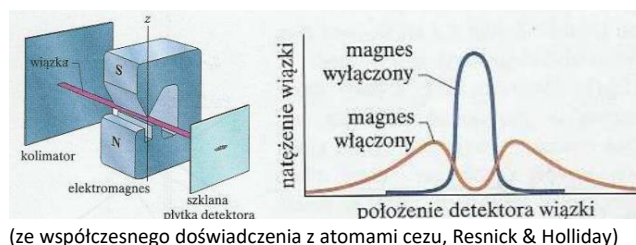


Wstęp do relatywistycznej mechaniki kwantowej

I. Rys historyczny (na gruncie fizyki nierelatywistycznej)

- (1) Doświadczenie Sterna- Gerlacha 1922, oryginalnie dla atomów złota przepuszczanych przez niejednorodne pole magnetyczne



dyskretny obraz jest generalnie potwierdzeniem kwantowania momentu pędu (inaczej to obraz byłby rozciągnięty dla ciągłych wartości momentu pędu elektronu), ale atomy Au lub Cs mają w stanie podstawowym zamknięte powłoki i jeden elektron na powłoce ns, orbitalny moment pędu jest $L=0$, zatem wynik świadczy o istnieniu wewnętrznego momentu pędu elektronu i związanego z nim momentu magnetycznego

Energia potencjalna momentu magnetycznego $\vec{\mu}$ w polu o indukcji $\vec{B}$ to $E = -\vec{\mu}\vec{B}$, a siła odchylająca dla B w kierunku z to $F = \mu_z \frac{dB}{dz}$ (siła = - grad. z potencjału);
tu: **tylko dwa rzuty m na oś z (!)**

(1) „Anomalny” efekt Zeemana – doświadczenie Uhlenbecka i Goudshmidta (1925)
zapropozowanie elektronu jako „punktowego magnesu” -
jeśli nie wiemy nic o spinie elektronu (formalnie dla atomów z całkowitym $S=0$) to rozszczepienie poziomów energetycznych atomu w polu B (a tym samym rozszczepienie linii widmowych) jest dyktowane rzutem orbitalnego momentu (liczba kwantowa m) magnetycznego na oś pola B (oś z),

$\Delta E = mB\mu_B$, a z-składowa orbitalnego momentu magnetycznego $\mu_{orb,z} = -m\mu_B$,
magneton Bohra $\mu_B = \frac{e\hbar}{2m}$ ($\sim 9 \cdot 10^{-24}$ J/T), m – liczba kwantowa orbitalnego momentu pędu

w ogólnym przypadku jest to bardziej skomplikowane, np. dla ciężkich atomów z jednym elektronem na powłoce s ($l=0$) wyjaśnienie doświadczenia wymaga wprowadzenia czynnika giromagnetycznego (Landego) $g_e=2$

$$\Delta E = m_s g_e B \mu_B,$$

wprowadzenie (w mechanice nierelatywistycznej) pojęcia spinu S , jako wewnętrznego momentu pędu elektronu i związanego z nim momentu magnetycznego zaproponował Pauli
 $\vec{\mu}_s = 2\mu_B \vec{S}$ z liczbą kwantową rzutu spinu na oś z $m_s = \pm 1/2$ i spinowym momentem magnetycznym $\mu_s = -2m_s\mu_B$;

wartość czynnika giromagnetycznego można wyprowadzić wówczas jako

$$g = 1 + \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}$$

(dla jednego elektronu na zewnętrznej powłoce, np. wodór, $l=0$, $s=1/2$,
czyli $L=0$, $S=1/2$, $J=1/2$, $g=2$)

II. Elementy teorii momentu pędu (przypomnienie)

Operator momentu pędu wprowadza się z zasady korespondencji

$$\mathbf{J} = \mathbf{r} \times \mathbf{p}$$

znając relacje komutacyjne dla składowych $\mathbf{r}$ i $\mathbf{p}$ $[r_j, p_k] = i\hbar\delta_{jk}$, otrzymuje się relacje komutacyjne dla składowych operatora $\mathbf{J}$

$$[J_m, J_n] = i\hbar\epsilon_{mnp}J_p,$$

operator momentu pędu pojawia się w definicji operatora obrotu $\mathbf{R}$ o kąt θ wokół osi zdefiniowanej wektorem jednostkowym $\mathbf{n}$

$$\mathbf{D}(\mathbf{n}, \theta) = e^{-i\theta\mathbf{n}\cdot\mathbf{J}/\hbar} = \mathbf{D}(\theta) = e^{-i\theta\cdot\mathbf{J}/\hbar}$$

gdzie θ – wektor obrotu;

jeśli układ fizyczny jest rotacyjnie niezmienniczy, to jego hamiltonian $\mathbf{H}$ komutuje zarówno z $\mathbf{D}$ jak i z $\mathbf{J}$ a także z $\mathbf{J}^2$ oraz za każdą składową $\mathbf{J}$, ale jednocześnie tylko z jedną składową np. J_z

wspólne funkcje własne (w których wszystkie operatory komutują i ich macierzowe reprezentacje są diagonalne) spełniają:

$$\begin{aligned} J^2\psi_j^m &= \lambda_j\psi_j^m \\ J_z\psi_j^m &= m\psi_j^m \end{aligned}$$

ponieważ ze względu na komutację $\mathbf{J}^2$ i J_z możemy mieć więcej funkcji własnych J_z do tej samej wartości λ_j numerowanej przez j ;

definiując operatory podnoszące i obniżające $J_{\pm} = J_x \pm iJ_y$ można prostą algebrą pokazać, że

$$\lambda_j = j(j+1)$$

dla każdego j , m zmienia się co 1 od $-j$ do j , $m = -j, -j+1, \dots, 0, \dots, j-1, j$ - $(2j+1)$ razy oraz że

$$J = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$$

Obroty tworzą grupę (ciągłą), reprezentacje grupy obrotów to macierze $\mathbf{D}$ - unitarne

$\mathbf{D}(\theta) = e^{-i\theta\mathbf{A}}$, gdzie (z dokładnością do stałej Plancka równej 1 w jedn. atom), $\mathbf{A}$ to hermitowskie macierze $\mathbf{J}$;

dla małych kątów, $\mathbf{D}$ po rozwinięciu w szereg jest $\mathbf{D}(\theta) = \mathbf{I} + i\theta\mathbf{A}$ i biorąc pod uwagę własności grupowe można wyprowadzić własności komutacyjne dla składowych macierzy $\mathbf{A}$ i są one takie same jak dla składowych $\mathbf{J}$,

zatem **macierze $\mathbf{A}$ można zbudować w bazach funkcji transformujących się przy obrotach podobnie jak funkcje własne orbitalnego momentu pędu – funkcje kuliste**, ale te tworzą bazy dla danego j i różnych m , których jest $2j+1$; mówimy o różnych reprezentacjach o wymiarach $2j+1$

przy obrocie wokół osi z , macierze $\mathbf{D}$ mają postać

$$\mathbf{D}_j(\phi) = \begin{vmatrix} e^{ij\phi} & 0 & 0 & \dots & 0 \\ 0 & e^{i(j-1)\phi} & 0 & \dots & 0 \\ 0 & 0 & e^{i(j-2)\phi} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & e^{-ij\phi} \end{vmatrix}$$

Tak jest dla j – całkowitego,

ale dla j połówkowego **nie istnieją funkcje** współrzędnych x, y, z (dwie, cztery, itp.), które transformowałyby się pomiędzy sobą przy obrotach o θ ;

i trzeba pozostać przy reprezentacjach macierzowych

III. Spin w teorii nierelatywistycznej

Dla $j=1/2$ reprezentacje mają wymiar $d=2$ – narzucającym się wyborem są macierze Pauliego

$$\mathbf{A}_x = \frac{1}{2}\sigma_x, \quad \mathbf{A}_y = \frac{1}{2}\sigma_y, \quad \mathbf{A}_z = \frac{1}{2}\sigma_z$$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

i powiązanie ich z operatorem spinu (tj. wewnętrznego momentu pędu elektronu)

$$\vec{\mathbf{S}} = \frac{1}{2} \hbar \vec{\sigma},$$

mają one wszystkie takie same własności komutacyjne jak $\mathbf{J}$;

i razem z macierzą jednostkową 2×2 $\mathbf{I}_2$, tworzą zbiór zupełny, w którym wszystkie macierze 2×2 nad ciałem liczb zespolonych mogą być przedstawione jako ich kombinacje liniowe;

uwaga:

to właśnie macierzowa postać $\mathbf{A}$ świadczy o wewnętrznym stopniu swobody, niezwiązanym z ruchem.

