory składowych pędu $-i\hbar \frac{\partial}{\partial x_k}$ działając na funkcję falową dają zero, to znaczy, że ta funkcja nie zależy ani od x, ani od y, ani od z, a więc że jest stała. Funkcja (skalarna, patrz dalej) mająca tę samą wartość w każdym punkcie przestrzeni nie zmienia się przy obrotach, więc działając na nią operatorem $-i\hbar \frac{\partial}{\partial \varphi}$ otrzymujemy zero. To jednak znaczy, że składowa z-owa krętu w stanie opisywanym przez tę funkcję falową jest równa zero. To samo dotyczy obrotów wokół osi y i z, a zatem i składowych krętu L_y i L_z. Trudność z tym wynikiem polega na tym, że jak wiadomo z doświadczenia, elektron z pędem zero ma różny od zera kręt. Podobnie jest dla wielu innych cząstek. Wprowadzamy więc nową wielkość: kręt cząstki w jej układzie spoczynkowym, który nazywamy spinem i oznaczamy $\vec{S}$. Jeżeli cząstka ma spin równy zero, to można ją opisywać za pomocą teorii z poprzedniego rozdziału. Jeżeli cząstka ma spin różny od zera, potrzebne jest uogólnienie teorii. Pełny kręt cząstki, który będziemy oznaczali $\vec{J}$, powinno się obliczać jako sumę krętu orbitalnego $\vec{L}$ i spinu $\vec{S}$. Wiemy już, że dla cząstki o spinie zero $\vec{J} = \vec{L}$ i że w układzie spoczynkowym cząstki $\vec{J} = \vec{S}$. Zwracamy uwagę, że cząstka o masie zero nie ma układu spoczynkowego, więc nasza dyskusja dotyczy tylko cząstek o masie różnej od zera.

Teoria z poprzedniego rozdziału została uzyskana przez kwantowanie klasycznego wzoru na kręt. Teraz potrzeba czegoś ogólniejszego. Zauważmy najpierw, że stwierdzenie: jeśli funkcja jest stała, to nie zmienia się przy obrotach, nie jest ogólnie prawdziwe. Weźmy jako przykład jednorodne pole sił, na przykład pole grawitacyjne Ziemi w niewielkim obszarze. Na dany punkt materialny w każdym punkcie działa taka sama siła. Siła jako funkcja położe-

nia jest więc stałą. Ta stała siła jest jednak skierowana w dół, nie jest więc prawdą, że jej składowe nie zmieniłyby się, gdyby obrócić układ. Zakładając, że funkcja falowa jest wektorem $\vec{\psi}(\vec{x})$, można opisać cząstkę ze spinem, ale ten opis nie pasuje do elektronu. Byłby to opis cząstki, dla której w jej układzie spoczynkowym rzut krętu (spinu) na oś z może przyjmować wartości $\hbar, 0, -\hbar$, podczas gdy dla elektronu obserwuje się tylko rzuty spinu $\frac{\hbar}{2}$ i $-\frac{\hbar}{2}$. W następnym rozdziale podamy ogólne rozwiązanie problemu: jaki spin może mieć cząstka i jakie funkcje falowe odpowiadają jakim spinom. Funkcje takie jak dla cząstek ze spinem zero nazywamy skalarnymi. Wspominaliśmy już możliwość użycia wektorowych funkcji falowych. Okaże się, że jest nieskończenie wiele możliwych spinów i każdemu z nich odpowiada inny typ funkcji falowej.

W dalszym ciągu tego rozdziału podanych będzie szereg założeń, twierdzeń i definicji. Żeby ułatwić orientację, twierdzenia i pojęcia definiowane zaznaczone są tłustym drukiem. Poza tym, odnoszące się do nich teksty zaczynają się: znakiem ♠ dla założeń, znakiem ♡ dla twierdzeń i znakiem ♢ dla definicji.

♠ Nie możemy zachować odpowiednika klasycznego wzoru na składowe krętu, który dyskutowaliśmy w poprzednim rozdziale, bo to daje teorię za mało ogólną. Potrzebujemy uogólnienia. Zauważmy, że dla cząstki o spinie zero operatory odpowiadające składowym krętu są blisko związane z operatorami bardzo małych obrotów. Rozważmy najpierw obroty wokół osi z. Z definicji, działanie na funkcję skalarną operatora obrotu o bardzo mały kąt $d\varphi$ dane jest wzorem

$$\hat{R}_z(d\varphi)f(\varphi) = f(\varphi + d\varphi) \tag{17.1}$$

W tym wzorze f oznacza dowolną funkcję, która oprócz φ może mieć jeszcze inne argumenty nie zmieniające się przy obrotach wokół osi z. Przykładami takich dodatkowych zmiennych są zmienne (r, θ) we współrzędnych sferycznych czy zmienne (r, z) we współrzędnych cylindrycznych. Z drugiej strony, z definicji pochodnej:

$$f(\varphi + d\varphi) = f(\varphi) + d\varphi \frac{\partial f(\varphi)}{\partial \varphi}. \tag{17.2}$$

Tu i w dalszym ciągu ograniczamy się do przybliżenia liniowego w bardzo małym parametrze $d\varphi$. Porównując te dwa wzory otrzymujemy

$$R_z(d\varphi)f(\varphi) = (1 + d\varphi \frac{\partial}{\partial \varphi})f(\varphi) \tag{17.3}$$

i stąd, wykorzystując wzór na składową krętu $\hat{L}_z$ z poprzedniego rozdziału,

$$R_z(d\varphi) = 1 + i\hbar^{-1} d\varphi \hat{L}_z. \tag{17.4}$$

Ponieważ cząstka opisywana przez skalarną funkcje falową ma spin zero, dla niej $J_i = L_i$ i powyższy wzór można zapisać w postaci

$$R_z(d\varphi) = 1 + i\hbar^{-1} d\varphi \hat{J}_z. \tag{17.5}$$

Zakładamy, że ten wzór i analogiczne wzory dla operatorów odpowiadających składowym J_x i J_y:

$$R_k(d\varphi) = 1 + i\hbar^{-1} d\varphi \hat{J}_k, \qquad k = x, y, z \tag{17.6}$$

są ogólnie prawdziwe dla wszystkich rodzajów funkcji falowych, czyli dla cząstek z dowolnymi spinami.

♡ **Twierdzenie** (bez dowodu): z przyjętego założenia o składowych wektora krętu wynika, że **reguły komutacyjne dla składowych J_i są takie jak znalezione w poprzednim rozdziale dla składowych L_i.** Te reguły są więc uniwersalne i stosują się dla cząstek z dowolnym spinem. Wynika z nich, jak w poprzednim rozdziale, że kwadrat wektora krętu

$$\hat{\vec{J}}^2 = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \tag{17.7}$$

komutuje ze wszystkimi operatorami składowych $\hat{J}_i$.

♢ Wartości własne operatorów $\hat{J}_z$ i $\hat{\vec{J}}^2$ oznacza się podobnie jak w poprzednim rozdziale, z tym że w wartości własnej kwadratu krętu zamiast l pisze się j. W dalszym ciągu ważną rolę, podobnie jak dla cząstek o spinie zero, będą odgrywać **wspólne wektory własne**

$$\hat{J}_z|j,m\rangle = \hbar m|j,m\rangle, \tag{17.8}$$

$$\hat{\vec{J}}^2|j,m\rangle = \hbar^2 j(j+1)|j,m\rangle; \qquad j \geq 0. \tag{17.9}$$

W szczególności z takich wektorów utworzymy bazę w przestrzeni stanów cząstki o danej energii i pędzie zero, ale na razie nie wiemy ile różnych stanów o danych (j,m) trzeba uwzględnić – może się okazać potrzebny dodatkowy wskaźnik, żeby takie stany rozróżniać. Nie wiemy też ilu i jakich wartości wskaźników (j,m) trzeba będzie użyć. Wybrany sposób zapisania wartości własnych operatora $\hat{\vec{J}}^2$ nie stanowi żadnego ograniczenia, bo ta wartość własna musi być nieujemna, a dla każdej liczby nieujemnej g^2 można znaleźć takie $j \geq 0$, że $g^2 = \hbar^2 j(j+1)$. Zakładamy dodatkowo, że $j \geq 0$. W ten sposób j jest określone jednoznacznie przez wartość własną.

♠ Rozważmy teraz w szczególności stany danej cząstki, o danej energii i z pędem równym zero. Dla cząstki o spinie zero był tylko jeden taki stan. Naogół jest ich więcej. Na mocy zasady superpozycji te stany stanowią przestrzeń wektorową. **Zakładamy, że ta przestrzeń jest niezmiennicza względem grupy obrotów.** Grupa obrotów, to jest zbiór wszystkich możliwych obrotów. Niezmienniczość przestrzeni względem grupy obrotów oznacza, że dokonując dowolnego obrotu dowolnego wektora należącego do tej przestrzeni otrzymamy wektor, ten sam lub inny, też należący do tej przestrzeni. Przykładem przestrzeni niezmienniczej względem obrotów jest przestrzeń zwykłych wektorów w przestrzeni trójwymiarowej (zaczepionych w początku układu). Jakkolwiek obrócimy taki wektor otrzymamy znów wektor z tej samej przestrzeni trójwymiarowej. Przykładem przestrzeni, która nie jest niezmiennicza względem grupy obrotów trójwymiarowych, jest przestrzeń wektorów leżących w płaszczyźnie (x,y) (też zaczepionych w początku układu). Obrót takiego wektora wokół osi leżącej w płaszczyźnie (x,y) przekształca go naogół w wektor, który już nie leży w płaszczyźnie (x,y). Nasuwają się pytania: czy to założenie nie jest oczywiste, a jeśli nie jest oczywiste, to czy jest prawdziwe. Na nieoczywistość tego założenia wskazuje fakt, że bardzo podobne twierdzenie jest fałszywe. Jeżeli stan neutrina poddamy odbiciu lustrzanemu w płaszczyźnie równoległej do pędu cząstki, to otrzymamy stan, który nie występuje w przyrodzie, a więc jest niemożliwy. Obecnie nie

są znane żadne fakty świadczące o tym, że przez obrót stanu możliwego można otrzymać stan niemożliwy, więc założenie, które tu omawiamy jest rozsądne. Jest też powszechnie przyjmowane.

♢ Każda przestrzeń niezmiennicza względem grupy obrotów ma dwie "trywialne" podprzestrzenie niezmiennicze: całą tę przestrzeń i podprzestrzeń złożoną z samego tylko wektora zerowego. Jeżeli **przestrzeń niezmiennicza** nie ma "nietrywialnych" podprzestrzeni niezmienniczych, to mówimy że jest **nieredukowalna**.

♠ **Przyjmujemy założenie, że przestrzeń stanów cząstki o danej energii i pędzie równym zero jest** nie tylko niezmiennicza, ale i **nieredukowalna.** Zobaczymy w dalszym ciągu, jak dużo daje to założenie. Jego interpretację fizyczną dyskutujemy w następnym rozdziale.

♢ Często stosowany jest jeszcze inny zapis założenia (17.5). W teorii grup wprowadza się **generatory** $\hat{G}_i$, związane z transformacjami bardzo bliskimi jedności wzorami, które w naszym przypadku przyjmują postać

$$R_k(d\varphi) = 1 + id\varphi\hat{G}_k; \qquad k = x, y, z. \tag{17.10}$$

Operatory składowych wektora krętu wiążą się więc z generatorami grupy obrotów prostym wzorem

$$\hat{J}_k = \hbar\hat{G}_k, \qquad k = x, y, z. \tag{17.11}$$

Z tego wzoru wynika, że generatory $\hat{G}_i$ są operatorami hermitowskiemi.

♢ Operator

$$\hat{\vec{G}}^2 = \hat{G}_x^2 + \hat{G}_y^2 + \hat{G}_z^2 = \hbar^{-2}\hat{\vec{J}}^2 \tag{17.12}$$

bywa często nazywany **operatorem Casimira**. Zauważmy, że baza $|j, m\rangle$ może być równie dobrze uważana za zbiór wspólnych wektorów własnych operatora Casimira i operatora $\hat{G}_z$ jak i za zbiór wspólnych wektorów własnych operatorów kwadratu krętu i $\hat{J}_z$

♡ **Twierdzenie**: **oznaczmy przez $\mathcal{H}$ dowolną przestrzeń wektorową niezmienniczą i nieredukowalną względem grupy obrotów. Wszystkie wektory w tej przestrzeni są wektorami własnymi operatora Casimira do tej samej wartości** j. Jako bazę przestrzeni $\mathcal{H}$ można przyjąć wektory $|j, \gamma\rangle$, gdzie γ oznacza wszystkie wskaźniki potrzebne oprócz j do jednoznacznego określenia wektora bazowego. Wybieramy $j = j_1$ odpowiadające jednemu z tych wektorów bazowych. Wektory własne operatora Casimira z $j = j_1$ stanowią przestrzeń wektorową (Rozdział 10). Z faktu, że operator Casimira komutuje ze wszystkimi trzema generatorami wynika, że komutuje także z bardzo małymi obrotami wokół każdej z osi kartezjańskiego układu współrzędnych (x, y, z). Ale z takich obrotów można złożyć każdy obrót, więc dla dowolnego obrotu $\hat{R}$ mamy

$$\hat{\vec{G}}^2\hat{R}|j_1, \gamma\rangle = \hat{R}\hat{\vec{G}}^2|j_1, \gamma\rangle = j_1(j_1 + 1)\hat{R}|j_1, \gamma\rangle. \tag{17.13}$$

To znaczy, że wektor obrócony $\hat{R}|j_1, \gamma\rangle$ odpowiada tej samej wartości własnej operatora Casimira, co wektor $|j_1, \gamma\rangle$. Udowodniliśmy więc, że przestrzeń wektorów własnych operatora Casimira do wartości j_1 jest niezmiennicza. Ale ta przestrzeń jest podprzestrzenią niezmienniczej i nieredukowalnej przestrzeni $\mathcal{H}$.

Ponieważ nie sprowadza się do wektora zerowego, musi więc być identyczna z całą tą przestrzenią. A zatem wszystkie wektory z przestrzeni $\mathcal{H}$ są wektorami własnymi operatora Casimira do tej samej wartości $j = j_1$, co było do udowodnienia.

Pozostaje nam już tylko sprawdzić, jakie wartości j są dozwolone, jakie wartości m są dopuszczalne przy danej wartości j i czy są potrzebne dalsze wskaźniki. Jak pokażemy w następnym rozdziale, odpowiedzi te pytania wynikają z reguł komutacji i z założenia o nieredukowalności.

Rozdział 18

Spin

Podsumujmy uzyskane dotąd informacje na temat spinu. Przez spin cząstki rozumiemy tu kręt tej cząstki w jej układzie spoczynkowym. Dlatego reguły komutacji dla składowych spinu są takie same jak dla składowych całkowitego krętu. Powyższa definicja spinu ma sens tylko dla cząstek o masie różnej od zera. Dla cząstki o masie zero, na przykład dla fotonu, nie da się przejść do układu spoczynkowego i spin trzeba definiować inaczej. Stany cząstki w jej układzie spoczynkowym stanowią przestrzeń wektorową. Zakładamy, że ta przestrzeń jest niezmiennicza względem grupy obrotów i nieredukowalna. Założenie niezmienniczości omawialiśmy już w poprzednim rozdziale. Sens założenia o nieredukowalności wyjaśnia następujący przykład.

Przypuśćmy, że ktoś chciałby uznać za cząstkę atom wodoru na pierwszym poziomie wzbudzonym ($n = 2$). Jeżeli rozmiar atomu jest dużo mniejszy od wszystkich innych wielkości o wymiarze długości występujących w problemie i jeśli z jakichś przyczyn energia atomu nie ulega zmianie, to taką propozycję można rozważać. Pierwszy poziom wzbudzony atomu wodoru jest czterokrotnie zwyrodniały, więc przestrzeń stanów (w układzie spoczynkowym atomu) jest czterowymiarowa. Ta przestrzeń jest niezmiennicza względem grupy obrotów, ale okazuje się, że zawiera dwie nietrywialne podprzestrzenie niezmiennicze względem grupy obrotów: jednowymiarową i trójwymiarową (rozdziały 15 i 16). W myśl przyjętego założenia trzeba więc uznać, że mamy tu dwie różne cząstki o tej samej energii w układzie spoczynkowym atomu. Jedna z nich jest opisywana wektorem z przestrzeni jednowymiarowej, druga wektorami z przestrzeni trójwymiarowej. Okaże się, że pierwsza z ma spin zero, a druga spin jeden. Te dwie cząstki mogą ze sobą interferować. Założenie o nieredukowalności przestrzeni stanów w układzie spoczynkowym jest więc elementem przyjętej tu definicji cząstki. Wiele cząstek, jak na przykład elektrony czy nukleony, spełnia to założenie. W przypadkach bardziej skomplikowanych obiektów, jak omówiony tu atom wodoru, trzeba wprowadzić dwie lub więcej cząstek mogących ze sobą interferować.

W rozważanej przestrzeni wprowadzamy bazę złożoną z wspólnych wektorów własnych kwadratu krętu i rzutu krętu na oś z. Będziemy oznaczali te wektory $|j, m\rangle$, choć na razie nie wykluczyliśmy jeszcze, że jakieś dodatkowe liczby kwantowe oprócz j i m będą potrzebne do jednoznacznego określenia takiego wektora bazowego. Liczbę wektorów bazowych, czyli wymiar przestrzeni, będziemy oznaczali d. Zakładamy, że d jest liczbą skończoną. Reprezentantem wektora stanu

jest więc uporządkowany zbiór d liczb, czyli wektor d-wymiarowy, a reprezentantem operatora, macierz hermitowska o d kolumnach i d wierszach.

Wiemy już, że w danej przestrzeni niezmienniczej i nieredukowalnej wszystkie wektory bazowe mają tę samą wartość liczby kwantowej j. Pozostaje do zbadania: jakie wartości może przyjmować j, czy j jednoznacznie określa przestrzeń nieredukowalną, jakie wartości może przyjmować liczba m dla danej wartości j i czy są potrzebne jakieś dalsze liczby kwantowe?

Elementami macierzy reprezentującej dowolny operator $\hat{O}$ w przestrzeni niezmienniczej i nieredukowalnej względem grupy obrotów są liczby $\langle j, m'|\hat{O}|j, m\rangle$. Takie reprezentacje są znane jako reprezentacje macierzowe. Reprezentacje macierzowe działające w przestrzeniach nieredukowalnych są określane jako reprezentacje nieredukowalne. Przykładem reprezentacji nie macierzowej jest reprezentacja $\hat{J}_z = -i\hbar\frac{\partial}{\partial\varphi}$ wprowadzona w rozdziale 16. Znajdziemy wszystkie macierzowe, skończenie wymiarowe, nieredukowalne reprezentacje dla operatorów $\hat{J}_x, \hat{J}_y, \hat{J}_z, \hat{\vec{J}}^2$. Wiemy już, że każdy wektor z przestrzeni odpowiadającej danemu j jest wektorem własnym operatora kwadratu krętu do wartości własnej $\hbar^2 j(j+1)$. To znaczy, że działanie operatorem kwadratu krętu na stan jest równoważne z pomnożeniem wektora stanu przez liczbę $\hbar^2 j(j+1)$. Wobec tego

$$\vec{J}^2 = \hbar^2 j(j+1)\mathbf{1}, \tag{18.1}$$

gdzie $\mathbf{1}$ oznacza d-wymiarową macierz jednostkową. Poza tym wiemy, że operator $\hat{J}_z$ jest diagonalny i że na diagonali ma wszystkie liczby $\hbar m$. Pozostałe informacje musimy uzyskać z reguł komutacji.

Wzory robią się bardziej przejrzyste, jeżeli zamiast składowych krętu używać proporcjonalnych do nich generatorów $\hat{G}_i$ i jeśli zamiast generatorów $\hat{G}_x$ i $\hat{G}_y$ stosować ich kombinacje liniowe

$$\hat{G}_\pm = \hat{G}_x \pm i\hat{G}_y, \tag{18.2}$$

gdzie i jest jednostką urojoną. Zauważmy, że operatory $\hat{G}_\pm$ nie są hermitowskie, spełniają natomiast relację

$$\hat{G}_+^\dagger = \hat{G}_-. \tag{18.3}$$

Za pomocą reguł komutacji dla generatorów $\hat{G}_x, \hat{G}_y, \hat{G}_z$ otrzymuje się łatwo reguły komutacji

$$[\hat{G}_z, \hat{G}_\pm] = \pm\hat{G}_\pm, \tag{18.4}$$
$$[\hat{G}_+, \hat{G}_-] = 2\hat{G}_z. \tag{18.5}$$

Operator Casimira może być zapisany w postaci

$$\hat{\vec{G}}^2 = \hat{G}_-\hat{G}_+ + \hat{G}_z^2 + \hat{G}_z. \tag{18.6}$$

Z pierwszej pary reguł komutacji wynika, że

$$\hat{G}_z\hat{G}_\pm = \hat{G}_\pm\hat{G}_z \pm \hat{G}_\pm = \hat{G}_\pm(\hat{G}_z \pm 1). \tag{18.7}$$

Działając obu stronami tej równości na wektor $|j, m\rangle$ otrzymujemy

$$\hat{G}_z\hat{G}_\pm|j,m\rangle = (m \pm 1)\hat{G}_\pm|j,m\rangle. \tag{18.8}$$

To znaczy, że każdy z wektorów $\hat{G}_\pm|j,m\rangle$ jest: albo wektorem własnym operatora $\hat{G}_z$ do wartości własnej odpowiednio $m+1$ lub $m-1$, albo wektorem zerowym. Oba te przypadki można zapisać łącznie wzorami

$$\hat{G}_+|j,m\rangle = x_j^m|j,m+1\rangle, \tag{18.9}$$

$$\hat{G}_-|j,m+1\rangle = x_j^{m*}|j,m\rangle. \tag{18.10}$$

Relacja między współczynnikami liczbowymi po prawej stronie wynika, jak zaraz zobaczymy, z relacji (18.3). Mnożąc powyższe wzory stronami przez wektory dualne $\langle j,n|$ i korzystając z ortogonalności wektorów $|j,m\rangle$ dla różnych wartości wskaźnika m, stwierdzamy, że jedynymi różnymi od zera elementami macierzowymi operatorów $\hat{G}_\pm$ są

$$\langle j,m+1|\hat{G}_+|j,m\rangle = x_j^m, \tag{18.11}$$

$$\langle j,m|\hat{G}_-|j,m+1\rangle = x_j^{m*}. \tag{18.12}$$

Każdy element (m,m') pierwszej macierzy może być otrzymany przez sprzężenie zespolone z elementu (m',m) drugiej. Spełniona jest więc relacja (18.3). Zbiór dopuszczalnych wartości liczby kwantowej m musi mieć wartość maksymalną, oznaczymy ją max, bo gdyby do każdego dopuszczalnego m można było znaleźć większe od niego m też dopuszczalne, to wymiar przestrzeni byłby nieskończony wbrew założeniu. Z tej samej przyczyny zbiór wartości m musi mieć wartość najmniejszą min. Wiemy już jednak, że dopuszczalne wartości m zmieniają się krokami $\triangle m = 1$. Wynika z tego, że liczba różnych dopuszczalnych wartości m jest $max - min + 1$, które wobec tego musi być liczbą całkowitą. Z tej dyskusji widać też, że dla zapewnienia niezmienniczości przestrzeni względem grupy obrotów nie potrzeba wprowadzać więcej niż jednego wektora dla danych (j,m). W taki razie, z założenia nieredukowalności wynika, że nie wolno tego robić.

Korzystając ze wzoru na operator Casimira (18.6) i wiedząc jak działają na stany $|j,m\rangle$ operatory $\hat{G}_\pm, \hat{G}_z$, otrzymujemy na diagonalne elementy operatora Casimira, o których wiemy, że są równe $j(j+1)$, wzór

$$|x_j^m|^2 + m^2 + m = j(j+1). \tag{18.13}$$

Określony jest tu tylko moduł współczynnika x_j^m. Tak musi tak być, bo faza tego współczynnika zależy od różnicy faz między stanami $|j,m\rangle$ i $|j,m+1\rangle$, której nie określiliśmy. Zwykle przyjmuje się konwencję Condona i Shortley'a, że liczby x_j^m są rzeczywiste i nieujemne. Wtedy otrzymujemy

$$x_j^m = \sqrt{j(j+1) - m^2 - m} = \sqrt{(j-m)(j+m+1)}. \tag{18.14}$$

Z ostatniego wyrażenia widać, że $max = j$, $min = -j$, a zatem, że 2j jest liczbą całkowitą.

Podsumujmy uzyskane wyniki. Sformułujemy je dla spinu, chociaż jak wynika z dyskusji na początku rozdziału, można je stosować i do całkowitego krętu, jeśli założenie o nieredukowalności przestrzeni stanów jest spełnione.

- Bazę w przestrzeni niezmienniczej i nieredukowalnej stanów cząstki o pędzie zero stanowią wektory $|s,m\rangle$. Liczby s i m są często nazywane spinem cząstki i rzutem spinu na oś z, choć możliwe wyniki pomiaru są odpowiednio $\hbar\sqrt{s(s+1)}$ i $\hbar m$.

- Liczba kwantowa s jest jednoznacznie określona dla danej przestrzeni i może mieć dowolną wartość ze zbioru $s = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$. To jest ważne przewidywanie fizyczne. Nie istnieją cząstki o spinach $\frac{1}{3}$, $\frac{3}{4}$ i tak dalej. Wynik dotyczy cząstek w przestrzeni trójwymiarowej. W dwu wymiarach możliwe są dowolne spiny. Wprowadzono nawet dla takich cząstek nazwę anyony od angielskiego *any* – jakikolwiek.

- Dla danej wartości liczby s liczba m może przyjmować wartości od $+s$ do $-s$ krokami co jeden. Wynika z tego, że przestrzeń ma wymiar d=2s+1. Ten wynik nie przenosi się na cząstki o masie zero. Na przykład foton ma spin jeden, a możliwe rzuty spinu na kierunek pędu są tylko $m = \pm 1$.

