

Romuald Kotowski

PODSTAWY KLASYCZNEJ TEORII POLA

4/1990

WARSZAWA 1990

<http://rcin.org.pl>

Praca wpłynęła do Redakcji 3 października 1990 r.



56784



Na prawach rękopisu

Instytut Podstawowych Problemów Techniki PAN

Nakład 140 egz. Ark. wyd. 4,4 Ark.druk. 5,5

Oddano do drukarni w lutym 1990 r.

Nr zamówienia 52/90

Warszawska Drukarnia Naukowa, Warszawa
ul. Śniadeckich 8

Romuald Kotowski
 Samodzielna Pracownia
 Fizyki Ośrodków Strukturalnych

PODSTAWY KLASYCZNEJ TEORII POLA

W pracy omówiono podstawowe pojęcia z teorii pola, a mianowicie formalizm lagrangowski, pola: skalarne, wektorowe, Maxwella, Diraca oraz superpole. Wprowadzono podstawowe pojęcia z teorii cechowania. W dodatku zebrano definicje i twierdzenia z teorii grup, przydatne w studiowaniu niniejszego opracowania.

WSTĘP

Polem nazywamy wielkość fizyczna określona w całej czasoprzestrzeni i zmieniająca się od punktu do punktu. Teoria pola zajmuje się badaniem pól w przestrzeni fizycznej. Różnych pól jest oczywiście wielka różnorodność. Mówimy np. o polu elektromagnetycznym, o polu temperatury, o polu prędkości... W praktyce, mówiąc teorią pola, ma się na myśli jednak jej najnowsza i najbardziej doskonała postać, a mianowicie kwantową teorię pola (KTP), czyli teorię oddziaływań elementarnych. Dziś właściwie możemy mówić o dwu nurtach KTP - jeden zajmuje się teorią cząstek elementarnych, a drugi zjawiskami w ośrodkach skondensowanych.

Sukces KTP jest spowodowany wyjątkowo szczęśliwym połączeniem dwu elementów: *formalizmu lagrangowskiego* i pojęcia pola, czyli sposobu oddziaływania obiektów fizycznych oddalonych od siebie na skończone odległości. Formalizm lagrangowski (bardziej szczegółowo powiemy o nim później) ma kontrkandydata w postaci tzw. *formalizmu kanonicznego*, czyli *formalizmu hamiltonowskiego*.

W formalizmie hamiltonowskim wielkościami podstawowymi są uogólnione współrzędne, uogólnione pędy i zbudowana za ich pomocą tzw. funkcja Hamiltona. Przy wszystkich swoich zaletach (np. możliwość porównania z mechaniką klasyczną) formalizm kanoniczny ma jedną istotną wadę: czas pełni rolę wyróżnioną i tym samym z punktu widzenia teorii względności Einsteina formalizm traci tak pożądaną niezmienniczość.

Max Planck swą hipotezę, która dziś formułujemy mówiąc, że kwantami pola elektromagnetycznego są fotony, wygłosił w roku 1901. Jednakże kwantowa elektrodynamika oparta na kowariantnej teorii zaburzeń została sformułowana dopiero w roku 1950. Sukces nowej teorii był wspaniały: za jej pomocą można było np. wyznaczyć anomalny moment magnetyczny elektronu z dokładnością do szóstego miejsca po przecinku.

Pojawił się jednakże pewien problem, a mianowicie dualizm korpuskularno-falowy. Ruch elektronu (cząstki) opisywano za pomocą falowego równania Schrödingera. Przestało być zatem całkiem jasne, na czym polega różnica pomiędzy cząstkami a polami: cząstki punktowe (elektrony) opisuje się równaniem falowym, a pole elektromagnetyczne, które w fizyce klasycznej było traktowane jako ośrodek rozciągliwy, otrzymało właściwości cząstek (fotonów). Poszukiwania odpowiedzi na pytanie, jaka jest istotna różnica pomiędzy polami a cząstkami doprowadziło do wypracowania następującego poglądu. Istnieje pewien ciąg stanów materii, odpowiadających tzw. cząstkom elementarnym (np. e, μ , p, Ω , Σ itd.) oraz pole, za pomocą którego te cząstki oddziałują ze sobą. Mówiąc inaczej, mamy do czynienia z cząstkami i z ich oddziaływaniami.

Fizyka współczesna zna cztery rodzaje oddziaływań: *grawitacyjne*, *elektromagnetyczne*, *słabe* i *silne*. Jak powiedzieliśmy powyżej, kwantami pola elektromagnetycznego są fotony. Z fizyki klasycznej znane jest nam jeszcze oddziaływanie grawitacyjne, które opisywane jest za pomocą *ogólnej teorii względności*. Kwantowanie pola grawitacyjnego, analogicznie jak w teorii elektromagnetyzmu, powinno dać kwanty pola sił ciężkości, czyli *grawitony*. Jak na razie nie udało

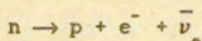
się znaleźć doświadczalnie przekonujących dowodów ich istnienia. Trudności są również natury matematycznej i pojęciowej. Samo równanie Einsteina jest bardziej skomplikowane od równań Maxwella, a ponadto jest jeszcze nieliniowe. Pole elektromagnetyczne rozciąga się w czasoprzestrzeni, natomiast w teorii Einsteina ciężenie jest wynikiem zakrzywienia samej czasoprzestrzeni. W jakim więc sensie mamy rozumieć samo pojęcie kwantowania? Na problemy te nie znaleziono jeszcze rzetelnych odpowiedzi, chociaż pewną nadzieję stwarza nieabelowa teoria cechowania, która powinna mieć zastosowanie również w teorii grawitacji. Siły grawitacji są na szczęście najslabsze ze znanych oddziaływań i dlatego w teorii cząstek elementarnych mogą być i są pomijane.

Oddziaływania silne i słabe pojawiają się przy rozpatrywaniu sił jądrowych. Początkowo wydawało się, zgodnie z teorią Yukawy, że za oddziaływanie silne proton-neutron (p-n) w jądrze atomowym, odpowiedzialny jest mezon π^+ o masie $140 \text{ MeV}/c^2$ (mezon - ponieważ ma masę pośrednią pomiędzy masą elektronu a masą protonu). Dalsze badania natrafiły jednak na spore trudności. Okazało się, że przy wysokich energiach wymiana pionu nie dawała zadowalających wyników, że oddziaływania pomiędzy samymi pionami nie da się wytłumaczyć wymianą jednego pionu (zw względu na niezachowanie parzystości), że klasyfikacja cząstek na podstawie grupy $SU(3)$ stwarza istotne trudności: do oktetu wchodzi trzy piony π^+ , π^- i π^0 , a ponadto jeszcze mezony K i η , które są 'zwykłymi' cząstkami. Trudności te udało się pokonać dopiero z chwilą wprowadzenia kwarków. W tej nowej teorii mezony są stanami związanymi kwark-antykwar - a więc ich uprzywilejowane położenie przestało mieć rację bytu (foton na przykład nie jest zbudowany z kwarków) i tym samym pion przestał być kandydatem na kwant pola.

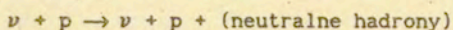
Pojawia się nowe pytanie: jak kwarki oddziałują ze sobą? Obecnie uważa się, że kwark jest nieobserwowalny w stanie swobodnym i może występować tylko w stanach związanych. Ponadto jest on wyposażony w liczbę kwantową analogiczną do ładunku elektrycznego, lecz jednocześnie różniącą się od niego tym, że ma trzy stopnie swobody a

nie dwa (trzy rodzaje ładunku). Ta nowa liczba kwantowa nosi nazwę *koloru*, a stopnie swobody nazywa się odpowiednio kolorem białym (White), czerwonym (Red) i niebieskim (Blue). Kolor, podobnie jak w elektromagnetyzmie ładunek, powoduje istnienie pola kwantowego, za pomocą którego kwarki oddziałują. Kwanty tego pola noszą nazwę *gluonów*. Teoria oddziaływania kwarków za pomocą gluonów nosi nazwę *chromodynamiki kwantowej* (analogicznie do elektrodynamiki kwantowej). Nie dziwi teraz fakt, że siły oddziaływania p-n były opisywane za pomocą wymiany pionu tylko w przybliżeniu - w rzeczywistości siła ta jest wielkością dość skomplikowaną jako wypadkowa oddziaływań kwarków.

Pojęcie *oddziaływań słabych* związane jest z tak zwanym *rozpadem beta*. Ta ostatnia nazwa pochodzi z fizyki jądrowej, gdzie rozpatrywano *promieniowanie alfa* - czyli strumień jader helu, *promieniowanie beta* - strumień elektronów i *promieniowanie gamma* - strumień fotonów odpowiedniej energii. W teorii Fermiego rozpad beta następował punktowo, np. rozpad neutronu na proton, elektron i antyneutrino elektronowe



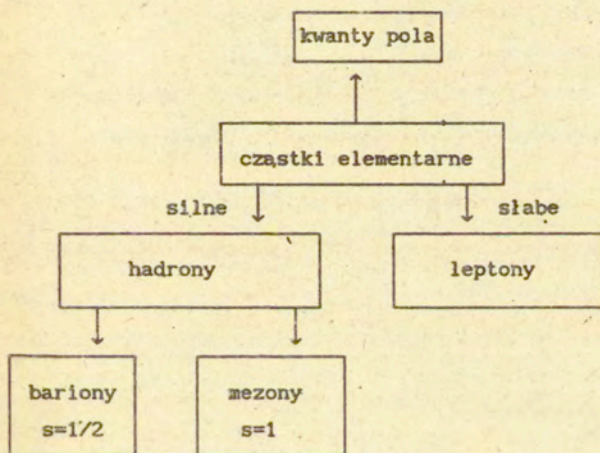
i nie było potrzeby wprowadzania żadnego kwantu oddziaływań. Teoria Fermiego była zadowalająca przez wiele lat, niemniej uważano ją za nieprawdziwą, gdyż nie poddawała się *renormalizacji*. W latach sześćdziesiątych S.L. Glashow, S. Weinberg i A. Salam przedstawili teorię, w której w nietrywialny sposób udało im się połączyć oddziaływania elektromagnetyczne i słabe. Kwantami słabego oddziaływania w tej nowej teorii są *bozony W i Z*, około osiemdziesiąt razy cięższe od protonu. Przepowiedziano ponadto reakcje z tak zwanymi *pradami neutralnymi*, na przykład



Najważniejszy był jednak fakt, że teoria ta łączyła w sobie dwie niezależne poprzednio teorie, wprowadzając tak zwane *oddziaływania elektroslabe*. Następnym logicznym krokiem jest tak zwana *wielka unifikacja* (grand unification) z teorią oddziaływań silnych. Na razie na tę teorię jeszcze czekamy.

Na zakończenie tej części wstępu trzeba jeszcze dodać parę słów na temat terminologii. Cząstki elementarne (nie czasteczki) dzieli się na trzy podstawowe grupy: te które podlegają silnym oddziaływaniom nazywa się *hadronami*, a te które w tych oddziaływaniach nie uczestniczą nazywa się *leptonami*; trzecia grupa stanowią kwanty pola odpowiedzialne za oddziaływania. Hadrony dzieli się na *bariony*, czyli te cząstki które mają spin połowkowy i *mezony* o spinie całkowitym.

Eksperymentalnie za pomocą metod fizyki wysokich energii (akceleratory, komory pęcherzykowe, emulsje fotograficzne) znaleziono już setki barionów i mezonów. Coraz trudniej więc można o nich mówić jako o cząstkach elementarnych. Znanych leptonów jest na razie pięć typów: e , ν_e , μ , ν_μ i τ . Ze względów estetycznych oczekuje się, że istnieje również neutrino ν_τ . Jako kwanty pola znane są fotony, słabe bozony W i Z, a problematyczne jest jak na razie istnienie gluonów.



Drugi nurt w KTP zajmuje się analiza możliwości powstawania struktur uporządkowanych w fizyce ośrodków skondensowanych. Podstawa

teorii jest słynne twierdzenie Goldstonea, które praktycznie sprowadza się do stwierdzenia, że istnienie stanów uporządkowanych jest możliwe dzięki kondensacji ciągłych (czyli mających ciągle widmo) bozonów Goldstonea. Fakt ten jest szaroko znany w fizyce ciała stałego, gdzie spotykamy się na przykład z fononami w kryształach lub z magnonami w ferromagnetykach. Charakter kondensacji bozonów jest związany z tak zwanym spontanicznym złamaniem symetrii. Tym zjawiskiem i wynikającymi z niego konsekwencjami zajmujemy się dokładnie przy innej okazji.

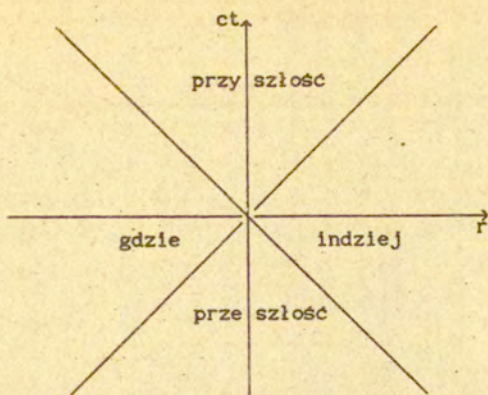
ROZDZIAŁ I

WPROWADZENIE

Dowolna teoria stawiająca sobie za cel opisanie podstawowej struktury materii powinna być zgodna z dwiema uznanymi teoriami, a mianowicie z teorią względności Einsteina i z mechaniką kwantową. Wprowadzimy teraz oznaczenia powszechnie w tych teoriach stosowane.

Rozpatrzmy dwa zdarzenia (ct, x, y, z) i $(ct+cdt, x+dz, y+dy, z+dz)$ w pewnym inercjalnym układzie odniesienia O w czasoprzestrzeni. Odległość pomiędzy tymi zdarzeniami oznaczmy przez ds . Aby ta odległość była taka sama dla wszystkich inercjalnych obserwatorów, powinna być ona niezmiennicza względem transformacji czasoprzestrzeni na przykład względem transformacji Lorentza lub szerszej transformacji Poincaré. Z sadania tego wynika, że powinna mieć ona postać

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = (cdt - dr)(cdt + dr). \quad (I.1)$$



Mozna by oczywiście przyjąć i inną postać, a mianowicie $ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2$, ale definicja (I.1) jest z różnych względów wygodniejsza. Wykres odpowiadający równaniu (I.1) w prostokątnym układzie współrzędnych ct i r nazywa się *stozkiem świetlnym*. Stozek świetlny ma znaczenie lokalne: odnosi się on do tego punktu czasoprzestrzeni, z jakiego został wystawiony. Tak więc pojęcia przyszłość, przeszłość i teraźniejszość mają sens lokalny.

Przy przyjętym oznaczeniu (I.1), interwały

dla których $ds^2 > 0$ nazywamy *czasopodobnymi*,

dla których $ds^2 < 0$ nazywamy *przestrzennopodobnymi*, a te

dla których $ds^2 = 0$ nazywamy *światopodobnymi*.

W zwykłej przestrzeni trójwymiarowej wielkość $dr^2 = dx^2 + dy^2 + dz^2$ jest niezmiennicza względem obrotów. Przy przejściu do czasoprzestrzeni, odcinek niezmienniczy przestaje być dodatnio określony. Wektor, którego długość (kwadrat skalarny) jest równy zeru nosi nazwę wektora *izotropowego*. Z (I.1) wynika, że składowe wektora izotropowego mogą być różne od zera. Wprowadzamy oznaczenia

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z), \quad (I.2)$$

$$x_\mu = (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z),$$

i korzystając z zasady sumowania Einsteina po jednym górnym i po

Jednym dolnym indeksie otrzymujemy, iż

$$ds^2 = \sum_{\mu=0}^3 dx^{\mu} dx_{\mu} = dx^{\mu} dx_{\mu} = c^2 dt^2 - dx^2 - dy^2 - dz^2. \quad (I.3)$$

Zdefiniowany powyżej 4-wektor x^{μ} z górnym indeksem nazywamy wektorem kontrawariantnym, a 4-wektor z dolnym indeksem nazywamy wektorem kowariantnym. Można znaleźć związek pomiędzy wektorami kontra- i kowariantnymi wprowadzając tensor metryczny $g_{\mu\nu}$, a mianowicie

$$x_{\mu} = g_{\mu\nu} x^{\nu} = g_{\mu 0} x^0 + g_{\mu 1} x^1 + g_{\mu 2} x^2 + g_{\mu 3} x^3, \quad (I.4)$$

Ze wzoru (I.2) widać, że $x_0 = x^0$, $x_1 = -x^1$ itd., więc tensor $g_{\mu\nu}$ można zapisać w postaci następującej macierzy kwadratowej

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (I.5)$$

Przestrzeń Riemanna z tak określonym tensorem metrycznym we współrzędnych kartezjańskich nazywa się *przestrzenią Minkowskiego*.

Wyznacznik macierzy (I.5) jest różny od zera, więc istnieje macierz odwrotna $g^{\mu\nu}$, czyli taka, że $g_{\mu\nu} g^{\mu\nu} = 1$ i ma ona postać

$$g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (I.6)$$

Cała informacja o geometrii czasoprzestrzeni jest zawarta w tensorze metrycznym. W specjalnej teorii względności Einsteina tensor metryczny odgrywa rolę pasywną i w zasadzie można go wcale nie wprowadzać. Inaczej jest w ogólnej teorii względności. Tam gra on rolę aktywną, ponieważ w tym przypadku geometria czasoprzestrzeni nie jest z góry określona. Wyznacza ją dopiero rozkład materii.