Spełnione będą te same relacje dla funkcji własnych

$$s^2 \chi^m = \frac{3}{4} \chi^m$$

$$s_z \chi^m = m \chi^m, \quad m = \pm 1/2$$

i χ^m są dwukomponentowymi spinorami $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$;

- w hamiltonianie nierelatywistycznym nie ma spinu, więc wszystkie funkcje jednocząstkowe (orbitale atomowe, molekularne, krystaliczne) powinny być pomnożone przez spinory spinowe χ^m ,

$$\psi = \varphi(\vec{r})\chi^m$$

a funkcja układu wielu cząstek powinna być zbudowana z takich spinorbitali ;

- jeśli chcemy mieć funkcje własne całkowitego, wypadkowego spinu $\mathbf{S}$, to składamy go tak samo jak orbitalne momenty dwóch (kilku) elektronów do całkowitego $\mathbf{L}$ i tak samo budujemy funkcje (spinory) składając z odpowiednimi współczynnikami Clebsha-Gordona - jak w ogólnym przypadku, np. składając dwa momenty pędu

$$\Psi_J^M = \sum_{m_1, m_2} C(j_1 j_2 j; m_1, m_2 M) \psi_{j_1}^{m_1} \psi_{j_2}^{m_2}$$

gdzie ψ_j^m to albo funkcje przestrzenne albo spinory; albo składając $\mathbf{l}$ i $\mathbf{s}$ do całkowitego wypadkowego momentu pędu $\mathbf{J} = \mathbf{l} + \mathbf{s}$

- dla ułamkowego $s > 1/2$ spinory będą więcej wymiarowe;

- włączając oddziaływanie orbitalnego momentu magnetycznego z zewnętrznym polem magnetycznym o indukcji $\mathbf{B}$, tzn. $-\mu_z B$, nie możemy zapomnieć o oddziaływaniu z polem $\mathbf{B}$ magnetycznego momentu spinowego, a w zasadzie momentu magnetycznego pochodzącego od wypadkowego momentu pędu $\mathbf{J}$;

Formalnie, wkład pola $\mathbf{B}$ do hamiltonianu realizuje się przez podejście półklasyczne, przez potencjał wektorowy $\mathbf{A}$, (podobnie jak dla przejść kwantowych pod wpływem fali elektromagnetycznej)

$$\mathbf{B} = \text{rot } \mathbf{A} = \nabla \times \mathbf{A}$$

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}$$

potencjał wektorowy nie jest określony jednoznacznie, lecz z dokładnością do gradientu dowolnej (niezależnej od czasu) funkcji skalarnej f , gdyż z definicji $\nabla \times \nabla f = 0$

pola $\mathbf{B}$ i $\mathbf{E}$ pozostają niezmiennicze; ale pęd swobodnej cząstki ulega zmianie

$$\mathbf{p} \rightarrow \mathbf{p} - e\mathbf{A}$$

(siła z prawa Newtona: $F = dp/dt$, siła od pola $\mathbf{E}$ działająca na ładunek $-e$: $F = -eE = e dA/dt \Rightarrow p$ (od pola) $= eA$, ale A wyznacza też pole $\mathbf{B}$)

a zatem hamiltonian dla elektronu w dowolnym potencjale V i polu opisanym przez potencjały φ i $\mathbf{A}$ ma postać

$$H = \frac{1}{2m}(\mathbf{p} + e\mathbf{A})^2 - e\varphi + V = H_0 + H', \quad H_0 = \frac{\mathbf{p}^2}{2m} + V$$

$$H' = \frac{e}{2m}(\mathbf{p} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{p}) + \frac{e^2}{2m}\mathbf{A}^2 - e\varphi$$

potraktujemy jako małe zaburzenie i:

- a) zaniedbamy wyraz kwadratowy,
- b) rozważymy przypadek tylko zewnętrznego pola magnetycznego o indukcji $\mathbf{B}$
- c) wybierzemy takie cechowanie $\mathbf{A}$, żeby $[\mathbf{p}, \mathbf{A}] = 0$, co prowadzi do $\mathbf{pA} = \mathbf{Ap}$, czyli

$$H' = \frac{e}{m}\mathbf{Ap}$$

- d) dla jednorodnego pola możemy wybrać $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$ co spełnia $\mathbf{B} = \text{rot } \mathbf{A}$ oraz $\text{div} \mathbf{A} = 0$,
zatem

$$H' = \frac{e}{m}\mathbf{pA} = \frac{e}{2m}\mathbf{p}(\mathbf{B} \times \mathbf{r}) = \frac{e}{2m}\mathbf{B}(\mathbf{r} \times \mathbf{p}) = \frac{e}{2m}\mathbf{B}\mathbf{L}$$

co odpowiada oddziaływaniu orbitalnego momentu magnetycznego $\vec{\mu}_l = \mu_B \vec{l}$ z polem o indukcji $\mathbf{B}$;

dołożenie oddziaływania spinowego momentu magnetycznego z polem prowadzi do

$$H' = \frac{e}{2m}\mathbf{B}(\mathbf{L} + g_e\mathbf{S}), \quad g_e = 2$$

dla układu wielu elektronów operatory $\mathbf{L}$ i $\mathbf{S}$ zastępujemy wypadkowymi całkowitymi $\mathbf{L}$ i $\mathbf{S}$,
a licząc wartości oczekiwane H' z funkcjami $|SLJM\rangle$ ($\mathbf{L}$ i $\mathbf{S}$ złożone do wypadkowego $\mathbf{J}$) dostaniemy

$$\langle SLJM | H' | SLJM \rangle = \frac{e}{2m} B g_J M$$

gdzie

$$g_J = 1 + (g_e - 1) \left(\frac{S(S+1) - L(L+1) + J(J+1)}{2J(J+1)} \right)$$

- można też wprowadzić do hamiltonianu człon odpowiedzialny za oddziaływanie magnetycznych momentów orbitalnego i spinowego – oddziaływanie spin-orbita – jako wyraz precesji Larmora momentu spinowego, wokół pola magnetycznego wytworzonego przez orbitujący elektron.

IV. Relatywistyczna mechanika kwantowa - podstawy

Dlaczego powinniśmy się martwić że nierelatywistyczny hamiltonian nie opisuje dobrze układów, w których cząstki poruszają się z dużymi prędkościami?

- a) Oszacowanie prędkości elektronu w jonie wodoropodobnym

z twierdzenia o wirale dla pola kulombowskiego $2E_{\text{kin}} = -E_p$, ale $E_c = E_{\text{kin}} + E_p$ to $E_c = -2E_{\text{kin}}$, energia jonu wodoropodobnego (w j.at.) $E_c = -Z^2 / 2$, dla $Z=1$ (wodór) to 13,6 eV, czyli ok. 2×10^{-18} J, dla $Z=100$ E_c jest 10000 x większa

klasycznie $E_{\text{kin}} = mv^2 / 2$, $m=10^{-27}$ kg,

dla $Z=1$ klasyczna prędkość elektronu jest rzędu 44.000 m/s = 44 km/s to jest $\sim 10^{-4}$ c – prędkości światła,

dla $Z=100$, v elektronu na powłoce 1s jest już rzędu $\sim 10^{-2}$ c - ZNACZĄCA ,
masa relatywistyczna

$$m_r = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

wzrasta już o ok 0.0001, podczas gdy dla wodoru jest to wzrost o 10^{-8} m_0 zaniedbywalny

b) Niezmienniczość względem transformacji Lorentza

przy tak dużych prędkościach v efekty relatywistyczne nie są zaniedbywalne, a nierelatywistyczne równanie Schroedingera nie jest niezmiennicze wzgl. transformacji Lorentza

$$x' = \gamma(x - vt), \quad t' = \gamma\left(t - \frac{xv}{c^2}\right), \quad \gamma = \left(1 - \left(\frac{v}{c}\right)^2\right)^{-1/2}$$

jest tak dlatego, że czas wchodzi do równania pierwszą pochodną (liniowo), a współrzędne przestrzenne drugą pochodną (T_{kin}),
np. dla swobodnej cząstki

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi$$

podczas gdy w teorii relatywistycznej

współrzędna czasowa jest traktowana równorzędnie z przestrzennymi czasowymi; punkty w czasoprzestrzeni Minkowskiego zdefiniowane są jako $x = [x^0, x^1, x^2, x^3]$, gdzie x^i to trzy współrzędne przestrzenne, a $x^0 = ct$.

c) Przy olbrzymich energiach (relatywistycznych prędkościach) następują procesy mogące zmieniać liczbę cząstek; poza tym, w rzeczywistości niemożliwe jest dokładne określenie lokalizacji cząstki w przestrzeni w czasie (zasada nieoznaczoności dla energii i czasu);
potrzebna jest inna definicja gęstości – oparta o czterowektor (x,y,z,ct), a także równanie ciągłości...