- Macierz reprezentująca kwadrat spinu jest iloczynem liczby $\hbar^2 s(s+1)$ i macierzy jednostkowej o $2s+1$ wierszach i $2s+1$ kolumnach.

- Macierz reprezentująca z-ową składowa spinu jest diagonalna i ma na diagonali liczby $\hbar m$ zmieniające się krokami $\hbar\triangle m = \hbar$ od $\hbar s$ do $-\hbar s$.

- Jedynymi różnymi od zera elementami macierzy odpowiadającej operatorowi $\hat{S}_x = \frac{\hbar}{2}(\hat{G}_+ + \hat{G}_-)$ są $\langle s, m+1|\hat{S}_x|s,m\rangle = \langle s,m|\hat{S}_x|s,m+1\rangle = \frac{\hbar}{2}\sqrt{(s-m)(s+m+1)}$.

- Jedynymi różnymi od zera elementami macierzy odpowiadającej operatorowi $\hat{S}_y = \frac{\hbar}{2i}(\hat{G}_+ - \hat{G}_-)$ są $\langle s, m+1|\hat{S}_y|s,m\rangle = -\langle s,m|\hat{S}_y|s,m+1\rangle = \frac{\hbar}{2i}\sqrt{(s-m)(s+m+1)}$.

Najprostszym przykładem jest cząstka o spinie zero, dla której $s = 0$. Wynika stąd, że $m = 0$, $d = 1$; macierze $S_x, S_y, S_z, \vec{S}^2$ są macierzami o jednym wierszu, jednej kolumnie i o jedynym elemencie równym zero. Wektory stanu są jednokomponentowe i nie zmieniają się przy obrotach. Takie funkcje nazywamy skalarnymi (względem obrotów). Reguły komutacji są oczywiście spełnione, ale przypadek jest niezbyt ciekawy. Potwierdza się, że spin cząstki opisywanej przez skalarną funkcje falową jest równy zero. Ciekawszy przypadek cząstki o spinie $s = \frac{1}{2}$ omówimy w następnym rozdziale. Przestrzeń stałych funkcji wektorowych jest trójwymiarowa, niezmiennicza względem obrotów i, jak łatwo sprawdzić, nieredukowalna, więc $2s+1 = 3$ i $s = 1$. Cząstki opisywane przez wektorowe funkcje falowe, mówi się czasem krótko cząstki wektorowe, to są cząstki o spinie jeden.

Rozdział 19

Spin $\frac{1}{2}$

Cząstki ze spinem $\frac{1}{2}$ są interesujące między innymi z dwu przyczyn. Prawie cała materia w naszym otoczeniu składa się z nukleonów i elektronów, czyli z cząstek o spinie $\frac{1}{2}$. Funkcje falowe cząstek o spinie $\frac{1}{2}$ nie są ani skalarami, ani wektorami, więc stanowią dobry przykład mniej znanych własności transformacyjnych funkcji falowych cząstek z różnym od zera spinem.

Jeżeli $j = \frac{1}{2}$, to $d = 2$ i możliwe są wartości $m = \pm\frac{1}{2}$. Jedynym różnym od zera współczynnikiem jest $x_{1/2}^{-1/2} = 1$. Ze wzorów podanych w poprzednim rozdziale otrzymujemy

$$S_x = \frac{\hbar}{2}\sigma_x; \qquad S_y = \frac{\hbar}{2}\sigma_y; \qquad S_z = \frac{\hbar}{2}\sigma_z. \tag{19.1}$$

Macierze reprezentujące operatory piszemy bez daszków. Występujące w tych wzorach macierze Pauli'ego σ_i są zdefiniowane wzorami

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{19.2}$$

Wektory bazowe odpowiadające $m = \frac{1}{2}$ i $m = -\frac{1}{2}$ zapisujemy w postaci

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \qquad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{19.3}$$

Dualne do nich są wektory

$$\alpha^\dagger = (1, 0); \qquad \beta^\dagger = (0, 1). \tag{19.4}$$

Łatwo sprawdzić (ćwiczenie), że wektory α i β stanowią układ ortonormalny i że są wektorami własnymi macierzy S_z odpowiednio do wartości własnych $\pm\frac{\hbar}{2}$. Na podstawie wyników z poprzedniego rozdziału macierz $\vec{S}^2$ jest iloczynem liczby $\frac{3}{4}\hbar^2$ i dwuwymiarowej macierzy jednostkowej. Ten sam wynik można otrzymać podnosząc do kwadratu i dodając do siebie macierze reprezentujące operatory składowych krętu. To, że wektory α i β są wektorami własnymi macierzy $\vec{S}^2$, jest więc oczywiste, bo każdy niezerowy wektor jest wektorem własnym macierzy jednostkowej. Dalej można sprawdzić (ćwiczenie), że otrzymane macierze spełniają wszystkie reguły komutacji otrzymane dla operatorów składowych krętu.

Często można w dobrym przybliżeniu założyć, że hamiltonian nie zależy od spinów cząstek. Wtedy dla jednej cząstki, jeśli funkcja falowa $\psi(\vec{x})$ jest funkcją własną tego hamiltonianu, to są niemi także funkcje

$$\psi_+(\vec{x}) = \psi(\vec{x})\alpha, \qquad \psi_-(\vec{x}) = \psi(\vec{x})\beta, \tag{19.5}$$

bo z punktu widzenia rozwiązywania niezależnego od spinu równania Schrödingera, wektory α i β są stałymi współczynnikami, które można zawsze wprowadzić. Z punktu widzenia przewidywanych wyników pomiaru, czynnik α lub β uzupełnia informację zawartą w skalarnej funkcji falowej informacją o rzucie krętu na oś z w układzie spoczynkowym cząstki. Jeżeli hamiltonian wielocząstkowy, pełny lub efektywny, nie zawiera oddziaływań miedzy cząstkami, to rozwiązaniami są odpowiednio zsymetryzowane iloczyny jednocząstkowych funkcji falowych. Tego przybliżenia używaliśmy dla elektronów w atomie, omawiając układ okresowy pierwiastków w rozdziale dwunastym. Dla fermionów wielocząstkowa funkcja falowa musi być antysymetryczna względem wymian cząstek. Wynika z tego zakaz Pauli'ego, ale w rozdziale dwunastym podaliśmy, że ponieważ elektrony mają spin $\frac{1}{2}$, to w każdym stanie teorii bez spinu można umieścić dwa elektrony. Możemy teraz pokazać jak się to dzieje. Jednocząstkowej funkcji falowej dla cząstki bezspinowej $\psi_0(\vec{x})$ przypisujemy stan dwucząstkowy

$$\psi_s(\vec{x}_1, \vec{x}_2) = \psi_0(\vec{x}_1)\psi_0(\vec{x}_2)\frac{\alpha_1\beta_2 - \beta_1\alpha_2}{\sqrt{2}}. \tag{19.6}$$

Czynnik spinowy nie wpływa na to, że ta funkcja falowa jest funkcją własną hamiltonianu, ale zapewnia antysymetrię wymaganą przez statystykę Fermiego-Diraca. Zauważmy, że gdyby nie spin, taki stan dwucząstkowy byłby niemożliwy, bo część niezależna od spinu (przestrzenna) jest symetryczna, a nie antysymetryczna. Powyższe rozumowanie możemy uogólnić na przypadek dowolnego hamiltonianu niezależnego od spinu, zastępując zsymetryzowany iloczyn jednocząstkowych przestrzennych funkcji falowych dowolną, symetryczną względem wymiany cząstek, funkcją zmiennych przestrzennych $(\vec{x}_1, \vec{x}_2)$. Tak jest zbudowany, w bardzo dobrym przybliżeniu, stan podstawowy atomu helu. Dla stanów wzbudzonych atomu helu sytuacja jest bardziej skomplikowana. Przypuśćmy, że dwa elektrony zajmują dwa różne stany przestrzenne opisywane przez funkcje falowe $\psi_1(\vec{x})$ i $\psi_2(\vec{x})$. Wtedy przy uwzględnieniu spinu można zbudować cztery stany dwucząstkowe

$$\psi_s(\vec{x}_1, \vec{x}_2) = \frac{1}{2}\left(\psi_1(\vec{x}_1)\psi_2(\vec{x}_2) + \psi_2(\vec{x}_1)\psi_1(\vec{x}_2)\right)(\alpha_1\beta_2 - \beta_1\alpha_2), \tag{19.7}$$

$$\psi_{a+}(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}}\left(\psi_1(\vec{x}_1)\psi_2(\vec{x}_2) - \psi_2(\vec{x}_1)\psi_1(\vec{x}_2)\right)\alpha_1\alpha_2, \tag{19.8}$$

$$\psi_{a0}(\vec{x}_1, \vec{x}_2) = \frac{1}{2}\left(\psi_1(\vec{x}_1)\psi_2(\vec{x}_2) - \psi_2(\vec{x}_1)\psi_1(\vec{x}_2)\right)(\alpha_1\beta_2 + \beta_1\alpha_2), \tag{19.9}$$

$$\psi_{a-}(\vec{x}_1, \vec{x}_2) = \frac{1}{\sqrt{2}}\left(\psi_1(\vec{x}_1)\psi_2(\vec{x}_2) - \psi_2(\vec{x}_1)\psi_1(\vec{x}_2)\right)\beta_1\beta_2. \tag{19.10}$$

Ponieważ z punktu widzenia niezależnego od spinów równania Schrödingera, i to nawet zawierającego oddziaływania między elektronami, czynniki spinowe w tych funkcjach są tylko stałymi współczynnikami, wartości średnie energii w trzech stanach zapisanych ze wskaźnikiem a muszą być identyczne. Wartość

energii odpowiadająca stanowi ze wskaźnikiem s może być natomiast inna, bo inna jest przestrzenna część funkcji falowej. Tu też można uogólnić rozumowanie wprowadzając ogólne funkcje położeń symetryczne i antysymetryczne względem wymiany cząstek zamiast zsymetryzowanych i zantysymetryzowanych iloczynów funkcji jednocząstkowych. Dla atomu helu wynika stąd, że powinny występować dwa rodzaje poziomów energetycznych: singlety, do których należy stan podstawowy i triplety. Ten fakt był odkryty doświadczalnie zanim powstała mechanika kwantowa i nie dawał się zrozumieć na gruncie starej teorii kwantów. Wyjaśnienie go było jednym z wczesnych efektownych sukcesów mechaniki kwantowej. Analogiczne rozumowanie wyjaśnia fakt przedstawiony w rozdziale piętnastym, że stan podstawowy dwu atomów wodoru po uwzględnieniu oddziaływania między tymi atomami rozszczepia się na singlet i triplet. Można nawet zrozumieć, dla czego stan związany odpowiada singletowi. Po to, żeby dwa atomy wodoru związały się w cząsteczkę, elektrony muszą z dużym prawdopodobieństwem znaleźć się w niewielkim obszarze między dwoma protonami, żeby je przyciągać i tym samym ekranować ich kulombowskie odpychanie. W stanach a przestrzenna część funkcji falowej jest antysymetryczna względem przestawienia elektronów. W związku z tym jest równa zero dla $\vec{x}_1 = \vec{x}_2$ i z ciągłości nieduża dla $\vec{x}_1 \approx \vec{x}_2$. To zmniejsza prawdopodobieństwo znalezienia się obu elektronów naraz w niewielkim obszarze. Część przestrzenna funkcji ψ_s jest symetryczna względem przestawienia elektronów i ta trudność tam nie występuje. Wobec tego ekranowanie jest skuteczniejsze i energia niższa. Podobnie należy rozumieć używany niekiedy przez chemików zwrot, że po to żeby powstało wiązanie chemiczne potrzebne są dwa elektrony o przeciwnych spinach. Nie chodzi tu o jakieś siły zależne od spinu, tylko o to, żeby część przestrzenna funkcji falowej była symetryczna względem przestawienia elektronów. Zauważmy jeszcze, że przeciwne spiny oznaczają w tym stwierdzeniu tylko funkcję spinową występującą w funkcji ψ_s. To, że w funkcji ψ_{a0} rzuty spinu elektronu na oś z też są przeciwne, nie liczy się.

Zbadamy na koniec własności transformacyjne wektorów α i β przy obrotach. W rozdziale siedemnastym pokazaliśmy, że dla obrotu o bardzo mały kąt $d\varphi$ wokół osi $k = x, y, z$

$$\hat{R}_k(\varphi) = 1 + id\varphi\hat{G}_k. \tag{19.11}$$

Obrót o skończony kąt φ możemy interpretować jako N kolejnych obrotów o kąty $\frac{\varphi}{N}$. Jeżeli N jest bardzo duże i co za tym idzie kąt $\frac{\varphi}{N}$ jest bardzo mały, możemy napisać

$$\hat{R}_k(d\varphi) = \left(1 + \frac{i\varphi}{N}\hat{G}_k\right)^N = e^{i\varphi\hat{G}_k}, \tag{19.12}$$

gdzie równości stają się dokładne przy $N \to \infty$. Przechodząc od ogólnych operatorów do macierzy reprezentujących je, dla cząstek o spinie $\frac{1}{2}$ otrzymujemy

$$R_k(\varphi) = e^{\frac{i\varphi}{2}\sigma_k}. \tag{19.13}$$

Ten wzór można przepisać w innej postaci używając znanego twierdzenia Eulera

$$R_k(\varphi) = \cos\left(\frac{\varphi}{2}\sigma_k\right) + i\sin\left(\frac{\varphi}{2}\sigma_k\right) \tag{19.14}$$

Rozwinięcie funkcji $\cos(x)$ w szereg potęgowy zawiera wyłącznie parzyste potęgi argumentu. Ponieważ, jak łatwo sprawdzić $\sigma_k^2 = 1$, czynnik σ_k w argumencie kosinusa można zastąpić macierzą jednostkową przed kosinusem. Rozwinięcie funkcji $\sin(x)$ w szereg potęgowy zawiera tylko nieparzyste potęgi argumentu, ale każda nieparzysta potęga σ_k jest równa σ_k, więc czynnik σ_k w argumencie sinusa można zastąpić czynnikiem σ_k mnożącym cały sinus. W końcu dostajemy dogodniejsze wyrażenie na macierz $R_k(\varphi)$:

$$R_k(\varphi) = \mathbf{1}\cos\left(\frac{\varphi}{2}\right) + i\sigma_k \sin\left(\frac{\varphi}{2}\right), \tag{19.15}$$

gdzie jedynka tłustym drukiem oznacza dwuwymiarową macierz jednostkową. W szczególności dla obrotów wokół osi z, podstawiając macierz σ_z i korzystając jeszcze raz z wzoru Eulera, otrzymujemy

$$R_z(\varphi) = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}. \tag{19.16}$$

Z tego wzoru widać, że mnożąc wektory α i β lewostronnie przez macierz $R_z(\varphi)$ otrzymujemy

$$R_z(\varphi)\alpha = e^{i\frac{\varphi}{2}}\alpha; \qquad R_z(\varphi)\beta = e^{-i\frac{\varphi}{2}}\beta. \tag{19.17}$$

Szczególnie ciekawy jest wynik obrotu o kąt 2π. Dla skalarów, wektorów czy tensorów taki obrót nic nie zmienia, co intuicyjnie jest dość oczywiste. Tu jednak zarówno wektor α, jak i wektor β zmienia znak. Ponieważ każdy wektor w przestrzeni stanów nieruchomej cząstki o spinie $\frac{1}{2}$ da się przedstawić jako kombinacja liniowa wektorów α i β, każdy wektor z tej przestrzeni zmienia znak pod wpływem obrotu wokół osi z o kąt 2π.

Każdy obrót w przestrzeni trójwymiarowej można przedstawić jako ciąg następujących trzech obrotów o "kąty Eulera": obrotu o kąt γ wokół osi z, obrotu o kąt β wokół osi y i obrotu o kąt α wokół osi z. Działanie operatorów obrotu wokół osi z już znamy. Żeby móc przeprowadzić dowolny obrót wystarczy jeszcze wiedzieć, jak działają obroty wokół osi y. Podstawiając we wzorze (19.15) dla $i = y$ macierz σ_y otrzymujemy

$$R_y(\varphi) = \begin{pmatrix} \cos\left(\frac{\varphi}{2}\right) & \sin\left(\frac{\phi}{2}\right) \\ -\sin\left(\frac{\phi}{2}\right) & \cos\left(\frac{\varphi}{2}\right) \end{pmatrix}. \tag{19.18}$$

Znów dla kąta obrotu 2π jest to minus macierz jednostkowa, więc każdy wektor pomnożony przez nią zmienia znak. Łatwo sprawdzić, że taka zmiana znaku następuje przy obrocie o kąt 2π wokół dowolnej osi, nie tylko wokół osi x, y, z. Wielkości, które przy obrotach transformują się tak jak tu opisaliśmy, są nazywane spinorami. Zauważmy, że iloczyn n spinorów obrócony o kąt 2π zyskuje czynnik fazowy $(-1)^n$, a zatem znak zmienia się wtedy i tylko wtedy, kiedy n jest nieparzyste.

Pozostają do przedyskutowania konsekwencje fizyczne zmiany znaku spinora przy obrocie o kąt 2π. Taki obrót nie powinien zmieniać stanu układu. To, że zmienia znak funkcji falowej, nie jest szkodliwe, bo stały czynnik fazowy w funkcji falowej jest nieobserwowalny. Problem mógłby powstać tylko wtedy, gdybyśmy rozważali sumę wektorów stanu, z których jeden, powiedzmy $|\varphi\rangle$,

zmienia znak przy obrocie o kąt 2π, a drugi, powiedzmy $|\psi\rangle$, nie zmienia znaku. Mielibyśmy wtenczas

$$R(2\pi)(|\psi\rangle + |\varphi\rangle) = |\psi\rangle - |\varphi\rangle. \tag{19.19}$$

To już nie jest zmiana fazy, tylko mierzalna zmiana stanu. Wyjściem z tej trudności jest przyjęcie postulatu, że nie wolno tworzyć kombinacji liniowych z wektorów stanu zmieniających znak przy obrocie o 2π i wektorów stanu nie zmieniających znaku przy takim obrocie. Ze względu na zasadę superpozycji oznacza to w szczególności, że jeden układ nie może mieć stanów tych dwu rodzajów. Nie może też być przejść między takimi stanami. Wynika stąd dalej, że w układzie zawierającym tylko cząstki o spinie $\frac{1}{2}$, takie cząstki mogą pojawiać się czy znikać tylko parami tak, żeby parzystość ich liczby nie uległa zmianie. Podaliśmy już tę regułę w rozdziale dwunastym, jako jeden ze sposobów stwierdzania, że cząstki są fermionami. Rzeczywiście, można udowodnić, że zmiana znaku przy obrocie o kąt 2π występuje dla wszystkich spinów połówkowych, czyli dla wszystkich fermionów, a nie występuje nigdy dla spinów całkowitych, czyli dla bozonów. Ogólną regułą jest więc, że dany układ nie może mieć stanów z różną parzystością liczby fermionów i, co za tym idzie, że parzystość liczby fermionów w układzie nie może się zmieniać. Jeżeli się przestrzega tej reguły, to zmiana znaku niektórych funkcji falowych przy obrocie o kąt 2π nie prowadzi do żadnych sprzeczności.

Rozdział 20

Równanie Schrödingera dla cząstki ze spinem 1/2

Funkcję falową cząstki o spinie $s = \frac{1}{2}$ można ogólnie zapisać w postaci

$$\Psi(\vec{x}, s_z) = \psi_+(\vec{x})\alpha + \psi_-(\vec{x})\beta = \begin{pmatrix} \psi_+(\vec{x}) \\ \psi_-(\vec{x}) \end{pmatrix}, \tag{20.1}$$

gdzie s_z oznacza rzut spinu na oś z. W niezależnym od czasu równaniu Schrödingera,

$$\hat{H}(\vec{x}, s_z)\Psi(\vec{x}, s_z) = E\Psi(\vec{x}, s_z), \tag{20.2}$$

operator $\hat{H}(\vec{x}, s_z)$ musi więc być macierzą o dwu wierszach i dwu kolumnach.

Najprostszy jest przypadek, kiedy hamiltonian nie zależy od spinu, to znaczy kiedy jest proporcjonalny do macierzy jednostkowej

$$\hat{H}(\vec{x}, s_z) = \begin{pmatrix} \hat{H}(\vec{x}) & 0 \\ 0 & \hat{H}(\vec{x}) \end{pmatrix}. \tag{20.3}$$

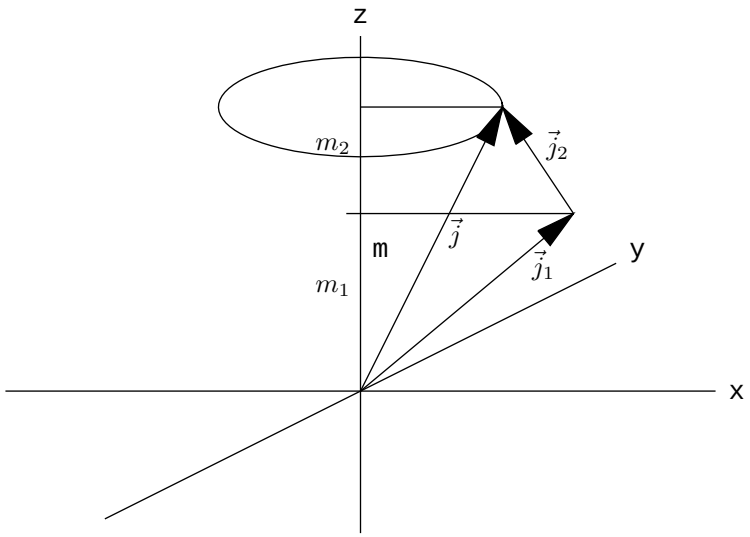
Wtedy równanie Schrödingera rozpada się na dwa identyczne równanie bezspinowe

$$\hat{H}(\vec{x})\psi_+(\vec{x}) = E\psi_+(\vec{x}), \tag{20.4}$$
$$\hat{H}(\vec{x})\psi_-(\vec{x}) = E\psi_-(\vec{x}). \tag{20.5}$$

Jako rozwiązania można przyjąć spinory

$$\Psi_+(\vec{x}, s_z) = \psi(\vec{x})\alpha, \qquad \Psi_-(\vec{x}, s_z) = \psi(\vec{x})\beta, \tag{20.6}$$

lub jakiekolwiek ich dwie, wzajemnie ortogonalne kombinacje liniowe. Pominęliśmy wskaźnik przy funkcji $\psi(\vec{x})$, bo w obu rozwiązaniach może to być ta sama funkcja. Musi tak nawet być, jeśli wartość własna E, rozumiana jako wartość własna równania $\hat{H}(\vec{x})\psi(\vec{x}) = E\psi(\vec{x})$, jest niezdegenerowna. Dla hamiltonianu niezależnego od spinu wylicza się więc funkcje własne tak, jakby cząstka miała spin zero, a następnie mnoży się otrzymaną funkcję przez spinor α lub β. W ten sposób każdemu rozwiązaniu bez spinu odpowiadają dwa rozwiązania spinowe o tej samej energii.



Rysunek 20.1: **Model wektorowy:** kręt j precesuje wokół osi z tak, że jego długość i rzut na oś z (m) są stałe. Wektory krętu j_1 i j_2 precesują wokół wektora j tak, że ich długości pozostają stałe i że ich suma wektorowa jest stale równa j. Rzuty wektorów j_1, j_2 na oś z (m_1, m_2) nie są ustalone, ale ich suma zawsze wynosi m.

Funkcje $\Psi_{\pm}(\vec{x}, s_z)$ zdefiniowane powyżej nie są naogół funkcjami własnymi całkowitego krętu układu. Przyczynę ilustruje model wektorowy przedstawiony na rysunku 1. Prostszą wersję tego modelu rozpatrywaliśmy w rozdziale szesnastym. Tu rozpatrujemy przypadek, kiedy całkowity kręt j jest sumą dwóch krętów j_1 i j_2. Na przykład kręt elektronu w atomie wodoru jest sumą jego krętu orbitalnego i jego spinu:

$$\vec{j} = \vec{l} + \vec{s}. \tag{20.7}$$

Nie będziemy omawiali teorii składanie krętów w mechanice kwantowej, ograniczając się do zwrócenia uwagi na fakty widoczne z modelu wektorowego (co oczywiście nie jest żadnym dowodem!). W modelu, wektor wypadkowy $\vec{j}$ precesuje wokół osi z tak, że jego rzut m na tę oś pozostaje stały. Mamy więc, jak to było omówione w rozdziale szesnastym, dobrze określone: długość wektora $\vec{j}$ i jego rzut na oś z. Rzuty wektora $\vec{j}$ na osie x i y są natomiast nieokreślone. Wektor $\vec{j}$ jest sumą wektorową wektorów $\vec{j}_1$ i $\vec{j}_2$. Wymaga to, według elementarnej geometrii, żeby były spełnione tak zwane nierówności trójkąta

$$|j_1 - j_2| \leq j \leq j_1 + j_2. \tag{20.8}$$

Symbole j, j_1, j_2 oznaczają tu liczby kwantowe występujące na przykład w definicji wektorów bazowych, a nie pierwiastki z wartości własnych operatorów kwadratów krętów (podzielone przez $\hbar$). Na przykład elektron P w atomie wodoru $(l = 1)$ może mieć tylko $j = \frac{1}{2}$ lub $j = \frac{3}{2}$. Wektory j_1, j_2 precesują wokół wektora j. Jak widać z rysunku, nie tylko rzuty tych wektorów na osie x i y, ale także i ich rzuty na oś z są nieokreślone. Zachodzi natomiast równość

$$m = m_1 + m_2. \tag{20.9}$$

Żeby uzyskać stan o danym kręcie równym j i jego rzucie na oś z równym m, czyli $|j, m\rangle$, należy zbudować kombinację liniową stanów iloczynowych $|j_1, m_1\rangle|j_2, m_2\rangle$, zawierającą naogół wszystkie składniki, dla których $m_1+m_2 = m$. Nie będziemy dalej omawiać tego problemu, choć jest to problem ważny. Dla sferycznie symetrycznego hamiltonianu można jednocześnie z energią określić wartości własne operatorów kwadratu całkowitego krętu i jego rzutu na oś z, podczas gdy analogiczne operatory zbudowane z krętu orbitalnego, lub spinu, naogół nie komutują z hamiltonianem.

