

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

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

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 sugerowałyby to potrzeba opisanie 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:

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

i ostatecznie

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

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 współmienniczą 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 $i\hbar \frac{\partial \Psi}{\partial t} - c\boldsymbol{\alpha} \mathbf{p} \Psi = \beta mc^2 \Psi$ 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

$$\begin{aligned} \rho &= \Psi^\dagger \Psi \\ \mathbf{j} &= \Psi^\dagger c\boldsymbol{\alpha} \Psi \end{aligned}$$

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 $\mathbf{r}$

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

to dla dowolnej składowej $\hat{r}$, 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 $\frac{d\hat{r}}{dt} = c\alpha \dots$

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 α i β 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

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

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

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

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

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

co, korzystając z tożsamości

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

dla dowolnych operatorów/wektorów o trzech składowych komutujących z $\boldsymbol{\sigma}$

prowadzi do

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

czyli nic nowego....

tyle, że nie możemy odrzucić ujemnego pierwiastka (!)

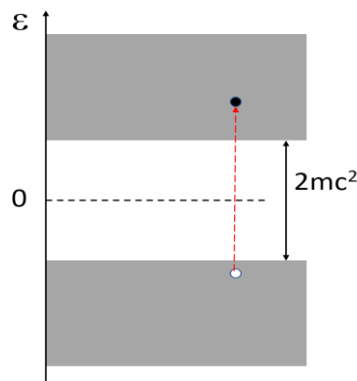
$$\varepsilon = \pm c \sqrt{m^2 c^2 + p^2} = \pm E_p.$$

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 zajęte



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!)

Zitterbewegung – 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

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

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 $\chi \approx \frac{c\sigma\mathbf{p}}{mc^2 - V + E'} \varphi$ gdyż V zależy na ogół od $\mathbf{r}$, a $\mathbf{p}$ może modyfikować V)

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

da (**)

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

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

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 $\chi \approx \frac{\boldsymbol{\sigma}\mathbf{p}}{2mc} \varphi$ mamy

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

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

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

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

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 pędność);

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 $\frac{1}{2}(\boldsymbol{\alpha} + \Lambda\boldsymbol{\alpha}\Lambda)$) to otrzymamy (bez wyprowadzenia)

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

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.