Równania niezmiennicze względem transformacji Lorentza

Są dwa podejścia: (i) albo szukamy równania (równań), w których pęd wchodzi w pierwszej pochodnej, tak jak czas, albo (ii) czas wchodzi w drugiej pochodnej tak jak pęd.

Zacznijmy od (ii)

Punktem wyjścia jest relatywistyczna (Einsteinowska) zależność między energią i pędem (zawierająca energię spoczynkową)

$$E^2 = \mathbf{p}^2 c^2 + m^2 c^4$$

Po przypisaniu operatorów pędowi i energii ($E \rightarrow H \rightarrow i\hbar \frac{\partial}{\partial t}$) dostajemy

$$\left[\frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} - \hbar^2 \nabla^2 + m^2 c^2 \right] \Psi = 0$$

jest to **równanie Kleina-Gordona**, (K-G) które w zapisie czterowektorowym (dla czteropędu zdefiniowanego jako $\mathbf{p} = (p_1, p_2, p_3, p_4 = i \frac{E}{c})$) przybiera postać

$$\left(\sum_{\mu=1}^4 p_{\mu}^2 + m^2 c^2 \right) = 0$$

wyprowadzając to równanie przypisaliśmy $E \rightarrow H \rightarrow i\hbar \frac{\partial}{\partial t}$, a zatem skorzystaliśmy z podstawienia na E nierelatywistycznego równania Schroedingera, w którym Ψ jest funkcją skalarną; nie może zatem opisywać elektronu (fermionów), gdzie funkcje powinny być spinorami

to równanie opisuje relatywistyczne cząstki bezspinowe.

równanie ciągłości $\frac{\partial \rho}{\partial t} + \text{div } \mathbf{j} = 0$ otrzymuje się mnożąc równanie K-G przez Ψ^* a równanie sprzężone przez Ψ i odejmując je od siebie, przy czym

$$\mathbf{j} = \frac{\hbar}{2mi} (\Psi^* \nabla \Psi - \Psi \nabla \Psi^*), \quad \rho = \frac{i\hbar}{2mc^2} \left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t} \right)$$

i w postaci kowariantnej

$$\sum_{\mu} \frac{\partial j_{\mu}}{\partial x_{\mu}} = 0, \quad j_4 = ic\rho$$

Równanie K-G jest równaniem różniczkowym drugiego rzędu w czasie, a to oznacza, że do znalezienia konkretnych rozwiązań musimy podać zarówno $\Psi(t=0)$ jak i $\frac{\partial \Psi}{\partial t} |_{t=0}$;

Pochodna może być dowolna (także ≤ 0), co generuje trudności w standardowej interpretacji ρ i $\mathbf{j}$ jako gęstości prawdopodobieństwa i prądu prawdopodobieństwa;

pomnożenie ρ i $\mathbf{j}$ przez ładunek e pozwala interpretować ρ i $\mathbf{j}$ jako gęstości ładunku i prądu elektrycznego: rozwiązania o ujemnym ρ mogą opisywać te same cząstki ale o przeciwnym ładunku, a $\rho = 0$ cząstki nienaładowane;

równanie K-G jest niezmiennicze względem obrotów w czteroprzestrzeni, niezmieniających kwadratu długości czterowektora;

można przy tym pokazać (AD), że przy operacji inwersji przestrzennej funkcje mogą zarówno:

$$\Psi(-\mathbf{r}, t) = \Psi(\mathbf{r}, t), \quad \text{lub} \quad \Psi(-\mathbf{r}, t) = -\Psi(\mathbf{r}, t)$$

tę ostatnią nazywamy funkcją *pseudoskalarną*, są one charakterystyczne dla cząstek o spinie zero.

Przejście do nierelatywistycznego równania Schroedingera (z funkcją falową φ) uzyskuje się poprzez transformację unitarną

$$\Psi(\mathbf{r}, t) = \varphi(\mathbf{r}, t)e^{-imc^2t/\hbar}$$

$\frac{\partial \Psi}{\partial t} = \left(\frac{\partial \varphi}{\partial t} - \frac{imc^2}{\hbar} \varphi \right) e^{-imc^2t/\hbar}$ co przy założeniu, że nierelatywistyczna energia całkowita mało różni się od spoczynkowej: tzn. $\left| i\hbar \frac{\partial \varphi}{\partial t} \right| \sim E' \varphi < mc^2 \varphi$, zatem $\frac{\partial \varphi}{\partial t}$ można zastąpić przez $\frac{mc^2}{i\hbar} \varphi$, to druga pochodna da $\frac{\partial^2 \Psi}{\partial t^2} \approx - \left(\frac{i2mc^2}{\hbar} \frac{\partial \varphi}{\partial t} + \frac{m^2 c^4}{\hbar^2} \varphi \right)$ i przyrównanie do równania K-G prowadzi do nierelatywistycznego równania Schroedingera $i\hbar \frac{\partial \varphi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \varphi$.

Równanie Diraca

W 1928 r. Dirac zaproponował, żeby

(a) równanie było liniowe we wszystkich współrzędnych przestrzennych i czasie,

(b) gęstość była dodatnio określona i

(c) równanie opisywało także spin elektronu;

ten ostatni warunek wskazuje, że funkcja powinna być spinorem wieloskładnikowym

$$i\hbar \frac{\partial \Psi}{\partial t} = H_D \Psi = [c\boldsymbol{\alpha} \mathbf{p} + \beta mc^2] \Psi$$

gdzie $\boldsymbol{\alpha} \mathbf{p} = \sum_{k=1}^3 \alpha_k p_k$, a p_k – współrzędne operatora pędu;

α_k i β muszą być macierzami kwadratowymi hermitowskimi (gdyż $\mathbf{p}$ jest hermitowski);

podnosząc H do kwadratu i porównując z hamiltonianem K-G (żeby zweryfikować poprawność relatywistycznego związku pomiędzy E i p) dostajemy

$$H^2 = c^2 \left[\sum_{i,k} (\alpha_i \alpha_k + \alpha_k \alpha_i) p_i p_k + mc \sum_i (\alpha_i \beta + \beta \alpha_i) p_i + \beta^2 m^2 c^2 \right] \Rightarrow c^2 [\sum_i p_i p_i + m^2 c^2]$$

zatem

$$\begin{aligned} \alpha_i \alpha_k + \alpha_k \alpha_i &= 2\delta_{ik} \\ \alpha_i \beta + \beta \alpha_i &= 0 \\ \alpha_i^2 &= \beta^2 = 1 \end{aligned}$$

tzn. cztery antykomutujące macierze o wartościach własnych $= \pm 1$,

antykomutacja prowadzi do $\beta \alpha_i \beta = -\alpha_i$, a zatem $\text{Tr}(\beta \alpha_i \beta) = -\text{Tr}(\alpha_i)$, ale pod śladem macierze można cyklicznie przestawiać, więc $\text{Tr}(\beta \alpha_i \beta) = \text{Tr}(\alpha_i \beta^2) = \text{Tr}(\alpha_i)$ i podobnie dla

β , zatem $\text{Tr}(\beta) = \text{Tr}(\alpha_i) = 0$, co przy wartościach własnych ± 1 wskazuje na parzysty wymiar...

wymiar $d=2$ nie wchodzi w grę (choć sugerowałaby to potrzeba opisu spinu) gdyż nie może być czterech niekomutujących macierzy o wymiarze 2 (!)