Wprowadźmy teraz operatory różniczkowe

$$\partial_{\mu} = \frac{\partial}{\partial x^{\mu}} = (\partial_0, \partial_1, \partial_2, \partial_3) = \left[\frac{1}{c} \frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right] = \left[\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right], \quad (1.7)$$

$$\partial^{\mu} = g^{\mu\nu} \partial_{\nu} = \left[\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right].$$

Ich iloczyn skalarny daje operator różniczkowy drugiego rzędu, niezmienniczy względem transformacji Lorentza

$$\square = \partial^{\mu} \partial_{\mu} = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) = \left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right], \quad (1.8)$$

który nosi nazwę operatora d'Alamberta (d'alambertian). Czterowektor energii-pędu cząstki ma postać

$$p^{\mu} = \left(\frac{E}{c}, \vec{p} \right), \quad p_{\mu} = \left(\frac{E}{c}, -\vec{p} \right), \quad (1.9)$$

skąd otrzymujemy niezmiennik

$$p^2 = p^{\mu} p_{\mu} = \frac{E^2}{c^2} - \vec{p} \cdot \vec{p} = m^2 c^2. \quad (1.10)$$

Będziemy również stosować oznaczenie $p \cdot x$ dla wielkości $p_{\mu} x^{\mu}$

$$p \cdot x = p_{\mu} x^{\mu} = Et - \vec{p} \cdot \vec{r}. \quad (1.11)$$

ROZDZIAŁ II

GRUPY PRZEKSZTAŁCEN CZASOPRZESTRZENI I SYMETRII WEWNĘTRZNYCH

Rozpatrzmy transformacje czasoprzestrzeni

$$x'^{\mu} = a^{\mu}_{\nu} x^{\nu} + d^{\mu}, \quad (II.1)$$

pozostawiająca niezmienniczą formę kwadratową (I.1)

$$ds^2 = g_{\mu\nu} dx^{\mu} dx^{\nu} = g_{\mu\nu} dx'^{\mu} dx'^{\nu}. \quad (II.2)$$

Wielkości x^{μ} i x'^{μ} są to współrzędne tego samego zdarzenia w dwu inercjalnych układach odniesienia O i O' . Niejednorodne

przekształcenia (II.1) tworzą grupę, która nazywa się *grupa Poincaré*. Przekształcenia jednorodne (czyli z $d^\mu = 0$) tworzą podgrupę grupy Poincaré, która nazywa się *pełna grupa Lorentza*. Przekształcenia Poincaré łączy w sobie dwa przekształcenia: 4-przemieszczenia określone parametrami d^μ i 4-obróty określone parametrami a^μ_ν . Pokażemy teraz, że te przekształcenia rzeczywiście tworzą grupę. Będziemy korzystali z możliwości podnoszenia i opuszczania wskaźników za pomocą tensora metrycznego, na przykład

$$a_{\mu\nu} = g_{\mu\lambda} a^\lambda_\nu, \quad a^{\mu\nu} = a^\mu_\lambda g^{\lambda\nu}. \quad (\text{II.3})$$

Z warunku niezmienniczości formy kwadratowej (II.2) wynika, że¹

$$a_{\lambda\mu} a^{\lambda\nu} = \delta^\nu_\mu, \quad a_{\mu\lambda} a^{\nu\lambda} = \delta^\nu_\mu. \quad (\text{II.4})$$

Przekształcenie (II.1) odpowiada przejściu z O do O' . Dokonajmy następnego przejścia z O' do O'' ,

$$x''^\rho = a'^\rho_\mu x'^\mu + d'^\mu. \quad (\text{II.5})$$

Wstawiając (II.1) otrzymujemy

$$\begin{aligned} x''^\rho &= a'^\rho_\mu (a^\mu_\nu x^\nu + d^\mu) + d'^\mu = \\ &= a'^\rho_\mu a^\mu_\nu x^\nu + a'^\rho_\mu d^\mu + d'^\rho. \end{aligned} \quad (\text{II.6})$$

Zamiast powyższego przejścia, możemy dokonać bezpośredniej transformacji z O do O'' ,

$$x''^\rho = a''^\rho_\nu x^\nu + d''^\rho. \quad (\text{II.7})$$

Tak więc

$$a''^\rho_\nu = a'^\rho_\mu a^\mu_\nu, \quad d''^\rho = a'^\rho_\mu d^\mu + d'^\rho. \quad (\text{II.8})$$

¹ Rzeczywiście,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = g_{\mu\nu} a^\mu_\lambda dx^\lambda a^\nu_\kappa dx^\kappa = g_{\lambda\kappa} dx^\lambda dx^\kappa.$$

Mamy więc

$$g_{\lambda\kappa} = g_{\mu\nu} a^\mu_\lambda a^\nu_\kappa = a_{\nu\lambda} a^{\nu\alpha}_\kappa = a_{\nu\lambda} a^{\nu\alpha} g_{\alpha\kappa},$$

skąd wynika, że

$$a_{\nu\lambda} a^{\nu\alpha} = \delta^\alpha_\lambda.$$

Łatwo można pokazać, że jeśli współczynniki a^μ_ν i a'^ρ_μ spełniają warunki (II.4), to będą je również spełniać i współczynniki a''^ρ_ν . Świadczy to o grupowym charakterze przekształceń Poincaré i Lorentza. Związek (II.8) są definicją działań w grupie Poincaré. Z warunków (II.4) wynika², że

$$\det(a^\mu_\nu) = \pm 1. \quad (\text{II.9})$$

Ponadto, z (II.4) wynika również³, że

$$(a^0_0)^2 = 1 + a^0_k a^0_k, \quad (\text{II.10})$$

gdzie $k = 1, 2, 3$, a więc

$$|a^0_0| \geq 1. \quad (\text{II.11})$$

Przekształcenia, dla których $a^0_0 \geq 1$ także tworzą grupę i nazywają się ortochronna grupa Lorentza.

Nieskończenie małe przekształcenia Poincaré rozpadają się na nieskończenie małe (infinitesimalne) translacje

$$x'^\mu = x^\mu + \epsilon^\mu, \quad (\text{II.12})$$

gdzie ϵ^μ jest 4-wektorem nieskończenie małych przemieszczeń, i infinitesimalne 4-obroty

² Rzeczywiście, korzystając z tożsamości $\det(AB) = \det(A) \det(B)$ i z faktu, że $\det(g_{\mu\nu}) = -1$, mamy

$$\det(a_{\rho\mu} a^{\rho\nu}) = \det(\delta^\nu_\mu) = 1,$$

$$\det(a_{\rho\mu}) \det(a^{\rho\nu}) = \det(a_{\rho\mu}) \det(g^{\lambda\rho} g^{k\nu} a_{\lambda k}) = \det^2(a_{\rho\nu}) = 1,$$

skąd wynika (II.9).

³ Z (II.4) mamy, że

$$\delta^\nu_\mu = a_{\mu\rho} a^{\nu\rho} = g_{\alpha\mu} g^{\beta\nu} a^\alpha_\rho a^\rho_\beta.$$

Kładąc $\mu = \nu = 0$ otrzymujemy

$$\begin{aligned} \delta^0_0 = 1 &= g_{\alpha 0} g^{\beta 0} a^\alpha_\rho a^\rho_\beta = a^0_\rho a^0_\rho = a^0_\rho g_{0\alpha} g^{\rho\beta} a^\alpha_\beta = g^{\rho\beta} a^0_\rho a^0_\beta = \\ &= a^0_0 a^0_0 + g^{k1} a^0_k a^0_1 = (a^0_0)^2 - \sum_{k=1}^3 a^0_k a^0_k, \end{aligned}$$

skąd już bezpośrednio wynika (II.10).

$$x'^{\mu} = x^{\mu} + \epsilon^{\mu}_{\nu} x^{\nu}, \quad (\text{II.13})$$

gdzie $\epsilon_{\mu\nu} = g_{\mu\lambda} \epsilon^{\lambda}_{\nu}$ jest nieskończenie mała antysymetryczna macierza $\epsilon_{\mu\nu} = -\epsilon_{\nu\mu}$.

Z pełnej grupy Lorentza można wydzielić podgrupę, nazywaną właściwą grupą Lorentza. W podgrupie tej, dowolne przekształcenie można otrzymać z przekształcenia tożsamościowego zmieniając w sposób ciągły elementy macierzy $a_{\mu\nu}$, a więc poprzez ciąg nieskończenie małych przekształceń Lorentza. Ponieważ $\det(a_{\mu\nu})$ przyjmuje tylko dwie dyskretne wartości (zgodnie z (II.9)), to dla dowolnego przekształcenia właściwej grupy Lorentza mamy

$$\det(a_{\mu\nu}) = 1 \quad a^0_0 \geq 1. \quad (\text{II.14})$$

Rozważmy teraz, jak wraz ze zmianą układu odniesienia zmienia się funkcja pola $\psi_1(x)$. Zadając, by związek pomiędzy składowymi tej samej funkcji pola, ale w różnych układach odniesienia ($\psi(x)$ w O i $\psi'(x')$ w O') był liniowy i jednorodny, możemy napisać

$$\psi'_1(x') = S_1^j (a^{\mu}_{\nu}) \psi_j(x), \quad (\text{II.15})$$

lub w skrócie $\psi'(x') = S(a)\psi(x)$. $S(a)$ jest pewną macierzą, zależną od współczynników a^{μ}_{ν} przekształcenia Lorentza. Przekształcenia (II.15) tworzą grupę.⁴ Grupa ta nosi nazwę *reprezentacji grupy Lorentza*. Tak więc, funkcja pola przy przejściu od jednego układu odniesienia do innego zmienia się zgodnie z pewną reprezentacją grupy Lorentza. Pojawia się możliwość klasyfikacji pól według odpowiadających im reprezentacji grupy Lorentza.

Przy infinitezymalnych przekształceniach właściwej grupy Lorentza macierz $S(a)$ można rozłożyć w szereg potęgowy względem nieskończenie małych parametrów $\epsilon_{\mu\nu}$ i pozostawić tylko człony liniowe

⁴Przy przejściu od układu O' do O'' mamy $\psi''(x'') = S(a')\psi(x')$, a przy przejściu od O do O'' $\psi''(x'') = S(a)\psi(x)$. Wykorzystując (II.15), a następnie porównując otrzymane wyniki stwierdzamy, że $S(a'') = S(a')S(a)$, gdzie $a'' = a'a$.

$$S(a) = 1 - \frac{1}{2} \epsilon^{\mu\nu} \Sigma_{\mu\nu}, \quad (\text{II.16})$$

gdzie $\Sigma_{\mu\nu} = -\Sigma_{\nu\mu}$ są pewnymi macierzami, nie zależącymi od $\epsilon_{\mu\nu}$ i nazywanymi *infinitesimalnymi operatorami*. Operatory te definiują reprezentacje właściwej grupy Lorentza.

Funkcje pola mogą podlegać również pewnym transformacjom nie związanym z transformacjami czasoprzestrzeni i równocześnie pozostawiającym bez zmian lagrangian

$$\psi_i(x) \rightarrow \psi'_i(x) = U_i^j(\lambda) \psi_j(x), \quad (\text{II.17})$$

Mozna łatwo pokazać, że te transformacje również tworzą grupę⁵, którą będziemy nazywać grupą symetrii wewnętrznych. Macierz U może zależeć od całego zbioru ciągłych parametrów λ^a , $a = 1, 2, \dots, m$. Zakładamy, że odwzorowaniu tożsamościowemu $U = I$ odpowiadają zerowe wartości parametrów $\lambda = 0$, a odwzorowaniu odwrotnemu $U^{-1}(\lambda)$, pewne inne wartości parametrów λ , które oznaczymy jako λ^{-1} . Mamy więc

$$U(0) = I, \quad U^{-1}(\lambda) = U(\lambda^{-1}). \quad (\text{II.18})$$

Z uwagi⁵ i z (II.18) wynika, że

$$\varphi(\lambda, \lambda^{-1}) = 0, \quad \varphi(\lambda, 0) = \varphi(0, \lambda) = \lambda. \quad (\text{II.19})$$

Łączność prawa mnożenia nakłada ponadto warunek, aby

$$\varphi(\lambda_1, \varphi(\lambda_2, \lambda_3)) = \varphi(\varphi(\lambda_1, \lambda_2), \lambda_3). \quad (\text{II.20})$$

Widzimy więc, że wzór

$$\lambda_3 = \varphi(\lambda_2, \lambda_1), \quad (\text{II.21})$$

jest prawem mnożenia parametrów grupy. Grupy, których elementy zależą od pewnego zbioru ciągłych parametrów, noszą nazwę grup *Lie*.

⁵ Jeśli lagrangian jest niezmienniczy względem przekształceń (II.17) z macierzami $U(\lambda_1)$ i $U(\lambda_2)$, to będzie on również niezmienniczy względem transformacji z macierzą $U(\lambda_2)U(\lambda_1)$. Można więc uważać, że istnieją takie parametry λ , dla których $U(\lambda_2)U(\lambda_1) = U(\lambda_3)$, przy czym $\lambda_3 = \varphi(\lambda_2, \lambda_1)$.

ROZDZIAŁ III

PODSTAWY FORMALIZMU, LAGRANGOWSKIEGO

W rozdziale tym chcemy rozpocząć rozważania dotyczące ogólnych własności pól fizycznych. Rozważania będą dotyczyły pól klasycznych, nie będziemy więc zwracali uwagi na ich możliwe własności kwantowe. Różnorakie cechy pól będziemy wyprowadzać z jednej ogólnej zasady wariacyjnej Hamiltona, bazującej na istnieniu jednej rzeczywistej funkcji tzw. *gęstości funkcji Lagrangea* \mathcal{L} . Będziemy zakładali często, że \mathcal{L} zależy od współrzędnych $x^\mu = (t, x^k)$ tylko pośrednio, poprzez funkcje pola $\psi_i(x^\mu)$ i ich pierwsze pochodne $\partial_\mu \psi_i(x^\nu)$, czyli że

$$\mathcal{L} = \mathcal{L}(\psi_i(x), \partial_\mu \psi_i(x)), \quad i = 1, \dots, N. \quad (\text{III.1})$$

Z gęstości funkcji Lagrangea, budujemy *funkcję Lagrangea (lagrangian)*

$$L(t) = \int_{\Omega(t)} d^3x \mathcal{L}(\psi_i(x^\nu), \partial_\mu \psi_i(x^\nu)), \quad (\text{III.2})$$

i działanie

$$S = \int_1^2 dt L(t) = \int_{\Omega} d^4x \mathcal{L}(\psi_i, \partial_\mu \psi_i), \quad (\text{III.3})$$

będące całką po 4-objętości Ω ograniczonej dwiema przestrzennopodobnymi hiperpowierzchniami σ_1 i σ_2 . Widać, że działanie jest funkcjonałem zależnym od funkcji pola ψ_i i od powierzchni σ_1 i σ_2 .

Lagrangianem będziemy często nazywali również gęstość funkcji Lagrangea, mając nadzieję, że nie doprowadzi to do nieporozumień.

Zasada stacjonarnego działania Hamiltona głosi, że całkowita informacja o układzie fizycznym jest zawarta w gęstości funkcji Lagrangea $\mathcal{L}(\psi, \partial_\mu \psi)$ i że rzeczywiste procesy fizyczne przebiegają w taki sposób, że działanie osiąga ekstremum. Fakt ten wyraża się matematycznie poprzez zadanie, aby zniknęła pierwsza wariacja działania

$$\delta S = 0, \quad (\text{III.4})$$

dla dowolnych, swobodnych wariacji pola $\psi_1(x)$, ale takich, że $\delta\psi_1$ znikają na granicach obszaru całkowania. Obszar ten w ogólności bierze udział w procesie fizycznym i może ulegać zmianom (wtedy \mathcal{L} zależy w sposób jawny od x). Tak więc wariacjom będzie podlegało zarówno pole ψ_1 jak i obszar całkowania. W związku z tym wariację działania rozbijemy na dwie części

$$\delta S = \delta_\psi S + \delta_\sigma S, \quad (\text{III.5})$$

Mamy więc

$$\delta_\psi S = \int_\Omega d^4x \left[\frac{\partial \mathcal{L}}{\partial \psi_1} \delta\psi_1 + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} \delta \partial_\mu \psi_1 \right], \quad (\text{III.6})$$

Ponieważ $\delta(\partial_\mu \psi_1) = \partial_\mu(\delta\psi_1)$ i ponadto zachodzi tożsamość

$$\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} \partial_\mu(\delta\psi_1) = \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} \delta\psi_1 \right] - \delta\psi_1 \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)}, \quad (\text{III.7})$$

otrzymujemy, iż

$$\delta_\psi S = \int_\Omega d^4x \left[\frac{\partial \mathcal{L}}{\partial \psi_1} - \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} \partial_\mu \right] \delta\psi_1 + \int_\Omega d^4x \partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} \delta\psi_1 \right]. \quad (\text{III.8})$$

Wyrażenie podcałkowe w drugiej całce w (III.8) jest pełną dywergencją, możemy więc skorzystać z twierdzenia Gaussa

$$\int_\Omega d^4x \partial_\mu G^\mu = \int_{\sigma_2} d\sigma_\mu G^\mu(x) - \int_{\sigma_1} d\sigma_\mu G^\mu(x), \quad (\text{III.9})$$

i zamienić całkę po hiperobjętości Ω na całkę po powierzchniach ograniczających obszar Ω

$$\delta_\psi S = \int_\Omega d^4x \delta\psi_1 \left[\frac{\partial \mathcal{L}}{\partial \psi_1} - \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} \partial_\mu \right] + \int_{\sigma_2} d\sigma_\mu \delta\psi_1 \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)} - \int_{\sigma_1} d\sigma_\mu \delta\psi_1 \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi_1)}. \quad (\text{III.10})$$

Przez $d\sigma_\mu = n_\mu d\sigma$ oznaczyliśmy zorientowany element hiperpowierzchni, n_μ jest czsopodobnym wektorem jednostkowym normalnym do powierzchni σ_μ w punkcie x i leżącym w górnej części stożka świetlnego o wierzchołku

w punkcie x . W wyniku wariacji powierzchni, σ_1 i σ_2 przechodzą w niewiele od nich różniące się powierzchnie σ'_1 i σ'_2 . W tym przypadku wariacja $\delta_{\sigma} S$ będzie równa całce z \mathcal{L} po czteroobjętości zawartej pomiędzy powierzchniami σ'_1 i σ_1 oraz σ'_2 i σ_2 . Elementy tych objętości wynoszą $d\sigma_{\mu} \delta x^{\mu}(x)$, gdzie $\delta x^{\mu}(x)$ jest wariacją współrzędnych brzegu objętości Ω . Tak więc

$$\delta_{\sigma} S = \int_{\sigma_2} d\sigma_{\mu} \delta x^{\mu} \mathcal{L} - \int_{\sigma_1} d\sigma_{\mu} \delta x^{\mu} \mathcal{L}. \quad (\text{III.11})$$

Ostatecznie otrzymujemy, iż wariacja działania wynosi

$$\delta S = \int_{\Omega} d^4 x \delta\psi_1 \left[\frac{\partial \mathcal{L}}{\partial \psi_1} - \partial_{\mu} \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \psi_1)} \right] + G(\sigma_2) - G(\sigma_1), \quad (\text{III.12})$$

gdzie

$$G(\sigma_1) = \int_{\sigma_1} d\sigma_{\mu} \left[\frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \psi_1)} \delta\psi_1 + \mathcal{L} \delta x^{\mu} \right]. \quad (\text{III.13})$$

Z zasady stacjonarnego działania, przy dodatkowym warunku że granice obszaru całkowania Ω nie podlegają wariacji ($\delta x^{\mu} = 0$), i że wzoru (III.12) wynikają tzw. równania Eulera-Lagranga

$$\frac{\partial \mathcal{L}}{\partial \psi_1} - \partial_{\mu} \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \psi_1)} = 0. \quad (\text{III.14})$$

Są to równania ruchu dla pól. Równania te przybierają konkretną postać dopiero po ustaleniu konkretnej postaci lagrangianu \mathcal{L} .

Z przeprowadzonych rozważań wynika, że jeśli w lagrangianie da się wydzielić jego część będącą pełną dywergencją, to ta część da się w działaniu sprowadzić do całki powierzchniowej i zniknie na podstawie warunków brzegowych, nie dając wkładu do równań Eulera-Lagranga. Widać więc, że lagrangian nie jest zdefiniowany jednoznacznie, gdyż zawsze można do niego dodać pełną dywergencję

$$\mathcal{L}(x) \longrightarrow \mathcal{L}(x) + \partial_{\mu} F^{\mu}(x), \quad (\text{III.15})$$

nie zmieniając równań ruchu. Fakt ten czasami wykorzystuje się, zakładając, że F^{μ} jest pewną funkcją zmiennych polowych

$$F^{\mu}(x) = F_1^{\mu}(\psi), \quad (\text{III.16})$$

gdz człony te mogą mieć wpływ na wynikające z lagrangianu prawa zachowania.