Ciekawe są przypadki, kiedy hamiltonian zależy od spinu. Na przykład dla atomu wodoru używa się często hamiltonianu

$$\hat{H}(\vec{x}, s_z) = \mathbf{1}\hat{H}_0(\vec{x}) - \frac{e\hbar}{2m_e c}\vec{B}\cdot\vec{\sigma} + \frac{\hbar}{4m_e^2c^2}\frac{1}{r}\frac{\partial V(r)}{\partial r}\vec{\sigma}\cdot\vec{L}. \tag{20.10}$$

Pierwszy wyraz po prawej stronie jest zwykłym bezspinowym hamiltonianem zawierającym energię kinetyczną elektronu, energię oddziaływania kulombowskiego elektronu z protonem i energię oddziaływania poruszającego się elektronu z zewnętrznym polem magnetycznym. W szczególności, jeżeli elektron ma kręt orbitalny $\vec{L}$, to energia oddziaływania jego ruchu orbitalnego z polem magnetycznym, dla nie za silnych pól, dana jest znanym z klasycznej fizyki wzorem

$$E_{1magn} = -\frac{e}{2mc}\vec{L}\cdot\vec{B}. \tag{20.11}$$

W tym wzorze e i m oznaczają odpowiednio ładunek elektronu ($e < 0$) i masę elektronu. Można też powiedzieć, że krążący elektron, jak każdy krążący ładunek elektryczny, wytwarza magnetyczny moment dipolowy

$$\vec{\mu}_{orb} = \frac{e}{2mc}\vec{L}, \tag{20.12}$$

którego energia oddziaływania z polem magnetycznym dana jest zwykłym wyrażeniem $-\vec{\mu}\cdot\vec{B}$.

Drugi człon – znany jako człon Pauli'ego – daje dodatkowe oddziaływanie elektronu z zewnętrznym polem magnetycznym. Składowymi wektora $\vec{\sigma}$ są macierze Pauli'ego, więc z definicji iloczynu skalarnego

$$\vec{B}\cdot\vec{\sigma} = \begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{pmatrix}. \tag{20.13}$$

Szczególnie prosty jest przypadek, kiedy trzeci człon można zaniedbać, na przykład dla tego, że pole magnetyczne jest bardzo silne. Wybierając oś z wzdłuż pola magnetycznego mamy[1] $B_x = B_y = 0$ i wobec tego niezależne od czasu równanie Schrödingera przybiera postać

$$\left(\hat{H}_0(\vec{x}) + \mu B\right)\psi_+(\vec{x}) = E\psi_+(\vec{x}), \tag{20.14}$$

$$\left(\hat{H}_0(\vec{x}) - \mu B\right)\psi_-(\vec{x}) = E\psi_-(\vec{x}). \tag{20.15}$$

[1] Ponieważ pole magnetyczne musi spełniać równanie Maxwella $\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0$, założenie $B_x = B_y = 0$ oznacza, że $B_z = \mathrm{Const}$, czyli że pole magnetyczne jest jednorodne.

Wprowadziliśmy tu wartość liczbową momentu magnetycznego elektronu

$$\mu = \frac{|e|\hbar}{2m_e c}. \tag{20.16}$$

Ta liczba jest znana jako magneton Bohra. Zastępując masę elektronu przez masę nukleonu, otrzymalibyśmy inną, blisko dwa tysiące razy mniejszą, jednostkę nazywaną magnetonem jądrowym. Widać, że poprawka do energii od członu Pauli'ego rozszczepia poziom energetyczny na dwa, przesunięte odpowiednio o $\pm|\mu B|$. Efekt jest taki, jakby elektron posiadał oprócz momentu magnetycznego związanego z ruchem orbitalnym $\vec{\mu}_{orb}$ jeszcze związany ze spinem moment magnetyczny wielkości μ, który się ustawia równolegle lub antyrównolegle do pola magnetycznego. Formalnie można napisać

$$\vec{\mu} = \frac{e\hbar}{2mc}\vec{\sigma} = \frac{e}{mc}\vec{S}, \tag{20.17}$$

gdzie $\vec{S}$ oznacza spin elektronu. Zauważmy, że współczynnik przy wektorze spinu w tym wzorze jest dwa razy większy niż współczynnik przy kręcie orbitalnym we wzorze (20.12). Nie da się więc przedstawić energii oddziaływania z polem magnetycznym jako iloczynu jakiegoś współczynnika liczbowego i całkowitego krętu $\vec{J} = \vec{L} + \vec{S}$. Dla danego J energie stanów o różnych L i S są różnie przesuwane pod wpływem pola magnetycznego. Prowadzi to do rozszczepień linii widm atomowych znanych jako efekt Zeemana. Ściślej mówiąc, jest to normalny efekt Zeemana. Zwykle w praktyce obserwuje się anomalny efekt Zeemana, gdzie przyczynki do przesunięć poziomów od drugiego i trzeciego członu w hamiltonianie (20.10) są porównywalne. Komplikuje to oczywiście teorię.

Jeżeli pole magnetyczne nie jest jednorodne, to interesujące jest obliczenie siły, zależnej od tego pola, działającej na elektron. Różniczkując człon Pauli'ego po z, zmieniając znak i średniując po stanie spinowym otrzymujemy z-ową składową siły

$$F_z = \mu\frac{\partial B_z}{\partial z}\langle\sigma_z\rangle, \tag{20.18}$$

gdzie założyliśmy, że $\langle\sigma_x\rangle = \langle\sigma_y\rangle = 0$. Ten wzór stanowi podstawę analizy doświadczenia Sterna - Gerlacha.

Trzeci człon ma taką postać tylko jeśli potencjał $V(\vec{x})$ jest sferycznie symetryczny, co zaznaczyliśmy pisząc $V(r)$. Jest to tak zwane sprzężenie spin-orbita, które może być zinterpretowane następująco. Elektron w swoim ruchu orbitalnym wytwarza pole magnetyczne. To pole oddziałuje ze spinowym momentem magnetycznym elektronu. Powodowane przez ten człon rozszczepienia linii widmowych atomu wodoru są nazywane strukturą subtelną linii. Rzeczywiście, te rozszczepienia są znacznie mniejsze niż różnice między wartościami własnymi bezspinowego hamiltonianu. Analogiczne rozszczepienia są natomiast duże i ważne w fizyce jądrowej.

Jest jeszcze wiele innych efektów spinowych, których nie uwzględniliśmy w hamiltonianie (20.10). Wspomnimy jeszcze jeden. Energia oddziaływania momentu magnetycznego elektronu z momentem magnetycznym jądra daje tak zwaną strukturę nadsubtelną linii. Na przykład stan podstawowy atomu wodoru rozszczepia się na dwa stany o energiach tak bliskich, że przejściom między niemi odpowiada promieniowanie o długości fali około 21cm, co odpowiada

częstości około 1420 megaherców. To promieniowanie jest wykorzystywane przez astronomów do oceny ilości wodoru w różnych obszarach kosmosu. Obserwując modyfikacje częstości tej fali, można nawet oceniać pola magnetyczne w obszarach, z których to promieniowanie pochodzi. Badania nadsubtelnych rozszczepień linii widmowych atomów dają też ważne informacje o strukturze jąder.

Rozdział 21

Równania ruchu – obrazy mechaniki kwantowej – równanie Schrödingera zależne od czasu.

W mechanice klasycznej, znając stan układu w jakiejś chwili $t = t_0$, można za pomocą równań ruchu przewidzieć stan tego układu w jakiejkolwiek późniejszej chwili $t > t_0$. W mechanice kwantowej sprawa jest bardziej skomplikowana, bo wektor stanu, określający stan układu, nie jest wielkością mierzalną. Mierzalne są tylko wartości średnie, czyli elementy macierzowe $\langle\psi|\hat{O}|\psi\rangle$, gdzie operator $\hat{O}$ odpowiada wielkości mierzalnej O, a $|\psi\rangle$ jest wektorem stanu układu. Można więc zrobić prawie dowolne założenie o zależności wektora stanu $|\psi\rangle$ od czasu, a następnie otrzymać poprawną zależność elementu macierzowego od czasu przez założenie odpowiedniej zależności operatora od czasu. Można też wybrać prawie dowolnie zależność operatora od czasu i uzyskać poprawną zależność elementu macierzowego od czasu przez dopasowanie zależności wektorów stanu od czasu. Z powodu tej dowolności istnieje wiele "obrazów" mechaniki kwantowej, które wszystkie dają te same przewidywania dla zależności od czasu wielkości mierzalnych, ale proponują bardzo różne równania ruchu dla wektorów stanu i operatorów. Różnica między mechaniką kwantową Schrödingera i mechaniką kwantową Heisenberga polegała głównie na tym, że używali różnych obrazów mechaniki kwantowej.

Załóżmy na przykład, że wektor stanu układu spełnia równanie

$$i\hbar\frac{d}{dt}|\psi\rangle = \hat{A}|\psi\rangle, \tag{21.1}$$

gdzie operator $\hat{A}$ jest dowolnym operatorem hermitowskim tak, że równanie dla dualnego wektora stanu ma postać

$$-i\hbar\frac{d}{dt}\langle\psi| = \langle\psi|\hat{A}. \tag{21.2}$$

Zakładamy teraz, że równanie ruchu dla operatorów ma postać

$$\frac{d\hat{O}}{dt} = \frac{i}{\hbar}\left[\hat{H} - \hat{A}, \hat{O}\right] + \frac{\partial \hat{O}}{\partial t}. \tag{21.3}$$

Różniczkując element macierzowy $\langle\psi|\hat{O}|\psi\rangle$ po czasie zgodnie ze zwykłą regułą różniczkowania iloczynu i podstawiając za pochodne wektorów i operatorów podane wyżej wzory otrzymujemy

$$\frac{d}{dt}\langle\psi|\hat{O}|\psi\rangle = \frac{i}{\hbar}\langle\psi|\left[\hat{H}, \hat{O}\right]|\psi\rangle + \langle\psi|\frac{\partial \hat{O}}{\partial t}|\psi\rangle. \tag{21.4}$$

Ten wynik w ogóle nie zależy od wyboru operatora $\hat{A}$. Powyższe równanie okazuje się poprawne, jeżeli operator $\hat{H}$ interpretować jako hamiltonian układu, i stanowi jedną z podstawowych zasad mechaniki kwantowej. Pokażemy dalej w tym rozdziale, jak ten fakt można próbować zrozumieć. Wynika z niego jednak, że podane na początku tego rozdziału równanie ruchu dla wektorów stanu i operatorów są poprawne przy każdym wyborze hermitowskiego operatora $\hat{A}$. Jeżeli potrafimy uzasadnić którekolwiek z nich, to tym samym uzasadniliśmy postulat (21.4). Chociaż formalnie każdy wybór $\hat{A}$ jest równie dobry, powstaje pytanie, czy w praktyce jakiś wybór nie jest dogodniejszy od innych. Według powszechnego przekonania nie ma jakiegoś najlepszego wyboru. W różnych sytuacjach różne wybory okazują się dogodne. Różnych obrazów mechaniki kwantowej opisano w literaturze wiele, ale szczególnie często używane są następujące trzy.

- **Obraz Heisenberga.** W tym obrazie $\hat{A} = 0$, a zatem wektory stanu nie zależą od czasu. Zależność od czasu operatorów jest natomiast dana dość skomplikowanym równaniem Heisenberga:

  $$\frac{d\hat{O}}{dt} = \frac{i}{\hbar}\left[\hat{H}, \hat{O}\right] + \frac{\partial \hat{O}}{\partial t}. \tag{21.5}$$

  Zastosowania obrazu Heisenberga zostaną omówione w rozdziale dwudziestym piątym.

- **Obraz Schrödingera.** W tym obrazie $\hat{A} = \hat{H}$. To znaczy, że tylko jawna zależność operatorów od czasu jest uwzględniona. $\frac{d\hat{O}}{dt} = \frac{\partial \hat{O}}{\partial t}$. Operatory będące funkcjami tylko współrzędnych, składowych wektorów pędu, spinów itd, w ogóle od czasu nie zależą. Wektor stanu natomiast, spełnia dość skomplikowane równanie Schrödingera zależne od czasu

  $$i\hbar\frac{d}{dt}|\psi\rangle = \hat{H}|\psi\rangle. \tag{21.6}$$

  Obraz Schrödingera zostanie użyty w tym rozdziale do uzasadnienia postulatu (21.4). Różne jego zastosowania omówimy w rozdziałach dwudziestym drugim, dwudziestym trzecim i dwudziestym czwartym. Z tego, że więcej miejsca poświęcamy obrazowi Schrödingera niż obrazowi Heisenberga, nie należy wyciągać wniosku, że obraz Heisenberga jest mniej ważny. Rzeczywiście, proste nierelatywistyczne problemy zwykle łatwiej jest rozwiązywać i dyskutować w obrazie Schrödingera, ale w bardziej skomplikowanych problemach, zwłaszcza relatywistycznych, często obraz Heisenberga okazuje się dogodniejszy.

- **Obraz oddziaływania**, nazywany też obrazem Diraca lub obrazem Tomonagi. Ten obraz stosuje się tylko w sytuacji, kiedy hamiltonian układu $\hat{H}$ jest sumą dwu operatorów hermitowskich $\hat{H}_0$ i $\hat{H}_1$. Zwykle operator $\hat{H}_0$ jest hamiltonianem niezaburzonym, a operator $\hat{H}_1$ zaburzeniem. W obrazie oddziaływania $\hat{A} = \hat{H}_1$ i równania ruchu przybierają postać

$$i\hbar \frac{d}{dt}|\psi\rangle = \hat{H}_1|\psi\rangle, \tag{21.7}$$

$$\frac{d\hat{O}}{dt} = \frac{i}{\hbar}\left[\hat{H}_0, \hat{O}\right] + \frac{\partial \hat{O}}{\partial t}. \tag{21.8}$$

  Ewolucja operatorów jest więc taka, jaka byłaby w obrazie Heisenberga, gdyby operator $\hat{H}_1$ był zaniedbywalny, a ewolucja funkcji falowej jest taka, jaka byłaby w obrazie Schrödingera, gdyby hamiltonianem był operator $\hat{H}_1$. W sytuacji, kiedy stosuje się rachunek zaburzeń, operator $\hat{H}_1$ jest "mały" i wektor stanu powoli zmienia się z czasem. Ten obraz jest szczególnie wygodny przy wyprowadzaniu wzorów metody zaburzeń dla procesów zależnych od czasu.

Ponieważ wektory stanu i operatory zmieniają się z upływem czasu inaczej w każdym z obrazów, powstaje pytanie, jak porównywać stany i operatory w różnych obrazach. Zwykle przyjmuje się konwencję, że w chwili $t = 0$ wektory stanu i operatory mają jakieś dane wartości, a następnie z upływem czasu zmieniają się, w każdym obrazie inaczej.

Pokażemy teraz uzasadnienie postulatu (21.4). Nie jest to żaden dowód – postulatów się nie dowodzi – ale rozumowanie pomagające zrozumieć, skąd się taki postulat wziął. Jak już wspominaliśmy, założenie (21.4) jest równoważne z równaniami ruchu w którymkolwiek obrazie. Wybieramy obraz Schrödingera i zakładamy, że hamiltonian jest operatorem lokalnym (rozdział siódmy). Wtedy zależne od czasu równanie Schrödingera w reprezentacji położeń przybiera postać[1]

$$i\hbar \frac{\partial \psi(\vec{x})}{\partial t} = \hat{H}(\vec{x})\psi(\vec{x}). \tag{21.9}$$

To równanie jest podobne do równań dla operatorów składowych pędu. Na przykład

$$-i\hbar \frac{\partial \psi(\vec{x})}{\partial x} = \hat{p}_x(\vec{x})\psi(\vec{x}). \tag{21.10}$$

Niema w tym podobieństwie nic dziwnego, bo jak wiadomo z teorii względności energia i wektor pędu, podobnie jak czas i wektor położenia, tworzą razem jeden czterowektor. Pozostaje tylko do wyjaśnienia różnica w znaku. Otóż, żeby takimi czterowektorami wolno było operować jak zwykłymi wektorami euklidesowymi, trzeba przyjąć, że czwarta składowa jest urojona. Mamy więc czterowektory $(\vec{x}, it)$ oraz $(\vec{p}, iE)$. Powyższemu wzorowi dla składowych (trój)pędu odpowiada więc wzór dla czwartych składowych $i\hat{E} = -i\hbar\frac{\partial}{\partial it}$ czyli $\hat{E} = +i\hbar\frac{\partial}{\partial t}$, co było do wykazania.

[1] Pełna pochodna po czasie zamienia się na cząstkową, bo $\langle\vec{x}|\frac{d}{dt}|\psi\rangle = \frac{\partial}{\partial t}\langle\vec{x}|\psi\rangle$ — różniczkowanie po czasie ma być prowadzone przy stałym $\vec{x}$.

Można się przekonać o tej różnicy znaków jeszcze prościej, jeżeli przyjąć, że klasyczny wzór na pole fali płaskiej $e^{-i\omega t+i\vec{k}\cdot\vec{x}}$ jest też wzorem na funkcję falową cząstki swobodnej. Żeby uzyskać czynnik równy składowej pędu, trzeba tę funkcję zróżniczkować po odpowiedniej współrzędnej, pomnożyć przez $-i\hbar$ i skorzystać ze wzoru de Broglie'a. Żeby uzyskać czynnik równy energii, trzeba tę funkcję zróżniczkować po czasie, pomnożyć przez $+i\hbar$ i skorzystać ze wzoru Einsteina. Widać stąd, że równanie Schrödingera zależne od czasu jest naturalnym odpowiednikiem wzoru na operatory składowych pędu, co oczywiście nie jest dowodem jego poprawności. Dowodem takim jest ogromna liczba przewidywań zgodnych z doświadczeniem, które za pomocą tego równania uzyskano.

Rozdział 22

Stany stacjonarne i pakiety falowe

Równanie Schrödingera zależne od czasu można łatwo rozwiązać analitycznie, jeśli spełnione są następujące dwa warunki.

- Hamiltonian nie zależy explicite od czasu, to znaczy

$$\frac{\partial \hat{H}(\vec{x})}{\partial t} = 0. \tag{22.1}$$

- Rozwiązania niezależnego od czasu równania Schrödingera z tym hamiltonianem

$$\hat{H}(\vec{x})\psi_n(\vec{x}) = E_n\psi_n(\vec{x}) \tag{22.2}$$

są znane.

W takim przypadku, wprowadzając oznaczenie

$$\psi_n(\vec{x}, t) = \psi_n(\vec{x})e^{-i\omega_n t}, \tag{22.3}$$

gdzie jak zwykle $\omega_n = \frac{E_n}{\hbar}$, łatwo sprawdzić przez podstawienie, że

$$i\hbar\frac{\partial \psi_n(\vec{x}, t)}{\partial t} = \hat{H}(\vec{x})\psi_n(\vec{x}, t). \tag{22.4}$$

Otrzymujemy więc obszerną klasę rozwiązań zależnego od czasu równania Schrödingera.

Stany układu odpowiadające tym rozwiązaniom są znane jako stany stacjonarne i mają szereg ciekawych własności. Ponieważ dla każdej wartości t funkcja falowa $\psi_n(\vec{x}, t)$ jest funkcją własną hamiltonianu odpowiadającą wartości własnej E_n, układ w stanie stacjonarnym ma energię równą E_n, ściśle określoną i niezależną od czasu. Prawdziwe jest też twierdzenie odwrotne: jeżeli układ ma ściśle określoną energię E_n stałą w czasie, co znaczy, że dla każdej chwili t funkcja falowa jest wektorem własnym hamiltonianu do wartości własnej E_n, to stan układu jest stanem stacjonarnym.

Dla chwili $t = 0$ czynnik wykładniczy w funkcji falowej jest równy jeden i otrzymujemy

$$\psi_n(\vec{x},0) = \psi_n(\vec{x}). \tag{22.5}$$

Rozwiązanie niezależnego od czasu równania Schrödingera jest tu więc warunkiem początkowym dla równania Schrödingera zależnego od czasu. Dla równania różniczkowego pierwszego rzędu względem czasu, jak zależne od czasu równanie Schrödingera, warunek początkowy jednoznacznie określa rozwiązanie. Ponieważ, oczywiście, także rozwiązanie określa jednoznacznie warunek początkowy – znaleźć wszystkie możliwe rozwiązania równania znaczy to samo co znaleźć rozwiązania odpowiadające wszystkim możliwym warunkom początkowym. Stany stacjonarne są rozwiązaniami dla szczególnych warunków początkowych (22.5). Są to jedyne rozwiązania przy takich warunkach początkowych. Stany o ustalonej energii, jak stan podstawowy atomu wodoru czy piąty stan wzbudzony oscylatora harmonicznego, są stanami stacjonarnymi. W stanie stacjonarnym

$$|\psi_n(\vec{x},t)|^2 = |\psi_n(\vec{x})|^2. \tag{22.6}$$

Gęstość prawdopodobieństwa znalezienia cząstki w przestrzeni jest więc niezależna od czasu i taka, jak wynika z rozwiązania niezależnego od czasu równania Schrödingera.

W ogólnym przypadku funkcja wyrażająca warunek początkowy: $\psi(\vec{x},0)$, nie jest funkcją własną hamiltonianu, ale ponieważ wektory własne hamiltonianu stanowią układ zupełny, zawsze możemy napisać

$$\psi(\vec{x},0) = \sum_n c_n \psi_n(\vec{x}). \tag{22.7}$$

Jak zwykle piszemy sumę, ale pamiętamy, że jej część lub całość może zostać zastąpiona całką. Rozwiązaniem równania Schrödingera zależnego od czasu jest w tym wypadku funkcja falowa

$$\psi(\vec{x},t) = \sum_n c_n \psi_n(\vec{x}) e^{-i\omega_n t}, \tag{22.8}$$

co łatwo sprawdzić przez podstawienie. Ponieważ takie rozwiązanie potrafimy zbudować dla każdego warunku początkowego, jest to ogólne rozwiązanie zależnego od czasu równania Schrödingera w przypadku, kiedy hamiltonian nie zależy explicite od czasu. Jest oczywiście problem praktyczny – żeby korzystać z tego rozwiązania trzeba znać funkcje własne hamiltonianu: $\psi_n(\vec{x})$. Ważnym zastosowaniem tych rozwiązań jest badanie ewolucji w czasie pakietów falowych: w chwili $t = 0$ budujemy pakiet falowy, który ma takie własności jak chcemy, a następnie budujemy odpowiadające mu rozwiązanie równania Schrödingera zależnego od czasu i patrzymy, co się z pakietem dzieje w miarę upływu czasu. Oczywiście, żeby taką analizę przeprowadzić, trzeba mieć dany hamiltonian $\hat{H}(\vec{x})$.

Przykłady pakietów falowych omówimy w następnym rozdziale. Tu chcemy zwrócić uwagę na pewną ogólną ich własność. Przypuśćmy, że wybraliśmy warunek początkowy dla cząstki w jednym wymiarze

$$\psi(x,0) = \frac{1}{\sqrt[4]{2\pi R^2}} e^{-\frac{x^2}{4R^2}}, \tag{22.9}$$

gdzie R^2 jest liczbą dodatnią. Znając gęstość prawdopodobieństwa $|\psi(x,0)|^2$, łatwo znaleźć (porównaj następny rozdział) średnie: $\langle x\rangle = 0$ i $\langle x^2\rangle = R^2$. Jeżeli

więc parametr R^2 jest mały, ten pakiet opisuje cząstkę dobrze zlokalizowaną w otoczeniu punktu $x = 0$. Na pytanie, jak ten pakiet będzie ewoluował w miarę upływu czasu, nie da się bez dalszych założeń odpowiedzieć. Jeżeli hamiltonian jest hamiltonianem oscylatora harmonicznego ze stałą siłową dobraną tak, że $\psi_0(x) = \psi(x, 0)$, to odpowiadające temu warunkowi początkowemu rozwiązanie zależnego od czasu równania Schrödingera jest stanem stacjonarnym i rozkład prawdopodobieństwa znalezienia cząstki w przestrzeni nie będzie się w ogóle zmieniał. Jeśli hamiltonian jest hamiltonianem cząstki swobodnej, to pakiet jest skomplikowaną superpozycją fal płaskich o różnych częstościach. Każda taka fala z osobna dałaby jednakową gęstość prawdopodobieństwa znalezienia cząstki w każdym punkcie przestrzeni ($-\infty < x < \infty$); ale przy $t = 0$, w superpozycji, dla $x \approx 0$ interferencja jest konstruktywna i prawdopodobieństwo znalezienia cząstki jest znaczne, wszędzie gdzie indziej interferencja jest destruktywna i prawdopodobieństwo znalezienia cząstki jest bliskie zera. Wybierzmy z tej superpozycji dwie fale o częstościach $\omega_1 > \omega_2$. W chwili $t = 0$ te dwie fale interferują konstruktywnie, czyli mają w przybliżeniu takie same fazy. Po upływie czasu t powstanie między nimi różnica faz $(\omega_1 - \omega_2)t \equiv \triangle E \triangle t/\hbar$, gdzie różnicę energii $E_1 - E_2$ oznaczyliśmy $\triangle E$ i czas t od powstania pakietu oznaczyliśmy $\triangle t$. Póki ta różnica jest mała, interferencja jest dalej konstruktywna. Interferencja destruktywna może się pojawić dopiero kiedy

$$\triangle E \triangle t \geq \hbar. \tag{22.10}$$

Ten wynik można stosować do całego pakietu, jeśli ΔE jest odpowiednio uśrednione po parach fal płaskich składających się na pakiet.