zatem $d=4$, wybór niejednoznaczny ze względu na możliwe transformacje unitarne, jedną z możliwości jest:

$$\alpha = \begin{pmatrix} 0 & \sigma \\ \sigma & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbf{I}_2 & 0 \\ 0 & -\mathbf{I}_2 \end{pmatrix}$$

i ostatecznie

$$i\hbar \frac{\partial \Psi}{\partial t} = [c\alpha \mathbf{p} + \beta mc^2] \Psi$$

gdzie Ψ jest czterekomponentowym spinorem $\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$

współrzędne przestrzenne i czas nie wchodzą symetrycznie do równania

jeśli położymy $\gamma^0 = \beta$ i $\gamma^i = \beta \alpha_i$, $i=0,1,2,3,4$, to dostaniemy równanie w zwartej postaci

$$i\hbar \gamma^\mu \frac{\partial}{\partial x^\mu} \Psi = mc \Psi$$

gdzie $(x^\mu) = (ct, x, y, z)$

(LA) definiując kontrawariantne składowe czteropędu $p_\mu = i\hbar \frac{\partial}{\partial x^\mu}$, dostajemy w pełni współzmienniczą postać równania Diraca

$$\gamma^\mu p_\mu \Psi = mc \Psi$$

Gęstość prawdopodobieństwa i równanie ciągłości

Równanie ciągłości otrzymujemy w taki sam sposób jak to się robi w nierelatywistycznej albo

dla równania K-G: równanie Diraca $i\hbar \frac{\partial \Psi}{\partial t} - c\alpha \mathbf{p} \Psi = \beta mc^2 \Psi$ mnożymy lewostronnie przez

$$\Psi^\dagger, \text{ pamiętając, że } \Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}, \text{ a } \Psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*),$$

następnie równanie sprzężone po hermitowsku ($\dagger$) mnożymy lewostronnie przez Ψ i oba równania odejmujemy od siebie, pamiętając, że $\mathbf{p} = -i\hbar \nabla$, a po sprzężeniu dostaniemy $i\hbar \nabla$ to w rezultacie dostaniemy

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \mathbf{j}$$

gdzie

$$\rho = \Psi^\dagger \Psi$$

$$\mathbf{j} = \Psi^\dagger c \boldsymbol{\alpha} \Psi$$

uwaga: równanie ciągłości wszędzie w fizyce wyraża taką samą zasadę zachowania mówiącą, że zmiana w czasie pewnej „substancji” (tu gęstości prawdopodobieństwa ρ) w danej objętości V równa się wypływowi (lub wpływowi) tej „substancji” (w postaci prądu $\mathbf{j}$) przez powierzchnię otaczającą V (tu gęstości prądu prawdopodobieństwa); $\mathbf{j}$ oznacza zwykle $\rho \mathbf{v}$, gdzie $\mathbf{v}$ to prędkość

(uwaga: w teorii nierelatywistycznej czy dla równania K-G, mamy $|\mathbf{j}| \sim \left| \frac{\mathbf{p}}{2m} \right| \sim v$)

tu natomiast dostajemy „relatywistyczną” prędkość jako $\mathbf{v} = c \boldsymbol{\alpha}$ (interpretacja ?);

rzeczywiście, jeśli wyjść od równania ruchu dla operatora położenia $\hat{\mathbf{r}}$

$$\frac{d\hat{\mathbf{r}}}{dt} = \frac{1}{i\hbar} [\hat{\mathbf{r}}, H_D]$$

to dla dowolnej składowej $\hat{r}_x$, np. x

mamy $[x, H_D] = c\alpha_x(xp_x - p_x x) = c\alpha_x[x, p_x] = c\alpha_x i\hbar$, zatem $\frac{dx}{dt} = c\alpha_x$ i $\frac{d\hat{\mathbf{r}}}{dt} = c \boldsymbol{\alpha} \dots$

problem leży w tym, że to nie są odpowiedniki nierelatywistycznych wielkości położenia i prędkości

Cząstka swobodna; duże i małe składowe; rozwiązania o ujemnej energii; pozyton

Dla stacjonarnych rozwiązań równania

$$H_D \Psi = \varepsilon \Psi$$

dla cząstki swobodnej,

zależność od czasu jest określona przez $i\hbar \frac{\partial \Psi}{\partial t} = H_D \Psi$ tzn. $\Psi(\mathbf{r}, t) = \Psi(\mathbf{r}) e^{-i\frac{\varepsilon t}{\hbar}}$

Blokowa postać macierzy $\boldsymbol{\alpha}$ i $\boldsymbol{\beta}$ wskazuje, że czteroskładnikowe funkcje Ψ możemy zapisać w postaci

$$\Psi(\mathbf{r}) = \begin{pmatrix} \varphi \\ \chi \end{pmatrix} e^{i\mathbf{p}\cdot\mathbf{r}} \frac{1}{(\sqrt{2\pi\hbar})^3}$$

gdzie φ i χ są spinorami wymiaru 2;

rozpisując na te składowe równanie zagadnienia własnego dla H_D dostajemy

$$c \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \mathbf{p} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} + mc^2 \begin{pmatrix} \mathbf{I}_2 & 0 \\ 0 & -\mathbf{I}_2 \end{pmatrix} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \varepsilon \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$$

tzn.

układ równań na spinory φ i χ

$$c\boldsymbol{\sigma}\mathbf{p}\chi + mc^2\varphi = \varepsilon\varphi$$

$$c\boldsymbol{\sigma}\mathbf{p}\varphi - mc^2\chi = \varepsilon\chi$$

ponieważ mówimy o stanach własnych o określonej wartości pędu (cząstka swobodna) możemy zapisać je jako (*)

$$\begin{aligned} c\vec{\sigma}\vec{p}\chi + (mc^2 - \varepsilon)\varphi &= 0 \\ (-mc^2 - \varepsilon)\chi + c\vec{\sigma}\vec{p}\varphi &= 0 \end{aligned}$$

gdzie $\vec{p}$ to już wartość własna operatora pędu $\mathbf{p}$, który ma rozwiązanie gdy znika wyznacznik, czyli

$$\varepsilon^2 = m^2c^4 + c^2(\vec{p}\vec{\sigma})^2$$

co, korzystając z tożsamości

$$(\vec{\sigma}\mathbf{A})(\vec{\sigma}\mathbf{B}) = \mathbf{A}\mathbf{B} + i\vec{\sigma}[\mathbf{A} \times \mathbf{B}]$$

dla dowolnych operatorów/wektorów o trzech składowych komutujących z $\vec{\sigma}$ prowadzi do

$$\varepsilon^2 = p^2c^2 + m^2c^4$$

czyli nic nowego....

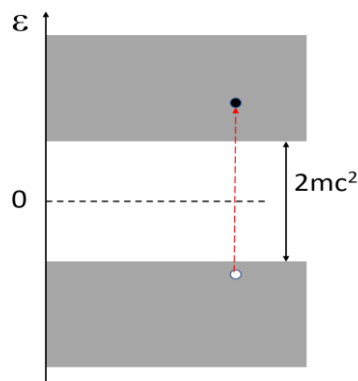
tyle, że nie możemy odrzucić ujemnego pierwiastka (!)

$$\varepsilon = \pm c\sqrt{m^2c^2 + p^2} = \pm E_p.$$

Rozwiązania o ujemnej energii to połowa widma, a połowa funkcji nie tworzy układu zupełnego (!) Ale stany o energii dodatniej nie będą w zasadzie nigdy osiągalne...

Interpretacja P. Diraca:

dla $p=0$ (cząstki spoczywającej) mamy dwa rozwiązania $\varepsilon = \pm mc^2$ i wszystkie stany o $E < -mc^2$ powinny być wypełnione



to z kolei rodzi problemy z nieskończoną liczbą obsadzonych stanów (rozwiązuje się w QED przez renormalizację);

wzbudzenie ze stanu o $E < -mc^2$ powoduje powstanie „dziury” w „morzu” Diraca (próżni Diraca) – o dodatnim ładunku, masie elektronu, i spinie $\frac{1}{2}$ - Dirac w 1928 r. zaproponował interpretację takiego wzbudzenia jako pojawienia się antycząstki – pozytonu;

pozyton został odkryty w 1932 r. przez Andersona;

zapełnienie „dziury” musi odbywać się z emisją kwantu gamma γ o energii $\hbar\omega \geq 2mc^2$ (obserwowane przy zderzeniach $e^+ + e^-$), ale jednocześnie anihilacja kwantu promieniowania γ może prowadzić do spontanicznego powstawania par $e^+ + e^-$ (obserwowane), a to w zasadzie oznacza, że nie istnieje teoria „jednego” elektronu Diraca... (zauważmy też podobieństwo do struktury energetycznej półprzewodników!)