Zakładamy, że wszystkie inercjalne układy odniesienia są równoważne. W związku z tym żądamy, by funkcja Lagrange'a była niezmiennicza względem grupy Poincaré, czyli by zachodziła równość

$$\mathcal{L}(\psi(x), \partial_\mu \psi(x)) = \mathcal{L}(\psi'(x'), \partial'_\mu \psi'(x')) , \quad (\text{III.17})$$

gdzie $\psi'(x') = S(\lambda)\psi(x)$ i $x' = ax + d$. Z niezmienniczości funkcji Lagrange'a wynika niezmienniczość działania

$$S(\psi(x); \sigma_1, \sigma_2) = S(\psi'(x'); \sigma'_1, \sigma'_2) , \quad (\text{III.18})$$

gdzie $\psi'(x) = S(\lambda)\psi(a^{-1}x - a^{-1}d)$, a hiperpowierzchnie σ'_1 i σ'_2 otrzymuje się z σ_1 i σ_2 w rezultacie transformacji Poincaré. Ponieważ lagrangian pola jest relatywistycznym skalarem, to równania pola dla $\psi(x)$ i $\psi'(x')$ będą miały identyczną postać. Mówiąc inaczej, równania pola są niezmiennicze względem transformacji Poincaré.

Słynne twierdzenie Noether głosi, że z każdą jednoparametrową grupą symetrii związane jest pewne prawo zachowania. Nie chcąc tu teraz wchodzić w dosyć znużone i szczegółowe rachunki, wypiszemy tylko wyniki końcowe.

Z grupą Poincaré związane jest prawo zachowania energii-pędu i gęstości momentu pędu:

$$\partial_\nu T^{\mu\nu}(x) = 0 , \quad (\text{III.19})$$

$$\partial_\rho M^{\mu\nu;\rho}(x) = 0 , \quad (\text{III.20})$$

gdzie w przypadku pola skalarnego

$$T^{\mu\nu} = \partial^\mu \psi \frac{\partial \mathcal{L}}{\partial (\partial_\nu \psi)} - \mathcal{L} g^{\mu\nu} , \quad (\text{III.21})$$

$$M^{\mu\nu;\rho} = x^\mu T^{\nu\rho} - x^\nu T^{\mu\rho} . \quad (\text{III.22})$$

Dla symetrii wewnętrznych mamy

$$\partial_\mu J_a^\mu = 0 , \quad (\text{III.23})$$

gdzie

$$J_a^\mu = -i \frac{\partial \mathcal{L}}{\partial (\partial_\nu \psi)} T_a \psi , \quad (\text{III.24})$$

nazywają się uogólnionymi 4-prądami związanymi z symetrią wewnętrzną, a T_a są generatorami grupy symetrii wewnętrznej.

ROZDZIAŁ IV

POLE SKALARNE

Przegląd konkretnych pól fizycznych rozpoczynamy od pola skalarnego. Jest to pole opisywane przez jedną zespoloną funkcję skalarną $\varphi(x)$, transformująca się przy właściwych przekształceniach Lorentza $x \rightarrow x' = ax$ zgodnie ze wzorem

$$\varphi(x) \rightarrow \varphi'(x') = \varphi(x) \quad (\text{IV.1})$$

Zespolone pole skalarne opisuje cząstki ze spinem zero, ale obdarzone ładunkiem. Lagrangian dla tego pola, niezmienniczy względem przekształcenia (IV.1) ma postać

$$\mathcal{L} = \partial_{\mu} \varphi \partial^{\mu} \varphi^{*} - m^2 c^2 \varphi \varphi^{*} \quad (\text{IV.2})$$

Równanie Eulera-Lagranga przyjmuje następującą postać:

- jeśli położymy $\psi_1 = \varphi$, to

$$\frac{\partial \mathcal{L}}{\partial \varphi} = -m^2 c^2 \varphi^{*}, \quad \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi)} = \partial^{\mu} \varphi^{*}, \quad \text{więc otrzymujemy}$$

$$\partial^{\mu} \partial_{\mu} \varphi^{*} + m^2 c^2 \varphi^{*} = 0 \quad (\text{IV.3})_1$$

- jeśli położymy $\psi_2 = \varphi^{*}$, to

$$\frac{\partial \mathcal{L}}{\partial \varphi^{*}} = -m^2 c^2 \varphi, \quad \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} \varphi^{*})} = \partial_{\mu} \varphi \quad \text{i otrzymujemy}$$

$$\partial^{\mu} \partial_{\mu} \varphi + m^2 c^2 \varphi = 0 \quad (\text{IV.3})_2$$

Równania (IV.3) noszą nazwę równań Kleina-Gordona.

Parametr m interpretujemy jako masę cząstki swobodnej. Wynika to z faktu, że jeśli rozwiązania równania (IV.3) przedstawimy w postaci fali płaskiej

$$\varphi(x) = \varphi_0 e^{-i p x}, \quad (\text{IV.4})$$

gdzie $p x = E t - \vec{p} \cdot \vec{x}$, to otrzymujemy następujący związek dyspersyjny

$$\frac{E^2}{c^2} = \vec{p}^2 + m^2 c^2. \quad (\text{IV.5})$$

Na równanie Kleina-Gordona można spojrzeć jeszcze z innego punktu widzenia. Jeśli przyjmiemy, że 4-wektor energii-pędu ma postać

$$p^\mu = \left[\frac{E}{c}, p \right], \quad p_\mu = \left[\frac{E}{c}, -p \right], \quad (\text{IV.6})$$

to iloczyn $p^\mu p_\mu$ będzie niezmiennikiem

$$p^2 = p^\mu p_\mu = \frac{E^2}{c^2} - p p = m^2 c^2. \quad (\text{IV.7})$$

Jest to prawo zachowania energii w teorii względności Einsteina. Jeśli postąpimy teraz tak jak w mechanice kwantowej, czyli przyporządkujemy energii E -i pędowi p odpowiednie operatory różniczkowe

$$E \longrightarrow i\hbar \frac{\partial}{\partial t}, \quad p \longrightarrow -i\hbar \nabla, \quad (\text{IV.8})$$

to powyższy związek odtwarza nam równanie Kleina-Gordona

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right] \varphi + \frac{m^2 c^2}{\hbar^2} \varphi = 0, \quad (\text{IV.9})$$

lub w układzie jednostek, w którym $\hbar = c = 1$ z uwzględnieniem postaci d'Alambertianu \square

$$(\square + m^2) \varphi = 0. \quad (\text{IV.10})$$

Jeżeli zamiast wzoru (IV.5) wykorzystamy przybliżenie nierelatywistyczne

$$E = p^2/2m, \quad (\text{IV.11})$$

i dokonamy odpowiedniego podstawienia operatorów różniczkowych, to otrzymamy równanie Schrödingera dla cząstki swobodnej

$$-i\hbar \frac{\partial \varphi}{\partial t} = \frac{\hbar^2}{2m} \nabla^2 \varphi. \quad (\text{IV.12})$$

Wniosujemy stąd, że równanie Schrödingera jest nierelatywistycznym przybliżeniem równania Kleina-Gordona.

Zwróćmy teraz uwagę na problemy interpretacyjne związane z równaniem Kleina-Gordona. W równaniu Schrödingera mamy następujące wyrażenie dla gęstości prawdopodobieństwa

$$\rho = \varphi \varphi^* , \quad (IV.13)$$

i dla prądu prawdopodobieństwa

$$j = -\frac{i\hbar}{2m} (\varphi^* \nabla \varphi - \varphi \nabla \varphi^*) . \quad (IV.14)$$

Wielkości te spełniają równanie ciągłości

$$\frac{\partial \rho}{\partial t} + \nabla j = \frac{\partial}{\partial t} (\varphi \varphi^*) - \frac{i\hbar}{2m} (\varphi^* \nabla^2 \varphi - \varphi \nabla^2 \varphi^*) = \quad (IV.15)$$

$$\varphi^* \left(\frac{\partial \varphi}{\partial t} - \frac{i\hbar}{2m} \nabla^2 \varphi \right) + \varphi \left(\frac{\partial \varphi^*}{\partial t} + \frac{i\hbar}{2m} \nabla^2 \varphi^* \right) = 0 .$$

Ostatnie dwa człony, to równanie Schrödingera i równanie sprzężone do niego.

W przypadku równania Kleina-Gordona, ρ nie jest skalarem, ale transformuje się jak składowa czasowa 4 - wektora, którego składowe przestrzenne dane są wzorem (IV.14). Mamy więc

$$\rho = \frac{i\hbar}{2m} (\varphi^* \frac{\partial \varphi}{\partial t} - \varphi \frac{\partial \varphi^*}{\partial t}) . \quad (IV.16)$$

Wykorzystując wyrażenie

$$j^\mu = (\rho, j) = \frac{i\hbar}{m} \varphi^* (\overset{\rightarrow}{\partial}_\sigma, -\overset{\rightarrow}{\nabla}) \varphi = \frac{i\hbar}{m} \varphi^* \overset{\rightarrow}{\partial}_\mu \varphi , \quad (IV.17)$$

gdzie zgodnie z definicją

$$A \overset{\rightarrow}{\partial}{}^\mu B = \frac{1}{2} \left[A \partial^\mu B - (\partial^\mu A) B \right] , \quad (IV.18)$$

otrzymujemy równanie ciągłości

$$\partial_\mu j^\mu = \frac{i\hbar}{2m} (\varphi^* \square \varphi - \varphi \square \varphi^*) = 0 . \quad (IV.19)$$

Jak wiemy, φ^* też spełnia równanie Kleina-Gordona. Tak więc ρ i j jawią się jako poszukiwana gęstość i prąd prawdopodobieństwa. Jednakże taka interpretacja od razu napotyka się na trudność, gdyż w odróżnieniu od ρ z równania Schrödingera (IV.13), wyrażenie dla ρ z równania Kleina-Gordona (IV.16) nie jest wielkością dodatnio

określona. Równanie Kleina-Gordona jest równaniem drugiego rzędu, więc wielkości φ i $\partial\varphi/\partial t$ mogą być w danym momencie tak dobrane, by wielkość ρ była ujemna. Tym samym interpretacja ρ jako gęstości prawdopodobieństwa jest niemożliwa. Niemożliwa jest również interpretacja równania Kleina-Gordona jako równania jednocząstkowego z funkcją falową φ . Dopiero w wyniku kwantowania, możliwa staje się interpretacja wielocząstkowa.

Zwróćmy ponadto uwagę, że funkcja φ przyjmuje również wartości zespolone. Jeśli traktować φ jako funkcję rzeczywistą, to wielkość ρ dana wzorem (IV.16) jak i prąd j z (IV.14) zerują się. Można natomiast pokazać, że φ da się interpretować jako funkcja opisująca cząstki naładowane. Rzeczywiste funkcje φ odpowiadają cząstkom neutralnym. W tym przypadku ρ i j są gęstościami ładunku i prądu, a nie prawdopodobieństwa.

Pojawia się jeszcze jeden problem związany z równaniem Kleina-Gordona. Chodzi mianowicie o równanie (IV.5), które jeśli traktowane jako równanie na energię E , przyjmuje postać

$$E = \pm \sqrt{m^2 c^4 + p^2 c^2} \quad (IV.20)$$

Widać stąd, że równanie Kleina-Gordona oprócz rozwiązań z dodatnimi energiami, ma również rozwiązania odpowiadające ujemnym energiom. W przypadku cząstki swobodnej, dla której energia jest stała, trudność tę można przezwyciężyć, ponieważ tylko od nas zależy, jaką energię uważamy za dodatnią, to stany o energii ujemnej możemy odrzucić. Gorzej jest w przypadku cząstek oddziałujących. Cząstka oddziałująca z innymi cząstkami może wymieniać się energią z otoczeniem, a nie ma niczego, co by zabroniło jej 'z walić się' w stan o nieskończonej energii ujemnej i wypromieniować przy okazji nieskończoną energię. To w naturze oczywiście nie zachodzi, więc równanie Kleina-Gordona napotyka się na istotne trudności. Ten kłopot interpretacyjny można pokonać, jeśli potraktować φ jako pole kwantowe.

Równanie Kleina-Gordona można sprowadzić do równania rzędu

pierwszego, jeśli oprócz pól φ i φ^* wprowadzić cztery wektory φ_μ i φ_μ^*

$$\varphi_\mu = \partial_\mu \varphi, \quad \varphi_\mu^* = \partial_\mu \varphi^*. \quad (\text{IV.21})$$

W tym przypadku równanie $(\partial_\mu \partial^\mu + m^2 c^2) \varphi = 0$ zostanie zastąpione przez układ równań

$$\partial^\mu \varphi_\mu + m^2 c^2 \varphi = 0 \quad (\text{IV.22})$$

$$\partial_\mu \varphi - \varphi_\mu = 0$$

Równaniom tym odpowiada Lagrangian

$$\mathcal{L} = \varphi_\mu \partial^\mu \varphi^* + \varphi_\mu^* \partial^\mu \varphi - \varphi_\mu^* \varphi^\mu - m^2 c^2 \varphi \varphi^*, \quad (\text{IV.23})$$

niezmienniczy względem transformacji Lorentza

$$\varphi(x) \longrightarrow \varphi'(x') = \varphi(x), \quad \varphi_\mu(x) \longrightarrow \varphi'_\mu(x') = a_\mu^{\nu} \varphi_\nu(x),$$

i transformacji fazowych

$$\varphi(x) \longrightarrow \varphi'(x) = e^{i\alpha} \varphi(x).$$

Mozna również chcieć operować tylko polami rzeczywistymi. W tym celu wystarczy położyć $\varphi = \varphi^1 + i\varphi^2$, $\varphi_\mu = \varphi_\mu^1 + i\varphi_\mu^2$ i traktować pola φ^1 i φ^2 jako pola rzeczywiste.

Tensor energii-pędu i wektor prądu odpowiadające lagrangianowi (IV.23) mają następującą postać

$$T^{\mu\nu} = (m^2 c^2 \varphi \varphi^* - \varphi_\lambda \varphi^{*\lambda}) g^{\mu\nu} + \partial^\mu \varphi^* \varphi^\nu + \partial^\mu \varphi \varphi^{*\nu}, \quad (\text{IV.24})$$

$$J^\mu = i(\varphi \partial^\mu \varphi^* - \varphi^* \partial^\mu \varphi).$$

Zauważmy, że dla rzeczywistego pola skalarnego prąd się zeruje.

Układ równań (IV.22) można zapisać również w postaci

$$(\beta_\mu \partial^\mu + mc) \Phi(x) = 0, \quad (\text{IV.25})$$

gdzie $\Phi(x)$ jest pięcioskładnikową funkcją pola

$$\Phi(x) = \begin{bmatrix} \sqrt{mc} \varphi(x) \\ \frac{1}{\sqrt{mc}} \varphi_\nu(x) \end{bmatrix}, \quad (\text{IV.26})$$

a β_μ są czterema pięciowierszowymi macierzami spełniającymi relacje

$$\beta_\mu \beta_\nu \beta_\rho + \beta_\rho \beta_\nu \beta_\mu = \varphi_{\mu\nu} \beta_\rho + \varphi_{\rho\nu} \beta_\mu. \quad (\text{IV.27})$$

ROZDZIAŁ V

POLE WEKTOROWE

Kolejnym polem jeśli chodzi o złożoność, jest wektorowe pole zespolone, opisane przez 4-wektor $\varphi_\mu(x)$, transformujące się względem właściwych transformacji Lorentza zgodnie z prawem

$$\varphi_\mu(x) \rightarrow \varphi'_\mu(x') = a_\mu^\nu \varphi_\nu(x). \quad (\text{V.1})$$

Pola $\varphi_\mu(x)$ i $\varphi_\mu^*(x)$ spełniają równania

$$(\partial_\mu \partial^\mu + m^2 c^2) \varphi_\nu = 0, \quad (\partial_\mu \partial^\mu + m^2 c^2) \varphi_\nu^* = 0, \quad (\text{V.2})$$

gdzie m jest masą cząstek związanych z polem. Równaniu temu czyni zadość również pewne pole skalarne $\varphi(x) = \partial^\mu \varphi_\mu$. Aby wykluczyć istnienie takiego pola nakładamy na pole φ_μ dodatkowy warunek (warunek Lorentza)

$$\partial^\mu \varphi_\mu(x) = 0. \quad (\text{V.3})$$

Zobaczymy, że $\varphi_\mu(x)$ z uwzględnieniem tego dodatkowego warunku opisuje cząstki ze spinem 1.

Od równań pola rzędu drugiego można przejść do równań pola rzędu pierwszego, jeśli wprowadzi się tensor antisymetryczny

$$\varphi_{\mu\nu} = \partial_\mu \varphi_\nu - \partial_\nu \varphi_\mu. \quad (\text{V.4})$$

W tym przypadku równania (V.2) przyjmą postać

$$\partial^\mu \varphi_{\mu\nu} - m^2 c^2 = 0,$$

lub równowaznie

(V.5)

$$\partial^\nu \varphi_{\mu\nu} - m^2 c^2 = 0,$$

i wraz z równaniem (V.4) tworzą układ równań pola. Równania te nazywają się równaniami Proca. Równaniom Proca odpowiada Lagrangian

$$\mathcal{L} = \frac{1}{2} \varphi^{*\mu\nu} (\partial_\nu \varphi_\mu - \partial_\mu \varphi_\nu) + \frac{1}{2} \varphi^{\mu\nu} (\partial_\nu \varphi_\mu^* - \partial_\mu \varphi_\nu^*) + \frac{1}{2} \varphi_{\mu\nu}^* \varphi^{\mu\nu} + m^2 c^2 \varphi_\mu^* \varphi^\mu, \quad (V.6)$$

niezmienniczy względem transformacji Lorentza

$$\varphi_\mu(x) \rightarrow \varphi'_\mu(x') = a_\mu^\nu \varphi_\nu(x),$$

$$\varphi_{\mu\nu}(x) \rightarrow \varphi'_{\mu\nu}(x') = a_\mu^\lambda a_\nu^\rho \varphi_{\lambda\rho}(x),$$

i względem transformacji fazy

$$\varphi_\mu(x) \rightarrow \varphi'_\mu(x) = e^{i\alpha} \varphi_\mu(x),$$

$$\varphi_{\mu\nu}(x) \rightarrow \varphi'_{\mu\nu}(x) = e^{i\alpha} \varphi_{\mu\nu}(x).$$

Tensor energii-pędu i wektor prądu, odpowiadający lagrangianowi (V.6) mają następującą postać

$$T^{\mu\nu} = - (m^2 \varphi_\lambda^* \varphi^\lambda - \frac{1}{2} \varphi_{\lambda\rho}^* \varphi^{\lambda\rho}) g^{\mu\nu} + \partial^\mu \varphi^\rho \varphi_\rho^{*\nu} + \partial^\mu \varphi^{*\rho} \varphi_\rho^\nu, \quad (V.7)$$

$$J^\mu = - i (\varphi^{*\rho} \varphi_\rho^\mu - \varphi^\rho \varphi_\rho^{*\mu}).$$

Rozwiązania układu równań (V.5), podobnie jak dla pola skalarnego, można poszukiwać w postaci fal płaskich

$$\varphi_\mu(x) = \frac{1}{\sqrt{V}} \sum_{p, \lambda=1}^3 \frac{1}{\sqrt{2p_0}} \left[a_{p\lambda} e^{i(\lambda)} e^{-ipx} + b_{p\lambda}^* e^{i(\lambda)} e^{-px} \right], \quad (V.8)$$

gdzie $px = p_0 t - \vec{p} \cdot \vec{x}$, $p_0 = +\sqrt{\vec{p}^2 + m^2}$, $c = 1$, $V = L_1 L_2 L_3$ - unormowana objętość wprowadzona w tym celu, aby można było rozważać dyskretne wektory falowe \vec{p} , $p_1 = \frac{2\pi n_1}{L_1}$, n_1 - liczby;

wprowadzenie czynnika $1/\sqrt{2p_0}$ ważne jest w teorii kwantowej; $e_{p\mu}^{(\lambda)}$ -
jednocząstkowe wektory polaryzacji, a $a_{p\lambda}$ i $b_{p\lambda}^*$ - współczynniki
rozkładu.