Zauważmy, że to rozumowanie, które przeprowadziliśmy na przykładzie swobodnego pakietu falowego, można zastosować przy dowolnym hamiltonianie, jeśli tylko na początku, w chwili $t = 0$, mamy pakiet dobrze zlokalizowany w przestrzeni, powstały dzięki interferencji składników o różnych energiach. To jest słynna zasada nieokreśloności Heisenberga dla energii i czasu: czas potrzebny na to, żeby pakiet się rozpłynął (zniknął), może być dowolnie duży czy nawet nieskończony, ale nie może być dużo mniejszy niż stosunek parametru $\hbar$ do typowej różnicy energii dla fal stanowiących pakiet. Parametr $\triangle E$ nazywa się często szerokością pakietu w energii. Zasada nieokreśloności dla czasu i energii jest wciąż jeszcze przedmiotem dyskusji i badań. Wzór jest zawsze taki jak podaliśmy powyżej, ale różne jego interpretacje prowadzą do różnych nierównoważnych, choć prawdziwych, twierdzeń.

Rozdział 23

Swobodne pakiety falowe

W tym rozdziale najpierw przedyskutujemy zachowanie pakietów falowych stanowiących odpowiednik, w mechanice kwantowej, klasycznej cząstki swobodnej o masie m, spoczywającej w punkcie $x = 0$. Hamiltonian takiej cząstki jest hamiltonianem bez oddziaływań

$$H = \frac{\vec{p}^2}{2m}. \tag{23.1}$$

Pędu cząstki i jej położenia nie da się w mechanice kwantowej jednocześnie ściśle określić, ale będziemy budowali pakiety, gdzie te wielkości są w dobrym przybliżeniu określone. Ograniczymy się do przypadku jednowymiarowego, bo wtedy wzory są najprostsze, a uogólnienie na więcej wymiarów jest łatwe i nic istotnie nowego nie wnosi. W poprzednim rozdziale wprowadziliśmy jednowymiarowy pakiet gaussowski, któremu dla chwili $t = 0$ odpowiada funkcja falowa

$$\psi(x,0) = \frac{1}{\sqrt[4]{2\pi R^2}} e^{-\frac{x^2}{4R^2}}. \tag{23.2}$$

Rozkład prawdopodobieństwa dla położenia cząstki jest dany przez kwadrat modułu funkcji falowej. Mamy więc

$$\rho(x,0) = \frac{1}{\sqrt{2\pi R^2}} e^{-\frac{x^2}{2R^2}}. \tag{23.3}$$

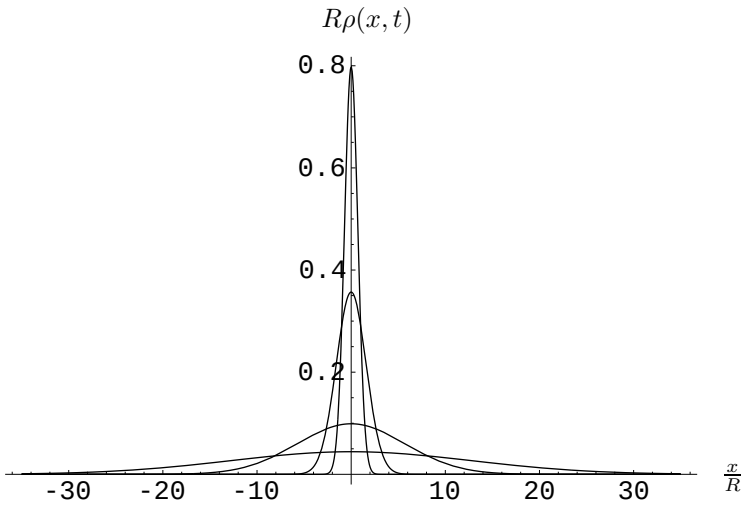
Ten rozkład jest pokazany na rysunku 1 (krzywa dla $t = 0$). Im mniejsze jest R^2, tym bardziej rozkład jest skupiony w otoczeniu punktu $x = 0$. Z symetrii widać, że wartość średnia $\langle x \rangle = 0$. Dobrą miarą szerokości rozkładu jest wariancja, która przy zerowej średniej dana jest wzorem

$$\langle x^2 \rangle = \int_{-\infty}^{+\infty} dx \rho(x,0) x^2. \tag{23.4}$$

Podstawiając gęstość ze wzoru (23.3) otrzymujemy

$$\langle x^2 \rangle (t = 0) = R^2. \tag{23.5}$$

Wąski rozkład możliwych położeń łatwo jest otrzymać dobierając odpowiednio małe R, ale chcielibyśmy jeszcze, żeby i rozkład pędu był wąski. Funkcję falową w reprezentacji pędów dla chwili $t = 0$ otrzymujemy wstawiając do wzoru na



Rysunek 23.1: Rozkład prawdopodobieństwa znalezienia cząstki opisywanej przez pakiet gaussowski omówiony w tekście. Kolejne, coraz szersze krzywe odpowiadają wartościom iloczynu $\triangle E\ t = 0, \frac{1}{2}\hbar, \hbar, \frac{3}{2}\hbar$

funkcję falową w reprezentacji pędów jedynkę złożoną z wektorów własnych operatora położenia

$$\psi^{(p)}(p) = \int_{-\infty}^{+\infty} dx \langle p|x\rangle\langle x|\psi(0)\rangle. \tag{23.6}$$

Pierwszy czynnik pod całką jest sprzężeniem zespolonym funkcji falowej w reprezentacji położeń cząstki o ustalonym pędzie p (rozdział dziewiąty). Drugi czynnik jest funkcją falową (23.2). Argument funkcji $\psi(0)$ oznacza, że wektor stanu jest liczony dla chwili $t = 0$. Mamy więc

$$\psi^{(p)}(p) = \int_{-\infty}^{+\infty} \frac{dx}{\sqrt{2\pi\hbar}\sqrt[4]{2\pi R^2}} e^{i\frac{px}{\hbar} - \frac{x^2}{4R^2}} = \sqrt[4]{\frac{2}{\pi}}\sqrt{\frac{R}{\hbar}} e^{-\frac{R^2p^2}{\hbar^2}}, \tag{23.7}$$

Łatwo sprawdzić, że kwadrat modułu tej funkcji jest znormalizowany do jedności, że wartość średnia pędu $\langle p\rangle = 0$ i otrzymać wariancję

$$\langle p^2\rangle = \frac{\hbar^2}{4R^2}. \tag{23.8}$$

Warunkiem na małą wariancję pędu jest więc duże R^2, w sprzeczności z warunkiem na małą wariancję położenia x. Nie ma w tym nic zaskakującego. Na mocy zasady nieokreśloności Heisenberga (rozdział ósmy) iloczyn wariancji położenia i pędu nie może być mniejszy niż $\frac{1}{4}\hbar^2$. Wymuszanie bardzo małej wariancji położenia, musi więc prowadzić do bardzo dużej wariancji pędu. W przypadkach makroskopowych, dzięki małemu współczynnikowi $\hbar^2$ we wzorze na wariancję pędu, można uzyskać jednocześnie małe wariancje i położenia, i pędu. Do takich przypadków odnosi się dyskusja w tym rozdziale. Jeżeli jednak potrzebujemy na przykład lokalizacji elektronu w objętości rzędu objętości atomu, czy

nukleonu w objętości rzędu objętości jądra, to wariancja pędu robi się duża i cząstka z określonymi jednocześnie pędem i położeniem jest bardzo kiepskim obrazem rzeczywistości. W rozważanym tu przykładzie najprawdopodobniejsza jest wartość pędu zero, której odpowiada zerowa energia. Jako typową wartość energii, którą oznaczymy $\triangle E$, przyjmiemy jej wartość średnią $\frac{\langle p^2\rangle}{2m}$. Mamy więc

$$\triangle E = \frac{\hbar^2}{8mR^2}. \tag{23.9}$$

Ponieważ $\langle x\rangle = 0$ i $\langle p\rangle = 0$, rozważany pakiet jest odpowiednikiem klasycznej cząstki spoczywającej w punkcie $x = 0$. Klasycznie taki stan jest trwały. Zobaczymy teraz, co się z nim dzieje według mechaniki kwantowej. Funkcjami własnymi hamiltonianu i jednocześnie rozwiązaniami zależnego od czasu równania Schrödingera są funkcje

$$\psi_p(x,t) = \frac{1}{\sqrt{2\pi\hbar}} e^{i\frac{px}{\hbar} - i\omega_p t}, \tag{23.10}$$

gdzie

$$\omega_p = \frac{p^2}{2m\hbar}. \tag{23.11}$$

Funkcję falową opisującą stan początkowy pakietu możemy zapisać w postaci

$$\langle x|\psi(0)\rangle = \int_{-\infty}^{+\infty} dp \langle x|p\rangle\langle p|\psi(0)\rangle. \tag{23.12}$$

Pierwszy czynnik pod całką jest funkcją falową w reprezentacji położeń, cząstki która ma pęd p (rozdział dziewiąty), drugi jest funkcją falową w reprezentacji pędów, stanu początkowego pakietu (23.7). Podstawiając otrzymujemy

$$\psi(x,0) = \int_{-\infty}^{+\infty} dp \frac{1}{\sqrt{2\pi\hbar}} \sqrt[4]{\frac{2}{\pi}} \sqrt{\frac{R}{\hbar}} e^{-\frac{R^2p^2}{\hbar^2}} e^{i\frac{px}{\hbar}}. \tag{23.13}$$

Zgodnie z dyskusją z poprzedniego rozdziału, za pomocą wzoru (23.10) otrzymujemy funkcję falową pakietu ważną dla dowolnego czasu, zastępując czynnik $e^{i\frac{px}{\hbar}}$ czynnikiem $e^{i\frac{px}{\hbar} - i\omega_p t}$:

$$\psi(x,t) = \int_{-\infty}^{+\infty} dp \frac{1}{\sqrt{2\pi\hbar}} \sqrt[4]{\frac{2}{\pi}} \sqrt{\frac{R}{\hbar}} e^{-\frac{R^2p^2}{\hbar^2}} e^{i\frac{px}{\hbar} - i\omega_p t}. \tag{23.14}$$

To jest znowu całka Gaussa, którą łatwo jest wykonać i otrzymuje się

$$\psi(x,t) = \frac{\sqrt{\lambda(t)}}{\sqrt[4]{2\pi R^2}} e^{-\lambda(t)\frac{x^2}{4R^2}}, \tag{23.15}$$

gdzie

$$\lambda(t) = \frac{1}{1 + 16\left(\frac{t\triangle E}{\hbar}\right)^2}. \tag{23.16}$$

Dla $t = 0$ współczynnik $\lambda(t) = 1$ i otrzymujemy pakiet początkowy. W miarę upływu czasu współczynnik $\lambda(t)$ maleje i pakiet się rozmywa (rysunek 1). Zauważmy, że zgodnie z zasadą nieokreśloności dla czasu i energii rozmycie jest

zaniedbywalne, jeśli iloczyn $t\triangle E$ jest znacznie mniejszy od $\hbar$. Rozmycie dotyczy tylko położenia cząstki. Rozkład pędu zachowuje się w czasie. Iloczyn wariancji współrzędnej x i wariancji pędu rośnie nieograniczenie, ale od początku jest niemniejszy od granicznej wartości $\frac{\hbar^2}{4}$ wynikającej z zasady nieokreśloności dla pędu i współrzędnej.

W ogólnym przypadku jednowymiarowy, swobodny pakiet falowy można zapisać w postaci

$$\psi(x,t) = \int_{-\infty}^{+\infty} dp\varphi(p)e^{i\frac{px}{\hbar} - i\omega_p t}. \tag{23.17}$$

Zakładamy, że funkcja $\varphi(p)$ jest dobrana tak, że w chwili $t = 0$, w dobrym przybliżeniu $x \approx x_0$ i $p \approx p_0$. Szczególny przypadek, jedną z wielu możliwych takich funkcji dla $x_0 = 0$, $p_0 = 0$, rozważaliśmy powyżej. Rozkład pędu dany jest przez kwadrat modułu funkcji $\varphi(p)$ i nie zmienia się z upływem czasu. Zbadamy natomiast ewolucję rozkładu przestrzennego $|\psi(x,t)|^2$.

Funkcja $\omega_p \approx \omega_{p_0}$, bo $p \approx p_0$. Dokładniej, z tożsamości $p^2 = (p_0 + (p - p_0))^2 = p_0^2 + 2p_0(p - p_0) + (p - p_0)^2$ otrzymujemy

$$\omega_p \approx \omega_{p_0} + \frac{v_0}{\hbar}(p - p_0). \tag{23.18}$$

W tym wzorze pominęliśmy mały wyraz proporcjonalny do $(p-p_0)^2$ i wprowadziliśmy prędkość pakietu $v_0 = \frac{p_0}{m}$. Podstawiając do wzoru (23.17) i biorąc wartość bezwzględną otrzymujemy

$$|\psi(x,t)| = \left| \int_{-\infty}^{+\infty} dp\ \varphi(p)e^{i\frac{p(x-v_0 t)}{\hbar}} \right|. \tag{23.19}$$

Ta całka różni się od całki dającej funkcję $|\psi(x,0)|$ tylko podstawieniem $x - v_0 t$ w miejsce x. O całce $|\psi(x,0)|$ założyliśmy, że ma wartości istotnie różne od zera tylko, jeśli $x \approx x_0$. Wynika stąd, że całka $|\psi(x,t)|$ ma wartości istotnie różne od zera tylko jeśli $x - v_0 t \approx x_0$. Czyli, jeśli

$$x \approx x_0 + v_0 t. \tag{23.20}$$

To znaczy, że w rozpatrywanym tu przybliżeniu pakiet $\psi(x,t)$ przesuwa się bez zmiany kształtu, ruchem jednostajnym z prędkością v_0 wzdłuż osi x-ów. Jest to spodziewany wynik dla klasycznej cząstki, na którą nie działa żadna siła. Należy jednak pamiętać, że przedstawiony tu wynik jest tylko przybliżeniem. Pełna analiza, uwzględniająca zaniedbany tu wyraz rzędu $(p - p_0)^2$ wskazuje, że w miarę ruchu pakiet powoli się rozpływa.

Rozdział 24

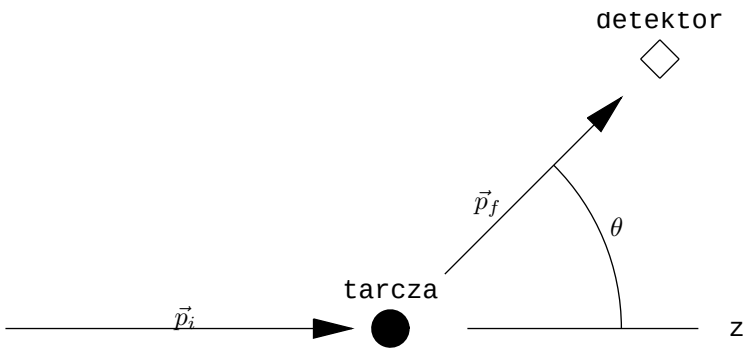
Rozpraszanie

Badając procesy rozpraszania, można uzyskać wiele cennych informacji o układach fizycznych. Ta metoda jest stosowana w różnych dziedzinach fizyki, od fizyki cząstek, przez fizykę jądrową, atomową i molekularną do fizyki fazy skondensowanej. Od strony doświadczenia podstawowym pojęciem jest przekrój czynny. Rozważamy cząstki padające na jakąś tarczę. Zakładamy, że każda z tych cząstek ma taki sam pęd skierowany wzdłuż osi z. Cząstki są niezależne i, pomijając oddziaływania z tarczą, każda z nich z jednakowym prawdopodobieństwem może trafić w każdy punkt płaszczyzny (x, y) prostopadłej do kierunku wiązki, czyli do pędu początkowego $\vec{p}_i$ padających cząstek. W tych warunkach dobrze określona jest wielkość n_i – średnia liczba cząstek padających na jednostkę powierzchni płaszczyzny prostopadłej do wiązki. Ta wielkość ma wymiar cm^{-2}. Na skutek oddziaływania z tarczą padające cząstki ulegają rozproszeniu. Procesy rozpraszania mogą być bardzo różne. Cząstka padająca może zmienić swój pęd, może zniknąć, mogą zostać wytworzone inne cząstki, może stać się coś z tarczą itd. Przypuśćmy, że mamy detektor rejestrujący interesujący nas rodzaj rozproszeń – na przykład rozproszenia elastyczne, w których końcowy pęd cząstki $\vec{p}_f$ tworzy z pędem cząstki padającej $\vec{p}_i$ kąt między θ i $\theta + d\theta$ i z półpłaszczyzną $(x > 0, z)$ kąt między φ i $\varphi + d\varphi$. Taki układ jest naszkicowany na rysunku 1. Oznaczmy N_D liczbę cząstek zarejestrowanych przez detektor. Oczywiście ta liczba jest bezwymiarowa. Wtedy przekrój czynny σ zdefiniowany jest wzorem

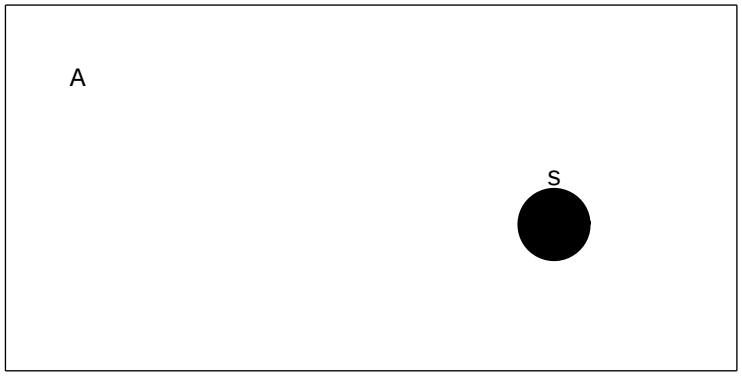
$$N_D = \sigma n_i. \tag{24.1}$$

Żeby wymiary się zgadzały, przekrój czynny musi mieć wymiar powierzchni. Na przykład można go liczyć w cm^2. W fizyce jądrowej i fizyce cząstek znacznie popularniejsza jest jednostka "barn"; 1 barn $= 10^{-24}\mathrm{cm}^2$. Zauważmy, że przekrój czynny jest zawsze przekrojem czynnym na jakiś proces, zdefiniowanym przez detektor. Ten proces może być bardzo ogólny. Na przykład rozpatruje się całkowity przekrój czynny, to znaczy przekrój czynny na jakiekolwiek oddziałanie cząstki padającej z tarczą.

Sens geometryczny przekroju czynnego wyjaśnia następujący przykład (rysunek 2). Na dużą powierzchnię o polu A, prostopadłą do pędu wiązki, pada $n_i A$ cząstek. Na powierzchni A znajduje się małe kółko o powierzchni s. Obliczymy



Rysunek 24.1: Cząstka nadlatuje z pędem $\vec{p}_i$ równoległym do osi z. Pod wpływem oddziaływania z tarczą, pęd cząstki ulega zmianie. Warunkiem trafienia do narysowanego detektora jest, żeby kierunek pędu końcowego był bliski kierunku $\vec{p}_f$. Kąt rozpraszania jest θ.



Rysunek 24.2: Prawdopodobieństwo, że cząstka padająca prostopadle na powierzchnię A trafi w czarny krążek o powierzchni s wynosi przy założeniach podanych w tekście $\frac{s}{A}$. Stąd przekrój czynny na trafienie w krążek wynosi s.

przekrój czynny na trafienie w to kółko. Liczba cząstek padających na centymetr kwadratowy powierzchni A wynosi n_i. Ponieważ każda z cząstek może z jednakowym prawdopodobieństwem trafić w każdy punkt powierzchni A, prawdopodobieństwo, że określona cząstka trafi w kółko wynosi $\frac{s}{A}$. Na mocy prawa wielkich liczb, dla bardzo dużej powierzchni A, i co za tym idzie dla bardzo dużej liczby padających cząstek, liczba cząstek, które trafią w kółko, jest równa iloczynowi liczby cząstek padających $n_i A$ i prawdopodobieństwa trafienia dla jednej cząstki $\frac{s}{A}$. To daje $N_D = n_i s$. Stąd przekrój czynny wynosi s czyli jest równy polu kółka. Istotne są tu założenia zrobione przy definiowaniu przekroju czynnego. Dobry strzelec mógłby za każdy strzałem trafiać w kółko, co nie dałoby żadnych informacji o powierzchni kółka, ale wtedy założenie, że każda para współrzędnych (x, y) jest równie prawdopodobna, nie byłoby spełnione. Inne byłyby też wyniki, gdyby trafienie w jakiś punkt (x, y) ułatwiało, lub utrudniało, następne trafienia w tej okolicy, czyli gdyby nie było spełnione założenie o niezależności padających cząstek.

Zadaniem dla teorii jest obliczenie przekroju czynnego na badany proces w oparciu o jakieś założenia fizyczne. Jeśli wyniki są zgodne z doświadczeniem, założenia są potwierdzone, jeśli nie, należy je odrzucić, lub zmodyfikować. Często założenia zawierają parametry, których wartości nie znamy. Wtedy porównanie z doświadczeniem daje odpowiedź na dwa pytania: czy wyliczony przekrój czynny jest poprawny przy jakichkolwiek rozsądnych wartościach parametrów i, jeżeli tak, to jakie są, najlepiej pasujące do danych, wartości parametrów. Klasycznym przykładem takiej analizy było oszacowanie przez Rutherforda promienia jądra atomu azotu, na podstawie danych o przekroju czynnym na elastyczne rozpraszanie cząstek alfa na azocie. Podstawowymi pojęciami teoretycznymi w takiej analizie są macierz S i amplituda rozpraszania, które teraz omówimy.

Przyjmujemy następujący model procesu rozpraszania. Cząstka padająca przylatuje z daleka i póki jest daleko od tarczy, może być w dobrym przybliżeniu opisywana jako cząstka swobodna z jakimś pędem $\vec{p}_{in}$, lub (ściślej!) jako swobodny pakiet falowy $\psi_{in}(\vec{x}, t)$. W jakimś skończonym przedziale czasu wokół chwili $t = 0$ cząstka padająca oddziałuje z tarczą i wtedy jej funkcja falowa $\psi(\vec{x}, t)$ może być bardzo różna od funkcji falowej cząstki swobodnej. Potem cząstka oddala się od tarczy i w końcu znów może być opisywana jako cząstka swobodna z jakąś funkcją falową $\psi_{out}(\vec{x}, t)$. Zauważmy, że to jest tylko pewien model procesu rozpraszania, który nie zawsze się stosuje. Na przykład, kiedy meteor trafia w Ziemię, jest to niewątpliwie proces rozpraszania, ale nie ma stanów $|\psi_{out}\rangle$. Kometa przelatująca obok Słońca może mieć tor paraboliczny. Parabola nie ma asymptot, więc ruch nie zdąża do ruchu prostoliniowego swobodnej cząstki, jakkolwiek duża jest odległość między kometą i Słońcem. Mimo istnienia takich kontrprzykładów, zakres stosowalności opisanego tu modelu jest tak duży, że ograniczymy się do niego.

Przez stany początkowe rozumiemy stany cząstek swobodnych, w które przechodzą prawdziwe stany cząstek padających, kiedy czas $t \to -\infty$. Stany początkowe $|\psi_{in}\rangle$ stanowią układ zupełny. Jako bazę w przestrzeni stanów początkowych wybiera się pakiety falowe z nieźle określonym pędem $\vec{p} \approx \vec{p}_i$. W doświadczeniu zwykle średni pęd cząstki padającej jest znany, ale nie znamy funkcji falowej pakietu. Na szczęście okazuje się, że przy rozsądnych założeniach wyniki nie zależą od takich szczegółów. Zrobimy więc (nieścisłe!) założenie, że stanami początkowymi są stany o określonym pędzie $|\vec{p}_i\rangle$. Podobnie definiujemy stany końcowe, jako stany w które przechodzą stany cząstek rozproszonych na tar-

czy, kiedy czas $t \to +\infty$. Z założenia są to stany cząstek swobodnych, które stanowią układ zupełny. Jako bazę w przestrzeni tych stanów wybieramy znów (nieścisłość!) stany cząstki swobodnej o dokładnie określonym pędzie $|\vec{p}_f\rangle$. Zwykle wiemy, jaki był pęd początkowy cząstki rozpraszanej, czyli stan $|\vec{p}_i\rangle$, i wiemy które ze stanów końcowych nas interesują, czyli które zapewniają trafienie cząstki do detektora. Cząstka w stanie $|\vec{p}_i\rangle$, na skutek oddziaływania z tarczą, przechodzi w superpozycję stanów końcowych z różnymi pędami $\vec{p}_f$. Opiszemy ten efekt jako działanie operatora $\hat{S}$ na stan $|\vec{p}_i\rangle$:

$$|\vec{p}_i\rangle \to \hat{S}|\vec{p}_i\rangle. \tag{24.2}$$

Stan po prawej stronie jest stanem końcowym stanowiącym superpozycję stanów $|\vec{p}_f\rangle$. Zgodnie z ogólnymi zasadami, amplituda prawdopodobieństwa na przejście ze stanu początkowego $|\vec{p}_i\rangle$ do danego stanu końcowego $|\vec{p}_f\rangle$ jest iloczynem skalarnym wektora $\hat{S}|\vec{p}_i\rangle$ i wektora $|\vec{p}_f\rangle$. Oba te wektory należą do tej samej przestrzeni stanów końcowych, więc iloczyn jest dobrze zdefiniowany. Operator $\hat{S}$ jest podstawowym pojęciem w teorii rozpraszania. Ze względów historycznych częściej niż o operatorze mówi się o macierzy S, której elementami są liczby

$$S_{fi} = \langle \vec{p}_f|\hat{S}|\vec{p}_i\rangle. \tag{24.3}$$

Zauważmy, że wektor $|\vec{p}_i\rangle$ jest z przestrzeni stanów początkowych, a stan $|\vec{p}_f\rangle$ z przestrzeni stanów końcowych.