Zitterbevegung – ruch drżący

Z zasady nieoznaczoności dla energii i czasu $\Delta E \Delta t \geq \hbar$ widać, że dla bardzo małych Δt nieokreśloność energii może być nawet $\sim 2mc^2$ a to może prowadzić do kreacji „wirtualnych” (nieobserwowalnych) par $e^+ + e^-$; elektron w czasie $< \Delta t$ może „doświadczać” w swoim ruchu „krótkich obecności” e^+ i e^- - jego ruch nie musi być prostoliniowy w klasycznym sensie, ale w obszarze Δx ($\Delta x \Delta p \geq \hbar$) poddany „drganom”; możemy mieć też problem z definicją położenia cząstki, podobnie jak z prędkością; **wróćmy do tego później.**

Generalnie, żeby uniknąć takich sytuacji powinniśmy „trzymać się” rozwiązań o dodatniej energii.

Zbadajmy teraz jawną postać spinorów φ i χ

z drugiego z równań (*) mamy

$$\chi = \frac{c\sigma\mathbf{p}}{mc^2 + \varepsilon} \varphi$$

wprowadzając **operator znaku** $\Lambda = \frac{H_D}{E_p}$, który komutuje oczywiście z H_D

i ma wartości własne $\lambda = \pm 1$ (rozdzielające rozwiązania o dodatnich i ujemnych energiach) możemy zapisać

$$\Psi(\mathbf{r}) = N \left(\frac{c\sigma\mathbf{p}}{mc^2 + \lambda E_p} u \right) e^{i\hbar\mathbf{p}\mathbf{r}} \frac{1}{(\sqrt{2\pi\hbar})^3}$$

a spinor u jako $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, (u_i – liczby), który musi być unormowany $uu^\dagger = 1$ i zawsze można przedstawić w bazie spinorów $u^{+1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $u^{-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, a zatem mamy dwie funkcje (zapisane bez czynnika fali płaskiej)

(*a)

$$\Psi_\pm = N \left(\frac{c\sigma\mathbf{p}}{mc^2 + \lambda E_p} u^\pm \right)$$

[stąd już blisko do spinu];

przyjrzyjmy się przybliżeniu nierelatywistycznemu,

dla dodatnich energii ($\lambda = 1$) o małych wartościach (ponad mc^2)

$$\varepsilon = E_p = mc^2 + E' \text{ i } E' \ll mc^2$$

mamy

$$\chi = \frac{c\sigma\mathbf{p}}{2mc^2 + E'}\varphi \approx \frac{\sigma\mathbf{p}}{2mc}\varphi \Rightarrow |\chi| \sim \frac{v}{c}|\varphi|, \quad |\chi| \ll |\varphi|$$

dlatego φ nazywamy **dużą składową** a χ **małą składową**
i odwrotnie dla rozwiązań z $\lambda = -1$;

To jest bardzo przydatne w poszukiwaniu przybliżonych równań dwuskładowych opisujących stany o $\lambda = 1$ - dające przybliżenia i granice nierelatywistyczne,

np. poprzez podstawienie χ do pierwszego z równań (*), gdy np. H_D zawiera jakiś potencjał V (musimy wtedy zostawić $\chi \approx \frac{c\sigma\mathbf{p}}{mc^2 - V + E'}$ φ gdyż V zależy na ogół od $\mathbf{r}$, a $\mathbf{p}$ może modyfikować V)

$$c\sigma\mathbf{p}\chi + (mc^2 - \varepsilon)\varphi + V\varphi = 0$$

da (**)

$$\left[\frac{1}{2m}(\sigma\mathbf{p}) \left(1 + \frac{E' - V}{2mc^2} \right)^{-1} (\sigma\mathbf{p}) + V - E' \right] \varphi = 0$$

co musi prowadzić do hamiltonianu kwadratowego w pędzie (czyli Schroedingera) z jakimiś dodatkami ... a dla stałego V da po prostu hamiltonian nierelatywistyczny, gdy skorzystamy z uprzedniej tożsamości $(\sigma\mathbf{p})(\sigma\mathbf{p}) = \mathbf{p}^2 + i\sigma[\mathbf{p} \times \mathbf{p}] = \mathbf{p}^2$

zapamiętajmy (**) – wrócimy do tego przy omawianiu przybliżenia Pauliego

Zobaczmy jak wygląda równanie ciągłości w takim przybliżeniu (dla $\lambda = 1$),
przy powyższym przybliżeniu (ale bez V w H_D), tzn. przyjmując $\chi \approx \frac{\sigma\mathbf{p}}{2mc}\varphi$ mamy

$$\rho = \Psi^\dagger \Psi = (\varphi^\dagger \varphi + \chi^\dagger \chi) \sim \left(1 + \frac{(\sigma\mathbf{p})^2}{4m^2c^2} \right) \varphi^\dagger \varphi = \left(1 + \frac{\mathbf{p}^2}{4m^2c^2} \right) \varphi^\dagger \varphi$$

dla „nierelatywistycznych” cząstek liczy się tylko „1” czyli $\rho \sim \varphi^\dagger \varphi$

natomiast gęstość prądu prawdopodobieństwa:

$$\mathbf{j} = c\Psi^\dagger \boldsymbol{\alpha} \Psi = c(\varphi^\dagger \chi^\dagger) \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \frac{\hbar}{2mi} [\varphi^\dagger \boldsymbol{\sigma}(\boldsymbol{\sigma}\nabla)\varphi - (\boldsymbol{\sigma}\nabla)^\dagger \boldsymbol{\sigma}\varphi]$$

(skorzystałem z faktu, że $\mathbf{p} = -i\hbar\nabla$, a po sprzężeniu $i\hbar\nabla$);

dalej, można pokazać (bez wyprowadzenia) korzystając z $(\boldsymbol{\sigma}\mathbf{A})(\boldsymbol{\sigma}\mathbf{B}) = \mathbf{AB} + i\boldsymbol{\sigma}[\mathbf{A} \times \mathbf{B}]$ i z własności komutacyjnych

$$\mathbf{j} = \frac{\hbar}{2mi} (\varphi^\dagger \nabla \varphi - \varphi \nabla \varphi^\dagger) + \frac{\hbar}{2mi} \text{rot}(\varphi^\dagger \boldsymbol{\sigma} \varphi)$$

(gdyż $\text{rot}(\mathbf{F}) = \nabla \times \mathbf{F}$)

pierwszy człon to odpowiednik nierelatywistyczny, ale drugi (ten, który zaniedbujemy w granicy nierelatywistycznej) zależy od ... spinu (!)

Parzystość i operatory rzutowe

Wcześniej zidentyfikowaliśmy problem z definicją / interpretacją pewnych wielkości (położenie i prędkość);

każde rozwiązanie Ψ równania Diraca (zagadnienia własnego dla H_D) ma dwie składowe, w stanach o ujemnej energii duża i mała składowa zamieniają się rolami; w przybliżeniu nierelatywistycznym powinniśmy „wyciąć” z rozwiązań części odpowiadające ujemnym energiom;

definiujemy dwa operatory

$$\Pi_+ = \frac{1}{2}(1 + \Lambda), \quad \Pi_- = \frac{1}{2}(1 - \Lambda), \quad \Lambda = \frac{H_D}{E_p}$$

hermitowskie, idempotentne,

są to operatory rzutowe na podprzestrzeń rozwiązań różniących się znakiem E_p

wszystkie operatory działające w przestrzeni Hilberta H_D możemy podzielić na dwie klasy