Warunki unormowania wektorów polaryzacji nie pozwalają na ich jednoznaczne określenie. Aby tę jednoznaczność osiągnąć, należy wprowadzić macierze spinu pola wektorowego. W tym celu rozpatrzmy prawo przekształcenia $\varphi_\mu(x)$ przy nieskończonej małych transformacjach Lorentza. Przekształcenie 4-wektora

$$\varphi_\mu(x) \rightarrow \varphi'_\mu(x') = \alpha_\mu^\nu \varphi_\nu(x),$$

w przypadku nieskończenie małych transformacji Lorentza

$$x'^\mu = x^\mu + \epsilon^\mu_\nu x^\nu,$$

ma postać

$$\varphi_\mu(x) \rightarrow \varphi'_\mu(x') = \varphi_\mu(x) - \frac{i}{2} \epsilon^{\lambda\nu} (\Sigma_{\lambda\nu})_\mu^\rho \varphi_\rho(x),$$

gdzie sześć czterowierszowych macierzy $\Sigma_{\lambda\nu} = -\Sigma_{\nu\lambda}$ ma elementy macierzowe

$$(\Sigma_{\lambda\nu})_\mu^\rho = i(g_{\lambda\mu} \delta_\nu^\rho - g_{\nu\mu} \delta_\lambda^\rho) \quad (V.9)$$

Macierze spinu Σ^i pola wektorowego są związane z macierzmi $\Sigma_{\mu\nu}$ następującym związkiem

$$\Sigma^i = \frac{1}{2} \epsilon^{ikl} \Sigma_{kl} \quad (V.10)$$

Z (V.9) wynika, że

$$\Sigma^1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad \Sigma^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad \Sigma^3 = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

(V.11)

Mozna sprawdzić, że macierze (operatory) spinu spełniają relacje komutacyjne dla operatora momentu pędu

$$[\Sigma^i, \Sigma^k] = -i\epsilon^{ikl} \Sigma_l, \quad \Sigma_l = -\Sigma^l \quad (V.12)$$

i ze jego kwadrat skalarny jest niezmiennikiem grupy, czyli komutuje ze wszystkimi generatorami

$$\Sigma^1 \Sigma^1 = s(s+1), \quad (V.13)$$

gdzie s jest spinem cząstki. Ponieważ

$$\Sigma^1 \Sigma^1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (V.14)$$

to na podstawie (V.13) wnioskujemy, że spin cząstek związanych z polem jest równy 1 lub 0. Spin zerowy jest wykluczony przez odpowiednie warunki narzucone na wektor polaryzacji

$$e_{p\mu}^{(\lambda)} p^\mu = 0, \quad \lambda = 1, 2, 3. \quad (V.15)$$

Wektory polaryzacji $e_{p\mu}^{(\lambda)}$ określa się za pomocą związków

$$(\vec{\Sigma} \vec{n}) e_p^{(1)} = e_p^{(1)}, \quad (\vec{\Sigma} \vec{n}) e_p^{(2)} = -e_p^{(2)}, \quad (\vec{\Sigma} \vec{n}) e_p^{(3)} = 0, \quad (V.16)$$

gdzie $\vec{n} = \vec{p}/|p|$. Ponieważ $(\vec{\Sigma})_{\alpha 0} = (\vec{\Sigma})_{0\alpha} = 0$, to te związki nie definiują składowej czasowej wektorów $e_{p\mu}^{(\lambda)}$. Składową czasową można określić korzystając ze związku (V.15).

Równanie Proca (V.5) można napisać również w następującej postaci

$$(\beta_\mu \partial^\mu + mc) \Phi(x) = 0, \quad (V.17)$$

gdzie $\Phi(x)$ jest 10-składnikowa funkcja pola

$$\Phi(x) = \begin{pmatrix} \sqrt{mc} \varphi_\mu(x) \\ \frac{1}{\sqrt{mc}} \varphi_{\mu\nu}(x) \end{pmatrix}, \quad (V.18)$$

składająca się z czterech składowych funkcji $\varphi_\mu(x)$ i z sześciu składowych funkcji $\varphi_{\mu\nu}(x)$, a β_μ są to cztery dziesięciowierszowe macierze spełniające te same relacje (IV.27) co i macierze β_μ wchodzące do równania dla cząstek ze spinem zerowym. Równania (V.17) noszą nazwę równań *Duffina-Kemmera*.

ROZDZIAŁ VI

POLE MAXWELLA

Zarówno fotony, kwanty pola elektromagnetycznego nie posiadające masy i opisane przez równania Maxwella, jak i czastki W^+ odpowiedzialne za słabe oddziaływania i opisane przez równania Proca, są obdarzone spinem 1.

Pokazemy teraz, że tak jak i równania Proca tak i równania Maxwella dadzą się zapisać w jawnej postaci Lorentzowsko-nieziemienniczej. Dobrze wiemy, że równania Maxwella są niezmiennicze względem transformacji Lorentza, boprzecież zauważenie przez Einsteina tej ich cechy doprowadziło do powstania teorii względności. My tutaj poszukamy takich oznaczeń, w których ta niezmienniczość przejawia się możliwie najwyraźniej. Równania Maxwella zapiszemy w układzie jednostek Heaviside'a-Lorentza, w którym $\alpha = e^2/4\pi\hbar c = \frac{1}{137}$; α - stała subtelnej struktury magnetycznej.

$$(a) \operatorname{div} \vec{B} = 0,$$

$$(b) \operatorname{rot} \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0,$$

$$(c) \operatorname{div} \vec{E} = \rho,$$

$$(d) \operatorname{rot} \vec{B} - \frac{\partial \vec{E}}{\partial t} = \vec{j}.$$

(VI.1)

Równanie (a) informuje nas, że nie ma ładunków magnetycznych. Równanie (b) wyraża prawo Faradaya mówiące, że zmianie pola magnetycznego towarzyszy powstanie pola elektrycznego. Równanie (c) to twierdzenie Gaussa-Ostrogradzkiego: całkowity ładunek umieszczony wewnątrz zamkniętej powierzchni można otrzymać całkując składową normalną wektora \vec{E} po tej powierzchni. Równanie (d) wyraża prawo Ampera $\operatorname{rot} \vec{B} = \vec{j}$, do którego wprowadzono dodatkowy człon maxwellowski

$\partial \vec{E} / \partial t$ uwzględniający tę okoliczność, że zmianie pól elektrycznych towarzyszy powstanie pól magnetycznych. Równania (a) i (b) są jednorodne, a równania (c) i (d) są niejednorodne.

Jeśli wprowadzić 4-wektorowy potencjał

$$A^\mu = (\varphi, \vec{A}), \quad (\text{VI.2})$$

taki, aby były spełnione związki

$$\vec{B} = \text{rot } \vec{A}, \quad \vec{E} = \frac{\partial \vec{A}}{\partial t} - \nabla \varphi, \quad (\text{VI.3})$$

to równia (a) i (b) będą spełnione automatycznie, ze względu na tożsamości

$$\text{div rot}(\cdot) \equiv 0, \quad \text{rot grad}(\cdot) \equiv 0.$$

Zauważmy, że prawe strony związków (VI.3) są składowymi 4-rotacji, określonej równościami

$$F^{\mu\nu} = -F^{\nu\mu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (\text{VI.4})$$

Rotacja ta ma następujące składowe (przypominamy, że $\partial^1 = -\partial_1$)

$$F^{01} = \partial^0 A^1 - \partial^1 A^0 = \left(\frac{\partial \vec{A}}{\partial t} + \nabla \varphi \right)_1 = -E^1, \quad (\text{VI.5})$$

$$F^{1j} = \partial^1 A^j - \partial^j A^1 = -\epsilon^{1jk} B_k. \quad (\text{VI.6})$$

Związki (VI.5) i (VI.6) można przedstawić w postaci macierzy

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \quad (\text{VI.7})$$

Tensor $F^{\mu\nu}$ nosi nazwę *tensora pola elektromagnetycznego*. Przy przekształceniach Lorentza zachowuje się on jak antysymetryczny tensor rzędu drugiego,

$$F^{\mu\nu} \rightarrow F'^{\mu\nu} = a_{\alpha}^{\mu} a_{\beta}^{\nu} F^{\alpha\beta}.$$

Mozna sprawdzić bezpośrednim rachunkiem, że dwa pozostałe równania Maxwella (c) i (d) są zawarte w kowariantnym równaniu

$$\partial_{\mu} F^{\mu\nu} = j^{\nu}, \quad (\text{VI.8})$$

$$j = (\rho, \vec{j}). \quad (\text{VI.9})$$

Rzeczywiście, kładąc $\nu = 0$, otrzymujemy

$$\partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30} = j^0,$$

$$\text{div } \vec{E} = \rho,$$

czyli równanie (c). Kładąc $\nu = 1$, otrzymujemy

$$\partial_0 F^{01} + \partial_2 F^{21} + \partial_3 F^{31} = j^1,$$

$$-\frac{\partial E^1}{\partial t} + \frac{\partial B^3}{\partial x_2} - \frac{\partial B^2}{\partial x_3} = j^1,$$

czyli pierwszą składową równania (d).

Chociaż pola elektryczne i magnetyczne są określone przez potencjały \vec{A} i φ związkami (VI.3), to taka definicja nie jest jednak jednoznaczna. Można się o tym przekonać, poddając pola \vec{A} i φ tzw. przecechowaniu.

$$\varphi \rightarrow \varphi + \frac{\partial \chi}{\partial t}, \quad \vec{A} \rightarrow \vec{A} - \nabla \chi, \quad (\text{VI.10})$$

lub w postaci kowariantnej

$$A^{\mu} \rightarrow A^{\mu} + \partial^{\mu} \chi, \quad (\text{VI.11})$$

gdzie χ jest dowolną wielkością skalarną. Po przecechowaniu, pola \vec{E} i \vec{B} nie ulegają zmianie. Równoważna forma tego faktu sprowadza się do zadania, by nie ulegał zmianie także tensor $F^{\mu\nu}$:

$$F^{\mu\nu} \rightarrow F^{\mu\nu} + (\partial^{\mu} \partial^{\nu} - \partial^{\nu} \partial^{\mu}) \chi = F^{\mu\nu}, \quad (\text{VI.12})$$

co zachodzi (bo $\partial^{\mu} \partial^{\nu}$ jest przemienne dla χ).

Podstawiając (VI.4) w (VI.8) widzimy, że 4-wektor A^μ spełnia równania

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (\text{VI.4})$$

$$\partial_\mu F^{\mu\nu} = j^\nu \Rightarrow \partial_\mu \partial^\mu A^\nu - \partial_\mu \partial^\nu A^\mu = j^\nu; \quad (\text{VI.13})$$

$$\square A^\nu - \partial^\nu (\partial_\mu A^\mu) = j^\nu. \quad (\text{VI.14})$$

Korzystając ze swobody, jaka daje nam równanie (VI.11) wiemy, jak wybrać funkcję χ , aby przekształcony 4-wektor A^μ spełniał warunek Lorentza

$$\partial_\mu A^\mu = \frac{\partial \varphi}{\partial t} + \nabla \vec{A} = 0, \quad (\text{VI.15})$$

Przy takim doborze 'skali', równanie (VI.14) przyjmie postać

$$\square A^\mu = j^\mu, \quad (\text{VI.16})$$

czyli jest kowariantną reprezentacją znanych równań

$$\frac{\partial^2 \varphi}{\partial t^2} - \nabla^2 \varphi^2 = \rho, \quad (\text{VI.17})$$

$$\frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} = \vec{j},$$

Rozwiązaniami tych równań są potencjały *Lienarda-Wiecherta*.

W próżni równanie (VI.16) przyjmuje postać

$$\square A^\mu = 0. \quad (\text{VI.18})$$

Widzimy, że pole elektromagnetyczne, jeśli uwzględnić jego kwantową naturę, odpowiada cząstkom bez masy poruszającym się z prędkością światła (ze względu na brak masy, a stąd wynikła zasada względności Einsteina).

Przedstawimy teraz równanie Maxwella w jawnej postaci kowariantnej. Równania jednorodne (a) i (b) połączyły się w równaniu (VI.4). Niejednorodne równania (c) i (d) połączyły się w równaniu (VI.8). Pokażemy jak w prosty sposób można połączyć równania (VI.4) i

(VI.8) bez potrzeby wykorzystywania w sposób jawny potencjału A^μ

Z równania (VI.4) wynika, że

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0. \quad (\text{VI.19})$$

Wprowadzimy teraz tensor dualny $\tilde{F}^{\mu\nu}$, określony wzorem

$$\tilde{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\rho\delta} F_{\rho\delta}, \quad (\text{VI.20})$$

($\epsilon^{0123} = 1$). Elementy tego tensora mają następującą postać

$$\tilde{F}^{\mu\nu} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & E^1 & 0 \end{pmatrix} \quad (\text{VI.21})$$

Z antysymetrii tensora $\epsilon^{\mu\nu\rho\delta}$ wynika, że równanie

$$\partial_\mu \tilde{F}^{\mu\nu} = 0, \quad (\text{VI.22})$$

jest równoważne równaniu (VI.19), a z kolei z postaci tensora (VI.21) wynika, że daje ono równania (a) i (b). Ostatecznie możemy zapisać równania Maxwella w zwartej postaci

$$\partial_\mu F^{\mu\nu} = j^\nu; \quad \partial_\mu \tilde{F}^{\mu\nu} = 0. \quad (\text{VI.23})$$

Chcąc opisać cząstki obdarzone masą i ze spinem i w formalizmie wprowadzonym dla równań Maxwella, musimy równania Maxwella uogólnić

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (\text{VI.24})$$

$$\partial_\mu \tilde{F}^{\mu\nu} + m^2 A^\nu = 0.$$

Jeśli weźmiemy dywergencję drugiego z powyższych równań, to otrzymamy

$$m^2 \partial_\nu A^\nu = 0, \quad (\text{VI.25})$$

a ponieważ $m^2 \neq 0$, to otrzymujemy $\partial_\nu A^\nu = 0$. Tak więc warunek Lorentza jest spełniony zawsze, czyli straciliśmy swobodę związaną z przecechowaniem, którą mieliśmy przy równaniach Maxwella. W rzeczywistości, ponieważ tensor $F^{\mu\nu}$ jest niezmienniczy ze względu na przecechowanie, to z (VI.24) wynika bezpośrednio, że równania dla cząstki obdarzonej masą i ze spinem 1 nie są niezmiennicze względem transformacji cechowania.

Podstawiając (VI.25) w (VI.24) otrzymujemy

$$(\square + m^2) A^\mu = 0, \quad (\text{VI.26})$$

wraz z

$$\partial_\mu A^\mu = 0, \quad (\text{VI.27})$$

czyli znane nam już równania Proca (V.2) i (V.3).

ROZDZIAŁ VII

POLE SPINOROWE DIRACA

Po omówieniu pól skalarnych i wektorowych ze spinem równym 0 i 1, przejdziemy teraz do zbadania pól opisujących cząstki ze spinem 1/2. Takie pola noszą nazwę *pól Diraca*. Jednym z ważniejszych przykładów jest pole *elektronowo-pozytonowe*, którego cząstki - elektrony i pozytony - mają spin równy 1/2. Na początku pokazemy, naśladowując Diraca, jak obniżając rząd równania Kleina - Gordona poprzez jego faktoryzację, można otrzymać równanie dla pola Diraca. Równanie Kleina-Gordona zapiszemy w postaci (kładziemy $c = 1$)

$$\square - m^2 = p_\mu p^\mu - m^2, \quad (\text{VII.1})$$

gdzie dla wygody zapisu, jak zwykle, stosujemy oznaczenie z mechaniki

kwantowej $p_\mu = i\partial_\mu$.

Operator (VII.1) jest kwadratowy względem pochodnych ∂_μ . Spróbujmy przedstawić go w postaci kanonicznej

$$\square - m^2 = (P - m)(P + m), \quad (\text{VII.2})$$

gdzie P jest pewną kombinacją operatorów p_μ ze współczynnikami γ^μ

$$P = p_\mu \gamma^\mu. \quad (\text{VII.3})$$

Aby zachodził związek (VII.2) musimy zażądać, aby było spełnione równanie

$$p_\mu p^\mu = p^2 = (p_\mu \gamma^\mu)^2. \quad (\text{VII.4})$$

Tak więc

$$\begin{aligned} (p_\mu \gamma^\mu)^2 &= p_\mu \gamma^\mu p_\nu \gamma^\nu = p_\mu p^\alpha g_{\alpha\nu} \gamma^\mu \gamma^\nu = \\ &= p_\mu p^\alpha g_{\alpha\nu} \left[\frac{1}{2} (\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu) + \frac{1}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \right] = \\ &= p_\mu p^\alpha g_{\alpha\nu} g^{\mu\nu} = p_\mu p^\alpha g_\alpha^\mu = p_\mu p^\mu. \end{aligned} \quad (\text{VII.5})$$

Z ostatniego wiersza widać, że iloczyn $p_\mu p^\nu$ z drugim składnikiem drugiego czynnika musi znikać, ponieważ

$$p_\mu p^\alpha g_{\alpha\mu} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) = p_\mu p^\alpha g_{\alpha\nu} \epsilon^{\mu\nu} = p_\mu p^\mu.$$

Z powyższych rachunków wnioskujemy, że aby taki rozkład był możliwy, to musi zachodzić związek

$$\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}. \quad (\text{VII.6})$$

Ponieważ wielkości γ^μ z różnymi wskaźnikami antykomutują, to γ^μ nie są liczbami i mogą być wybrane jako macierze. Tak więc operator Kleina - Gordona można przedstawić w postaci dwu komutujących operatorów macierzowych.

$$\square - m^2 = (i\gamma^\mu \partial_\mu + m)(i\gamma^\nu \partial_\nu - m). \quad (\text{VII.7})$$

Aby funkcja pola spełniała równanie Kleina - Gordona możemy zażądać, by spełniała ona także jedno z dwu równań 1-go rzędu

$$(i\gamma^\mu \partial_\mu + m)\psi(x) = 0 \quad \text{lub} \quad (i\gamma^\mu \partial_\mu - m)\psi(x) = 0. \quad (\text{VII.8})$$

Równania (VII.8) są mniej ogólne niż (VII.1) i choć każde z rozwiązań jednego z równań (VII.8) spełnia równanie Kleina - Gordona, to odwrotny związek nie zachodzi.