Szczególnie prosty jest przypadek, kiedy rozpraszania w ogóle nie ma, na przykład z braku tarczy, wtedy wektor $|\vec{p}_i\rangle$ nie zmienia się w procesie rozpraszania i macierz S jest macierzą jednostkową. Ściślej mówiąc, jest macierzą jednostkową, jeśli używamy wektorów stanu (znormalizowalnych pakietów falowych), a deltą Diraca, jeśli używamy wektorów uogólnionych. Dalej, dla uproszczenia wzorów, będziemy zakładali, że w tym przypadku macierz S jest macierzą jednostkową. Ponieważ w braku rozpraszania przekrój czynny na dowolny proces powinien być równy zero, wprowadza się oprócz operatora i macierzy S jeszcze amplitudę rozpraszania $\hat{T}$ zdefiniowaną wzorem

$$\hat{S} = \hat{1} + i\hat{T}. \tag{24.4}$$

Widać, że przy braku rozpraszania, kiedy macierz S jest macierzą jednostkową, amplituda rozpraszania znika. Działając operatorem $i\hat{T}$ na stan początkowy otrzymujemy "falę rozproszoną", która stanowi różnicę między falą w stanie końcowym i falą padającą. W ogólnym przypadku przekrój czynny jest z definicji proporcjonalny do kwadratu modułu odpowiedniego elementu macierzowego amplitudy rozpraszania. Wyprowadzenie współczynnika proporcjonalności w tym wzorze wymaga analizy kinematyki procesu, której tu nie będziemy omawiać. Zwrócimy natomiast uwagę, na ciekawe zjawisko ogólne.

Z zachowania liczby cząstek w rozpraszaniu elastycznym wynika, że działanie operatorem $\hat{S}$ nie może zwiększyć normy wektora. Może natomiast zmienić jego fazę. Przypadek $S = -1$ jest możliwy i odpowiada znanej z optyki zmianie fazy fali o π. W tym przypadku jednak

$$|T|^2 = |S - 1|^2 = 4. \tag{24.5}$$

To znaczy, że przekrój czynny jest cztery razy większy od przekroju geometrycznego, który odpowiada przypadkowi $S = 0$, kiedy fala rozproszona dokład-

nie znosi się z falą padającą i kiedy wszystkie cząstki, które oddziałały, są rejestrowane przez detektor. Taki duży przekrój czynny nie prowadzi do sprzeczności z doświadczeniem, bo dla $S = -1$ mamy tylko, rozpraszanie "w przód", kiedy cząstki rozproszone są dla detektorów nieodróżnialne od padających. Amplituda rozproszenia w przód, którą zwykle można oszacować przez ekstrapolację z pomiarów rozpraszania elastycznego pod bardzo małymi kątami, daje ciekawe informacje poprzez tak zwane twierdzenie optyczne. To twierdzenie pochodzi z optyki, gdzie oznacza, że osłabienie wiązki padającej (cień) jest dane przez ilość światła rozproszonego (nie w przód) i zaabsorbowanego. W normalizacji, w której funkcja falowa stanu końcowego przy odległości od tarczy r dążącej do nieskończoności dana jest wzorem (rozdział dziesiąty)

$$\psi(x) \to e^{i\frac{pz}{\hbar}} + \frac{f(\theta,\varphi)}{r} e^{i\frac{pr}{\hbar}}, \tag{24.6}$$

twierdzenie optyczne przybiera postać

$$\sigma^{tot} = \frac{p}{4\pi\hbar} \mathrm{Im} f(\theta = 0). \tag{24.7}$$

W tym wzorze σ^{tot} oznacza całkowity przekrój czynny, a $\mathrm{Im} f$ część urojoną amplitudy f. Argument $\theta = 0$ wskazuje, że występuje tu amplituda rozproszenia w przód. Doświadczalnie dość łatwo jest zmierzyć kwadrat modułu amplitudy f, równy różniczkowemu przekrojowi czynnemu na rozpraszanie elastyczne, i całkowity przekrój czynny σ^{tot}. Twierdzenie optyczne umożliwia uzyskanie informacji o fazie amplitudy rozpraszania elastycznego f dla $\theta = 0$.

Rozdział 25

Obraz Heisenberga

W obrazie Heisenberga wektory stanu nie zależą od czasu:

$$\frac{d}{dt}|\psi\rangle = 0, \tag{25.1}$$

a operatory spełniają równanie ruchu Heisenberga

$$\frac{d}{dt}\hat{O} = \frac{i}{\hbar}\left[\hat{H}, \hat{O}\right] + \frac{\partial \hat{O}}{\partial t}. \tag{25.2}$$

Do tego, że wielkości mierzalne, jak pęd czy położenie, zależą od czasu, jesteśmy przyzwyczajeni w mechanice klasycznej. Okazuje się, że w obrazie Heisenberga mechaniki kwantowej uzyskuje się szereg wyników wyglądających bardzo podobnie do klasycznych, chociaż oczywiście ich interpretacja jest inna. Zilustrujemy to przykładami.

Prawo zachowania energii.

Każdy operator komutuje sam ze sobą. Stosując równanie ruchu Heisenberga do hamiltonianu otrzymujemy

$$\frac{d\hat{H}}{dt} = \frac{\partial \hat{H}}{\partial t}. \tag{25.3}$$

Jeżeli hamiltonian nie zależy explicite od czasu, to jest stałą ruchu. Ten wynik bardzo przypomina klasyczną zasadę zachowania energii, ale należy pamiętać, że w mechanice kwantowej operator Hamiltona nie jest wielkością mierzalną, więc fakt, że w obrazie Heisenberga nie zależy od czasu, nie jest przewidywaniem nadającym się do bezpośredniego porównania z doświadczeniem. Zakładając, że hamiltonian nie zależy explicite od czasu, możemy natomiast napisać

$$\langle\psi|\frac{d\hat{H}}{dt}|\psi\rangle = \frac{d}{dt}\langle\psi|\hat{H}|\psi\rangle = 0. \tag{25.4}$$

Pierwsza równość jest poprawna tylko w obrazie Heisenberga, gdzie wektory $|\psi\rangle$ i $\langle\psi|$ nie zależą od czasu, więc można przesunąć operator $\frac{d}{dt}$. Druga równość jest poprawna niezależnie od obrazu, bo wartość średnia operatora odpowiadającego wielkości mierzalnej jest mierzalna i nie zależy od obrazu. Otrzymaliśmy więc twierdzenie nadające się do porównania z doświadczeniem: jeśli hamiltonian nie zależy explicite od czasu, to wartość średnia energii układu jest zachowywana. Jeżeli energia układu jest określona dokładnie lub prawie, sytuacja jest bardzo

podobna do tego, co znamy z fizyki klasycznej. Mierząc energię w różnych chwilach czasu (wymaga to użycia zespołu, a nie jednego układu, rozdział drugi), otrzymujemy zawsze wynik ten sam, lub prawie. Jeżeli natomiast rozkład energii układu ma dużą wariancję, wyniki kolejnych pomiarów mogą być bardzo różne, choć można pokazać[1], że ich rozkład jest stały w czasie.

Wzory Ehrenfesta.

Dla hamiltonianu jednocząstkowego w postaci

$$H = \frac{\vec{p}^2}{2m} + V(\vec{x}) \tag{25.5}$$

obliczymy pochodne po czasie operatorów współrzędnych i operatorów składowych pędu. Te operatory nie zależą explicite od czasu, więc problem sprowadza się do obliczenia odpowiednich komutatorów. Zaczynamy od składowej x-owej wektora położenia. Mamy

$$\left[\hat{H}, \hat{x}\right] = -\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}, \hat{x}\right], \tag{25.6}$$

bo operator $\hat{x}$ komutuje z operatorami współrzędnych, a więc i z operatorem energii potencjalnej, oraz z operatorami składowych pędu $\hat{p}_y$ i $\hat{p}_z$, a operatorem odpowiadającym x-owej składowej pędu jest $-i\hbar\frac{\partial}{\partial x}$. Żeby znaleźć komutator, działamy nim na dowolną funkcje $f(x)$. Otrzymujemy

$$\left[\frac{\partial^2}{\partial x^2}, \hat{x}\right] f(x) = \frac{\partial^2(xf(x))}{\partial x^2} - x\frac{\partial^2 f(x)}{\partial x^2} = 2\frac{\partial}{\partial x}f(x). \tag{25.7}$$

Te równości zachodzą dla dowolnych funkcji $f(x)$. Jeżeli dwa operatory działając na każdy wektor dają ten sam wynik, to muszą być równe (rozdział siódmy), więc

$$\left[\frac{\partial^2}{\partial x^2}, \hat{x}\right] = 2\frac{\partial}{\partial x}. \tag{25.8}$$

Podstawiając ten wynik do równania ruchu, i korzystając raz jeszcze z definicji operatora odpowiadającego x-owej składowej pędu, otrzymujemy

$$\frac{d\hat{x}}{dt} = \frac{\hat{p}_x}{m}. \tag{25.9}$$

Wygląda to na elementarny wzór z mechaniki: pochodna współrzędnej po czasie, czyli składowa prędkości, jest równa odpowiedniej składowej pędu podzielonej przez masę cząstki. Ale znów stosuje się dyskusja z poprzedniego przykładu. Żeby dostać wynik nadający się do porównania z doświadczeniem, trzeba wziąć wartości średnie operatorów występujących po obu stronach równości. Otrzymujemy

$$\frac{d}{dt}\langle\psi|\hat{x}|\psi\rangle = \frac{\langle\psi|\hat{p}_x|\psi\rangle}{m}. \tag{25.10}$$

I znów, jeśli mamy do czynienia ze stanem, w którym współrzędna i pęd są dość dokładnie określone, sytuacja jest bliska znanej z klasycznej fizyki. W ogólnym przypadku jednak, wyniki pomiarów mogą być bardzo dalekie od przewidywań

[1] Dowód wynika z uwagi, że zachowywana jest nie tylko wartość średnia hamiltonianu, ale i wartość średnia każdej potęgi hamiltonianu.

klasycznych. Analogiczne wyniki otrzymuje się dla pochodnych czasowych pozostałych dwu współrzędnych.

Dla składowej x-owej wektora pędu mamy

$$\frac{d\hat{p}_x}{dt} = \frac{i}{\hbar}\left[\hat{p}_x, \hat{V}(\vec{x})\right]. \tag{25.11}$$

Wykorzystaliśmy tu fakt, że operator $\hat{p}_x$ komutuje z operatorami odpowiadającymi wszystkim składowym wektora pędu, a więc i z operatorem energii kinetycznej. Podstawiając wyrażenie na operator $\hat{p}_x$ i rozumując jak w poprzednim przypadku otrzymujemy

$$\frac{d\hat{p}_x}{dt} = -\frac{\partial \hat{V}(\vec{x})}{\partial x} = \hat{F}_x, \tag{25.12}$$

gdzie siłę $\vec{F}$ zdefiniowaliśmy, jak w klasycznej fizyce, jako minus gradient potencjału. Otrzymane równanie wygląda jak klasyczne równanie Newtona, ale znów bezpośredni sens fizyczny ma tylko równość dla wartości średnich

$$\frac{d}{dt}\langle\psi|\hat{p}_x|\psi\rangle = \langle\psi|\hat{F}_x|\psi\rangle. \tag{25.13}$$

Równania opisane w tym punkcie są znane jako równania Ehrenfesta. Zapiszemy je w postaci wektorowej

$$\frac{d}{dt}\langle\psi|\hat{\vec{x}}|\psi\rangle = \frac{\langle\psi|\hat{\vec{p}}|\psi\rangle}{m}, \tag{25.14}$$

$$\frac{d}{dt}\langle\psi|\hat{\vec{p}}|\psi\rangle = \langle\psi|\hat{\vec{F}}|\psi\rangle. \tag{25.15}$$

Symetrie i prawa zachowania.

W fizyce klasycznej istnieją ważne powiązania między symetriami układu i prawami zachowania. Na przykład, załóżmy, że energia układu nie zmienia się przy przesunięciach układu równoległych do osi x. To jest pewna symetria układu. Fakt, że energia kinetyczna nie zmienia się przy przesunięciach, jest oczywisty. Ważne jest natomiast, że nie zmienia się także energia potencjalna. To znaczy, że energia potencjalna nie zależy od współrzędnej x. Wobec tego, x-owa składowa siły, która jest minus pochodną energii potencjalnej względem x, jest równa zero, a w takich warunkach, jak wynika z równań Newtona, x-owa składowa pędu jest zachowywana. Pokażemy teraz, jak się otrzymuje analogiczne wyniki w mechanice kwantowej.

Wprowadzamy operator $\hat{T}_x(a)$ przesunięcia o wektor a, równoległy do osi x. Ten operator z definicji działa na funkcje według wzoru

$$\hat{T}_x(a)f(x) = f(x+a). \tag{25.16}$$

Dla bardzo małego przesunięcia da, pomijając człony rzędu $(da)^2$, mamy

$$\hat{T}_x(da)f(x) = (1 + da\frac{\partial}{\partial x})f(x) = (1 + da\frac{i}{\hbar}\hat{p}_x)f(x). \tag{25.17}$$

To, że hamiltonian nie zmienia się przy przesunięciach, oznacza że dla dowolnej funkcji $\psi(\vec{x})$ działanie przesuniętym i nieprzesuniętym hamiltonianem daje ten sam wynik

$$\hat{T}_x(da)\hat{H}\hat{T}_x(-da)\psi(\vec{x}) = \hat{H}\psi(\vec{x}). \tag{25.18}$$

Zauważmy, że w wyrażeniu po lewej stronie operator $\hat{T}_x(da)$ przesuwa zarówno hamiltonian, jak i funkcję $\psi(\vec{x})$, ale dla funkcji $\psi(\vec{x})$ to przesunięcie jest skompensowane przeciwnym przesunięciem pochodzącym od operatora $\hat{T}_x(-da)$. Tak więc przesunięty jest tylko hamiltonian. Ponieważ powyższa równość zachodzi dla każdej funkcji $\psi(\vec{x})$, wynika z niej równość operatorowa $\hat{T}_x(da)\hat{H}\hat{T}_x(-da) = \hat{H}$ i stąd, mnożąc stronami prawostronnie przez $\hat{T}_x(da)$, zerowanie się komutatora

$$[\hat{T}_x(da), \hat{H}] = 0 \tag{25.19}$$

Na podstawie wzoru (25.17) ten wzór oznacza, że operator $\hat{p}_x$ komutuje z hamiltonianem, a zatem, na mocy równania ruchu Heisenberga, że x-owa składowa pędu jest zachowywana – formalnie zupełnie jak w klasycznej fizyce. Jak już wiemy, w mechanice kwantowej otrzymujemy stąd sprawdzalne doświadczalnie przewidywanie dla wielkości średnich

$$\frac{d}{dt}\langle\psi|\hat{p}_x|\psi\rangle = 0. \tag{25.20}$$

Podobnie jak w klasycznej mechanice ten wynik można znacznie uogólnić. Jeżeli hamiltonian nie zmienia się przy zmianie którejkolwiek współrzędnej, niekoniecznie kartezjańskiej, to kanonicznie sprzężony z tą współrzędną pęd jest zachowywany. Na przykład, jeśli hamiltonian nie zmienia się przy obrotach wokół osi z, czyli przy zmianach kąta φ, to składowa z-owa krętu jest zachowywana. Jeśli hamiltonian jest sferycznie symetryczny, to znaczy nie zmienia się przy żadnych obrotach, wszystkie składowe krętu są zachowywane, chociaż nie więcej niż jedna na raz jest określona.

Podobieństwo wzorów klasycznych do wzorów mechaniki kwantowej w obrazie Heisenberga odegrało zasadniczą rolę przy odkryciu mechaniki kwantowej. Heisenberg postulował relacje znane mu z mechaniki klasycznej, a następnie interpretował je kwantowo.

Rozdział 26

Metoda zaburzeń zależnych od czasu

Metoda zaburzeń zależnych od czasu, to jest metoda zaburzeń zastosowana do zależnego od czasu równania Schrödingera:

$$i\hbar\frac{d|\psi(t)\rangle}{dt} = \hat{H}|\psi(t)\rangle. \tag{26.1}$$

Jak zwykle w metodzie zaburzeń zakładamy, że pełny hamiltonian $\hat{H}$ da się zapisać w postaci

$$\hat{H} = \hat{H}_0 + \hat{H}_1, \tag{26.2}$$

gdzie zaburzenie $\hat{H}_1$ jest proporcjonalne do małego parametru λ. Będziemy niekiedy pisać

$$\hat{H}_1 = \lambda\hat{V}(t). \tag{26.3}$$

Parametr λ powinien być "mały"w sensie teorii zaburzeń, to znaczy taki, że szereg

$$|\psi(t)\rangle = \sum_{k=0}^{\infty}\lambda^k|\psi^{(k)}(t)\rangle \tag{26.4}$$

jest szybko zbieżny. Szersza dyskusja tego założenia była podana w rozdziale 13.

Zakładamy, że hamiltonian niezaburzony $\hat{H}_0$ nie zależy od czasu i ma znane, niezależne od czasu, wektory własne i wartości własne

$$\hat{H}_0|n\rangle = E_n|n\rangle. \tag{26.5}$$

Zauważmy zmianę oznaczeń w porównaniu do dyskusji metody zaburzeń dla równania Schrödingera niezależnego od czasu – teraz E_n oznacza wartość własną hamiltonianu niezaburzonego.

Będziemy szukali perturbacyjnego przybliżenia do rozwiązania zależnego od czasu równania Schrödingera (26.1) przy warunku początkowym

$$|\psi(0)\rangle = |k\rangle, \tag{26.6}$$

gdzie $|k\rangle$ jest jednym z wektorów własnych hamiltonianu $\hat{H}_0$. Przy tym wyborze warunku początkowego wzory wychodzą najprostsze, a ze względu na liniowość równań rozszerzenie na bardziej ogólne warunki początkowe nie przedstawia trudności. Ponieważ równanie Schrödingera jest pierwszego rzędu względem czasu, warunek początkowy jednoznacznie określa poszukiwane rozwiązanie.

Podobnie jak w metodzie zaburzeń niezależnych od czasu, podstawiamy rozwinięcie (26.4) do równania – w tym przypadku do zależnego od czasu równania Schrödingera – i przyrównujemy do zera współczynniki przy kolejnych potęgach parametru λ. W rzędzie λ^0 otrzymujemy

$$i\hbar \frac{d|\psi^{(0)}(t)\rangle}{dt} = H_0|\psi^{(0)}(t)\rangle. \tag{26.7}$$

To równanie dyskutowaliśmy już w rozdziale dwudziestym drugim. Jak łatwo sprawdzić przez podstawienie, rozwiązaniem spełniającym warunek początkowy (26.6) jest stan stacjonarny

$$|\psi^{(0)}(t)\rangle = e^{-i\frac{E_k t}{\hbar}}|k\rangle. \tag{26.8}$$

W tym przybliżeniu układ nigdy nie opuści stanu $|k\rangle$, co zresztą w przybliżeniu, w którym zaburzenie jest całkowicie zaniedbane nie powinno dziwić. Dla członów rzędu λ otrzymujemy równanie

$$i\hbar \frac{d\psi^{(1)}(t)\rangle}{dt} = \hat{H}_0|\psi^{(1)}(t)\rangle + \hat{V}(t)|\psi^{(0)}(t)\rangle. \tag{26.9}$$

Wektor $|\psi^{(0)}(t)\rangle$ już znamy. Wektora $|\psi^{(1)}(t)\rangle$ będziemy szukać w postaci

$$|\psi^{(1)}(t)\rangle = \sum_n c_n(t) e^{-i\frac{E_n t}{\hbar}}|n\rangle, \tag{26.10}$$

gdzie sumowanie przebiega po wszystkich n, dla których określone są wartości własne i wektory własne hamiltonianu niezaburzonego. Podstawiając do równania (26.9) i mnożąc powstałe równanie stronami przez wektor dualny $\langle m|$ otrzymujemy po prostych przekształceniach

$$\frac{dc_m(t)}{dt} = \frac{1}{i\hbar} e^{i\omega_{mk}t}\langle m|\hat{V}(t)|k\rangle, \tag{26.11}$$

gdzie wprowadziliśmy oznaczenie

$$\omega_{mk} = \frac{E_m - E_k}{\hbar}. \tag{26.12}$$

Warunek początkowy (26.6) daje

$$c_n(0) = \delta_{nk}. \tag{26.13}$$

Całkując stronami równanie (26.11) otrzymujemy dla $m \neq k$

$$c_m(t) = \frac{1}{i\hbar}\int_0^t dt'\; e^{i\omega_{mk}t'}\langle m|\hat{V}(t')|k\rangle, \tag{26.14}$$

Dolna granica całkowania jest tu wybrana tak, żeby warunki początkowe były spełnione.

Na mocy tego wzoru, prawdopodobieństwo że układ, który w chwili $t = 0$ był w stanie $|k\rangle$, znajdzie się w chwili t w stanie $|m\rangle$ różnym od stanu początkowego, wynosi

$$P_{k\to m}(t) = \hbar^{-2} \left| \int_0^t dt' \; e^{i\omega_{mk}t'} \langle m|\hat{H}_1(t')|k\rangle \right|^2 . \tag{26.15}$$

Czynnik λ, który zamienia operator $\hat{V}$ na operator $\hat{H}_1$, wynika stąd, że znaleziony przez nas wektor $|\psi^{(1)}(t)\rangle$ występuje w rozwinięciu (26.4) wektora $|\psi(t)\rangle$ pomnożony przez λ. Żeby podkreślić, że te prawdopodobieństwa przejść są rzędu λ^2, zapisuje się niekiedy wzór (26.15) w postaci

$$P_{k\to m}(t) = \frac{\lambda^2}{\hbar^2} \left| \int_0^t dt' \; e^{i\omega_{mk}t'} \langle m|\hat{V}(t')|k\rangle \right|^2 . \tag{26.16}$$

Znając prawdopodobieństwa przejść do wszystkich stanów różnych od stanu $|k\rangle$ i wiedząc, że suma prawdopodobieństw wszystkich możliwych przejść jest równa jeden, otrzymujemy

$$P_{k\to k}(t) = 1 - \sum_{m\neq k} P_{k\to m}(t) \tag{26.17}$$

Bezpośrednie obliczenie tego prawdopodobieństwa metodą zaburzeń byłoby trudniejsze. Dla współczynników c_m z $m \neq k$, pierwsze nie znikające człony rozwinięcia perturbacyjnego są rzędu λ i ich kwadraty dają pełne wkłady rzędu λ^2 do prawdopodobieństw przejścia $|c_m|^2$. Dla współczynnika c_k, rozwinięcie perturbacyjne zaczyna się od jedynki i, żeby obliczyć pełny wkład rzędu λ^2 do $|c_k|^2$, trzeba by obliczyć c_k też z dokładnością do członów rzędu λ^2.

W tym rozdziale wyprowadziliśmy podstawowe wzory na prawdopodobieństwa przejść w pierwszym przybliżeniu rachunku zaburzeń zależnych od czasu. W następnych rozdziałach omówimy szczególne przypadki i zastosowania tych wzorów.

Rozdział 27

Zaburzenie stałe w czasie

Rozpatrzymy teraz przypadek zaburzenie stałego w czasie, między chwilą włączenia $t = 0$ i chwilą wyłączenia $t = T$ (rysunek 1). Na przykład, w chwili $t = 0$ umieszczamy atomy wodoru w jednorodnym polu elektrycznym. Po czasie T wyłączamy pole. Pytanie jest, ile atomów wodoru zostało zjonizowanych? Jak zawsze w metodzie zaburzeń, wynik jest tym bardziej wiarygodny, im słabsze jest zaburzenie, ale w podanym tu przykładzie ważny jest stosunek pola zewnętrznego do typowego pola wewnątrz atomu. Jak dyskutowaliśmy w rozdziale trzynastym, całkiem silne pola, licząc w woltach na centymetr, są w tej skali słabe.

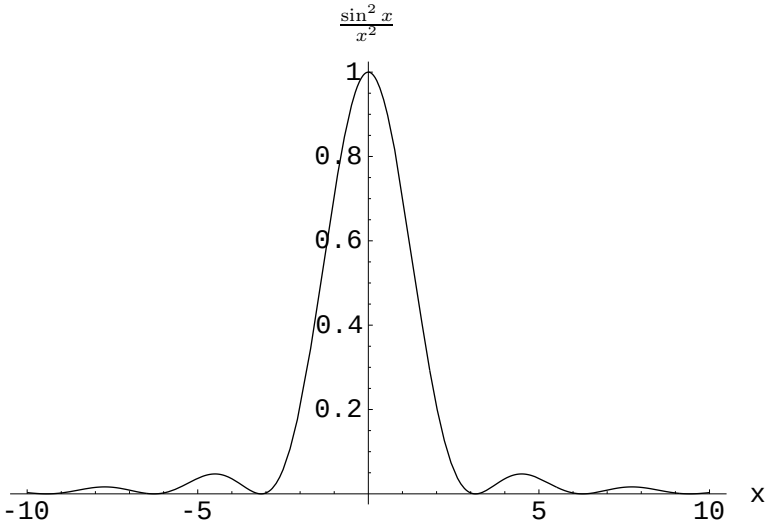
Zgodnie z wynikami z poprzedniego rozdziału, prawdopodobieństwo przejścia ze stanu początkowego $|k\rangle$ do jakiegoś wybranego stanu końcowego $|n\rangle$ dane jest dla czasów $t > T$ wzorem

$$P_{k\to n}(t) = \frac{\lambda^2 |V_{nk}|^2}{\hbar^2} \left| \int_0^T dt e^{i\omega_{nk}t} \right|^2 . \tag{27.1}$$

Wykorzystaliśmy tu założenie, że element macierzowy $V_{nk} = \langle n|\hat{V}|k\rangle$ znika dla $t > T$, a dla $0 < t < T$ nie zależy od czasu i wobec tego może być wyciągnięty przed całkę. Wykonując całkowanie otrzymujemy

$$\left| \int_0^T dt e^{i\omega_{nk}t} \right| = \left| \frac{e^{i\frac{1}{2}\omega_{nk}T}(e^{i\frac{1}{2}\omega_{nk}T} - e^{-i\frac{1}{2}\omega_{nk}T})}{i\omega_{nk}} \right| = \frac{2}{\omega_{nk}} \sin\left(\frac{1}{2}\omega_{nk}T\right). \tag{27.2}$$

Rysunek 27.1: Element macierzowy zaburzenia stałego w czasie $\hat{V}$.