- a) takie, które $[\mathbf{O}]\Psi_{\pm} = \Phi_{\pm}$ działając na funkcję nie wyprowadzają jej z podprzestrzeni stanów o danym znaku energii,

operatory parzyste

- b) i takie, które $\{\mathbf{O}\}\Psi_{\pm} = \Phi_{\mp}$, mieszają podprzestrzenie $E > mc^2$ i $E < -mc^2$ -

nieparzyste

oczywiście zachodzi $\langle \Psi_{\pm} | \{\mathbf{O}\} | \Psi_{\pm} \rangle$;

chcielibyśmy umieć wydzielać z dowolnego operatora $\mathbf{a} = [\mathbf{a}] + \{\mathbf{a}\}$ część parzystą ($[\mathbf{a}]$), gdyż proces kreacji par wiąże się z mieszaniem stanów o energii $\pm$, a my chcemy mieć przybliżenie jednocząstkowe;

ponieważ $\mathbf{a}\Psi_{\pm} = [\mathbf{a}]\Psi_{\pm} + \{\mathbf{a}\}\Psi_{\pm}$, zatem $\Lambda\mathbf{a}\Lambda\Psi_{\pm} = [\mathbf{a}]\Psi_{\pm} - \{\mathbf{a}\}\Psi_{\pm}$ (gdyż $\Lambda\Psi_{\pm} = \pm\Psi_{\pm}$) a to oznacza, że (poprzez dodanie stronami powyższych równości na $\mathbf{a}\Psi_{\pm}$ i $\Lambda\mathbf{a}\Lambda\Psi_{\pm}$)

$$\frac{1}{2}(\mathbf{a} + \Lambda\mathbf{a}\Lambda)\Psi_{\pm} = [\mathbf{a}]\Psi_{\pm}, \quad \frac{1}{2}(\mathbf{a} - \Lambda\mathbf{a}\Lambda)\Psi_{\pm} = \{\mathbf{a}\}\Psi_{\pm}$$

są operatorami wybierającymi część parzystą i część nieparzystą z danego operatora $\mathbf{a}$;

wracając do problematycznego określenia operatora prędkości $\frac{d\hat{\mathbf{r}}}{dt} = c\boldsymbol{\alpha}$ (który nie jest operatorem parzystym), to jeśli wytniemy z niego tylko część parzystą, tzn. policzymy $c[\boldsymbol{\alpha}]$, (trzeba zastosować $\frac{1}{2}(\boldsymbol{\alpha} + \boldsymbol{\Lambda}\boldsymbol{\alpha}\boldsymbol{\Lambda})$) to otrzymamy (bez wyprowadzenia)

$$c[\boldsymbol{\alpha}] = \frac{c^2 \mathbf{p} \boldsymbol{\Lambda}}{E_p}$$

zatem, biorąc wartości operatorów i licząc dla małych dodatnich $E' \ll mc^2$ (pamiętamy, że E_p zdefiniowaliśmy jako $E_p = mc^2 + E'$) dostaniemy $\sim$

$$\frac{c^2 p}{mc^2} = \frac{p}{m} \sim v$$

czyli prędkość zdefiniowaną nierelatywistycznie.

Spin

Rozpatrzmy wielkość $\frac{\hbar}{2} \boldsymbol{\Sigma} \mathbf{p}$, gdzie $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} = \mathbf{I}_2 \otimes \boldsymbol{\sigma}$, komutuje z hamiltonianem i jest stałą ruchu, ale dla swobodnej cząstki pęd $\mathbf{p}$ jest stałą ruchu i jeśli wybierzemy pęd w kierunku osi z to stałą ruchu będzie też

$$\frac{\hbar}{2} \Sigma = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

mający dwie wartości własne $S = \pm \frac{\hbar}{2}$, nazywamy go **operatorem rzutu spinu na kierunek pędu** nasze rozwiązania dla cząstki swobodnej mają już trzy indeksy $\Psi_{p,\lambda,s}(\mathbf{r})$

ze względu na blokową postać macierzy $\boldsymbol{\Sigma}$ możemy równie dobrze mówić o operatorze $\frac{\hbar}{2} \boldsymbol{\sigma} \mathbf{p}$;

niestety, funkcje $\Psi_{\pm}$ nie są funkcjami własnymi operatora $\boldsymbol{\Sigma}$, co łatwo sprawdzić,

$$\boldsymbol{\sigma} \mathbf{p} \Psi_{\pm} = N \begin{pmatrix} \boldsymbol{\sigma} \mathbf{p} u^{\pm} \\ \frac{cp^2}{mc^2 + \lambda E_p} u^{\pm} \end{pmatrix}$$

czyli $\boldsymbol{\Sigma}$ nie może być relatywistycznym operatorem spinu.

łatwo sprawdzić, że

$$[\boldsymbol{\sigma} \hat{\mathbf{n}}, H] = [\boldsymbol{\sigma} \hat{\mathbf{n}}, \boldsymbol{\alpha} \mathbf{p}] = 2i\boldsymbol{\alpha} \cdot \mathbf{n} \times \mathbf{p}$$

co oznacza, że komutator zniknie dla gdy $\mathbf{n} \times \mathbf{p} = 0$, a to oznacza, że może istnieć kombinacja liniowa $\Psi_{\pm}$, która będzie funkcją własną $\boldsymbol{\Sigma}$;

ale nam chodzi o funkcje własne H_D ,
wprowadzając dwa wektory $\hat{e}_1$ i $\hat{e}_2$ prostopadłe do $\mathbf{p}$
budujemy operator

$$\mathbf{O} = \boldsymbol{\sigma} \cdot \mathbf{p}\mathbf{p} + \beta \boldsymbol{\sigma} \cdot \hat{e}_1 \hat{e}_1 + \beta \boldsymbol{\sigma} \cdot \hat{e}_2 \hat{e}_2$$

mający jedną składową kartezjańską wzdłuż pędu, komutujący z H_D , ale jego składowe nie komutują między sobą (gdyż, np. $\hat{e}_1 \times \hat{e}_2 = \mathbf{p}$ i skorzystamy jak zwykle z tożsamości $(\boldsymbol{\sigma}\mathbf{A})(\boldsymbol{\sigma}\mathbf{B}) = \mathbf{AB} + i\boldsymbol{\sigma}[\mathbf{A} \times \mathbf{B}]$);

można pokazać, że

$$O_z \Psi_{\pm} = \pm \Psi_{\pm}$$

$\mathbf{O}$ nazywamy **relatywistycznym operatorem spinu**, lub „operatorem polaryzacji”;
można też sprawdzić, że wartość oczekiwana $\mathbf{O}$ w stanach $\Psi_{\pm}$ jest

$$\langle \Psi_{\pm} | \mathbf{O} | \Psi_{\pm} \rangle = \pm \hat{e}_z$$

jest tak ponieważ, $\boldsymbol{\sigma}$ działa tylko na $u^{\pm 1/2}$, które są wektorami własnymi σ_z , więc z każdego składnika $\mathbf{O}$ wycina tylko z-ową składową (rzyt tego składnika na kierunek z);

Fakt, że wartość średnia $\mathbf{O}$ leży na osi z wynika z tego, że w $\Psi_{\pm}$ mamy w dużych i małych składowych φ i χ spinory $u^{+1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ i $u^{-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, które są spinorami własnymi σ_z .

Cząstka w polu elektromagnetycznym (granica nierelatywistyczna)

Wiemy już (FK2 i powyżej), że wpływ zewnętrznych pól elektromagnetycznych opisujemy w mechanice kwantowej *via* potencjał wektorowy, który w H_D modyfikuje operator pędu

$$\mathbf{p} \rightarrow \boldsymbol{\pi} = \mathbf{p} - e\mathbf{A}$$

jeśli mamy dodatkowo stałe pole ϕ to do energii wnosi ono energię potencjalną równą $-e\phi$

$$H_D(\boldsymbol{\pi})\Psi = (\varepsilon - e\phi)\Psi$$

zatem w przedstawieniu na duże i małe składowe $\Psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$ równanie przyjmuje postać