Należy podkreślić, że równania (VII.8) są istotnie nowymi równaniami, gdyż funkcje które je spełniają realizują spinorową reprezentację grupy Lorentza. Można więc oczekiwać, że równania (VII.8) zawierają bardziej szczegółową informację niż równania Kleina - Gordona. Jak wiemy skądinąd, ten fakt ma rzeczywiście miejsce, ponieważ właśnie za pomocą równań (VII.8) Diracowi udało się po raz pierwszy opisać spin elektronu równy 1/2. Równania (VII.8) noszą nazwę równań Diraca, a macierze γ^μ zdefiniowane wzorem (VII.6) noszą nazwę macierzy Diraca.

Ze względu na macierzowy charakter operatorów (VII.6) funkcja falowa ψ , nazywana bispinorem, jest wieloskładnikowa, przy czym liczba jej składników zależy od wymiaru macierzy γ . Można wykazać, że macierze te tworzą algebrę nad ciałem liczb zespolonych, a wymiar nieprzywiedlnej reprezentacji macierzowej wynosi $m = 4$. W związku z tym istnieje $m^2 = 16$ liniowo niezależnych macierzy, które wybiera się w następujący sposób:

macierz jednostkowa I ,

cztery macierze γ^μ , $\mu = 0, 1, 2, 3$,

sześć macierzy $\sigma^{\mu\nu} = \frac{1}{2} [\gamma^\mu, \gamma^\nu] = i \gamma^\mu \gamma^\nu$, $\mu < \nu$,

jedna macierz $\gamma^5 = -i \gamma^0 \gamma^1 \gamma^2 \gamma^3$,

cztery macierze $D^\mu = \gamma^\mu \gamma^5$. (VII.9)

Symbol γ^4 rezerwuje się zazwyczaj dla macierzy różniacej się od γ^0 czynnikiem i tzn. $\gamma^4 = i\gamma^0$. Macierz γ^0 jest hermitowska, a γ^k ,

$k = 1, 2, 3$ są macierzami antyhermitowskimi. Wynika stąd, że macierze γ^μ są unitarne. Warunek hermitowskości piszemy w postaci

$$\gamma^{\mu\dagger} = \gamma^\nu = \gamma_\mu, \quad \mu = 0, 1, 2, 3, 5 \quad (\text{VII.10})$$

(przez sprzężenie hermitowskie a^\dagger macierzy a , rozumiemy wykonanie sprzężenia zespolonego, a następnie transpozycji wierszy i kolumn, tzn. $a^\dagger_{\alpha\beta} = a^*_{\beta\alpha}$).

Zwykle stosuje się taką reprezentację macierzy Diraca, w której γ^0 jest macierzą diagonalną

$$\gamma^0 = \begin{bmatrix} \sigma^0 & 0 \\ 0 & \sigma^0 \end{bmatrix}, \quad \gamma^1 = \begin{bmatrix} 0 & \sigma^1 \\ -\sigma^1 & 0 \end{bmatrix}, \quad \gamma^2 = \begin{bmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{bmatrix}$$

$$\gamma^3 = \begin{bmatrix} 0 & \sigma^3 \\ -\sigma^3 & 0 \end{bmatrix}, \quad \gamma^5 = \begin{bmatrix} 0 & -\sigma^0 \\ -\sigma^0 & 0 \end{bmatrix}, \quad (\text{VII.11})$$

gdzie σ^μ są to tzw. macierze Pauli'ego

$$\sigma^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \sigma^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

(VII.12)

Wszystkie inne postaci reprezentacji macierzy Diraca możemy otrzymać z powyższej poprzez transformacje unitarne

$$\gamma^\mu \rightarrow U \gamma^\mu U^{-1}, \quad (\text{VII.13})$$

gdzie U jest dowolną nieosobliwą macierzą unitarną.

Okazuje się, że układ macierzy Diraca jest niezmienniczy względem znaku γ (transformacja (VII.13) z $U = \gamma^5$), a więc i znak przed masą w operatorze Diraca też nie jest istotny. Zwykle podstawowe równanie Diraca zapisuje się w postaci

$$(i\gamma^\mu \partial_\mu - m)\psi(x) = 0. \quad (\text{VII.14})$$

Inne, sprzężone równanie Diraca otrzymuje się sprzęgając hermitowsko równanie (VII.14) i uwzględniając przy tym (VII.10). Po pewnych przekształceniach otrzymujemy

$$({}_1\tilde{\gamma}^{\mu}\partial_{\mu} + m)\bar{\psi}(x) = 0, \quad (\text{VII.15})$$

lub

$$\bar{\psi}(x)({}_1\tilde{\gamma}^{\mu\kappa}\partial_{\mu} + m) = 0. \quad (\text{VII.16})$$

przy czym symbol $\tilde{\gamma}^{\mu}$ oznacza transponowaną macierz γ^{μ} , a symbol $\tilde{\delta}_{\mu}$ oznacza, że różniczkowanie dotyczy funkcji stojącej po lewej stronie symbolu różniczkowania. Uwzględniono tu fakt, że macierz może działać na funkcję zarówno z lewej jak i z prawej strony

$$(\Lambda\psi)_{\rho} = \Lambda_{\rho\delta}\psi_{\delta}, \quad (\psi\Lambda)_{\rho} = \psi_{\delta}\Lambda_{\delta\rho}.$$

Wynika stąd, że

$$\Lambda\psi = \tilde{\psi}\tilde{\Lambda}. \quad (\text{VII.17})$$

Wchodząca do równań funkcja

$$\bar{\psi}(x) = \psi^{\circ}(x)\gamma^0, \quad (\text{VII.18})$$

nazywa się funkcją sprzężoną diracowsko względem funkcji ψ . Wybór γ^0 w (VII.18) jest zwyczajowy. Wprowadzenie $\bar{\psi}$ pozwala na skonstruowanie lagrangianu.

Zbadamy teraz kowariantność równania Diraca względem transformacji Lorentza. Przy translacjach czasowo - przestrzennych

$$x^{\mu} \rightarrow x'^{\mu} + d^{\mu}, \quad (\text{VII.19})$$

operator Diraca nie ulega zmianie

$${}_1\tilde{\gamma}^{\mu}\partial_{\mu} - m = {}_1\tilde{\gamma}^{\mu}\partial'_{\mu} - m, \quad (\text{VII.20})$$

wiec możemy położyć $\psi'(x') = \psi(x)$, gdyż w nowych zmiennych równanie ma starą postać

$$({}_1\tilde{\gamma}^{\mu}\partial'_{\mu} - m)\psi'(x') = 0, \quad (\text{VII.21})$$

czyli pozostaje kowariantnym. Przy nieskończonej małych obrotach $x \rightarrow x' = x + \delta x$, lub inaczej

$$x'^{\mu} = a^{\mu}_{\nu} x^{\nu} = (\delta^{\mu}_{\nu} + \epsilon^{\mu}_{\nu}) x^{\nu}, \quad (\text{VII.22})$$

równanie Diraca będzie kowariantne jeśli

$$\psi(x) \rightarrow \psi'(x') = S(a)\psi(x), \quad (\text{VII.23})$$

a 4-wymiarowa macierz $S(a)$ będzie spełniała związek

$$S^{-1}(a) \gamma^{\mu} S(a) = a^{\mu}_{\nu} \gamma^{\nu}. \quad (\text{VII.24})$$

Z równań tych można wywnioskować, że jeśli $S(a)$ przedstawić w postaci

$$S(a) = 1 - \frac{1}{2} \epsilon_{\mu\nu} \Sigma^{\mu\nu}, \quad \Sigma^{\mu\nu} = -\Sigma^{\nu\mu}, \quad (\text{VII.25})$$

gdzie $\epsilon_{\mu\nu}$ to nieskończenie mała macierz antysymetryczna, a $\Sigma^{\mu\nu}$ to operatory infinitesimalne nie zależące od $\epsilon_{\mu\nu}$, to macierze $\Sigma^{\mu\nu}$ spełniają relacje

$$[\Sigma^{\mu\nu}, \gamma^{\lambda}] = g^{\mu\lambda} \gamma^{\nu} - g^{\nu\lambda} \gamma^{\mu}. \quad (\text{VII.26})$$

Z powyższego warunku i z warunku, że $\text{Tr } S = 0$, który z kolei wynika z zadania by $\det S = 1$ otrzymujemy

$$\Sigma^{\mu\nu} = \frac{1}{4} [\gamma^{\mu}, \gamma^{\nu}]. \quad (\text{VII.27})$$

Ponieważ funkcja falowa transformuje się zgodnie z wzorem (VII.23), a dla pełnej grupy Lorentza zachodzi

$$\gamma^0 S(a)^{\dagger} \gamma^0 = S(a)^{-1}, \quad (\text{VII.28})$$

to prawo transformacji funkcji sprzężonej dla tej grupy ma postać

$$\bar{\psi}'(x') = \bar{\psi}(x) S^{-1}. \quad (\text{VII.29})$$

Równania Diraca (VII.14) i (VII.16) mogą być otrzymane za pomocą zasady wariacyjnej z następującego lagrangianu

$$\mathcal{L} = \frac{1}{2} \bar{\psi} (i \gamma^{\mu} \partial_{\mu} - m) \psi - \frac{1}{2} \bar{\psi} (i \gamma^{\mu} \overleftarrow{\partial}_{\mu} + m) \psi. \quad (\text{VII.30})$$

Na podstawie ogólnych prawideł możemy wypisać:

- tensor energii-pędu związany z niezmienniczością lagrangianu względem 4-translacji:

$$T^{\mu\nu}(x) = \frac{1}{2} \bar{\psi} \gamma^\nu \partial^\mu \psi - \frac{1}{2} \bar{\psi} \gamma^\nu \partial^\mu \psi, \quad (\text{VII.31})$$

- kanoniczny tensor momentu pędu związany z niezmienniczością lagrangianu względem 4 - obrotów

$$M^{\mu\nu;\rho} = x^\mu T^{\nu\rho} - x^\nu T^{\mu\rho} + \frac{1}{2} \{ \gamma^\rho, \Sigma^{\mu\nu} \} \psi, \quad (\text{VII.32})$$

- wektor prądu związany z niezmienniczością lagrangianu względem transformacji fazowych

$$J^\mu(x) = \bar{\psi} \gamma^\mu \psi. \quad (\text{VII.33})$$

Jeśli cząstka Diraca ma ładunek elektryczny, to oddziałuje ona z polem elektrycznym. Włączenie oddziaływania pola obdarzonego ładunkiem zachodzi drogą zamiany pochodnych ∂_μ na $\partial_\mu - ieA_\mu$. Nazywa się to zasadą minimalnego oddziaływania elektromagnetycznego; zagadnienie włączania do teorii oddziaływania cząstek z polami cechowania, zostanie omówione w rozdziale IX dotyczącym teorii cechowania. Równanie Diraca przyjmuje w tym przypadku postać

$$(i\gamma^\mu (\partial_\mu - ieA_\mu(x)) - m) \psi(x) = 0, \quad (\text{VII.34})$$

a równanie z nim sprzężone

$$(i\tilde{\gamma}^\mu (\partial_\mu + ieA_\mu(x)) + m) \bar{\psi}(x) = 0. \quad (\text{VII.35})$$

Równania (VII.34) i (VII.35) opisują równocześnie elektrony i pozytony, które różnią się tylko znakiem ładunku elektrycznego. Tak więc równanie Diraca winno być niezmiennicze ze względu na zmianę znaku ładunku $e \rightarrow -e$, czyli powinno zachodzić

$$A_\mu(x) \rightarrow A_\mu^C(x) = -A_\mu(x). \quad (\text{VII.36})$$

Powinna mieć również miejsce transformacja bispinora Diraca

$$\psi(x) \rightarrow \psi^C(x) = C \psi(x). \quad (\text{VII.37})$$

gdzie C jest pewna macierz. Wykazuje się, że macierz C spełnia relacje

$$C \gamma^\mu C^{-1} = -\gamma^\mu, \quad (\text{VII.38})$$

ze jest antysymetryczna

$$\tilde{C} = -C, \quad (\text{VII.39})$$

i że jest unitarna

$$C C^+ = I . \quad (\text{VII.40})$$

Przekształcenie C (VII.37) nosi nazwę *transformacji sprzężenia ładunkowego*, a funkcja $\psi^c(x)$ nazywa się *bispinorem sprzężonym ładunkowo*.

Bispinor $\bar{\psi}(x)$ transformuje się zgodnie ze wzorem

$$\bar{\psi}(x) \rightarrow \bar{\psi}^c(x) = C^{-1} \bar{\psi}(x) . \quad (\text{VII.41})$$

Jeśli ładunek jest równy zeru, to równanie Diraca dopuszcza rozwiązania, dla których bispinor $\psi(x)$ pokrywa się z bispinorem sprzężonym ładunkowo

$$\psi(x) = \psi^c(x) . \quad (\text{VII.42})$$

Taki bispinor nazywa się *bispinorem Majorany*. Mógłby on służyć do opisu neutrina, gdyby masa neutrina była różna od zera.

Bispinor Diraca zawsze można przedstawić w postaci

$$\psi = \psi_L + \psi_R , \quad (\text{VII.43})$$

gdzie

$$\psi_L = \frac{1+\gamma^5}{2} \psi , \quad \psi_R = \frac{1-\gamma^5}{2} \psi , \quad (\text{VII.44})$$

a $\frac{1 \pm \gamma^5}{2}$ nazywają się *operatorami rzutowymi* (rozbicie takie gra ważną rolę w teorii oddziaływania elektroslabego).

Dla bispinora Majorany

$$\psi_R = \frac{1-\gamma^5}{2} C\psi , \quad (\text{VII.45})$$

a uwzględniając że $C^{-1} \gamma^5 C = \tilde{\gamma}^5$ otrzymujemy

$$\psi_R = C \tilde{\gamma}^0 \left[\frac{1+\gamma^5}{2} \psi \right]^* = C \tilde{\gamma}^0 \psi_L^* = C \bar{\psi}_L ,$$

lub (VII.46)

$$\psi_R = \psi_L^c .$$

Okazuje się więc, że bispinor Majorany można przedstawić w postaci

$$\psi = \psi_L + \psi_L^c \quad (\text{VII.47})$$

Jezeli wykorzystac reprezentacje Weyla macierzy γ , to

$$\gamma^\mu = \begin{pmatrix} 0 & \bar{\sigma}^\mu \\ \sigma^\mu & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}, \quad (\text{VII.48})$$

gdzie $\bar{\sigma}^\mu = g_{\mu\nu} \sigma^\nu = (\sigma^0, -\sigma^k)$, to latwo sprawdzic, ze macierz C bedzie miala postac

$$C = i \begin{pmatrix} -\sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}, \quad i\sigma^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (\text{VII.49})$$

Tak wiec, jezeli zapisac bispinor Majorany w postaci $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$, to warunek $\psi^c = \psi$ doprowadza do rownania $\varphi = -i\sigma^2 \chi^*$ tak, ze

$$\psi = \begin{pmatrix} -i\sigma^2 \chi^* \\ \chi \end{pmatrix}, \quad (\text{VII.50})$$

a spinor χ speľnia rownanie

$$\sigma^\mu \partial_\mu \chi + m\sigma^2 \chi^* = 0.$$

• • •

Luźne fakty

- 1) Równania Weyla to równania Diraca dla $m = 0$
 $(\partial_0 - \sigma \partial_k) \varphi_1 = 0,$
 $(\partial_0 + \sigma \partial_k) \varphi_2 = 0.$
- 2) Pole spinorowe odpowiada cząstkom naładowanym z możliwymi wartościami rzutu spinu na zadaną oś równymi $\pm 1/2$ (patrz analiza rozwiązań równania Diraca).
- 3) *Kretem* nazywa się podwojony rzut spinu fermionu na kierunek jego pedu $\sigma \vec{p} / |\vec{p}|$.
- 4) Energia pola spinorowego w teorii klasycznej *nie jest* dodatnio określona.

• • •

Równanie Diraca ma rozwiązania w postaci fal płaskich

$$\psi(x) = u e^{-i p x}, \quad p x = p_0 t - \vec{p} \cdot \vec{x}, \quad (\text{VII.52})$$

gdzie bispinor u spełnia równanie

$$(\gamma^\mu p_\mu - m)u = 0.$$

Równanie to ma nietrywialne rozwiązania (wyznacznik musi być równy zero), jeśli

$$p^\mu p_\mu - m^2 = 0$$

lub

$$p_0 = \pm \sqrt{\vec{p}^2 + m^2}$$

Mamy więc dwa typy rozwiązań: jedno zależy od czsu jak $e^{-i p_0 t}$ (dodatnie częstości), a drugie jak $e^{i p_0 t}$ (ujemne częstości). W dalszym ciągu będziemy pisali, że $p = +\sqrt{\vec{p}^2 + m^2}$, i że rozwiązanie równania Diraca przybiera postać

$$\psi(x) = u^q(p) e^{-i q p x}, \quad q = \pm 1. \quad (\text{VII.53})$$

Bispinor $u^q(p)$ spełnia równanie

$$(q \hat{p} - m)u^q(p) = 0, \quad \hat{p} = \gamma^\mu p_\mu. \quad (\text{VII.54})$$

Widać, że bispinor $u^q(p)$ jest wektorem własnym macierzy \hat{p} o wartościach własnych q^m . Wartości własne są tylko dwie, a powinno ich być cztery, więc wartość własna q^m jest dwukrotnie zwyrodniała. Oznacza to, że z jedną wartością własną są związane dwa liniowo niezależne bispinory $u^q(p)$, które można uważać za wektory własne pewnej macierzy rzędu 4, komutującej z \hat{p} . Jako tę macierz wybieramy macierz $\vec{\Sigma} \cdot \vec{n}$, gdzie

$$\Sigma^i = \frac{1}{2} \epsilon^{i k l} \Sigma_{k l}, \quad \Sigma^{\mu\nu} = \frac{i}{4} [\gamma^\mu, \gamma^\nu],$$

$$\text{a } \vec{n} = \frac{\vec{p}}{|\vec{p}|}. \text{ Macierz ta komutuje z } \hat{p}: [\hat{p}, \vec{\Sigma} \cdot \vec{n}] = 0.$$

(komutator znika tylko wtedy, gdy $\vec{n} \parallel \vec{p}$). Tak więc bispinor u , który

teraz oznaczamy $u^{q\mu}(p)$ spełnia równania

$$(q\hat{p} - m) u^{q\mu}(p) = 0,$$

$$\vec{\Sigma} \vec{n} u^{q\mu}(p) = \mu u^{q\mu}(p),$$

(VII.55)

gdzie μ są wartościami własnymi macierzy $\vec{\Sigma} \vec{n}$. Okazuje się, że $(\vec{\Sigma} \vec{n})^2 = \frac{1}{4}$, więc $\mu = \pm \frac{1}{2}$. Macierz $\vec{\Sigma} \vec{n}$ jest infinytezymalnym operatorem grupy obrotów przestrzennych. Tak więc macierz $\vec{\Sigma} \vec{n}$ można interpretować jako operator rzutu spinu cząstek, związanych z polem ψ , na kierunek pędu \vec{p} . Wartości własne tej macierzy nazywamy *krętem* (spiralnością) cząstki. Z teorii kwantowej wynika, że wielkości q , określające wartości własne macierzy \vec{p} , można związać ze znakiem ładunku cząstki opisanej przez pole ψ . Tak więc widać, że zadając pęd cząstki p , otrzymujemy cztery stany, odpowiadające dwóm wartościom kwantowej liczby q i dwóm wartościom kwantowej liczby μ . Za pomocą pola ψ można więc opisywać zarówno elektrony jak i pozytony, dlatego też pole to nazywa się *polem elektronowo-pozytonowym*. Należy jednakże pamiętać, że pole Diraca może służyć do opisu i innych cząstek ze spinem 1/2.