Rysunek 27.2: Po podstawieniu $x = \frac{1}{2}\omega_{kn}T$ jest to wykres części zależnej od częstości ω_{nk} we wzorze na prawdopodobieństwo przejścia ze stanu $|k\rangle$ do stanu $|n\rangle$.

Druga równość wynika z tego, że dla każdego rzeczywistego x: $|e^{ix}| = 1$ i $e^{ix} - e^{-ix} = 2i\sin x$. Podstawiając do wzoru na prawdopodobieństwo otrzymujemy

$$P_{k\to n}(t) = \frac{\lambda^2|V_{nk}|^2}{\hbar^2}\,\frac{4\sin^2\left(\frac{1}{2}\omega_{nk}T\right)}{\omega_{nk}^2T^2}T^2. \tag{27.3}$$

Zauważmy, że prawdopodobieństwo przejścia nie zależy od czasu $t > T$, który można położyć nieskończoność, natomiast zależy od czasu działania zaburzenia T. Zależność prawdopodobieństwa przejścia od różnicy energii miedzy stanami początkowym i końcowym, proporcjonalnej do ω_{nk}, jest zawarta w drugim ułamku po prawej stronie. Wykres tej zależności pokazany jest na rysunku 2.

Jak widać z rysunku, najprawdopodobniejsze są przejścia zachowujące energie, to znaczy takie, dla których $\omega_{nk} = 0$. Dla ustalonej wartości parametru

$$x = \frac{1}{2}\omega_{nk}T, \tag{27.4}$$

prawdopodobieństwo przejścia zawsze zawiera czynnik T^2, ale współczynnik proporcjonalności jest bardzo mały, jeśli moduł x jest duży. Uznajmy za nieprawdopodobne te przejścia, dla których moduł x jest większy od jakiegoś x_0. Odpowiada to ograniczeniu na zmianę energii (z definicji dodatnią)

$$\triangle E < 2\hbar\frac{x_0}{T}. \tag{27.5}$$

A zatem dla długich czasów T energia jest prawie, że zachowywana. Iloczyn nieokreśloności energii $\triangle E$ i czasu T jest rzędu $\hbar$ zgodnie z zasadą nieokreśloności dla energii i czasu. Zauważmy, że nie ma tu żadnej korelacji między rozmyciem energii $\triangle E$ i siłą zaburzenia λ. Jakkolwiek delikatnie układ jest zaburzany, zawsze wywołuje to zmiany energii rzędu $\frac{\hbar}{T}$. Dla małego λ prawdopodobieństwo

oddziałania zaburzenia z układem może być znikome, ale jeżeli już to oddziaływanie zajdzie, to typowa zmiana energii nie zależy od λ. Jest to jeden z powodów, dla których według mechaniki kwantowej nie można przeprowadzić na układzie pomiaru nie zaburzając go.

W dalszym ciągu tego rozdziału będziemy dyskutowali przypadek, kiedy czas T jest bardzo długi. Podobnie jak przy dyskusji słabych pół (rozdział trzynasty) okazuje się, że czas bardzo długi w skali atomowej nie musi być długi licząc w sekundach. To jest ważne, bo gdyby ten czas był za długi, to zaburzenie byłoby duże i rachunek zaburzeń by się nie stosował. Będziemy więc formalnie rozpatrywali granicę $T \to \infty$, pamiętając jednak, że wyniki otrzymane w tej granicy mają zakres stosowalności sięgający też do niezbyt długich czasów.

W procesie jonizacji atomu wodoru przez stałe pole elektryczne, i w wielu innych ważnych przykładach, interesuje nas nie tyle prawdopodobieństwo przejścia do jakiegoś jednego ściśle określonego stanu końcowego, co prawdopodobieństwo przejścia do grupy stanów o zbliżonych energiach. Jak wynika z poprzedniej dyskusji, jeżeli prawdopodobieństwo przejścia nie ma być znikomo małe, to energie tych stanów muszą być bliskie energii stanu początkowego, tym bliższe im dłuższy jest czas T. Chcemy więc obliczyć

$$\sum_n P_{k\to n}(\infty) = \frac{\lambda^2}{\hbar^2} \int |V_{nk}|^2 \frac{\sin^2 x}{x^2} dn\, T^2. \tag{27.6}$$

Sumowanie po n zastąpiliśmy po prawej stronie całkowaniem. Zrobimy teraz trzy założenia upraszczające, które zwykle są bardzo dobrze spełnione

- Stany n, dla których prawdopodobieństwo przejścia nie jest zaniedbywalne, różnią się nieznacznie energią i zwykle są do siebie dość podobne. Zaniedbamy więc zależność elementów macierzowych od n.

- Rozkład stanów końcowych jest ciągły, lub prawie ciągły, w energii. Przyjmiemy więc, że

  $$dn = \rho(E_f) dE_f, \tag{27.7}$$

  gdzie gęstość stanów końcowych $\rho(E_f)$ jest stałą. W rzeczywistości gęstość stanów zależy od ich energii, ale przejścia możliwe przy dużych czasach T następują do tak wąskiego przedziału energii, że tę zależność możemy zaniedbać.

- Całkowanie po energii rozciągniemy od minus do plus nieskończoności. Poza wąskim zakresem energii odpowiadającym $\omega_{nk} \approx 0$ przejść praktycznie nie ma, więc to założenie nie zmienia w istotny sposób wyników.

Przy tych założeniach, zmieniając zmienne z n na E_f, a następnie z E_f na x, otrzymujemy

$$\sum_n P_{k\to n}(\infty) = \frac{\lambda^2 |V_{nk}|^2}{\hbar^2} \rho(E_f) \frac{2\hbar}{T} T^2 \int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2} dx. \tag{27.8}$$

Korzystając ze wzoru

$$\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2} dx = \pi, \tag{27.9}$$

którego nie będziemy tu wyprowadzać, otrzymujemy

$$\sum_n P_{k\to n}(\infty) = \frac{2\pi}{\hbar}|\lambda V_{nk}|^2\rho(E_f)T, \tag{27.10}$$

lub równoważnie

$$\frac{d}{dT}\sum_n P_{k\to n}(\infty) = \frac{2\pi}{\hbar}|\langle n|\hat{H}_1|k\rangle|^2\rho(E_f). \tag{27.11}$$

Ten ostatni wzór jest znany jako złota reguła Fermi'ego. Fermi go nie odkrył, ale w swoich znakomitych wykładach z mechaniki kwantowej tyle problemów za jego pomocą rozwiązywał, że nazywał go złotą regułą (regola d'oro) i ta nazwa przyjęła się.

Według złotej reguły Fermiego, prawdopodobieństwo przejścia na jednostkę czasu działania zaburzenia jest stałe. Na przykład, w danym polu elektrycznym, w ciągu dwu godzin ulegnie jonizacji dwa razy więcej atomów niż w ciągu jednej godziny. Ten wynik jest zgodny z intuicją. Zauważmy, że otrzymaliśmy go dzięki całkowaniu po stanach końcowych. Dla przejścia do jednego stanu n przy $|\omega_{nk}T| \ll 1$ prawdopodobieństwo przejścia jest proporcjonalne do T^2, a nie do T. Ten bardzo duży dodatkowy czynnik jest kompensowany znikomo małym współczynnikiem. Operowanie takimi wielkościami typu zero razy nieskończoność jest niewygodne i dla tego stosując metodę zaburzeń zależnych od czasu, jeżeli taki problem występuje, staramy się go uniknąć przez przejście do wielkości zcałkowanych. Inny przykład tego podejścia spotkamy w rozdziale trzydziestym trzecim.

Rozdział 28

Kinematyka rozpraszania elastycznego

W następnym rozdziale pokażemy, jak metodą zaburzeń z zaburzeniem stałym w czasie można wyprowadzić wzór na różniczkowy przekrój czynny w rozpraszaniu elastycznym. W tym rozdziale wprowadzimy potrzebne wiadomości z kinematyki. Z kinematyki, to znaczy te, do których nie są potrzebne informacje o tym, jakie oddziaływanie powoduje rozpraszanie.

Przez rozpraszanie elastyczne rozumiemy rozpraszanie, w którym ani pocisk, ani tarcza nie ulegają zmianie. W takich procesach energia kinetyczna jest zachowywana. W przypadku rozpraszania jednej cząstki na nieruchomym potencjale $V(\vec{x})$ znaczy to, że moduł pędu cząstki w stanie końcowym jest taki sam jak w stanie początkowym. Zmienić się może tylko kierunek pędu i ewentualnie spin, ale efektów spinowych nie będziemy tu dyskutować. Przykładami procesów nieelastycznych są wszystkie procesy z produkcją cząstek i procesy, w których liczba cząstek nie ulega zmianie, ale któraś z cząstek zostaje wzbudzona. W przypadku dwu cząstek rozpraszających się elastycznie na sobie, analiza jest bardzo podobna, jeśli prowadzić ją w układzie środka masy pary. Moduł pędu każdej z cząstek jest zachowywany i jako masa występuje masa zredukowana pary (rozdział jedenasty). Będziemy tu mówili o rozpraszaniu jednej cząstki na potencjale, ale należy pamiętać, że przeniesienie tych wyników na rozpraszanie dwucząstkowe jest bardzo proste.

Przy opisie rozpraszania elastycznego wygodnie jest wprowadzić detektory rejestrujące wszystkie cząstki rozproszone elastycznie[1], które trafiają do elementu kąta bryłowego $d\Omega$. Można równoważnie myśleć o obszarze $d\Omega$ na powierzchni odpowiedniej kuli jednostkowej w zwykłej przestrzeni lub w przestrzeni pędów. Liczba zliczeń N_D jest w tych warunkach funkcją powierzchni $d\Omega$, i to zarówno pola tej powierzchni, jak i jej położenia na powierzchni kuli. Jeżeli jednak średnica powierzchni $d\Omega$ jest dostatecznie mała, to zwykle liczba N_D dla danego położenia na powierzchni jest proporcjonalna do pola powierzchni $d\Omega$ i nie zależy od kształtu $d\Omega$. Gęstość padających cząstek n_i (rozdział dwudziesty czwarty) nie zależy od wyboru detektora, więc możemy zapisać przekrój czynny na rozpraszanie elastyczne do kąta bryłowego $d\Omega$ w postaci

[1] Sprawdzenie tego może wymagać pomiaru energii

$$\sigma(d\Omega) = \frac{d\sigma}{d\Omega} d\Omega. \tag{28.1}$$

Współczynnik przy $d\Omega$ po prawej stronie, który zależy od położenia detektora, ale nie zależy już ani od kształtu, ani od pola $d\Omega$, jest znany jako różniczkowy przekrój czynny rozpraszania elastycznego.

Jeżeli pęd cząstki padającej oznaczymy $\vec{p}_i$ i pęd cząstki rozproszonej $\vec{p}_f$, to różnica tych pędów wynosi

$$\hbar\vec{K} = \vec{p}_i - \vec{p}_f. \tag{28.2}$$

Tę różnicę nazywa się pędem przekazanym, bo to jest pęd, który utraciła cząstka padająca i, jeśli całkowity pęd jest zachowywany, zyskała tarcza. Podnosząc obie strony tego wzoru do kwadratu, pamiętając, że $\vec{p}_i \cdot \vec{p}_f = p^2 \cos\theta$, gdzie θ jest kątem między wektorami $\vec{p}_i$ i $\vec{p}_f$, czyli kątem rozpraszania, i gdzie p^2 oznacza kwadrat modułu pędu (wszystko jedno początkowego, czy końcowego) otrzymujemy

$$\hbar^2\vec{K}^2 = 2p^2(1-\cos\theta) = 4p^2\sin^2\frac{\theta}{2}. \tag{28.3}$$

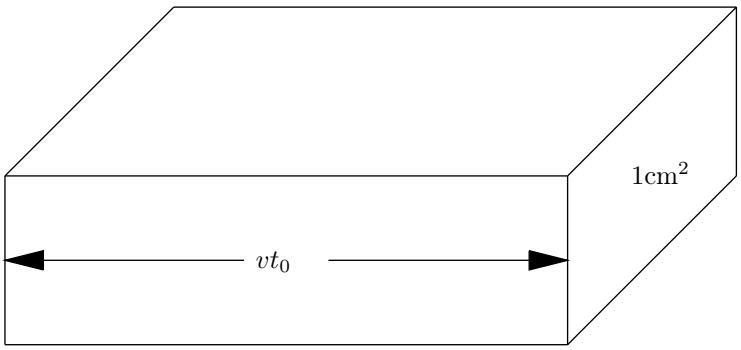
Potrzebne nam będą funkcje $\psi_{in}(\vec{x})$ opisujące cząstkę padającą jako cząstkę swobodną i funkcje $\psi_{out}(\vec{x})$ opisujące cząstkę rozproszoną jako cząstkę swobodną (rozdział dwudziesty czwarty). Przyjmiemy, że są to funkcje falowe cząstek z określonym pędem. Takie funkcje normalizowaliśmy do delty Diraca (rozdział szósty), ale tu chcielibyśmy pokazać, jak przy liczeniu różniczkowego przekroju czynnego różne czynniki, które mogą dążyć do zera czy do nieskończoności, upraszczają się. Tak powinno być zawsze przy liczeniu wielkości porównywalnych z doświadczeniem. Przyjmiemy więc, że te funkcje falowe znikają poza jakimś pudłem o bardzo dużej objętości V. Przy tym założeniu można stosować zwykłą normalizację do jedności i otrzymujemy

$$\psi_{in}(\vec{x}) = \frac{1}{\sqrt{V}} e^{i\frac{\vec{p}_i\cdot\vec{x}}{\hbar}}; \qquad \psi_{out}(\vec{x}) = \frac{1}{\sqrt{V}} e^{i\frac{\vec{p}_f\cdot\vec{x}}{\hbar}}. \tag{28.4}$$

Powyższa normalizacja, przy której $\psi_{in}(\vec{x})$ i $\psi_{out}(\vec{x})$ są zwykłymi wektorami, bo

$$\int_V |\psi(\vec{x})|^2 d^3x = 1, \tag{28.5}$$

jest dość często używana i nazywa się normalizacją w pudle (rozdział dziewiąty). Przy tej normalizacji musimy teraz obliczyć gęstość cząstek padających n_i. Rozważmy prostopadłościan (rysunek 1), którego podstawą jest 1cm^2 powierzchni prostopadłej do wiązki i którego wysokość jest równa vt_0. Jeżeli v jest prędkością padających cząstek ($\frac{|\vec{p}_i|}{m}$), to wszystkie cząstki, które znajdują się w tym prostopadłościanie w chwili $t = 0$ (i żadna inna!) przetną podstawę w czasie od $t = 0$ do $t = t_0$. Prawdopodobieństwo, że padająca cząstka znajduje się w tym prostopadłościanie, jest równe stosunkowi objętości prostopadłościanu do objętości V, bo w objętości V znajduje się na pewno dokładnie jedna cząstka i prawdopodobieństwo znalezienia tej cząstki w jakiejś objętości V' wewnątrz V jest równe V'/V, jako że kwadrat modułu funkcji falowej wynosi $\frac{1}{V}$, a więc jest stałą. Otrzymaliśmy



Rysunek 28.1: Prostopadłościan użyty w tekście przy obliczeniu gęstości cząstek n_i. Podstawa oznaczona 1cm^2 leży w płaszczyźnie prostopadłej do kierunku wiązki. Objętość prostopadłościanu jest $1\text{cm}^2 vt_0 = vt_0$.

$$n_i = \frac{vt_0}{V}. \tag{28.6}$$

Parametr t_0 oznacza czas, przez który włączony jest potencjał $V(\vec{x})$, i odpowiada parametrowi T z poprzedniego rozdziału. Oczywiście czas t_0 musi być taki, żeby prostopadłościan zmieścił się w objętości V, ale ponieważ ta objętość jest dowolnie duża, nie jest to żadne ograniczenie. Ściśle mówiąc, obliczone tu n_i jest prawdopodobieństwem – bardzo małą liczbą – a nie liczbą cząstek uderzających w wybrany 1cm^2 powierzchni, która może być tylko zero, lub jeden. Na mocy twierdzenia wielkich liczb jednak, przy wielokrotnym powtarzaniu doświadczenia, wartość średnia liczby cząstek będzie równa prawdopodobieństwu, któreśmy tu wyliczyli.

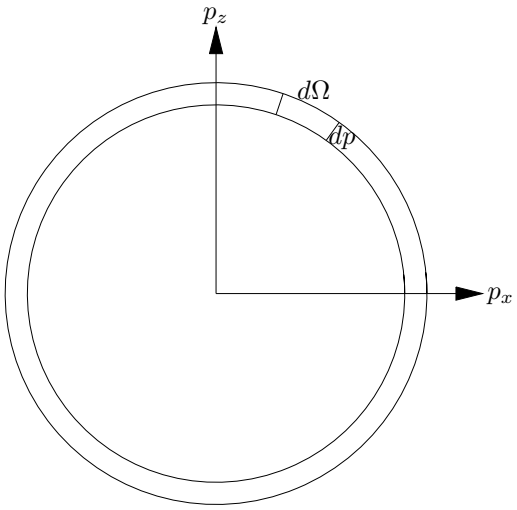
Żeby zastosować złotą regułę Fermi'ego, potrzebny jeszcze będzie wzór na gęstość stanów końcowych. Zaczniemy od prostszego przykładu, cząstki swobodnej w przestrzeni jednowymiarowej, w pudle o długości L. Wartości własne energii i funkcje falowe, przy $L = 1$ i warunkach brzegowych, że funkcja falowa przyjmuje wartość zero na końcach pudła ($\psi(0) = \psi(L) = 0$), znaleźliśmy już w rozdziale 10. Tu potrzeby nam będzie wynik przy dowolnym L i wygodniej jest przyjąć warunek, że funkcja falowa jest okresowa z okresem L, co oczywiście nie wyklucza innych, mniejszych czy większych okresów. Wtedy rozumując jak w rozdziale 10. otrzymujemy

$$E_n = \frac{4\pi^2\hbar^2}{2mL^2}n^2; \qquad n = 0, 1, 2, \ldots. \tag{28.7}$$

Każdej z tych wartości energii, za wyjątkiem E_0, odpowiadają dwa stany z pędami odpowiednio $\pm\sqrt{2mE_n}$. Otrzymujemy więc zbiór dozwolonych pędów, określających jednoznacznie stany,

$$p_n = \frac{2\pi\hbar}{L}n; \qquad n = 0, \pm1, \pm2, \ldots. \tag{28.8}$$

Te pędy są rozłożone równomiernie na osi pędów $-\infty < p < \infty$ tak, że każdemu stanowi można przypisać odcinek osi o długości $2\pi\hbar/L$. Objętość w przestrzeni fazowej przypadająca na jeden stan jest iloczynem objętości L w zwykłej przestrzeni



Rysunek 28.2: Element objętości w przestrzeni pędów. Warunek $p < |\vec{p}| < p + dp$ wycina sferę o powierzchni $4\pi p^2$ i grubości dp. Ograniczenie do pędów, które trafiają w kąt bryłowy $d\Omega$, redukuje powierzchnię $4\pi p^2$ do powierzchni $p^2 d\Omega$. Jeżeli powierzchnia $d\Omega$ jest ograniczona łukami $\theta = \text{Const}$ i $\varphi = \text{Const}$, dozwolony obszar jest w granicy bardzo małych rozmiarów prostopadłościanem o objętości $p^2 d\Omega dp = -p^2 d\cos\theta d\varphi dp$.

i objętości $2\pi\hbar/L$ w przestrzeni pędów, a więc wynosi $2\pi\hbar$. Można pokazać (twierdzenie Weyla), że przy $L \to \infty$ ten wynik nie zależy od wyboru warunków brzegowych.

Dla cząstki swobodnej w trzech wymiarach liczba stanów kwantowych też jest proporcjonalna do objętości przestrzeni fazowej, to znaczy sześciowymiarowej przestrzeni, której każdy punkt odpowiada określonym wartościom trzech składowych wektora położenia w zwykłej przestrzeni i trzem składowym wektora pędu cząstki. Podobnie jak w przypadku jednowymiarowym dowodzi się, że na każdą objętość $(2\pi\hbar)^3$ przestrzeni fazowej przypada jeden stan kwantowy. Dla cząstki, która może się znajdować gdziekolwiek w objętości V, ma moduł pędu z przedziału $[\, p, p + dp \,]$ i kierunek pędu odpowiadający elementowi kąta bryłowego $d\Omega$ (rysunek 2) mamy więc, korzystając z definicji gęstości stanów, liczbę stanów

$$\rho(E_f) dE_f = \frac{V p^2 dp d\Omega}{(2\pi\hbar)^3}. \tag{28.9}$$

Ponieważ $E_f = \frac{p^2}{2m}$, różniczka $dE_f = \frac{p}{m} dp$. Stąd otrzymujemy ostateczny wzór

$$\rho(E_f) = \frac{V m p}{(2\pi\hbar)^3} d\Omega. \tag{28.10}$$

Zestawmy dane o niebezpiecznych czynnikach, które muszą się poupraszczać, żeby powstał wzór nadający się do porównania z doświadczeniem. Objętość V, która może być dowolnie duża, występuje we wzorach na funkcje ψ_{in} i ψ_{out}, we wzorze na gęstość cząstek n_i i we wzorze na gęstość stanów $\rho(E_f)$. Czas

oddziaływania t_0, za który nie bardzo wiadomo co podstawić, występuje we wzorze na n_i. Element kąta bryłowego $d\Omega$, który powinien być jak najmniejszy, występuje we wzorze na gęstość stanów.

Rozdział 29

Pierwsze przybliżenie Borna

W tym rozdziale zastosujemy złotą regułę Fermi'ego do obliczenia różniczkowego przekroju czynnego na rozpraszanie cząstki bezspinowej o masie m na potencjale $V(\vec{x})$. Wystarczy w tym celu dokonać odpowiednich podstawień w ogólnym wzorze z rozdziału dwudziestego siódmego

$$\sum_f P_{i\to f} = \frac{2\pi}{\hbar} \left|\langle f|\hat{H}_1|i\rangle\right|^2 \rho(E_f) t_0. \tag{29.1}$$

Wprowadziliśmy tu trochę zmian w oznaczeniach, żeby się dostosować do zwyczajów panujących w teorii rozpraszania. Pominęliśmy czas oznaczany poprzednio t, który można zastąpić nieskończonością. Oznaczenie t_0 przyjęliśmy natomiast dla czasu oddziaływania oznaczanego przedtem T. Stan początkowy oznaczyliśmy $|i\rangle$. Jest to stan cząstki swobodnej o znanym pędzie $\vec{p}_i$. Grupa stanów końcowych $|f\rangle$ to są stany cząstki swobodnej z określonymi pędami $\vec{p}_f$ takimi, że gwarantują trafienie do detektora obejmującego kąt bryłowy $d\Omega$. Zaburzenie $\hat{H}_1(\vec{x})$ będzie też oznaczane $V(\vec{x})$, tym razem bez czynnika λ. Mamy więc

$$\langle f|\hat{H}_1|i\rangle = \int \psi^*_{out}(\vec{x}) V(\vec{x}) \psi_{in}(\vec{x}) d^3x \tag{29.2}$$

Przy pomocy wzorów z poprzedniego rozdziału możemy ten wynik przepisać jako

$$\langle f|\hat{H}_1|i\rangle = \frac{1}{V} \int e^{i\vec{K}\cdot\vec{x}} V(\vec{x}) d^3x \tag{29.3}$$

Przekrój czynny na rozproszenie elastyczne w kąt bryłowy $d\Omega$

$$\sigma(d\Omega) = \frac{\sum_f P_{i\to f}}{n_i} = \frac{V}{vt_0} \sum_f P_{i\to f}. \tag{29.4}$$

Podstawiając sumę prawdopodobieństw ze złotej reguły Fermi'ego i wzór na gęstość stanów z poprzedniego rozdziału, otrzymujemy:

$$\sigma(d\Omega) = \frac{V}{vt_0} \frac{2\pi}{\hbar} \frac{1}{V^2} \left| \int e^{i\vec{K}\cdot\vec{x}} V(\vec{x}) d^3x \right|^2 \frac{Vmp}{(2\pi\hbar)^3} d\Omega t_0. \tag{29.5}$$

Zauważmy, że objętość V i czas t_0 upraszczają się. Podstawiając za $\sigma(d\Omega)$ iloczyn $\frac{d\sigma}{d\Omega}d\Omega$ widzimy, że we wzorze na różniczkowy przekrój czynny upraszcza się także element kąta bryłowego $d\Omega$, i otrzymujemy po prostych redukcjach wzór bez "niebezpiecznych" parametrów

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{4\pi^2\hbar^4}\left|\int e^{i\vec{K}\cdot\vec{x}}V(\vec{x})d^3x\right|^2. \tag{29.6}$$

Ten wzór jest znany jako pierwsze przybliżenie Borna. Cały szereg perturbacyjny nazywa się w teorii rozpraszania szeregiem Borna. Wyższe przybliżenia Borna można wyliczać, licząc kolejne przybliżenia perturbacyjne. Jak widać wyliczenie różniczkowego przekroju czynnego w przybliżeniu Borna jest zwykle dość łatwe – sprowadza się do wykonania całki potrójnej. Naogół znacznie trudniej jest znaleźć odpowiedź na pytanie, jak wiarygodny jest ten wynik, czyli jak dalece się różni od wyniku dokładnego? Istnieje jednak mnóstwo wyników cząstkowych i eksperci na ich podstawie zwykle potrafią powiedzieć, gdzie w zakresie ich specjalności pierwsze przybliżenie Borna jest wiarygodne, a gdzie nie.