(***)

$$\begin{aligned} c\boldsymbol{\sigma}(\mathbf{p} - e\mathbf{A})\chi &= (\varepsilon - e\phi - mc^2)\varphi \\ (\varepsilon - e\phi + mc^2)\chi &= c\boldsymbol{\sigma}(\mathbf{p} - e\mathbf{A})\varphi \end{aligned}$$

postępując tak jak poprzednio, tzn. eliminując składową χ i zakładając energie nierelatywistyczne, czyli

$$(\varepsilon - e\phi + mc^2) = (mc^2 + E' - e\phi + mc^2) \sim 2mc^2$$

i tym samym

$$\chi \approx \frac{\boldsymbol{\sigma}(\mathbf{p} - e\mathbf{A})}{2mc} \varphi$$

co daje równanie

$$(E' - e\phi)\varphi = \frac{1}{2m} [\boldsymbol{\sigma}(\mathbf{p} - e\mathbf{A})]^2 \varphi$$

a korzystając ze znanej już tożsamości $(\boldsymbol{\sigma}\mathbf{A})(\boldsymbol{\sigma}\mathbf{B}) = \mathbf{AB} + i\boldsymbol{\sigma}[\mathbf{A} \times \mathbf{B}]$ dostaniemy

$$[\boldsymbol{\sigma}(\mathbf{p} - e\mathbf{A})]^2 = (\mathbf{p} - e\mathbf{A})^2 + i\boldsymbol{\sigma}[(\mathbf{p} - e\mathbf{A}) \times (\mathbf{p} - e\mathbf{A})]$$

ostatni wyraz nie znika – to są operatory a nie tylko wektory – zostają wyrazy mieszane, pamiętając o tym, że $\mathbf{p} = -i\hbar\nabla$ dostaniemy

$$\left[\frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 - \frac{\hbar e}{2m} \boldsymbol{\sigma}(\nabla \times \mathbf{A} + \mathbf{A} \times \nabla) + e\phi \right] \varphi = E' \varphi$$

przy odpowiednim cechowaniu $\mathbf{A}$ dla stałego pola $\mathbf{B}$ możemy przyjąć

$$(\nabla \times \mathbf{A} + \mathbf{A} \times \nabla) = \text{rot}\mathbf{A}$$

ale

$$\text{rot}\mathbf{A} = \mathbf{B}$$

zatem

$$\left[\frac{1}{2m} (\mathbf{p} - e\mathbf{A})^2 - \frac{\hbar e}{2m} \boldsymbol{\sigma}\mathbf{B} + e\phi \right] \varphi = E' \varphi$$

i to równanie nazywa się **przybliżeniem Pauliego**;

drugi wyraz wygląda jak oddziaływanie jakiegoś momentu magnetycznego z polem $\mathbf{B}$

$$\frac{\hbar e}{2m} \boldsymbol{\sigma}\mathbf{B} = \boldsymbol{\mu}_s \cdot \mathbf{B}$$

gdzie

$$\boldsymbol{\mu}_s = \frac{\hbar e}{2m} \boldsymbol{\sigma} = \frac{e}{m} \mathbf{s}$$

dla

$$\mathbf{s} = \frac{1}{2} \hbar \boldsymbol{\sigma}$$

Równanie Diraca w granicy nierelatywistycznej wprowadziło automatycznie spin i związany z nim moment magnetyczny (!)

Wybierając $\mathbf{A} = \mathbf{B} \times \frac{\mathbf{r}}{2}$ wyraz $(\mathbf{p} - e\mathbf{A})^2$ (przy zaniedbaniu małego $\mathbf{A}^2$) prowadzi do

$$\mathbf{p}^2 - 2e\mathbf{A}\mathbf{p} = \mathbf{p}^2 - 2e\frac{1}{2}(\mathbf{B} \times \mathbf{r})\mathbf{p} = \mathbf{p}^2 - e(\mathbf{r} \times \mathbf{p})\mathbf{B} = \mathbf{p}^2 - e\mathbf{L}\mathbf{B}$$

[$2e\mathbf{A}\mathbf{p}$ pojawia się ze względu na cechowanie $[\mathbf{p}, \mathbf{A}] = 0$, co daje $\mathbf{p}\mathbf{A} = \mathbf{A}\mathbf{p}$, a w rozwinięciu $(\mathbf{p} - e\mathbf{A})^2$ pojawi się $\mathbf{A}\mathbf{p}$ i $\mathbf{p}\mathbf{A}$]

ale to jest powyżej mnożone przez $\frac{1}{2m}$ (nb. $\mu_l = \frac{e}{2m}\mathbf{L}$, a $\mu_s = \frac{e}{m}\mathbf{S}$), zatem dostajemy:

$$\left[\frac{1}{2m}\mathbf{p}^2 - \frac{e}{2m}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B} + e\phi \right] \varphi = E'\varphi$$

czyli to co wprowadzaliśmy fenomenologicznie w teorii nierelatywistycznej.

Oddziaływanie spin-orbita i dalsze poprawki

Wróćmy do równań (***) na składowe φ, χ , ale $e\phi$ zapiszmy jako pewien potencjał V , np. pole centralne itp., oczywiście dla $E' \ll mc^2$

$$\begin{aligned} c\sigma(\mathbf{p} - e\mathbf{A})\chi &= (E' - V)\varphi \\ (E' - V + 2mc^2)\chi &= c\sigma(\mathbf{p} - e\mathbf{A})\varphi \end{aligned}$$

podobnie jak poprzednio wyznaczamy χ z drugiego równania i wstawiamy do pierwszego

$$E'\varphi = \frac{1}{2m}(\sigma\mathbf{p}) \left(1 + \frac{E' - V}{2mc^2} \right)^{-1} (\sigma\mathbf{p})\varphi + V\varphi$$

przygotowanie:

$$[\mathbf{p}, V(r)]f(r) = \mathbf{p}(V(r)f(r)) - V\mathbf{p}f(r) = (-i\hbar\nabla V)f(r) - i\hbar\nabla f(r)V - V\mathbf{p}f(r) = -i\hbar\nabla V f(r) + V\mathbf{p}f(r) - V\mathbf{p}f(r)$$

zatem $[\mathbf{p}, V] = -i\hbar\nabla V$, zatem $\mathbf{p}V = V\mathbf{p} - i\hbar\nabla V$

a nawias przybliżamy jako $[1/(1+x) \sim 1-x]$, czyli jako $\left(1 - \frac{E' - V}{2mc^2}\right)$, oraz $(\sigma\mathbf{p})(\sigma\mathbf{p}) = \mathbf{p}^2$
w rezultacie
(4*)

$$E'\varphi = \left[\left(1 - \frac{E' - V}{2mc^2} \right) \frac{\mathbf{p}^2}{2m} + V \right] \varphi - \frac{i\hbar}{4m^2c^2} (\sigma\nabla V)(\sigma\mathbf{p})\varphi$$

ale $(\sigma\mathbf{A})(\sigma\mathbf{B}) = \mathbf{A}\mathbf{B} + i\sigma[\mathbf{A} \times \mathbf{B}]$, czyli $(\sigma\nabla V)(\sigma\mathbf{p}) = (\nabla V)\mathbf{p} + i\sigma[\nabla V \times \mathbf{p}]$

zatem drugi wyraz po prawej stronie (4*)

$$-\frac{\hbar^2}{4m^2c^2}(\nabla V)(\nabla \varphi) + \frac{\hbar}{4m^2c^2}\boldsymbol{\sigma}(\nabla V \times \mathbf{p})$$

biorąc sferycznie symetryczny $V(r)$ zależny tylko od odległości (r) [jak to ma miejsce w atomach], dla którego $\nabla V \cdot \nabla = \frac{\partial V}{\partial r} \frac{\partial}{\partial r}$, a $\nabla V = \frac{\mathbf{r}}{r} \frac{\partial V}{\partial r}$, pamiętając, że $\mathbf{r} \times \mathbf{p} = \mathbf{L}$, oraz podstawiając $\frac{1}{2}\hbar\boldsymbol{\sigma} = \mathbf{S}$,

dostanemy

(5*)

$$E'\varphi = \left(\frac{\mathbf{p}^2}{2m} + V\right)\varphi - \left(\frac{E' - V}{2mc^2}\right)\frac{\mathbf{p}^2}{2m}\varphi - \frac{\hbar^2}{4m^2c^2}\frac{\partial V}{\partial r}\frac{\partial \varphi}{\partial r} + \frac{1}{2m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}(\mathbf{S} \cdot \mathbf{L})\varphi$$

ten ostatni wyraz to **oddziaływanie spin-orbita**,

a pierwszy wyraz to nierelatywistyczny hamiltonian,

a całość po prawej stronie to hamiltonian Pauliego H_P .

[Uwaga: to jest równanie tylko na dużą składową]

Dalsze poprawki

Zauważmy przede wszystkim, że przechodząc do granicy nierelatywistycznej:

- otrzymaliśmy poprawki takie jak oddziaływanie S-O, czy pojawienie się spinu i oddziaływanie jego momentu magnetycznego z zewnętrznym polem $\mathbf{B}$,
- efekty te **tkwią fundamentalnie w równaniu Diraca**, ale ich „klasyczne” postaci uwidaczniają się w granicy nierelatywistycznej,
- całkowicie nierelatywistyczne równanie dostaniemy przechodząc w przybliżeniu Pauliego z $c \rightarrow 0$.