Przy transformacjach Lorentza bispinor ψ transformuje się zgodnie ze wzorem (VII.22)

$$x^\mu \rightarrow x'^\mu = a^\mu_\nu x^\nu, \quad \psi(x) \rightarrow \psi'(x') = S(a) \psi(x).$$

Dla właściwej grupy Lorentza reprezentacja tej grupy za pomocą macierzy $S(a)$ rzędu 4, jest reprezentacją przywiedlną, tzn. że w pewnej bazie macierz $S(a)$ ma następującą strukturę

$$S(a) = \begin{bmatrix} T(a) & 0 \\ 0 & T'(a) \end{bmatrix} \quad (\text{VII.56})$$

gdzie $T(a)$ i $T'(a)$ są macierzami rzędu 2. Dla właściwych transformacji Lorentza $\det a = 1$, a więc ze wzoru

$$S^{-1}(a) \gamma^5 S(a) = \gamma^5 \det a, \quad (\text{VII.57})$$

wynika, że wszystkie macierze $S(a)$ komutują z γ^5 , a ponieważ γ^5 nie jest macierzą jednostkową, to wnioskujemy, że reprezentacja dana przez macierz $S(a)$ jest przywiedlna.

Rozpatrzmy taką bazę w której macierze γ^μ mają reprezentację Weyla, czyli

$$\gamma^\mu = \begin{pmatrix} 0 & \bar{\sigma}^\mu \\ \sigma^\mu & 0 \end{pmatrix}, \quad \bar{\sigma}^\mu = g_{\mu\nu} \sigma^\nu = (\sigma^0, -\sigma^k). \quad (\text{VII.58})$$

W tym przypadku

$$\gamma^5 = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix},$$

gdzie I jest macierzą jednostkową rzędu 2. Właśnie stąd wynika, że macierz rzędu czwartego $S(a)$, komutująca z γ^5 , powinna mieć strukturę (VII.56). Dla właściwej grupy Lorentza $S\gamma^0 S^\dagger = \gamma^0$, to wykorzystując (VII.58) znajdujemy, że

$$S(a) = \begin{pmatrix} T(a) & 0 \\ 0 & T^\dagger(a)^{-1} \end{pmatrix}. \quad (\text{VII.59})$$

W reprezentacji Weyla bispinor ψ zapisujemy w postaci

$$\psi = \begin{pmatrix} \varphi \\ \kappa \end{pmatrix}. \quad (\text{VII.60})$$

Przy właściwych transformacjach Lorentza dwuskładnikowe funkcje φ i κ transformują się niezależnie, np

$$\varphi_\alpha \rightarrow \varphi'_\alpha = T_{\alpha\beta} \varphi_\beta. \quad (\text{VII.61})$$

Wielkości transformujące się zgodnie z takimi wzorami nazywają się *spinorami*. Dwuskładnikowe wielkości transformujące się za pomocą zespolenie sprzężonej macierzy T^* nazywamy *spinorami kropkowanymi* i oznaczymy χ_μ . Transformują się one zgodnie z definicją

$$\dot{\varphi}^{\dot{\mu}} \rightarrow \dot{\varphi}'^{\dot{\mu}} = T^{\dot{\mu}}_{\dot{\sigma}} \dot{\varphi}^{\dot{\sigma}}, \quad T^{\dot{\mu}}_{\dot{\sigma}} = T^{*\mu}_{\sigma}, \quad (\text{VII.62})$$

$$\dot{\chi}_{\dot{\mu}} \rightarrow \dot{\chi}'_{\dot{\mu}} = \dot{\chi}_{\dot{\sigma}} (T^{-1})^{\dot{\delta}}_{\dot{\mu}}, \quad (T^{-1})^{\dot{\sigma}}_{\dot{\mu}} = (T^{-1})^{*\sigma}_{\mu},$$

i nazywają się odpowiednio kontrawariantnymi i kowariantnymi spinorami kropkowanymi rzędu pierwszego.

Można wprowadzić również tensory wyższego rzędu, które transformują się jak iloczyny odpowiednich spinorów rzędu pierwszego, np.:

$$\dot{\varphi}^{\dot{\lambda}\dot{\rho}} \rightarrow \dot{\varphi}'^{\dot{\lambda}\dot{\rho}} = T^{\dot{\lambda}}_{\dot{\alpha}} T^{\dot{\rho}}_{\dot{\beta}} \dot{\varphi}^{\dot{\alpha}\dot{\beta}}, \quad (\text{VII.63})$$

$$\dot{\varphi}_{\dot{\rho}}^{\dot{\lambda}} \rightarrow \dot{\varphi}'_{\dot{\rho}}^{\dot{\lambda}} = T^{\dot{\lambda}}_{\dot{\alpha}} \dot{\varphi}_{\dot{\beta}}^{\dot{\alpha}} (T^{-1})^{\dot{\beta}}_{\dot{\rho}}.$$

Rolę tensora uniwersalnego gra tensor $\varepsilon^{\alpha\beta} = -\varepsilon^{\beta\alpha}$, ($\varepsilon^{12} = 1$)

$$\varepsilon = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad (\text{VII.64})$$

który nie zmienia się przy transformacji spinorowej

$$\varepsilon^{\alpha\beta} \rightarrow \varepsilon'^{\alpha\beta} = T^{\alpha}_{\rho} T^{\beta}_{\lambda} \varepsilon^{\rho\lambda}, \quad (\text{VII.65})$$

a ponieważ $\det T = 1$, to $\varepsilon'^{\alpha\beta} = \varepsilon^{\alpha\beta}$. Za pomocą tego spinora można podnosić i opuszczać indeksy, np.:

$$\dot{\varphi}^{\dot{\alpha}} = \varepsilon^{\dot{\alpha}\dot{\beta}} \dot{\varphi}_{\dot{\beta}}. \quad (\text{VII.66})$$

Jak widzieliśmy, wielkości $\dot{\chi}_{\dot{\alpha}}$ transformują się zgodnie ze wzorem

$$\dot{\chi}'_{\dot{\alpha}} \rightarrow \dot{\chi}'_{\dot{\alpha}} = \dot{\chi}_{\dot{\beta}} (T^{-1})^{\dot{\beta}}_{\dot{\alpha}},$$

który bez uwzględnienia oznaczeń spinorowych można zapisać

$$x_{\alpha} \rightarrow x'_{\alpha} = (T^{\dagger})_{\alpha\beta}^{-1} x_{\beta} .$$

Ponieważ zgodnie z tym prawem transformuje się dwuskładnikowa funkcja x_{α} , tworząca wraz z φ czteroskładnikowy bispinor ψ , to można uważać, że czteroskładnikowa funkcja pola ψ łączy w sobie kontrawariantny spinor φ^{ρ} i kowariantny spinor kropkowany x_{λ}

$$\psi = \begin{bmatrix} \varphi^{\rho} \\ x_{\lambda} \end{bmatrix} . \quad (\text{VII.67})$$

ROZDZIAŁ VIII

SUPERPOLA

Pola, które rozpatrywaliśmy dotychczas, a więc pola skalarne, wektorowe i Diraca, związane są z określonymi reprezentacjami grupy Poincaré. Przy przekształceniach grupy Poincaré lub grupy symetrii wewnętrznych, pola tensorowe (bozonowe) i spinorowe (fermionowe) transformowały się niezależnie od siebie. Można jednakże wprowadzić bardziej ogólną grupę transformacji, tzw. *grupę transformacji supersymetrii*, która łączy pola bozonowe i fermionowe w taki sposób, że transformują się one łącznie 'mieszając' się ze sobą. Aby przekształcenia te miały sens fizyczny, zadamy, by lagrangian był niezmienniczy względem transformacji z tej grupy.

Pola badane przez nas poprzednio transformowały się przy infinitezymalnych transformacjach grupy Poincaré zgodnie z wzorami

$$x(x) \rightarrow x'(x') , \quad x_{\mu} \rightarrow x'_{\mu} = x_{\mu} + \epsilon_{\mu}^{\nu} x_{\nu} , \quad (\text{VIII.1})$$

gdzie

$$\delta\chi(x) = G \chi(x), \quad G = i \varepsilon_{\mu} \mathcal{P}^{\mu} - \frac{1}{2} \varepsilon_{\mu\nu} L^{\mu\nu}, \quad (\text{VIII.2})$$

a \mathcal{P}^{μ} i $L^{\mu\nu}$ są generatorami transformacji

$$\mathcal{P}_{\mu} = i \partial_{\mu}, \quad L_{\mu\nu} = \Sigma_{\mu\nu} + i(x_{\mu} \partial_{\nu} - x_{\nu} \partial_{\mu}). \quad (\text{VIII.3})$$

Generatory te spełniają następujące relacje komutacyjne

$$[\mathcal{P}_{\mu}, \mathcal{P}_{\nu}] = 0, \quad [\mathcal{P}_{\mu}, L_{\nu\rho}] = i(g_{\mu\nu} \mathcal{P}_{\rho} - g_{\mu\rho} \mathcal{P}_{\nu}), \quad (\text{VIII.4})$$

$$[L_{\mu\nu}, L_{\rho\lambda}] = i(g_{\mu\lambda} L_{\nu\rho} + g_{\nu\rho} L_{\mu\lambda} - g_{\mu\rho} L_{\nu\lambda} - g_{\nu\lambda} L_{\mu\rho}).$$

Rozszerzymy teraz grupę Poincaré, wprowadzając generatory dodatkowe, które w odróżnieniu od \mathcal{P}_{μ} i $L_{\nu\rho}$ będziemy uważali nie za wielkości tensorowe, ale za wielkości spinorowe. Będziemy je oznaczali Q_{α} i $Q_{\dot{\alpha}}$ (α i $\dot{\alpha}$ są to indeksy spinorowe, $\alpha = 1, 2$, $\dot{\alpha} = 1, 2$).

Aby miała miejsce niezmienniczość wynikająca z teorii względności założymy, że generatory \mathcal{P}_{μ} , $L_{\nu\rho}$, Q_{α} i $Q_{\dot{\alpha}}$ spełniają następujące relacje komutacyjne

$$\begin{aligned} [Q_{\alpha}, \mathcal{P}_{\mu}] &= [Q_{\dot{\alpha}}, \mathcal{P}_{\mu}] = 0, & \sigma^{\mu} &- \text{macierze Pauliego,} \\ \{Q_{\alpha}, Q_{\beta}\} &= \{Q_{\dot{\alpha}}, Q_{\dot{\beta}}\} = 0, & \sigma^{\mu\nu} &= \frac{1}{2} [\gamma^{\mu}, \gamma^{\nu}], \\ \{Q_{\alpha}, Q_{\dot{\beta}}\} &= 2\sigma^{\mu}_{\beta\alpha} \mathcal{P}_{\mu}, & & \\ [Q_{\alpha}, L_{\mu\nu}] &= \frac{1}{2} (\sigma_{\mu\nu})_{\alpha}^{\beta} Q_{\beta}, & & \\ [Q_{\dot{\alpha}}, L_{\mu\nu}] &= \frac{1}{2} (\bar{\sigma}_{\mu\nu})_{\dot{\alpha}}^{\dot{\beta}} Q_{\dot{\beta}}. & & \end{aligned} \quad (\text{VIII.5})$$

Relacje (VIII.4) muszą być oczywiście też spełnione. Współczynnik liczbowy dla antykomutatora $\{Q_{\alpha}, Q_{\dot{\beta}}\}$ zawsze można położyć równym dwa, jeśli odpowiednio unormuje się generatory Q_{α} i $Q_{\dot{\beta}}$.

Zwróćmy uwagę, że dla generatorów spinorowych, zamiast komutatorów używa się antykomutatorów. Związane jest to z faktem, że

pola spinorowe opisują fermiony spełniające *zasadę Pauliego*. Na pierwszy rzut oka wydaje się, że wprowadzenie antykomutatorów dla generatorów grupy jest niezgodne z teorią grup ciągłych, dla której

$$[T_a, T_b] = f_{ab}^c T_c.$$

W rzeczywistości sprzeczność nie powstaje, jeśli oprócz zwykłych c-liczbowych parametrów, określających elementy ciągłej grupy, wprowadzić tak zwane *zmienne spinorowe Grassmanna* ζ_α i ζ_α^\cdot , antykomutujące między sobą i z generatorami Q_α i Q_α^\cdot

$$\{\zeta_\alpha, \zeta_\beta\} = \{\zeta_\alpha^\cdot, \zeta_\beta^\cdot\} = \{\zeta_\beta, Q_\alpha\} = \{\zeta_\beta^\cdot, Q_\alpha^\cdot\} = \{\zeta_\beta, Q_\alpha^\cdot\} = 0,$$

(VIII.6)

i komutujące z pozostałymi generatorami.

Zauważmy, że związki (VIII.5) można zapisać w postaci

$$[Q_\zeta, Q_{\zeta'}] = [\bar{\zeta} \bar{Q}, \bar{\zeta}' \bar{Q}] = 0,$$

(VIII.7)

$$[\bar{\zeta} \bar{Q}, \zeta' Q'] = 2\bar{\zeta} \sigma \zeta' \mathcal{P}_\mu,$$

gdzie $Q_\zeta = Q_\alpha \zeta^\alpha$, $\bar{\zeta} \bar{Q} = \zeta_\alpha^\cdot Q_\alpha^\cdot$, ($Q_\alpha^\cdot \equiv \bar{Q}_\alpha$).

Rzeczywiście, na przykład

$$\begin{aligned} [Q_\zeta, Q_{\zeta'}] &= Q_\alpha \zeta^\alpha Q_\beta \zeta'^\beta - Q_\beta \zeta'^\beta Q_\alpha \zeta^\alpha = \\ &= -Q_\alpha Q_\beta \zeta^\alpha \zeta'^\beta + Q_\beta Q_\alpha \zeta'^\beta \zeta^\alpha = \\ &= Q_\alpha Q_\beta \zeta'^\beta \zeta^\alpha + Q_\beta Q_\alpha \zeta'^\beta \zeta^\alpha = \{Q_\alpha, Q_\beta\} \zeta'^\beta \zeta^\alpha = 0. \end{aligned}$$

Pokażemy teraz, że wielkości

$$\mathcal{F}(\zeta, a) = \exp G(\zeta, a),$$

(VIII.8)

tworzą grupę, gdzie

$$G(\zeta, c) = i(Q_\alpha \zeta^\alpha + \zeta_\alpha^\cdot Q_\alpha^\cdot + Q^\mu \mathcal{P}_\mu),$$

c to zmienne c-liczby, a ζ^α , $\zeta_\alpha^\cdot \equiv \bar{\zeta}^\alpha$ to zmienne Grassmanna.

Skorzystajmy w tym celu z tożsamości Campbell - Hausdorfa

$$e^{A+B} = e^A e^B e^{-[A,B]/2}, \quad (\text{VIII.9})$$

która jest prawdziwa jeśli operator $[A,B]$ komutuje z A i z B , czyli gdy $[A, [A,B]] = [B, [A,B]] = 0$. Jeśli ponadto uwzględnimy, że

$$-\frac{1}{2} [Q \zeta + \bar{\zeta} \bar{Q}, Q \zeta' + \bar{\zeta}' \bar{Q}] = \bar{\zeta}' \sigma^\mu \zeta \mathcal{P}_\mu - \bar{\zeta} \sigma^\mu \zeta' \mathcal{P}_\mu, \quad (\text{VIII.10})$$

co wynika ze wzorów (VIII.7), to

$$\mathcal{F}(\zeta, a) \mathcal{F}(\zeta', a') = \mathcal{F}(\zeta'', a''), \quad (\text{VIII.11})$$

jeśli

$$\zeta'' = \zeta + \zeta', \quad \bar{\zeta}'' = \bar{\zeta} + \bar{\zeta}',$$

$$(\text{VIII.12})$$

$$a''^\mu = a^\mu + a'^\mu + i(\bar{\zeta} \sigma^\mu \zeta' - \bar{\zeta}' \sigma^\mu \zeta).$$

Jak widzieliśmy, generatory grupy Poincaré \mathcal{P}_μ i $L_{\mu\nu}$ mogą być realizowane za pomocą operacji różniczkowania względem x^μ (VIII.3). Podobnie mogą być realizowane generatory Q_α i $\bar{Q}_{\dot{\alpha}}$, ale za pomocą różniczkowania względem pewnych zmiennych spinorowych e^α i $e^{\dot{\alpha}}$. Wielkości e^α i $e^{\dot{\alpha}}$ powinny być zmiennymi Grassmanna, antykomutującymi ze sobą i z innymi zmiennymi spinorowymi, np. ζ^α i $\zeta^{\dot{\alpha}}$. Mamy więc

$$\frac{\partial}{\partial e^\gamma} e^\alpha e^\beta e^{\dot{\alpha}} \dots = \frac{\partial e^\alpha}{\partial e^\gamma} e^\beta e^{\dot{\alpha}} \dots - e^\alpha \frac{\partial e^\beta}{\partial e^\gamma} e^{\dot{\alpha}} \dots + e^\alpha e^\beta \frac{\partial e^{\dot{\alpha}}}{\partial e^\gamma} \dots - \dots,$$

$$\frac{\partial e^\alpha}{\partial e^\gamma} = \delta_\gamma^\alpha, \quad \frac{\partial e^{\dot{\alpha}}}{\partial e^\alpha} = 0.$$

Analogiczne wzory obowiązują i dla pochodnej $\frac{\partial}{\partial e^{\dot{\alpha}}}$. Ze względu na to,

ze e^α i $e^{\dot{\alpha}}$ są zmiennymi Grassmanna, oraz że α i $\dot{\alpha}$ przyjmują tylko wartości 1 i 2, różne od zera są tylko następujące iloczyny

$$e^\alpha e^\beta, \quad \bar{e}^{\dot{\alpha}} \bar{e}^{\dot{\beta}}, \quad e^\alpha \bar{e}^{\dot{\beta}}, \quad \bar{e}^{\dot{\alpha}} e^\beta e^\gamma, \quad e^\alpha e^\beta \bar{e}^{\dot{\gamma}}, \quad e^\alpha e^\beta \bar{e}^{\dot{\lambda}} \bar{e}^{\dot{\rho}},$$

($\bar{e}^{\dot{\alpha}} = e^{\dot{\alpha}}$). Wynika to z faktu, że iloczyny zmiennych Grassmanna też

powinien być zmienna Grassmanna. Zauważmy ponadto, że ze wszystkich iloczynów $e^\alpha e^\beta$, różne od zera są tylko $e^1 e^2$ i $e^2 e^1 = -e^1 e^2$. Tak więc

$$e^\alpha e^\beta = -\frac{1}{2} \varepsilon^{\alpha\beta} (e e), \quad (\text{VIII.13})$$

gdzie $\varepsilon^{\alpha\beta}$ to antisymetryczny tensor rzędu drugiego, a

$$(e e) = e^\alpha e_\alpha \equiv e^\alpha \varepsilon_{\alpha\beta} e^\beta.$$

Podobnie

$$\bar{e}^\alpha \bar{e}^\beta = -\frac{1}{2} \varepsilon^{\alpha\beta} (\bar{e} \bar{e}), \quad (\bar{e} \bar{e}) = \bar{e}^\alpha \bar{e}_\alpha \equiv \bar{e}^\alpha \varepsilon_{\alpha\beta} \bar{e}^\beta. \quad (\text{VIII.14})$$