Jako przykład zastosowania wyliczymy, w pierwszym przybliżeniu Borna, różniczkowy przekrój czynny dla rozpraszania na potencjale Youkawy

$$V(\vec{x}) = \frac{ge^{-\mu r}}{r}. \tag{29.7}$$

Całkę

$$V_{fi} = \int e^{i\vec{K}\cdot\vec{x}}\frac{ge^{-\mu r}}{r}d^3x \tag{29.8}$$

wyliczymy we współrzędnych sferycznych wybierając oś z równolegle do wektora $\vec{K}$. Przy tym wyborze funkcja podcałkowa nie zależy od zmiennej φ, więc całkowanie po φ daje czynnik 2π. Całkowanie po $\cos\theta$ jest całkowaniem funkcji wykładniczej i otrzymujemy

$$V_{fi} = \frac{2\pi g}{iK}\int_0^\infty e^{-r(\mu-iK)} - e^{-r(\mu+iK)}dr. \tag{29.9}$$

Zauważmy, że gdybyśmy zamiast potencjału Youkawy wybrali jako przykład potencjał kulombowski, identyczny z potencjałem Youkawy dla $\mu = 0$, to otrzymalibyśmy całkę rozbieżną. Wykonując całkowania po r i porządkując wynik mamy

$$V_{fi} = \frac{4\pi g}{\mu^2 + K^2}. \tag{29.10}$$

Co po podstawieniu do wzoru Borna daje

$$\frac{d\sigma}{d\Omega} = \frac{4m^2K^2}{\hbar^4(\mu^2 + K^2)^2}. \tag{29.11}$$

Zauważmy, że ten wzór, podobnie jak i wzór na całkę V_{fi}, ma dobrze zdefiniowaną granicę przy $\mu \to 0$:

$$\frac{d\sigma}{d\Omega} = \frac{4m^2g^2}{(\hbar K)^4}. \tag{29.12}$$

Oznaczając stałą sprzężenia przez e^2 i korzystając ze wzoru na $(\hbar K)^2$ podanego w poprzednim rozdziale, dostajemy stąd poprawnie wzór Rutherforda na rozpraszanie kulombowskie:

$$\frac{d\sigma}{d\Omega} = \frac{m^2 e^4}{4p^4 \sin^4\left(\frac{1}{2}\theta\right)} \tag{29.13}$$

Ten wzór można wyprowadzić na wiele sposobów. Rutherford wyprowadził go klasycznie, co było możliwe, bo nie występuje tu stała Plancka. Podane w tekście wyprowadzenie można by uzasadnić następująco. Potencjał kulombowski rozciągający się na całą przestrzeń jest nierealistycznym potencjałem od ładunku punktowego. Prędzej, czy później pojawi się jakieś ekranowanie. Potencjał Youkawy z małym współczynnikiem μ może być interpretowany jako ekranowany potencjał kulombowski. Granica potencjału Youkawy przy $\mu \to 0$ może więc być lepszą aproksymacją realistycznego potencjału punktowego ładunku, niż po prostu potencjał kulombowski. Na szczęście nie musimy się zastanawiać, czy ten argument jest w pełni przekonywujący, bo na innej drodze niż tu podaliśmy, można ściśle wyprowadzić wzór Rutherforda od razu dla potencjału kulombowskiego.

Rozdział 30

Rozpady rezonansów

W tym rozdziale będziemy rozpatrywali przypadek zburzenia niezależnego od czasu, kiedy stan $|k\rangle$ jest rezonansem a przejście ze stanu $|k\rangle$ do grupy stanów $|n\rangle$ jest rozpadem tego rezonansu. Na mocy złotej reguły Fermi'ego prawdopodobieństwo, że rezonans nie rozpadł się do chwili $\triangle t$, można zapisać w postaci,

$$P_{k\to k}(\triangle t) = 1 - \frac{\triangle t}{\tau}, \tag{30.1}$$

gdzie

$$\frac{1}{\tau} = \frac{2\pi}{\hbar} \left|\langle n|\hat{H}_1|k\rangle\right|^2 \rho(E_f). \tag{30.2}$$

Czas $\triangle t$ nie może być za krótki, bo złota reguła Fermi'ego została wyprowadzona dla czasów długich, ale nie może też być za długi, bo w końcu prawdopodobieństwo wyliczone z tego wzoru stałoby się ujemne. Zresztą, już znacznie wcześniej straciłoby wiarygodność pierwsze przybliżenie metody zaburzeń. Zajmiemy się teraz wyprowadzeniem wzoru na prawdopodobieństwo dla czasów długich. Punkt wyjścia stanowi wzór

$$P_{k\to k}(t+dt) = P_{k\to k}(t)\left(1 - \frac{dt}{\tau}\right) \tag{30.3}$$

Ten wzór wymaga dyskusji. W języku teorii prawdopodobieństwa zdarzenie, że rezonans nie rozpadł się do chwili $t+dt$ jest iloczynem dwu zdarzeń: tego, że rezonans nie rozpadł się do chwili t i tego, że rezonans nie rozpadł się w przedziale czasowym $(t, t+dt)$. Prawdopodobieństwo rozpadu w danym odcinku czasu, jeśli wiadomo, że na początku tego okresu rezonans jeszcze się nie rozpadł, nie zależy od tego, jak długo przed tym żył rezonans, więc prawdopodobieństwo iloczynu zdarzeń jest iloczynem prawdopodobieństw czynników. Prawdopodobieństwo tego drugiego zdarzenia liczymy ze złotej reguły Fermi'ego. Jest oczywiście jakieś ograniczenie od dołu na czas dt, żeby ta reguła się stosowała, ale zakładamy, że praktycznie jest ono bez znaczenia. Taka sytuacja występuje w fizyce często. Na przykład, gęstość materiału w danym punkcie przestrzeni definiuje się jako stosunek masy materiału do jego objętości, przy objętości (ściślej średnicy objętości) dążącej do zera. Ta granica jest jednak rozumiana jako "fizyczna", to znaczy taka, że nawet w granicy objętość zawiera dużo

atomów. Podobnie jest i tu. Granica $dt \to 0$ oznacza granicę fizyczną, to znaczy taką, że nawet w granicy można stosować złotą regułę Fermi'ego. Z drugiej strony, ponieważ czas dt jest mały, zastosowanie metody zaburzeń jest wiarygodne. Przenosząc $P_{k\to k}(t)$ na drugą stronę równania, dzieląc obie strony równania przez dt i zastępując iloraz różnicowy przez pochodną, otrzymujemy równanie różniczkowe

$$\frac{dP_{k\to k}(t)}{dt} = -\frac{1}{\tau}P_{k\to k}(t); \qquad t \geq 0. \tag{30.4}$$

Zakładając, że w chwili $t = 0$, kiedy zaczynamy obserwację, rezonans się jeszcze nie rozpadł, mamy warunek początkowy

$$P_{k\to k}(0) = 1. \tag{30.5}$$

Jedynym rozwiązaniem równania z tym warunkiem początkowym jest

$$P_{k\to k}(t) = e^{-\frac{t}{\tau}}; \qquad t \geq 0. \tag{30.6}$$

To rozwiązanie jest ważne dla wszystkich czasów nieujemnych, z wyłączeniem bardzo krótkich, które nas tu nie interesują i, z przyczyn trudniejszych do prostego wyjaśnienia, bardzo długich, które zwykle też są praktycznie bez znaczenia. Jest to powszechnie znane prawo wykładniczego rozpadu. Tak się rozpadają jądra atomów promieniotwórczych, cząstki "elementarne", stany wzbudzone atomów i cząsteczek etc.

Powstaje pytanie, jakiej funkcji falowej, czy ogólniej amplitudzie, rezonansu odpowiada to prawo rozpadu. Prawdopodobieństwo jest kwadratem modułu amplitudy, a więc

$$|A_k(t)| = e^{-\frac{t}{2\tau}}; \qquad t \geq 0. \tag{30.7}$$

O fazie wzór na prawdopodobieństwo nie daje żadnej informacji, ale przyjmiemy założenie, naturalne przynajmniej dla rezonansów, które żyją dostatecznie długo, że faza jest taka, jak dla swobodnej cząstki o energii $E_k = \hbar\omega_k$. Wtedy

$$A_k(t) = e^{-\frac{t}{2\tau}}e^{-i\omega_k t}; \qquad t \geq 0. \tag{30.8}$$

Dla $t < 0$ musimy położyć $A_k(t) = 0$, żeby nam się rezonans za wcześnie nie rozpadł. Ze względu na pierwszy czynnik i na skok w chwili $t = 0$ nie jest to amplituda stanu o określonej energii. Musi być natomiast superpozycją stanów o określonej energii i dać się zapisać w postaci

$$A_k(t) = \int_{-\infty}^{+\infty} d\omega\varphi(\omega)e^{-i\omega t}. \tag{30.9}$$

Rozkład prawdopodobieństwa dla energii $\rho(E)$ i rozkład prawdopodobieństwa dla częstości $\tilde{\rho}(\omega)$ są związane wzorem $\rho(E)dE = \tilde{\rho}(\omega)d\omega$. Na mocy poprzedniego wzoru $\tilde{\rho}(\omega) = |\phi(\omega)|^2$ i ponieważ $dE = \hbar d\omega$,

$$\rho(E) = \hbar^{-1}|\varphi(\omega)|^2. \tag{30.10}$$

Żeby znaleźć funkcję $\varphi(\omega)$ można albo pomnożyć równość (30.9) obustronnie przez $e^{i\omega' t}$ i wykonać całkowanie po czasie po prawej stronie za pomocą

wzoru (6.15) z rozdziału szóstego, albo rozpoznać we wzorze (30.9) transformację Fouriera i skorzystać ze znanego wzoru na transformacje odwrotną. W każdym razie otrzymuje się

$$\varphi(\omega) = \frac{1}{2\pi}\int_0^\infty A_k(t)e^{i\omega t}dt = \frac{1}{2\pi}\frac{1}{\frac{1}{2\tau} + i(\omega_k - \omega)}. \tag{30.11}$$

Stąd rozkład energii jest

$$\rho(E) = \hbar^{-1}|\varphi(\omega)|^2 \sim \frac{1}{(\omega_k - \omega)^2 + \frac{1}{4\tau^2}}. \tag{30.12}$$

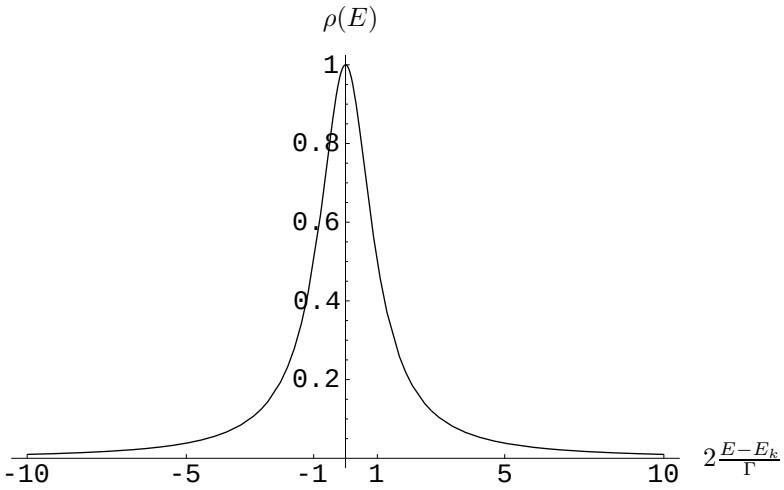
Jest to zarazem rozkład energii dla produktów rozpadu, z którego w praktyce wnioskujemy o energii uwolnionej w rozpadzie. Na przykład, jeśli wzbudzony atom rozpada się na atom w stanie podstawowym i kwant promieniowania elektromagnetycznego, to suma energii kwantu i energii odrzutu atomu w stanie podstawowym nie jest ustalona, ale ma rozkład (30.12). Mówi się w takim przypadku o profilu linii widmowej. W ostatnim wzorze pominęliśmy mało interesujący stały czynnik. Wprowadzając szerokość rezonansu Γ

$$\Gamma = \frac{\hbar}{\tau}, \tag{30.13}$$

otrzymujemy, znów pomijając czynnik normalizacyjny, wzór

$$\rho(E) \sim \frac{1}{(E_k - E)^2 + \frac{1}{4}\Gamma^2}. \tag{30.14}$$

Ten wzór jest tak ważny, że był wielokrotnie odkrywany i ma różne nazwy. Na przykład w fizyce cząstek i w fizyce jądrowej nazywa się wzorem Breita-Wignera. W optyce jest to wzór Lorentza, a w matematyce wzór Cauchy'ego. Wykres tej funkcji jest pokazany na rysunku 1. Jak widać ze wzoru i z rysunku, kiedy różnica energii $|E - E_k|$ jest duża, gęstość jest mała. Gęstość jest maksymalna dla $E = E_k$. W otoczeniu tego punktu rozciąga się obszar o szerokości rzędu Γ, w którym gęstość ma wartości porównywalne z gęstością maksymalną. Bardziej ilościowo, gęstość spada do połowy gęstości maksymalnej przy $E = E_k \pm \frac{1}{2}\Gamma$. Dla tego parametr Γ jest nazywany niekiedy szerokością połówkową. Zauważmy, że wobec prostego związku (30.13) między parametrami Γ i τ, jest wszystko jedno czy podajemy szerokość rezonansu Γ, czy czas życia rezonansu τ. Zwykle podaje się τ, jeśli czas życia jest dosyć długi, a Γ, jeśli jest bardzo krótki, ale to jest tylko zwyczaj. Zauważmy na koniec, że wariancja otrzymanego rozkładu jest rozbieżna do nieskończoności.



Rysunek 30.1: Rozkład Breita-Wignera

Rozdział 31

Oddziaływanie ładunku elektrycznego z polem elektromagnetycznym

Według fizyki klasycznej, na punktowy ładunek elektryczny e działa w polu elektromagnetycznym siła Lorentza

$$\vec{F} = e\left(\vec{E} + \frac{1}{c}\vec{v} \times \vec{B}\right). \tag{31.1}$$

Ta siła, przy danym ładunku e, zależy od natężenia pola elektrycznego $\vec{E}$, prędkości cząstki $\vec{v}$ i natężenia pola magnetycznego $\vec{B}$. Pola mogą być wyrażone przez potencjały. Jeżeli pole magnetyczne nie zależy od czasu, to można wprowadzić potencjał elektrostatyczny $\varphi(\vec{x})$ związany ze składowymi wektora pola elektrycznego wzorami

$$E_x = -\frac{\partial\varphi}{\partial x}; \quad E_y = -\frac{\partial\varphi}{\partial y}; \quad E_z = -\frac{\partial\varphi}{\partial z}. \tag{31.2}$$

Potencjał wektorowy $\vec{A}(\vec{x})$ jest związany ze składowymi pola magnetycznego wzorami

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}; \quad B_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}; \quad B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}. \tag{31.3}$$

W krótszym zapisie wektorowym

$$\vec{E} = -\vec{\nabla}\varphi(\vec{x}); \qquad \vec{B} = \vec{\nabla} \times \vec{A}. \tag{31.4}$$

Znając pole elektryczne, można wyznaczyć potencjał elektrostatyczny z dokładnością do stałej addytywnej. Znając pole magnetyczne, można wyznaczyć potencjał wektorowy tylko z dokładnością do tak zwanej transformacji cechowania

$$\vec{A}(\vec{x}) \to \vec{A}'(\vec{x}) = \vec{A}(\vec{x}) + \vec{\nabla}\Lambda(\vec{x}), \tag{31.5}$$

gdzie $\Lambda(\vec{x})$ jest dowolną dostatecznie regularną funkcją położenia. Potencjały wektorowe $\vec{A}$ i $\vec{A}'$ dają to samo pole magnetyczne, a za tym i tę samą siłę

działającą na cząstkę. Dla prostoty ograniczyliśmy się tu do szczególnego przypadku, kiedy ani potencjały, ani funkcja $\Lambda(\vec{x})$ nie zależą od czasu. Twierdzenie, że w teorii klasycznej siła zależy od potencjałów tylko poprzez pola, jest jednak ogólnie prawdziwe. Można więc potencjały całkowicie wyeliminować i rozwiązywać równania ruchu używając tylko pól: elektrycznego i magnetycznego.

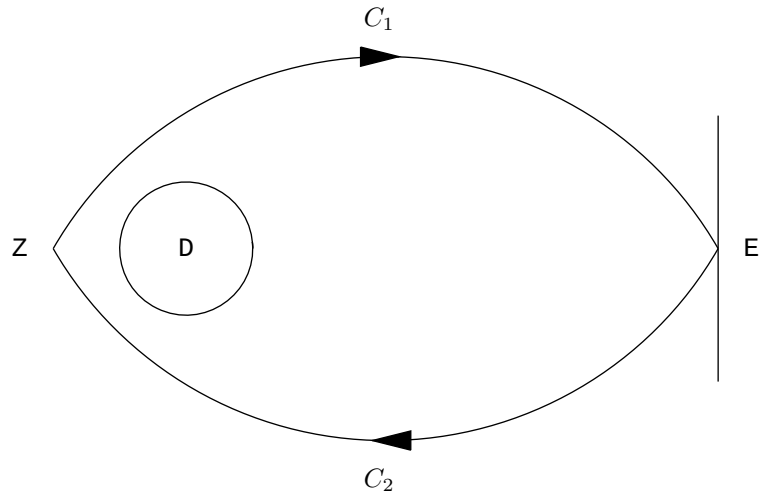
W teorii kwantowej niezależne od czasu równanie Schrödingera dla cząstki w polu elektromagnetycznym, którego wyprowadzenie pomijamy, ma postać[1]

$$-\frac{\hbar^2}{2m}\left(\vec{\nabla}-\frac{ie}{\hbar c}\vec{A}(\vec{x})\right)^2\psi(\vec{x})+e\phi(\vec{x})\psi(\vec{x})=E\psi(\vec{x}). \tag{31.6}$$

Kładąc $\vec{A}(\vec{x})=0$ i identyfikując iloczyn $e\phi(\vec{x})$ z potencjałem $V(\vec{x})$, otrzymujemy stąd zwykłe, niezależne od czasu, równanie Schrödingera bez pola magnetycznego. Jeżeli jednak potencjał wektorowy $\vec{A}(\vec{x})$ nie jest identycznie równy zero, to otrzymujemy równanie, z którego naogół nie da się wyeliminować potencjału wektorowego przez odpowiednie podstawienie pola magnetycznego. Można natomiast pokazać, że to równanie nie zmienia się przy transformacjach cechowania w tym sensie, że jeśli dwa potencjały wektorowe spełniają relację (31.5), to opisują taki sam układ. Fakt, że znajomość pól nie wystarczy do tego, żeby przewidzieć ruch cząstki, ilustruje słynny efekt Bohma-Aharonowa, który dyskutujemy poniżej.

Rozważmy następujący układ. Elektron wylatuje ze źródła Z i po drogach C_1 i C_2 dolatuje do punktu na ekranie E (rysunek 1, strzałki na razie ignorujemy). Zakładamy, że drogi cząstek są w płaszczyźnie rysunku. W tych warunkach na ekranie zachodzi interferencja i, na przykład jeśli ekran jest pokryty odpowiednią emulsją, można zobaczyć prążki interferencyjne. Jak dotąd jest to uproszczona wersja eksperymentu opisanego w rozdziale pierwszym. Przypuśćmy jednak, że między torami C_1 i C_2 znajduje się obszar D, narysowany jako kółko na rysunku, w którym jest jednorodne pole magnetyczne o natężeniu $\vec{B}$, skierowane prostopadle do płaszczyzny rysunku. Poza obszarem D pole magnetyczne jest identycznie równe zero. W szczególności jest równe zero w obszarach przebywanych przez elektrony. Klasycznie, skoro elektrony poruszają się w obszarze, gdzie pole jest zero, nie oddziałują z polem i to jakie pole panuje w obszarze D nie powinno wpływać na ich ruch. Tymczasem doświadczalnie stwierdza się, że położenia prążków na ekranie zależą od tego, jak silne jest pole B. Ten wynik, sprzeczny z klasyczną fizyką, jest znany jako efekt Bohma-Aharonowa. Wynika z oddziaływania elektronów z potencjałem wektorowym, ale nie zależy od wyboru wycechowania. Można go wyprowadzić z podanego powyżej równania Schrödingera. Łatwo się przekonać, że jakiekolwiek wybralibyśmy wycechowanie, potencjał wektorowy nie może znikać w całym obszarze, przez który lecą elektrony: Ze związku potencjału wektorowego z polem magnetycznym i z znanego z matematyki twierdzenia Stokesa wynika, relacja

[1] Wynika to stąd, że klasycznie, w obecności potencjału wektorowego $\vec{A}$, wektorowi położenia $\vec{x}$ cząstki o ładunku e odpowiada kanonicznie z nim sprzężony pęd $\vec{P}=\vec{p}+\frac{e}{c}\vec{A}$, gdzie $\vec{p}$ jest tym pędem, który wchodzi do wzoru na energię kinetyczną: $E_{kin}=\frac{\vec{p}^2}{2m}$. Pędowi $\vec{P}$ odpowiada zgodnie z ogólnymi regułami kwantowania operator $-i\hbar\vec{\nabla}$. Stąd pędowi do wzoru na energie odpowiada operator $\hat{\vec{p}}=-i\hbar\left(\vec{\nabla}-\frac{ie}{\hbar c}\vec{A}\right)$ i kwadrat tego operatora został wstawiony do równania Schrödingera.



Rysunek 31.1: Efekt Bohma-Aharonowa. Elektrony ze źródła Z drogami C_1 i C_2 docierają do ekranu E i tam dają prążki interferencyjne. Rozkład prążków zależy od pola magnetycznego w obszarze D, chociaż elektrony przez ten obszar nie przechodzą, a tam gdzie lecą elektrony pole magnetyczne jest zawsze równe zero. Krzywa zamknięta i zorientowana C, wspominana w tekście, składa się z łuków C_1 i C_2 przebieganych jak oznaczają strzałki.

$$-\oint_C \vec{A}\cdot d\vec{s} = \Phi(C). \tag{31.7}$$

Po lewej stronie tego wzoru jest całka okrężna po krzywej zamkniętej C, obieganej w kierunku zgodnym z ruchem wskazówek zegara. Po prawej stronie jest strumień pola magnetycznego przez dowolną powierzchnię ograniczoną krzywą C. Krzywa C ogranicza wiele różnych powierzchni, ale strumień pola magnetycznego przez każdą z nich jest taki sam. Weźmy jako krzywą C krzywą złożoną z toru C_1 przebieganego równolegle do prędkości elektronu i toru C_2 przebieganego w przeciwnym kierunku. Łącznie krzywa C jest krzywą zamkniętą i przebiega przez obszar, gdzie się poruszają elektrony. Wybierzmy jako powierzchnię, dla której liczymy strumień pola magnetycznego, obszar w płaszczyźnie rysunku 1 ograniczony krzywymi C_1 i C_2. W skład tego obszaru wchodzi obszar D z różnym od zera strumieniem pola magnetycznego i pozostałe obszary, których wkład do strumienia pola magnetycznego jest równy zero. Wynika z tego, że strumień całkowity po prawej stronie wzoru jest różny od zera. To z kolei dowodzi, że całka okrężna po lewej stronie musi być różna od zera, co nie byłoby możliwe, gdyby potencjał wektorowy $\vec{A}$ był w jakimkolwiek wycechowaniu równy zero wszędzie wzdłuż dróg elektronów. Niezależność tej całki od wycechowania można, zresztą, sprawdzić i bezpośrednio, bo przy dość słabych założeniach całka okrężna z gradientu dowolnej funkcji jest równa zero, a zatem zmiana wycechowania polegająca na dodaniu gradientu do potencjału wektorowego nie wpływa na całkę okrężną. Efekt Bohma-Aharonowa został przewidziany późno (dopiero w roku 1959) i zaskoczył wielu fizyków. Dziś rozumiemy, że ten efekt ilustruje jedną z głębokich różnic między fizyką klasyczna i fizyką kwantową.

Równanie Schrödingera zależne od czasu dla cząstki w polu elektromagne-

tycznym, które może być zależne od czasu, ma postać

$$i\hbar\frac{\partial\psi(\vec{x},t)}{\partial t} = -\frac{\hbar^2}{2m}\left(\vec{\nabla} - \frac{ie}{\hbar c}\vec{A}(\vec{x},t)\right)^2\psi(\vec{x},t) + e\Phi(\vec{x},t)\psi(\vec{x},t), \tag{31.8}$$

przy czym $\Phi(\vec{x},t)$ nie jest już naogół zwykłym potencjałem elektrostatycznym. To równanie upraszcza się nieco, jeśli potencjał wektorowy jest słaby w sensie teorii zaburzeń. W rozwinięciu

$$\left(\vec{\nabla} - \frac{ie}{\hbar c}\vec{A}(\vec{x},t)\right)^2 = \vec{\nabla}^2 - \frac{ie}{\hbar c}\left(\vec{\nabla}\cdot\vec{A}(\vec{x},t) + \vec{A}(\vec{x},t)\cdot\vec{\nabla}\right) + O(\vec{A}^2(\vec{x},t)) \tag{31.9}$$

możemy wtedy pominąć ostatni człon rzędu $\vec{A}^2(\vec{x},t)$. Dalsze uproszczenie można uzyskać, jeśli wybrać cechowanie, w którym

$$\vec{A}(\vec{x},t)\cdot\vec{\nabla} = \vec{\nabla}\cdot\vec{A}(\vec{x},t)). \tag{31.10}$$

Takie cechowanie zwykle da się znaleźć. W końcu wszystkie pozostałe człony zawierające potencjał wektorowy można zebrać w jeden operator zaburzenia

$$\hat{H}_1 = -\frac{e}{mc}\vec{A}(\vec{x},t))\cdot\hat{\vec{p}}. \tag{31.11}$$

Można więc próbować następującej strategii przy rozwiązywania problemów własnych z nie za silnym polem magnetycznym. Najpierw rozwiązać problem bez pola magnetycznego, a następnie wprowadzić to pole jako zaburzenie i stosować metodę zaburzeń.