Dwa środkowe wyrazy w (5*) też coś znaczą, ale najpierw:

funkcja φ nie jest unormowana, przypomnijmy normalizację, którą omawialiśmy przy równaniu ciągłości

$$\rho = \Psi^\dagger \Psi = (\varphi^\dagger \varphi + \chi^\dagger \chi) \sim \left(1 + \frac{\mathbf{p}^2}{4m^2c^2}\right) \varphi^\dagger \varphi$$

to powinniśmy brać nie φ , ale $\tilde{\varphi} = \left(1 + \frac{\mathbf{p}^2}{4m^2c^2}\right)^{1/2} \varphi = \mathbf{g}\varphi$, żeby uzyskać unormowanie do 1.

Ale jeśli transformujemy funkcję z φ do $\tilde{\varphi} = \mathbf{g}\varphi$, to powinniśmy też transformować równanie

$$E'\mathbf{g}\varphi = \mathbf{g}H_P\mathbf{g}^{-1}\mathbf{g}\varphi$$

rozwijając $\left(1 + \frac{\mathbf{p}^2}{8m^2c^2}\right)^{1/2}$ w szereg względem małych wyrazów $\sim v^2/c^2$, przybliżymy ten operator jako $1 + \frac{\mathbf{p}^2}{4m^2c^2} + \theta\left(\frac{v^4}{c^4}\right)$, wówczas

$$\tilde{H}_P = \mathbf{g}H_P\mathbf{g}^{-1} = \left(1 + \frac{\mathbf{p}^2}{8m^2c^2}\right)H_P\left(1 - \frac{\mathbf{p}^2}{8m^2c^2}\right) = H_P + \frac{1}{8m^2c^2}[\mathbf{p}^2, H_P] + \theta\left(\frac{v^4}{c^4}\right)$$

oddziaływanie S-O pozostaje bez zmian w H_P , ale inne wyrazy zawierające $V(r)$ będą wносиły swoje wkłady;
ostatecznie – bez wyprowadzenia – dostaniemy:

$$\tilde{H}_P = \frac{\mathbf{p}^2}{2m} + V - \frac{(E' - V)^2}{2mc^2} - \frac{\hbar^2}{8m^2c^2}\Delta V + \frac{1}{2m^2c^2}\frac{1}{r}\frac{\partial V}{\partial r}(\mathbf{s} \cdot \mathbf{L})$$

wszystkie poprawki są rzędu $\sim v^2/c^2$,

drugi wyraz, proporcjonalny w zasadzie do kwadratu energii kinetycznej, to tak zwana **poprawka na wariację masy**, wynikająca ze zmienności masy cząstki z prędkością,

trzeci wyraz to tzw. **poprawka Darwina** – związana jest z procesami kreacji i anihilacji, dla $V = -Ze/r$, $\Delta V = -4\pi\delta(r)$ – „oddziaływanie kontaktowe”; nie znika tylko w stanach atomowych s .

Atom wodoru

W potencjale kulombowskim, operatorami komutującymi z H_D są J^2 i J_z , gdzie $J=L+S$, rozwiązanie równania Diraca daje

$$E_{nj} = mc^2 \left[1 + \left(\frac{Z\alpha}{n - (j + \frac{1}{2}) + \sqrt{(j + \frac{1}{2})^2 - Z^2\alpha^2}} \right)^2 \right]^{-1/2} - mc^2$$

gdzie $\alpha = \frac{e^2}{\hbar c}$ nazywa się stałą struktury subtelnej (niemianowana $\sim 1/137$);

rozwiniecie E_{nj} wzgl. potęg $(Z\alpha)^2$ prowadzi do:

$$E_{nj} = -mc^2 \frac{1}{2} \left(\frac{Z\alpha}{n} \right)^2 \left[1 + \left(\frac{Z\alpha}{n} \right)^2 \left(\frac{n}{j + \frac{1}{2}} - \frac{3}{4} \right) \right] + \dots$$

pierwszy wyraz to

dokładnie nierelatywistyczna energia elektronu w wodorze: $-\frac{mZ^2e^4}{2\hbar^2n^2}$.

Zerowa masa spoczynkowa

Na początku lat 30. Zeszłego wieku, najpierw Pauli a potem Fermi zapostulowali istnienie neutrino, bezmasowej i bezładunkowej cząstki o spinie $\frac{1}{2}$ występującej w teorii rozpadu β . W 1956 r. po raz pierwszy wykryto neutrino (Reines i Cowan). Dziś wiemy, że neutrino ma masę, bardzo małą, ale nie znamy jej wartości. Początkowo sądzono, że równania Diraca dla cząstek o masie $m=0$ mogą opisywać neutrino.

Równania Diraca (na składowe φ, χ) dla swobodnej cząstki bezmasowej ($m=0$) o małej energii, mają postać

$$\begin{aligned} c\boldsymbol{\sigma}\mathbf{p}\chi &= E'\varphi \\ c\boldsymbol{\sigma}\mathbf{p}\varphi &= E'\chi \end{aligned}$$

definiując $\mathbf{n} = c\mathbf{p}/E'$ widać, że

$$(\boldsymbol{\sigma}\mathbf{n}) \begin{pmatrix} \varphi \\ \chi \end{pmatrix} = \begin{pmatrix} \chi \\ \varphi \end{pmatrix}$$

czyli

$$(\boldsymbol{\sigma}\mathbf{n}) = \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix} = \text{ozn.} = -\gamma_5$$

warto zauważyć, że z drugiej strony zachodzi związek

$$\gamma_5 = \gamma_1\gamma_2\gamma_3\gamma_4$$

gdzie, przypomnijmy, γ_i to macierze Diraca z współmienniczej postaci równania, tak zawiązane z macierzami $\boldsymbol{\alpha}$ i β :

$$\gamma = -i\beta\boldsymbol{\alpha}, \quad \gamma_4 = \beta ;$$

zamiast składowych φ, χ

wprowadźmy ich kombinacje liniowe

$$\begin{aligned} \Phi &= \frac{1}{2}(\varphi + \chi) = 1/2(1 + \boldsymbol{\sigma}\mathbf{n})\varphi \\ F &= \frac{1}{2}(\varphi - \chi) = 1/2(1 - \boldsymbol{\sigma}\mathbf{n})\varphi \end{aligned}$$

dodając i odejmując te równania dostajemy

$$\boldsymbol{\sigma}\mathbf{n}\Phi = \Phi, \quad \boldsymbol{\sigma}\mathbf{n}F = -F$$

Funkcje Φ i F są spinorami dwuskładnikowymi – funkcjami własnymi operatora $\boldsymbol{\sigma}\mathbf{n}$ - rzutu spinu na kierunek pędu (do wartości $+1/2$ i $-1/2$ [gdyż $\mathbf{s} = \frac{1}{2}\hbar\boldsymbol{\sigma}$]),

Tę własność nazywamy **skrętnością** (*helicity*)

Na koniec zauważmy, że nasze równanie

$$-i\hbar c\boldsymbol{\sigma}\nabla\Psi = E'\gamma_5\Psi,$$

jest – poza skrętnością –

podobne do równania jakie w metodzie ciasnego wiązania dla grfenu, w przybliżeniu pi-elektronowym, w którym funkcję wariacyjną buduje się tylko z jednego orbitala $2p_z$ na każdym

atomie węgla i które dla granicznych punktów strefy Brillouina (tzw. punkty **K** Diraca) wygląda tak:

$$-i\hbar v_F \boldsymbol{\sigma} \cdot \nabla \Psi(\mathbf{r}) = E \Psi(\mathbf{r}),$$

gdzie v_F prędkość elektronów na poziomie Fermiego [w szczycie stożka Diraca, stożki – gdyż energia $E(\mathbf{k})$ jest w punktach **K** liniowo proporcjonalna do pędu (wektora falowego $\mathbf{k}$)]

o takich elektronach w grafenie (układ 2D) mówimy, że są quasi-fermionami Diraca, a opisujące je równania pseudo-relatywistycznymi równaniami Diraca;

w pewnych egzotycznych układach grafenowych i w dwuwymiarowych izolatorach topologicznych mogą pojawiać się rozwiązania równań, które mogą posiadać cechy skrętności.