Wynika stąd, że

$$e^\alpha e^\beta e^\rho = -\frac{1}{2} \varepsilon^{\alpha\beta} (e e) e^\rho, \quad \bar{e}^\alpha \bar{e}^\beta \bar{e}^\rho = -\frac{1}{2} \varepsilon^{\alpha\beta} (\bar{e} \bar{e}) \bar{e}^\rho, \quad (\text{VIII.15})$$

$$e^\alpha e^\beta e^\alpha e^\beta = \frac{1}{4} \varepsilon^{\alpha\beta} \varepsilon^{\alpha\beta} (\bar{e} \bar{e}) (e e).$$

W ten sposób element jednostkowy, spinory e , \bar{e} i iloczyny spinorów Grassmanna

$$(e e), \quad (\bar{e} \bar{e}), \quad (e e) \bar{e}^\alpha, \quad (\bar{e} \bar{e}) e^\alpha, \quad (e e) (\bar{e} \bar{e}), \quad (\text{VIII.16})$$

tworzą bazę, która może być potraktowana jako baza pewnej 16-wymiarowej przestrzeni liniowej, w której realizuje się algebra Grassmanna, zdefiniowana prawem mnożenia spinorów e^α i \bar{e}^α . Zmienne e^α i \bar{e}^α nazywają się tworzącymi algebry Grassmanna (algebra Grassmanna z n tworzącymi realizuje się w przestrzeni 2^n wymiarowej). W przestrzeni zmiennych e i ζ można wprowadzić operację inwolucji. Będziemy ją oznaczać gwiazdką. W przypadku zwykłych c -liczb, operacja inwolucji odpowiada sprzężeniu zespolonemu.

$$e \rightarrow e^* = \bar{e}, \quad \zeta \rightarrow \zeta^* = \bar{\zeta}$$

$$\bar{e} \rightarrow \bar{e}^* = e, \quad \bar{\zeta} \rightarrow \bar{\zeta}^* = \zeta$$

(VIII.17)

$$(c e) \rightarrow (c e)^* = c^* e^*,$$

$$(c \zeta) \rightarrow (c \zeta)^* = c^* \zeta^*,$$

gdzie c jest liczbą zespoloną. Przy inwolucji iloczynu zmiennych

Grassmanna należy zmieniać kolejność czynników (tzn. należy zmieniać znak)

$$(e^\alpha e^\beta)^* = \bar{e}^\beta \bar{e}^\alpha = -\bar{e}^\alpha \bar{e}^\beta \quad (\text{VIII.18})$$

Relacje komutacyjne (VIII.5) będą spełnione, jeśli położymy

$$P_\mu = i \delta_\mu, \quad Q_\alpha = -i \frac{\partial}{\partial e^\alpha} + e^{\dot{\beta}} \sigma_{\beta\alpha}^\mu \delta_\mu, \quad (\text{VIII.19})$$

$$Q_\alpha = -i \frac{\partial}{\partial \bar{e}^\alpha} + \sigma_{\dot{\alpha}\beta}^\mu e^\beta \delta_\mu.$$

Są to infinitesimalne generatory badanej transformacji. Aby były spełnione relacje komutacyjne (VIII.5), generator $L_{\mu\nu}$ dany przez (VIII.3) musi ulec następującej modyfikacji

$$L_{\mu\nu} = \Sigma_{\mu\nu} + i(x_\mu \delta_\nu - x_\nu \delta_\mu) + \frac{1}{2} \left[e^\alpha (\sigma_{\mu\nu})_{\alpha\beta} \frac{\partial}{\partial e_\beta} + \bar{e}^\alpha (\bar{\sigma}_{\mu\nu})_{\alpha\beta} \frac{\partial}{\partial \bar{e}_\beta} \right] \quad (\text{VIII.20})$$

Po tej modyfikacji zmienne e^α i \bar{e}^β będą transformować się przy przekształceniach Lorentza jak relatywistyczne spinory

$$e^\alpha \rightarrow e'^\alpha = e^\alpha + \delta e^\alpha,$$

$$\delta e^\alpha = \frac{1}{2} \varepsilon^{\mu\nu} L_{\mu\nu} e^\alpha = \frac{1}{2} \varepsilon^{\mu\nu} (\sigma_{\mu\nu})^\alpha_\beta e^\beta.$$

Zauważmy, że zgodnie z (VIII.8) i (VIII.19) generatory Q_α i Q_α mają wymiar $\frac{1}{\sqrt{L}}$, a zmienne ζ , e mają wymiar \sqrt{L} , gdzie L jest długością.

Superpola, w odróżnieniu od zwykłych pól $\chi(x)$ zależą nie tylko od zmiennych przestrzenno-czasowych x^μ , ale również zmiennych spinorowych Grassmanna e^α , \bar{e}^α . Superpola będziemy więc oznaczać jako $\chi(x, e, \bar{e})$. Ze względu na nieznikające iloczyny (VIII.16), superpole można przedstawić w postaci

$$\begin{aligned} \chi(x, e, \bar{e}) = & A(x) + e^\alpha \psi_\alpha(x) + \bar{\varphi}_\alpha \bar{e}^\alpha + (e e)F(x) + \\ & + (\bar{e} \bar{e})G(x) + \bar{e} \sigma^\mu e B_\mu(x) + \\ & + (e e) \bar{\kappa}_\alpha(x) \bar{e}^\alpha + (\bar{e} \bar{e}) e^\alpha \lambda_\alpha(x) + (e e)(\bar{e} \bar{e}) D(x), \end{aligned} \quad (\text{VIII.21})$$

gdzie A, F, G, D są polami skalarnymi, $B_\mu(x)$ jest polem wektorowym, a $\psi, \bar{\varphi}, \bar{\kappa}, \lambda$ są polami spinorowymi. Superpole łączy więc w sobie cztery pola skalarne (A, F, G, D), cztery pola spinorowe ($\psi, \bar{\varphi}, \bar{\kappa}, \lambda$) i jedno pole wektorowe (B_μ).

Rozpatrzmy najprostszy przypadek; gdy superpole $\chi(x, \theta, \bar{\theta})$ określone w superprzestrzeni ośmiowymiarowej nie ma indeksów ani spinorowych ani tensorowych. W tym przypadku w generatorze $L_{\mu\nu}$ (VIII.20) można zaniedbać składnik $\Sigma_{\mu\nu}$. Generatory \mathcal{P}_μ tworzą zamkniętą algebrę, z którą związane jest następujące prawo transformacji superpola

$$\chi(x) \rightarrow \chi'(x'), \quad x \rightarrow x' = x + a,$$

gdzie (VIII.22)

$$\chi'(x) = \mathcal{F}(a)\chi(x) \equiv \chi(x - a),$$

$$\mathcal{F}(a) = \exp(i\mathcal{P}_\mu a^\mu),$$

które pokazuje, że przy translacjach czasoprzestrzeni superpole transformuje się jak skalar.

Chcąc znaleźć prawo transformacji superpola związanego z algebra zadana przez generatory Q_α, \bar{Q}_α i \mathcal{P}_μ rozpatrzmy superpole

$$\chi'(x, \theta, \bar{\theta}) = \mathcal{F}(\zeta, a) \chi(x, \theta, \bar{\theta}) \mathcal{F}^{-1}(\zeta, a). \quad \text{(VIII.23)}$$

Wykorzystując związek

$$e^G A e^{-G} = A + [G, A] + \frac{1}{2} [G, [G, A]] + \dots \quad \text{(VIII.24)}$$

zauważamy, że

$$\mathcal{F} x \mathcal{F}^{-1} = x + [\mathcal{F}, x] + \frac{1}{2} [G, [G, x]],$$

$$\mathcal{F} \theta \mathcal{F}^{-1} = \theta + [G, \theta], \quad \text{(VIII.25)}$$

$$\mathcal{F} \bar{\theta} \mathcal{F}^{-1} = \bar{\theta} + [G, \bar{\theta}],$$

gdzie zgodnie z (VIII.19)

$$[G, x^\mu] = -a^\mu + i \bar{\theta} \sigma^\mu \zeta - i \bar{\zeta} \sigma^\mu \theta,$$

$$[G, [G, x^\mu]] = i \bar{\zeta} \sigma^\mu \zeta - i \bar{\zeta} \sigma^\mu \zeta = 0,$$

(VIII.26)

$$[G, e] = -\zeta,$$

$$[G, \bar{e}] = -\bar{\zeta},$$

a wszystkie wyższe komutatory znikają. Tak więc

$$\mathfrak{S} x^\mu \mathfrak{S}^{-1} = x^\mu - a^\mu + i(\bar{e} \sigma^\mu \zeta - \bar{\zeta} \sigma^\mu e),$$

$$\mathfrak{S} e \mathfrak{S}^{-1} = e - \zeta, \quad (\text{VIII.27})$$

$$\mathfrak{S} \bar{e} \mathfrak{S}^{-1} = \bar{e} - \bar{\zeta}.$$

Stąd wynika, że

$$\mathfrak{S}(\zeta, a) \chi(x, e, \bar{e}) = \chi(x - a + i(\bar{e} \sigma \zeta - \bar{\zeta} \sigma e), e - \zeta, \bar{e} - \bar{\zeta}) = \chi'(x, e, \bar{e}'). \quad (\text{VIII.28})$$

Okazuje się więc, że superpole $\chi(x, e, \bar{e})$ można traktować jako superskalar

$$\chi(x, e, \bar{e}) \rightarrow \chi'(x', e, \bar{e}) = \chi(x, e, \bar{e}), \quad (\text{VIII.29})$$

jeśli uważać, że przy supertransformacjach wielkości x , e i \bar{e} transformują się zgodnie z wzorami

$$x \rightarrow x' = x + a - i(\bar{e} \sigma \zeta - \bar{\zeta} \sigma e),$$

$$e \rightarrow e' = e + \zeta, \quad (\text{VIII.30})$$

$$\bar{e} \rightarrow \bar{e}' = \bar{e} + \bar{\zeta}.$$

Wprowadzimy teraz operacje różniczkowania kowariantnego \mathcal{D}_α i \mathcal{D}_α^\cdot względem zmiennych spinorowych, które nie powodują zmiany charakteru transformacji superpolea przy supertransformacjach:

$$\mathcal{D}_\alpha = \frac{\partial}{\partial e^\alpha} - i e^\beta \sigma_{\beta\alpha}^\mu \partial_\mu, \quad (\text{VIII.31})$$

$$\mathcal{D}_\alpha^\cdot = - \frac{\partial}{\partial \bar{e}^{\dot{\alpha}}} + i \sigma_{\dot{\alpha}\beta}^\mu \bar{e}^\beta \partial_\mu.$$

Z definicji (VIII.19) generatorów spinorowych Q_α i Q_α^\cdot wynika, że

$$\{\mathcal{D}_\alpha, Q_\beta\} = \{\mathcal{D}_\alpha^\cdot, Q_\beta^\cdot\} = \{\mathcal{D}_\alpha, Q_\beta^\cdot\} = \{\mathcal{D}_\alpha^\cdot, Q_\beta\} = 0. \quad (\text{VIII.32})$$

Oprócz tego, operatory \mathcal{D}_α , \mathcal{D}_α^\cdot antykomutują z parametrami ζ^α i $\zeta^{\dot{\alpha}}$, więc zgodnie z definicją (VIII.8) operatora \mathfrak{S} mamy

$$[\mathcal{D}_\alpha, \mathcal{F}] = [\mathcal{D}_\alpha, \mathcal{F}] = 0, \quad (\text{VIII.33})$$

Wynika stąd, że pochodne kowariantne superpól

$$\mathcal{D}_\alpha \chi(\kappa, \vartheta, \bar{\vartheta}), \quad \mathcal{D}_\alpha \chi(\kappa, \vartheta, \bar{\vartheta}),$$

też są superpolami (transformują się tak samo). Jeżeli $\chi(x, \vartheta, \bar{\vartheta})$ jest superpolem skalarnym, to pola $\mathcal{D}_\alpha \chi(x, \vartheta, \bar{\vartheta})$, $\mathcal{D}_\alpha \chi(x, \vartheta, \bar{\vartheta})$ będą superpolami spinorowymi (ze względu na transformację Lorentza). Ponieważ $[\partial_\mu, \mathcal{F}] = 0$, to pole $\partial_\mu \chi(x, \vartheta, \bar{\vartheta})$ będzie superpolem wektorowym.

Zauważmy również, że operatory kowariantnego różniczkowania spinorowego spełniają takie same relacje komutacyjne jak generatory Q_α i Q_α :

$$\{\mathcal{D}_\alpha, \mathcal{D}_\beta\} = \{\mathcal{D}_\alpha, \mathcal{D}_\beta\} = 0,$$

(VIII.34)

$$\{\mathcal{D}_\alpha, \mathcal{D}_\beta\} = 2 \sigma^\mu_{\beta\alpha} \mathcal{P}_\mu.$$

Aby superpole $\chi(x, \vartheta, \bar{\vartheta})$ było nieprzywiedlne należy zażądać, aby był spełniony warunek

$$\mathcal{D}_\alpha \chi(x, \vartheta, \bar{\vartheta}) = 0, \quad (\text{VIII.35})$$

lub

$$\mathcal{D}_\alpha \chi(x, \vartheta, \bar{\vartheta}) = 0. \quad (\text{VIII.36})$$

Ponieważ związki (VIII.34) i (VIII.35) spełnione są również dla pól χ' , to powodują one, że z przestrzeni superpól zawierających dziewięć pól niezależnych (VIII.20) można wyodrębnić podprzestrzenie niezmiennicze zawierające tylko trzy niezależne superpola.

Pomocne w przekonaniu się o tym jest wprowadzenie nowych reprezentacji superpól, tzw. plus- i minus-reprezentacji

$$\chi_\pm(x, \vartheta, \bar{\vartheta}) = e^{\mp i \bar{\theta} \sigma^\mu_{\theta\theta} \mu} \chi(x, \vartheta, \bar{\vartheta}). \quad (\text{VIII.37})$$

Operatory A_+ i A_- w plus- i minus-reprezentacji związane są z wyjściowym operatorem A w następujący sposób

$$A_{\pm} = e^{\mp i \bar{\theta} \sigma^{\mu} \theta} A_{\pm} e^{\pm i \bar{\theta} \sigma^{\mu} \theta} \quad (\text{VIII.38})$$

Wykorzystując rozkład (VIII.23) można otrzymać następujące wyrażenie dla operatorów \mathcal{D} i Q w plus-reprezentacji

$$\mathcal{D}_{\alpha}^{+} = \frac{\partial}{\partial \theta^{\alpha}} - 2i (\bar{\theta} \sigma^{\mu})_{\alpha} \partial_{\mu}, \quad \mathcal{D}_{\dot{\alpha}}^{+} = - \frac{\partial}{\partial \theta^{\dot{\alpha}}}, \quad (\text{VIII.39})$$

$$Q_{\alpha}^{+} = -i \frac{\partial}{\partial \theta^{\alpha}}, \quad Q_{\dot{\alpha}}^{+} = i \frac{\partial}{\partial \theta^{\dot{\alpha}}} - 2(\sigma^{\mu} \theta)_{\dot{\alpha}} \partial_{\mu},$$

i w minus-reprezentacji

$$\mathcal{D}_{\alpha}^{-} = \frac{\partial}{\partial \theta^{\alpha}}, \quad \mathcal{D}_{\dot{\alpha}}^{-} = - \frac{\partial}{\partial \theta^{\dot{\alpha}}} + 2i (\sigma^{\mu} \theta)_{\dot{\alpha}} \partial_{\mu}, \quad (\text{VIII.40})$$

$$Q_{\alpha}^{-} = -i \frac{\partial}{\partial \theta^{\alpha}} + 2(\bar{\theta} \sigma^{\mu})_{\alpha} \partial_{\mu}, \quad Q_{\dot{\alpha}}^{-} = i \frac{\partial}{\partial \theta^{\dot{\alpha}}}.$$

Superpola χ'_{\pm} i χ_{\pm} w plus- i minus-reprezentacji transformują się zgodnie z wzorami

$$\chi'_{\pm} = \mathfrak{S}_{\pm} \chi_{\pm}, \quad \mathfrak{S}_{\pm} = \exp G_{\pm}, \quad (\text{VIII.41})$$

gdzie

$$G_{\pm} = i (Q^{\pm} \zeta + \bar{\zeta} \bar{Q}^{\pm} + a^{\mu} p_{\mu}).$$

Równanie (VIII.35) zgodnie z (VIII.40) w minus-reprezentacji przyjmuje prostą postać

$$\mathcal{D}_{\alpha}^{-} \chi_{-} = \frac{\partial}{\partial \theta^{\alpha}} \chi_{-} = 0, \quad (\text{VIII.42})$$

a więc χ_{-} nie zależy od θ , czyli ma następującą strukturę

$$\chi_{-}(x, \bar{\theta}) = A_{-}(x) + \psi_{- \alpha}(x) \theta^{\alpha} + (\theta^{\alpha} \theta_{\dot{\alpha}}) F_{-}(x) \quad (\text{VIII.43})$$

Widzimy więc, że superpole $\chi_{-}(x, \bar{\theta})$ jest określone tylko przez trzy niezależne pola A_{-} , F_{-} , ψ_{-} - dwoma polami skalarnymi i jednym polem spinorowym, jeśli spełnia ono warunek (VIII.35). Wyjściowa reprezentacja (VIII.21) przyjmuje następującą nieprzywiedlną postać

$$\chi(x, \theta, \bar{\theta}) = e^{-i \bar{\theta} \sigma^{\mu} \theta} \chi_{-}(x, \bar{\theta}) =$$

$$= A_-(x) + \bar{\psi}_-(x) \bar{\theta} + (\bar{\theta} \bar{\theta}) F_-(x) - i \bar{\theta} \sigma^\mu \theta \partial_\mu A_-(x) \quad (\text{VIII.44})$$

$$+ \frac{1}{2} (\bar{\theta} \bar{\theta}) \partial_\mu \bar{\psi}(x) \sigma^\mu \theta + \frac{1}{4} (\theta \theta) (\bar{\theta} \bar{\theta}) \partial_\mu \partial^\mu A_-(x).$$

Uwzględniliśmy w tym wzorze, że $(\bar{\theta} \sigma^\mu \theta) \bar{\theta}^P = -\frac{1}{2} (\bar{\theta} \bar{\theta}) (\sigma^\mu \theta)^P$. Analogiczna sytuacja mamy i dla superpola χ_+ , gdzie w plus-reprezentacji otrzymujemy

$$D_\alpha^+ \chi_+ = -\frac{\partial}{\partial \theta^\alpha} \chi_+ = 0. \quad (\text{VIII.45})$$

Stąd wynika, że superpole χ_+ nie zależy od $\bar{\theta}$, czyli, że ma następującą strukturę

$$\chi_+(x, \theta) = A_+(x) + \theta^\alpha \psi_{+\alpha}(x) + (\theta^\alpha \theta_\alpha) F_+(x). \quad (\text{VIII.46})$$

W reprezentacji wyjściowej, polu $\chi(x, \theta)$ odpowiada pole

$$\chi(x, \theta, \bar{\theta}) = A_+(x) + \theta^\alpha \psi_{+\alpha}(x) + (\theta^\alpha \theta_\alpha) F_+(x) +$$

$$- i \bar{\theta} \sigma^\mu \theta \partial_\mu A_+(x) - \frac{1}{2} (\theta \theta) \bar{\theta} \sigma^\mu \partial_\mu \psi_+(x) + \quad (\text{VIII.47})$$

$$+ \frac{1}{4} (\theta \theta) (\bar{\theta} \bar{\theta}) \partial_\mu \partial^\mu A_+(x).$$

To superpole jest oczywiście też nieprzywiedlne. Superpola $\chi_+(x, \theta)$ (zależne tylko od θ) i $\chi_-(x, \bar{\theta})$ (zależne tylko od $\bar{\theta}$) nazywają się *superpolami chiralnymi* względem transformacji Lorentza.