Żeby móc przeprowadzić konkretne rachunki, trzeba znać potencjał wektorowy pola. Dla przejść dipolowych elektrycznych w atomach, dyskutowanych w rozdziale trzydziestym trzecim, $\Phi(\vec{x},t) = 0$, ale będzie nam potrzebny, wzór na potencjał wektorowy płaskiej fali elektromagnetycznej o częstości kołowej ω i wektorze falowym $\vec{k}$. Ogólny wzór ma postać

$$\vec{A}(\omega,\vec{x},t) = \vec{A_0}(\omega)\cos(\vec{k}\cdot\vec{x} - \omega t). \tag{31.12}$$

Jak widać wychodzimy tu poza klasę potencjałów wektorowych niezależnych od czasu, ale wzór na operator zaburzenia $\hat{H}_1$ pozostaje w mocy. Wektor $\vec{A_0}(\omega)$ jest wektorem stałym w czasie i w przestrzeni, który ze względu na wybrane wycechowanie musi spełniać warunek $\vec{k}\cdot\vec{A_0}(\omega) = 0$. Długość wektora $\vec{A_0}(\omega)$ jest związana z częściej używanym parametrem — z natężeniem fali przy częstości ω oznaczanym $I(\omega)$ — wzorem

$$|\vec{A_0}(\omega)|^2 = \frac{8\pi c}{\omega^2}I(\omega)d\omega \tag{31.13}$$

Dalsze uproszczenie można uzyskać, jeśli zauważyć, że w argumencie kosinusa pierwszy człon jest bardzo mały. Oszacujmy ten człon. Operator zaburzenia będzie średniowany po stanie atomu. Istotne są więc wektory położenia z $|\vec{x}| \sim a_0$, gdzie $a_0 \approx 0.53\times10^{-8}$cm oznacza, jak zwykle, promień Bohra. Wektor falowy $\vec{k}$ zależy od długości emitowanej fali. Dla światła widzialnego o długości fali λ rzędu 10^{-5} cm mamy $|\vec{k}| \equiv \frac{2\pi}{\lambda} \approx 6\times10^5\text{cm}^{-1}$. Iloczyn skalarny $\vec{k}\cdot\vec{x}$ jest

więc w praktyce rzędu 0.003 radiana. Taka poprawka do argumentu kosinusa jest zwykle bez znaczenia i można stosować przybliżenie

$$\vec{A}(\omega) \approx \vec{A}_0(\omega)\cos(\omega t) = \frac{1}{2}\vec{A}_0(\omega)\left(e^{i\omega t} + e^{-i\omega t}\right). \tag{31.14}$$

To przybliżenie, jak pokażemy w rozdziale trzydziestym trzecim, prowadzi do wzorów na prawdopodobieństwa przejść dipolowych elektrycznych w atomach. Często się zdarza, zwykle w związku z jakąś symetrią, że dla danych dwu stanów atomu prawdopodobieństwo przejścia dipolowego elektrycznego jest zero – linia jest wzbroniona. W takich sytuacjach małe wyrazy, które pomijamy dla przejść dozwolonych, stają się istotne i otrzymuje się przejścia wzbronione: dipolowe magnetyczne, kwadrupolowe elektryczne itd.

W stworzeniu teorii i doświadczalnym odkryciu takich przejść zasadniczą rolę odegrali polscy fizycy: teoretycy Wojciech Rubinowicz i Jan Blaton, oraz eksperymentator Henryk Niewodniczański. Trudność doświadczalna polegała na tym, że czasy życia odpowiadające przejściom wzbronionym są długie i zwykle zanim atom zdąży wypromieniować kwant elektromagnetyczny i przejść do stanu podstawowego, uderza w inny atom lub w ściankę i traci swoją energię bez wypromieniowania kwantu. Łatwo jest zredukować wpływ ścianek zwiększając ciśnienie gazu, lub zredukować wpływ zderzeń zmniejszając ciśnienie gazu. Nie było jednak wiadomo, jak wyeliminować obie te przeszkody jednocześnie. Pomysł Henryka Niewodniczańskiego polegał na tym, żeby rozcieńczyć gaz promieniujących atomów gazem doskonałym. Atomy gazów doskonałych nie mają nisko leżących stanów wzbudzonych i dla tego ich zderzenia z innymi atomami są zwykle elastyczne — zderzające się atomy zmieniają tylko pędy, zachowując bez zmiany swoje stany elektronowe. Z drugiej strony, takie zderzenia równie dobrze chronią przed dotarciem do ścianek jak inne. Mieszając atomy par ołowiu z atomami argonu, rzeczywiście udało mu się zaobserwować linie widmowe zakazane przez reguły rządzące promieniowaniem dipolowym elektrycznym – linie wzbronione. W tym przypadku było to promieniowanie dipolowe magnetyczne, co udało się udowodnić badając rozszczepienie linii w polu magnetycznym (efekt Zeemana). Potem odkryto jeszcze wiele linii wzbronionych. Jak dziś wiemy, niektóre są nawet dobrze widoczne gołym okiem w zorzach polarnych. Tam panuje niskie ciśnienie, ale nie ma problemu ze ściankami. Linie wzbronione tlenu i azotu zaobserwowano w roku 1864 w widmie Wielkiej Mgławicy w Orionie. Ponieważ w warunkach ziemskich te linie nie były obserwowane, przypisano je nowemu pierwiastkowi, który nazwano nebulium. Jak hel na Słońcu, tak nebulium miało znajdować się w mgławicach. Sprawa została definitywnie wyjaśniona dopiero metodami mechaniki kwantowej w drugiej połowie lat dwudziestych XX wieku. W tych wykładach ograniczamy się do przejść dipolowych elektrycznych.

Rozdział 32

Zaburzenie harmonicznie zależne od czasu

W tym rozdziale omówimy inny ważny przykład zaburzenia zależnego od czasu – zaburzenie harmonicznie zależne od czasu. Przyjmujemy, że element macierzowy zaburzenia ma postać

$$\langle n|\hat{H}_1|k\rangle = \lambda V_{nk} e^{-i\omega t}; \qquad \omega \neq 0. \tag{32.1}$$

Współczynnik V_{nk} w tym wzorze jest stałą i dzięki temu całkę z rozdziału dwudziestego szóstego

$$P_{k\to n}(t) = \hbar^{-2} \left| \int_0^t e^{i\omega_{nk}t'} \langle n|\hat{H}_1|k\rangle dt' \right|^2 \tag{32.2}$$

łatwo jest wykonać. Całkując jak w rozdziale dwudziestym siódmym otrzymujemy

$$P_{k\to n}(t) = \frac{\lambda^2|V_{nk}|^2}{\hbar^2} \left| \int_0^t e^{i(\omega_{nk}-\omega)t'} dt' \right|^2 = \frac{\lambda^2|V_{nk}|^2}{\hbar^2} \frac{\sin^2\left(\frac{1}{2}(\omega_{nk}-\omega)t\right)}{\left(\frac{1}{2}(\omega_{nk}-\omega)t\right)^2} t^2. \tag{32.3}$$

Widać, że dla $\omega = 0$ ten wzór przechodzi w wzór dla zaburzenia niezależnego od czasu z rozdziału dwudziestego siódmego. Stosuje się też rysunek 2 z tego rozdziału z tą tylko zmianą, że teraz

$$x = \frac{1}{2}(\omega_{nk} - \omega)t. \tag{32.4}$$

Tak więc, bieżący czas t zastępuje czas trwania zaburzenia T i różnica $\omega_{nk} - \omega$ wielkość ω_{nk}. Przenosząc tu dyskusję tamtego rysunku widzimy, że dla dużych czasów t możliwe są tylko przejścia, dla których

$$\omega_{nk} \approx \omega. \tag{32.5}$$

Interpretacja tego wyniku zależy od znaku parametru ω. Zacznijmy od przypadku $\omega > 0$. Wtedy warunkiem przejścia jest

$$E_n - E_k = \hbar\omega > 0. \tag{32.6}$$

Po wpływem tego zaburzenia możliwe jest tylko przejście ze stanu k do stanu o energii wyższej, mniej więcej o $\hbar\omega$. Dla $\omega < 0$ odpowiedni warunek ma postać

$$E_k - E_n = \hbar|\omega| > 0. \tag{32.7}$$

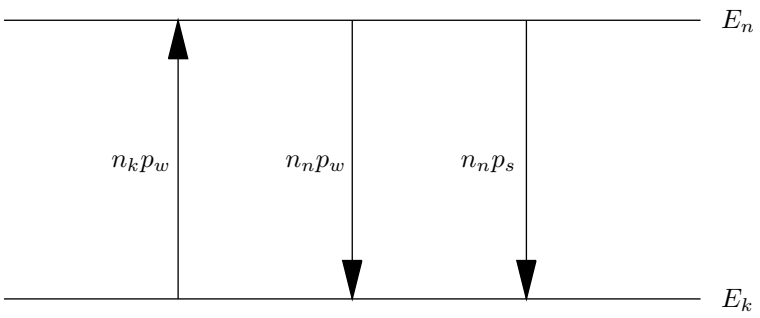
W tym przypadku możliwe jest więc tylko przejście do stanu o energii niższej o $\hbar|\omega|$.

Rozpatrzmy konkretny przykład. Gaz składa się z atomów o poziomach energetycznych $E_0 < E_1 < E_2 < \ldots$ Przez ten gaz przechodzi wiązka światła. Jak omawialiśmy w poprzednim rozdziale, odpowiada to zaburzeniu, którego operator jest iloczynem operatora niezależnego od czasu i czynnika

$$\cos(\omega t) = \frac{1}{2}\left(e^{i\omega t} + e^{-i\omega t}\right); \qquad \omega > 0. \tag{32.8}$$

Ściślej mówiąc, tak byłoby, gdyby wiązka była dokładnie monochromatyczna. Realistyczna wiązka światła ma pewien rozkład częstości $\rho(\omega)$. Jak pokażemy w dalszej części tego rozdziału, prawdopodobieństwa przejść pochodzące od różnych wartości i znaków parametru ω można obliczać osobno i dodawać do siebie. Przypuśćmy teraz, że temperatura jest niska i praktycznie wszystkie atomy są w stanie o energii E_0. Wtedy $k = 0$ i mogą zachodzić tylko przejścia z $E_n > E_k$. Wobec tego człony zaburzające proporcjonalne do $e^{i\omega t}$; $\omega > 0$, czyli $e^{-i\omega t}$; $\omega < 0$ w poprzedniej notacji, nie dają przejść. Człony zaburzające proporcjonalne do $e^{-i\omega t}$; $\omega > 0$ mogą powodować przejścia, ale tylko dla częstości spełniających w przybliżeniu warunek (32.6). Co zaobserwujemy badając wiązkę, która przeszła przez gaz? Przepuszczając wychodzącą wiązkę przez pryzmat o dostatecznie dużej zdolności rozdzielczej zobaczymy w widmie czarne prążki. Brakujące częstości będą powiązane z energiami wzbudzeń atomów wzorem (32.6). Stanowią więc wizytówkę gazu, przez który światło przeszło. Zaobserwowanie takich prążków w świetle słonecznym pozwoliło na ustalenie składu chemicznego atmosfery Słońca i było początkiem współczesnej astrofizyki. Myśląc o wiązce jako o roju fotonów, można zadać sobie pytanie, co się z niemi dzieje po zaabsorbowaniu. Zwykle po chwili zostają znów wyemitowane, ale naogół już nie w kierunku wiązki, więc obserwator patrzący wzdłuż wiązki ich nie widzi. To co tu opisaliśmy, to są widma absorbcyjne i efekt rezonansowy polegający na tym, że tylko niektóre częstości są zdolne do wzbudzania atomów. Patrząc na tę samą wiązkę w gazie z boku, widzimy tylko fotony weymitowane przez atomy, które je pochłonęły. Ze względu na opisany powyżej mechanizm, będą to tylko fotony o określonych częstościach charakterystycznych dla rozpraszającego gazu. Robiąc analizę rozkładu częstości można więc i z tego światła wyznaczyć skład promieniującego gazu. Nazywa się to badaniem widm emisyjnych.

Nasuwa się pytanie, co powoduje, że gaz emituje fotony. Na pierwszy rzut oka odpowiedź wydaje się bardzo prosta. W emisji stan początkowy ma wyższą energię niż stan końcowy, więc emisje powodują człony zaburzające typu $e^{i\omega t}$; $\omega > 0$. Jest tu jednak jedna istotna trudność, na którą zwrócił uwagę Einstein. Ponieważ w fali płaskiej (32.8) człony $e^{+i\omega t}$ i $e^{-i\omega t}$ występują z równymi współczynnikami, prawdopodobieństwa przejść ze stanu k do stanu n i ze stanu n do stanu k, liczone na jeden atom, wychodzą takie same. Liczba atomów przechodzących ze stanu k do stanu n jest równa iloczynowi tego prawdopodobieństwa



Rysunek 32.1: Równowaga między stanami $|k\rangle$ i $|n\rangle$ wymaga, żeby średnie liczby cząstek w tych stanach, n_k i n_n były zachowywane, przy czym wiadomo, że $n_k > n_n$. Prawdopodobieństwa przejść wymuszonych między rozważanymi stanami są równe i wynoszą p_w na atom i na jednostkę czasu. Prawdopodobieństwo przejścia samorzutnego ze stanu $|n\rangle$ do stanu $|k\rangle$ wynosi p_s na atom i jednostkę czasu. Przejść samorzutnych ze stanu $|k\rangle$ do stanu $|n\rangle$ nie ma. Warunkiem równowagi jest $n_k p_w = n_n(p_w + p_s)$.

i liczby atomów w stanie k. Liczba atomów przechodzących w tym samym czasie ze stanu n do stanu k jest równa iloczynowi tego samego prawdopodobieństwa i liczby atomów w stanie n. Jeżeli układ jest w równowadze, to liczba atomów przechodzących ze stanu k do stanu n powinna być równa liczbie atomów przechodzących w tym samym czasie ze stanu n do stanu k. Sprzeczność staje się widoczna, jeżeli się wie, że według mechaniki statystycznej w stanie o niższej energii znajduje się w stanie równowagi więcej atomów niż w stanie o wyższej energii. Einstein rozwiązał problem postulując, że chociaż wzbudzanie atomów (absorbcja światła) jest poprawnie opisywana przez mechanikę kwantową, to istnieją dwa mechanizmy emisji: emisja wymuszona, taka jak się liczy w mechanice kwantowej i dodatkowa emisja samorzutna, której jest akurat tyle, żeby spełniony był warunek równowagi (rysunek 1). To znaczy, że nawet przy zupełnym braku pola zewnętrznego, kiedy zaburzenie jest zero i nie powoduje żadnych przejść, wzbudzony atom ma skończony czas życia. Odwrotnością tego czasu życia, pomnożoną przez $\hbar$, jest tak zwana szerokość własna linii. Dziś wiadomo, jak prawdopodobieństwa samorzutnych przejść w atomach liczyć metodami elektrodynamiki kwantowej. Jest godne uwagi, że Einstein zdołał je wyliczyć, używając tylko mechaniki kwantowej i fizyki statystycznej.

Rachunek jest bardzo prosty. W stanie równowagi (rysunek 1)

$$n_k(T)I(\omega,T)\frac{dP^{(1)}_{k\to n}}{dt} - n_n(T)I(\omega,T)\frac{dP^{(1)}_{n\to k}}{dt} = n_n\frac{dP^{(S)}_{n\to k}}{dt}. \tag{32.9}$$

Prawdopodobieństwa przejść wymuszonych napisaliśmy tak, że cała zależność od temperatury znajduje się w czynniku $I(\omega,T)$ (rozdział 33). Natężenie promieniowania w równowadze $I(\omega,T)$ otrzymuje się mnożąc gęstość energii ze wzoru Plancka (rozdział 1) przez prędkość światła c

$$I(\omega,T) = \frac{\hbar\omega^3}{\pi^2c^2}\frac{1}{e^{\frac{\hbar\omega}{k_BT}}-1}. \tag{32.10}$$

Średnie populacje stanów dane są wzorem Boltzmanna (rozdział 8)

$$n_i = \frac{N}{Z} e^{-\frac{E_i}{k_B T}}; \qquad i = k, n, \tag{32.11}$$

gdzie N/Z jest stałą normalizacyjną. Podstawiając i rozwiązując na prawdopodobieństwo przejścia samorzutnego na jednostkę czasu, otrzymujemy

$$\frac{dP^{(S)}_{n\to k}}{dt} = \frac{\hbar\omega^3}{\pi^2 c^2} \frac{dP^{(1)}_{n\to k}}{dt} = \frac{4e^2\omega^3}{3\hbar c^3} |\langle n|\vec{x}|k\rangle|^2, \tag{32.12}$$

gdzie druga równość wynika ze wzoru, wyprowadzonego w rozdziale 33, na prawdopodobieństwa przejść wymuszonych na jednostkę czasu, w przybliżeniu dipolowym elektrycznym.

W granicy długich czasów można uprościć wyrażenie na prawdopodobieństwo przejścia przy zaburzeniu harmonicznym, używając matematycznego twierdzenia

$$\lim_{t\to\infty} \frac{\sin^2(xt)}{\pi x^2 t} = \delta(x). \tag{32.13}$$

Podstawiając ten związek do wzoru na prawdopodobieństwo przejścia na jednostkę czasu, otrzymujemy przybliżenie dobre dla bardzo dużych t

$$\frac{d}{dt} P_{k\to n}(t) = \frac{\lambda^2 |V_{nk}|^2 \pi}{\hbar^2} \delta(\frac{\omega_{nk} - \omega}{2}). \tag{32.14}$$

W tej granicy dwie różne częstości ω nie mogą interferować ze sobą, bo dają wkłady do przejść do różnych stanów końcowych. Widać też, że po to, żeby ten wzór porównywać z doświadczeniem, trzeba się pozbyć delty Diraca. W poprzednim rozdziale osiągnęliśmy to przez całkowanie po energii stanu końcowego, czyli, z dokładnością do stałego czynnika, po ω_{nk}. W następnym rozdziale wykorzystamy w tym celu całkowanie po ω. Zapiszemy jeszcze wynik dla zaburzenia, gdzie zależność od czasu jest typu $\lambda\hat{V}\cos(\omega t)$. Wobec braku interferencji otrzymujemy

$$\frac{d}{dt} P_{k\to n}(t) = \frac{\lambda^2 |V_{nk}|^2 \pi}{4\hbar^2} \left(\delta(\frac{\omega_{nk} - \omega}{2}) + \delta(\frac{\omega_{nk} + \omega}{2}) \right). \tag{32.15}$$

Wyniki dotyczące rezonansowej absorbcji i emisji mają prostą interpretację, jeżeli rozumieć światło (ogólniej falę elektromagnetyczną) jako zbiór fotonów o energiach $\hbar\omega$. W pierwszym przybliżeniu rachunku zaburzeń możliwe są procesy dwu typów: atom absorbuje kwant i zwiększa swoją energię o $\hbar\omega$ lub atom emituje kwant i zmniejsza swoją energię o $\hbar\omega$. Tak więc w obu procesach energia jest zachowywana. W wyższych przybliżeniach rachunku zaburzeń otrzymuje się procesy wielokwantowe, gdzie atom absorbuje, czy emituje, dwa lub więcej kwantów i zmienia swoją energię tak, że całkowita energia jest zachowywana. W zakresie stosowalności metody zaburzeń, to znaczy przy nie za silnych polach elektromagnetycznych, prawdopodobieństwo takiego procesu szybko spada ze wzrostem liczby łącznie absorbowanych, lub emitowanych, kwantów.

Rozdział 33

Przejścia dipolowe elektryczne w atomach

W tym rozdziale wyprowadzimy wzór na prawdopodobieństwa przejść dipolowych elektrycznych w atomach. Będą nam potrzebne elementy macierzowe hamiltonianu zburzenia wyprowadzonego w rozdziale trzydziestym pierwszym

$$\langle n|\hat{H}_1|k\rangle = -\frac{e}{mc}\vec{A}_0(\omega)\cdot\langle n|\hat{\vec{p}}|k\rangle\cos(\omega t). \tag{33.1}$$

W całym obecnym rozdziale przyjmujemy konwencję $\omega > 0$. Element macierzowy pędu przekształcimy wykorzystując wzory Ehrenfesta i równanie ruchu Heisenberga

$$\langle n|\hat{\vec{p}}|k\rangle = m\langle n|\frac{d\hat{\vec{x}}}{dt}|k\rangle = \frac{im}{\hbar}\langle n|\hat{H}\hat{\vec{x}} - \hat{\vec{x}}\hat{H}|k\rangle. \tag{33.2}$$

Długość wektora $|\vec{A}_0|$ gra w tym problemie rolę małego parametru λ. Element macierzowy (33.1) obliczamy w pierwszym przybliżeniu rachunku zaburzeń, czyli z dokładnością do członów rzędu λ. Z tego wynika, że we wzorze (33.2) możemy zaniedbać operator $\hat{H}_1$, który jest rzędu λ i po podstawieniu do wzoru (33.1) dałby wyraz rzędu λ^2. We wzorze (33.2) możemy więc zastąpić operator $\hat{H}$ przez operator $\hat{H}_0$. Jest to istotne uproszczenie, bo stany $|k\rangle$ i $|n\rangle$ są stanami własnymi operatora $\hat{H}_0$ do wartości własnych odpowiednio E_k i E_n. Wynika stąd, że wzór (33.2) można przepisać w postaci

$$\langle n|\hat{\vec{p}}|k\rangle = im\omega_{nk}\langle n|\hat{\vec{x}}|k\rangle, \tag{33.3}$$

gdzie, jak zwykle, wprowadziliśmy oznaczenie

$$\omega_{nk} = \frac{E_n - E_k}{\hbar}. \tag{33.4}$$

Mamy więc

$$|\langle n|\hat{H}_1|k\rangle|^2 = \frac{e^2\omega_{nk}^2}{c^2}|\vec{A}_0(\omega)\cdot\langle n|\hat{\vec{x}}|k\rangle|^2. \tag{33.5}$$

Ostatni czynnik po prawej stronie jest kwadratem iloczynu skalarnego dwu wektorów. Iloczyn skalarny można zapisać jako iloczyn trzech czynników: długości

pierwszego wektora, długości drugiego wektora i kosinusa kąta między tymi dwoma wektorami. Przyjmiemy, że kierunek wektora $\vec{A}_0$ jest ustalony, a orientacje atomów w gazie są rozdzielone losowo tak, że wszystkie orientacje wektora $\langle n|\hat{\vec{x}}|k\rangle$ są jednakowo prawdopodobne. Wartość średnia funkcji $\cos^2\theta$ po rozkładzie sferycznie symetrycznym jest równa jednej trzeciej. Mamy więc dla średniego kwadratu elementu macierzowego

$$\overline{|\langle n|\hat{H}_1|k\rangle|^2} = \frac{e^2\omega_{nk}^2}{3c^2}|\vec{A}_0(\omega)|^2|\langle n|\hat{\vec{x}}|k\rangle|^2 \tag{33.6}$$

lub podstawiając wzór z rozdziału trzydziestego pierwszego

$$\overline{|\langle n|\hat{H}_1|k\rangle|^2} = \frac{8\pi e^2\omega_{nk}^2}{3c\omega^2}|\langle n|\hat{\vec{x}}|k\rangle|^2 I(\omega)d\omega. \tag{33.7}$$

Podstawiając ten wynik do wzoru na pochodną po czasie prawdopodobieństwa przejścia (32.15), otrzymujemy

$$\frac{dP_{k\to n}}{dt} = \frac{\pi}{4\hbar^2}\int_0^{+\infty} d\omega \frac{8\pi e^2\omega_{nk}^2}{3c\omega^2}|\langle n|\hat{\vec{x}}|k\rangle|^2 I(\omega)\left(\delta(\frac{\omega_{nk}-\omega}{2}) + \delta(\frac{\omega_{nk}+\omega}{2})\right). \tag{33.8}$$

Ze względu na obecność delt Diraca pod całką możemy wprowadzić uproszczenia:

$$\frac{\omega_{nk}^2}{\omega^2} = 1; \qquad I(\omega) = I(|\omega_{nk}|). \tag{33.9}$$

Powstaje więc wzór

$$\frac{dP_{k\to n}}{dt} = \frac{4\pi^2 e^2}{3\hbar^2 c}|\langle n|\hat{\vec{x}}|k\rangle|^2 I(|\omega_{nk}|)\int_0^{+\infty} d\frac{\omega}{2}\left(\delta(\frac{\omega_{nk}-\omega}{2}) + \delta(\frac{\omega_{nk}+\omega}{2})\right). \tag{33.10}$$

Zmieniliśmy zmienną całkowania z ω na $\frac{1}{2}\omega$, żeby ułatwić całkowanie delt Diraca. Niezależnie od tego, czy opisujemy absorbcję promieniowania ($\omega_{nk} > 0$), czy emisję ($\omega_{nk} < 0$), jedna z delt pod całką nie daje przyczynku, a druga daje całkę równą jeden. Mamy więc końcowy wzór

$$\frac{dP_{k\to n}}{dt} = \frac{4\pi^2 e^2}{3\hbar^2 c}|\langle n|\hat{\vec{x}}|k\rangle|^2 I(|\omega_{nk}|) \tag{33.11}$$

stosujący się równie dobrze do emisji, jak i do absorbcji. Równość tych dwu prawdopodobieństw była już wykorzystywana w poprzednim rozdziale przy uzasadnianiu istnienia przejść samorzutnych.

Otrzymane tu prawdopodobieństwo jest proporcjonalne do kwadratu modułu elementu macierzowego

$$\langle n|e\hat{\vec{x}}|k\rangle \equiv \int dx\psi_n^*(\vec{x})e\vec{x}\psi_k(\vec{x}). \tag{33.12}$$

Jeżeli dodatni ładunek $|e|$ jest skoncentrowany w początku układu współrzędnych, a ujemny ładunek e ma gęstość $\rho(\vec{x})$ znormalizowaną do jedności, to taki układ ma moment dipolowy elektryczny

$$\vec{d} = \int d^3x e\vec{x}\rho(\vec{x}). \tag{33.13}$$

Ze względu na podobieństwo tego wzoru do wzoru (33.12), przejścia, które tu omawiamy, nazywa się przejściami dipolowymi elektrycznymi.

Przejście danego typu między danymi stanami atomu jest możliwe tylko, jeśli spełnione są odpowiednie warunki, znane jako reguły wyboru. Na przykład badając całkę po prawej stronie wzoru (33.12) można udowodnić, że w atomie wodoru przejście dipolowe elektryczne ze stanu $|n, l, m, s_z\rangle$ do stanu $|n', l', m', s'_z\rangle$ jest możliwe tylko, jeśli $s'_z = s_z$, $m' = m$ lub $m' = m \pm 1$ i $l' = l \pm 1$.