Na zakończenie tego paragrafu zajmiemy się skonstruowaniem równań ruchu nieprzywiedlnego superpola chiralnego $\chi_+(x, \theta)$. W tym celu zbudujemy lagrangian zależny od pól A , F i ψ , przy czym będzie on zawierał tylko pierwsze pochodne tych pól względem współrzędnych x^μ . Będziemy żądali, by równanie pola dla ψ było rzędu pierwszego. Oznacza to, że pochodne $\partial_\mu \psi$ będą wchodziły do lagrangianu tylko liniowo.

Będziemy poszukiwali takiej postaci lagrangianu $\mathcal{L}(A, F, \psi)$, aby była ona niezmiennicza względem transformacji

$$\chi'(x, \theta, \bar{\theta}) = \mathcal{F}(\zeta, a) \chi(x, \theta, \bar{\theta}).$$

Zauważmy, że dla dowolnego superpola $\Phi_+(x, \theta, \bar{\theta})$ w plus-reprezentacji, składowe zawierające maksymalną liczbę zmiennych Grassmana θ i $\bar{\theta}$, tzn. $(\bar{\theta} \bar{\theta}) (\theta \theta) D(x)$, transformują się przy supertransformacjach jak

$$\delta D(x) = \delta_{\mu} P_{\theta}^{(4)} (-2i \bar{\zeta}^{\mu} \sigma^{\mu} \theta) \Phi_{+}(x, \theta, \bar{\theta}), \quad (\text{VIII. 48})$$

gdzie

$$P_{\theta}^{(4)} = \frac{1}{4} \frac{\partial}{\partial \theta^{\alpha}} \frac{\partial}{\partial \theta} \frac{\partial}{\partial \theta} \frac{\partial}{\partial \bar{\theta}^{\beta}} \frac{\partial}{\partial \bar{\theta}}.$$

Tak więc $\delta D(x)$ jest dywergencją pewnego pola wektorowego. Związek ten wynika z prawa transformacji superpola $\Phi(x, \theta, \bar{\theta})$

$$\begin{aligned} \delta \Phi_{+}(x, \theta, \bar{\theta}) &= G_{+} \Phi_{+}(x, \theta, \bar{\theta}) = \\ &= i \left(-1 \frac{\partial}{\partial \theta^{\alpha}} \zeta^{\alpha} + i \bar{\zeta}^{\alpha} \frac{\partial}{\partial \bar{\theta}^{\alpha}} - 2 \bar{\zeta}^{\mu} \sigma^{\mu} \theta \frac{\partial}{\partial \theta} \right) \Phi_{+}(x, \theta, \bar{\theta}), \end{aligned}$$

i z tego, że

$$D(x) = P_{\theta}^{(4)} \Phi_{+}(x, \theta, \bar{\theta}).$$

W przypadku superpola chiralnego składowe $(\theta \theta)F(x)$, zawierające maksymalną liczbę zmiennych θ , transformują się przy supertransformacjach zgodnie ze wzorem

$$\delta F(x) = \delta_{\mu} P_{\theta}^{(2)} (-2i \bar{\zeta}^{\mu} \sigma^{\mu} \theta) \Phi_{+}(x, \theta, \bar{\theta}),$$

gdzie

(VIII. 170)

$$P_{\theta}^{(4)} = \frac{1}{2} \frac{\partial}{\partial \theta^{\alpha}} \frac{\partial}{\partial \theta} \frac{\partial}{\partial \theta^{\alpha}}.$$

czyli $\delta F(x)$ jest też dywergencją pewnego pola wektorowego.

W związku z tym, przy budowaniu supernieziemniczego lagrangianu $\mathcal{L}(A, F, \psi)$ powinniśmy utworzyć różne kombinacje $\chi_{+}(x, \theta)$ i wyodrębnić z nich człony zawierające maksymalną liczbę zmiennych $\theta, \bar{\theta}$. Sumę takich kombinacji można przyjąć jako poszukiwany lagrangian superpola $\chi_{+}(x, \theta)$. Przy supertransformacjach lagrangian będzie się transformować jak

$$\delta \mathcal{L} = \delta_{\mu} B^{\mu}, \quad (\text{VIII. 50})$$

gdzie B^{μ} jest pewnym 4-wektorem. Najprostsza kombinacja daje lagrangian zaproponowany przez Wessa i Zumino, który ma postać

$$\mathcal{L} = -\frac{1}{2} P_{\theta}^{(4)} \chi_{+}(x, \theta) e^{-2i \bar{\theta} \sigma^{\mu} \theta} \chi_{+}^{*}(x, \theta) +$$

$$+ \frac{m}{2} P_{\theta}^{(2)} \chi_+^2(x, \theta) + \frac{m}{2} P_{\theta}^{(2)} \chi_+^{*2}(x, \theta) + \quad (\text{VIII.51})$$

$$+ \frac{2}{3} g P_{\theta}^{(2)} \chi_+^3(x, \theta) - \frac{2}{3} g P_{\theta}^{(2)} \chi_+^{*3}(x, \theta) ,$$

gdzie g jest stałą bezwymiarową grającą rolę stałej oddziaływania, a m jest dowolną stałą mającą wymiar masy. Uwzględnienie iloczynów wyższego rzędu niż trzy związane byłoby z wprowadzeniem stałych oddziaływania posiadających wymiar. W wyrażeniu (VIII.51) wchodzi operatory różniczkowania $P_{\theta}^{(2)}$ i $P_{\theta}^{(4)}$ względem zmiennych Grassmanna.

Pojęcie całki na algebrze Grassmanna wprowadził Berezin. Formalnie całka Grassmanna jest zdefiniowana następującymi formułami:

$$\int d\theta_{\alpha} = 0 \quad , \quad \int d\theta_{\beta} \theta^{\alpha} = \delta_{\beta}^{\alpha} . \quad (\text{VIII.52})$$

Dla funkcji $f(\theta, \bar{\theta})$ całkowanie w algebrze Grassmanna jest równoważne różniczkowaniu

$$\int d\theta_{\alpha} f(\theta, \bar{\theta}) = \frac{\partial}{\partial \theta^{\alpha}} f(\theta, \bar{\theta}) , \quad (\text{VIII.53})$$

$$\int d\bar{\theta}_{\alpha} f(\theta, \bar{\theta}) = \frac{\partial}{\partial \bar{\theta}^{\alpha}} f(\theta, \bar{\theta}) .$$

Całki wielokrotne definiujemy powtarzając całkowanie, czyli

$$\int d\theta_1 d\theta_2 f(\theta, \bar{\theta}) = \frac{\partial}{\partial \theta^1} \frac{\partial}{\partial \theta^2} f(\theta, \bar{\theta}) = P_{\theta}^{(2)} f(\theta, \bar{\theta}) , \quad (\text{VIII.54})$$

$$\int d\theta_1 d\theta_2 d\bar{\theta}_1 d\bar{\theta}_2 f(\theta, \bar{\theta}) = \frac{\partial}{\partial \theta^1} \frac{\partial}{\partial \theta^2} \frac{\partial}{\partial \bar{\theta}^1} \frac{\partial}{\partial \bar{\theta}^2} f(\theta, \bar{\theta}) = P_{\theta}^{(4)} f(\theta, \bar{\theta})$$

Różniczki $d\theta_{\alpha}$, $d\bar{\theta}_{\beta}$ są to antykomutujące między sobą różniczki Grassmanna, więc w wyniku otrzymujemy, że

$$d\theta_1 d\theta_2 = \frac{1}{2} \epsilon^{\alpha\beta} d\theta_{\alpha} d\theta_{\beta} = d^2\theta , \quad (\text{VIII.55})$$

$$d\bar{\theta}_1 d\bar{\theta}_2 = \frac{1}{2} \epsilon^{\dot{\alpha}\dot{\beta}} d\bar{\theta}_{\dot{\alpha}} d\bar{\theta}_{\dot{\beta}} = d^2\bar{\theta} .$$

Tak więc możemy napisać

$$\begin{aligned} \mathcal{L} = & \int d^4\theta \left\{ \frac{m}{4} \bar{\theta} \bar{\theta} \chi_+^2(x, \theta) - \frac{m}{4} \theta \theta \chi_+^{*2}(x, \theta) \right. \\ & - \frac{1}{2} \chi_+(x, \theta) e^{-2i\bar{\theta}\sigma^\mu\theta} \chi_+^*(x, \theta) + \\ & \left. + \frac{g}{3} ((\bar{\theta} \bar{\theta}) \chi_+^3(x, \theta) - (\theta \theta) \chi_+^{*3}(x, \theta)) \right\}, \end{aligned} \quad (\text{VIII.56})$$

gdzie $d^4\theta = d^2\theta d^2\bar{\theta}$. Wyrażenia w dwu pierwszych liniijkach opisują swobodne superpola, a wyrażenie w liniijkce trzeciej opisuje oddziaływanie pomiędzy polami A , F i ψ . Równanie ruchu otrzymane na podstawie tego lagrangianu ma postać

$$\frac{1}{4} \bar{\mathcal{D}}_{\dot{\alpha}} \bar{\mathcal{D}}^{\dot{\alpha}} \bar{\chi}_+(x, \theta, \bar{\theta}) = m \chi_+(x, \theta) + 2g \chi_+^2(x, \theta), \quad (\text{VIII.57})$$

gdzie

$$\bar{\chi}_+(x, \theta, \bar{\theta}) = e^{-2i\bar{\theta}\sigma^\mu\theta} \chi_+^*(x, \theta).$$

Równanie to ma postać jawnie kowariantna, bo wchodzi w nie pochodne kowariantne pola $\bar{\chi}_+$. Równania ruchu dla pól A , F , i ψ wchodzących w skład superpola $\chi_+(x, \theta)$ mają następującą postać

$$\begin{aligned} i \partial_{\mu} \bar{\psi} \sigma^{\mu} - m \psi &= 4g A \psi, \\ \partial_{\mu} \partial^{\mu} A^* &= m F + 4g A F - g(\psi\psi), \\ F^* &= -mA - 2gA^2, \end{aligned} \quad (\text{VIII.58})$$

a równania do nich sprzężone w sposób zespolony

$$\begin{aligned} -i \sigma_{\mu} \partial^{\mu} \bar{\psi} - m \bar{\psi} &= 4g A^* \bar{\psi}, \\ \partial_{\mu} \partial^{\mu} A &= m F^* + 4g A^* F^* - g(\bar{\psi} \bar{\psi}), \\ F &= -mA^* - 2gA^{*2}. \end{aligned} \quad (\text{VIII.59})$$

Dla pola swobodnego ($g = 0$) równania (VIII.58) przyjmują postać

$$\begin{aligned} i \partial_{\mu} \bar{\psi} \sigma^{\mu} - m \psi &= 0, \\ \partial_{\mu} \partial^{\mu} A^* &= m F, \\ F^* &= mA. \end{aligned} \quad (\text{VIII.60})$$

Pierwsze równanie (VIII.60) jest równaniem dla spinora Majorany, a równania drugie i trzecie dają równanie Kleina-Gordona dla zespolonego pola skalarnego.

Równania (VIII.58) i (VIII.59) można otrzymać bezpośrednio z lagrangianu (VIII.56), jeśli wykonać całkowania po θ i $\bar{\theta}$. Mamy wtedy

$$\begin{aligned} \mathcal{L} = & 2 \partial_{\mu} A \partial^{\mu} A^{*} + 2m(AF + A^{*}F^{*}) + 2FF^{*} - \frac{1}{2} \partial_{\mu} \bar{\psi} \sigma^{\mu} \psi + \\ & + \frac{1}{2} \bar{\psi} \sigma^{\mu} \partial_{\mu} \psi - \frac{m}{2}(\psi \psi) + \frac{m}{2}(\bar{\psi} \bar{\psi}) + \\ & + 2g(2A^2F + 2A^{*2}F^{*} - A(\psi \psi) + A^{*}(\bar{\psi} \bar{\psi})). \end{aligned} \quad (\text{VIII.61})$$

Rozpatrywany przez nas przykład lagrangianu dla skalarnego chiralnego superpola $\chi_{+}(x,0)$ pokazuje, że zadanie superniezmienniczości daje w wyniku nie tylko połączenie pól spinorowych i skalarnych w jedno superpole, ale że ponadto samooddziaływanie pola skalarnego i jego oddziaływanie z polem spinorowym opisane jest tą samą bezwymiarową stałą oddziaływania, a masy cząstek skalarnych i fermionowych są identyczne.

ROZDZIAŁ IX

PODSTAWOWE POJĘCIA TEORII CECHOWANIA

Lagrangian swobodnego pola elektronowo-pozytonowego

$$\mathcal{L}(\psi, \bar{\psi}) = \frac{1}{2} \bar{\psi} (i \gamma^{\mu} \partial_{\mu} - m) \psi - \frac{1}{2} \bar{\psi} (i \gamma^{\mu} \partial_{\mu} + m) \psi,$$

jest niezmienniczy względem transformacji fazy funkcji pola

$$\psi(x) \rightarrow \psi'(x) = e^{-i\varepsilon} \psi(x); \quad \bar{\psi}(x) \rightarrow \bar{\psi}'(x) = e^{i\varepsilon} \bar{\psi}(x), \quad (\text{IX.1})$$

gdzie ε jest stałą rzeczywistą, czyli względem transformacji z grupy abelowej $U(1)$. W przypadku infinitezymalnych transformacji

$$\delta\psi = \psi'(x) - \psi(x) = -i\epsilon\psi(x), \quad (IX.2)$$

$$\delta\bar{\psi} = \bar{\psi}'(x) - \bar{\psi}(x) = i\epsilon\bar{\psi}(x).$$

Ponieważ $\psi(x)$ jest funkcją czasu i współrzędnych przestrzennych, można by zażądać, aby lagrangian był niezmienniczy względem bardziej ogólnego przekształcenia, a mianowicie takiej transformacji (IX.1), w której ϵ byłby też funkcją czasu i współrzędnych przestrzennych. Celem tego jednak nie da się osiągnąć używając tylko funkcji ψ bez wprowadzenia dodatkowego pola. W przypadku pola elektronowo-pozytonowego, tym dodatkowym polem jest pole elektromagnetyczne $A_\mu(x)$. Żadamy dodatkowo, by pole A_μ przy transformacjach (IX.1) pola ψ podlegało tzw. transformacji cechowania

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \delta_\mu\omega(x), \quad (IX.3)$$

gdzie $\epsilon(x) = -e\omega(x)$. Niezmienniczość lagrangianu zostanie osiągnięta, jeśli występującą w nim zwykłą pochodną $\partial_\mu\psi(x)$ zastąpimy pochodną kowariantną

$$\partial_\mu\psi(x) \rightarrow D_\mu\psi(x) = (\partial_\mu - ieA_\mu(x))\psi(x), \quad (IX.4)$$

i wtedy

$$\begin{aligned} \mathcal{L}(\psi, \bar{\psi}, A) = & \frac{1}{2} \psi [i\gamma^\mu(\partial_\mu - ieA_\mu) - m]\psi \\ & - \frac{1}{2} \bar{\psi} [i\gamma^\mu(\partial_\mu + ieA_\mu) + m]\bar{\psi}. \end{aligned} \quad (IX.5)$$

Fakt ten ma miejsce dlatego, że przy transformacjach (IX.1) i (IX.4) pochodna kowariantna $D_\mu\psi(x)$ transformuje się tak samo jak pole $\psi(x)$. Porównajmy:

$$\begin{aligned} \partial_\mu\psi'(x) = \partial_\mu(e^{-i\epsilon(x)}\psi(x)) &= (-i\partial_\mu\epsilon(x)\psi(x) + \partial_\mu\psi(x))e^{-i\epsilon(x)} \\ &= e^{-i\epsilon(x)}(\partial_\mu + ie\delta_\mu\omega)\psi(x). \end{aligned} \quad (IX.6)$$

$$D_\mu\psi'(x) = -(\partial_\mu - ie(A_\mu(x) + \delta_\mu\omega))\psi(x)e^{-i\epsilon(x)}$$

$$\begin{aligned}
&= e^{-i\varepsilon(x)} (\partial_\mu + i e \delta_\mu^\omega - i e A_\mu - i e \delta_\mu^\omega) \psi(x) = \\
&= e^{-i\varepsilon(x)} (\partial_\mu - i e A_\mu) \psi(x) = e^{-i\varepsilon(x)} D_\mu \psi(x)
\end{aligned}$$

Pokazemy teraz jak pomysł zastosowania pola elektromagnetycznego w celu zbudowania niezmienniczego lagrangianu pola $\psi(x)$ względem transformacji fazowych ze zmienną fazą $\varepsilon(x)$ można uogólnić na przypadek dowolnych pól $\psi(x)$ posiadających pewną wewnętrzną symetrię.

Rozpatrzmy infinitezymalne transformacje, dla których przy transformacji pola

$$\begin{aligned}
\psi &\rightarrow \psi' = U\psi, \\
U &= 1 - i\varepsilon^a T_a, \quad (IX.7)
\end{aligned}$$

$$\delta\psi = \psi' - \psi = -i\varepsilon^a T_a \psi(x), \quad \varepsilon^a = -g\omega^a,$$

i gdzie $\varepsilon^a(\omega^a)$ są to nieskończenie małe wielkości, a macierze T_a spełniają relacje komutacyjne

$$[T_a, T_b] = i f_{ab}^c T_c, \quad (IX.8)$$

Bez ograniczenia ogólności możemy założyć, że pole $\psi(x)$ jest rzeczywiste. W tym przypadku macierze T_a są czysto urojone.

Warunek niezmienniczości lagrangianu $\mathcal{L}(\psi, \partial_\mu \psi)$ względem transformacji (IX.7) ma postać

$$\frac{\partial \mathcal{L}}{\partial \psi} T_a \psi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} T_a \partial_\mu \psi = 0, \quad (IX.9)$$

wynika stąd, a również i z równań Eulera-Lagranga,

$$\frac{\partial \mathcal{L}}{\partial \psi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} = 0,$$

prawo zachowania uogólnionych strumieni J_a^μ , ponieważ

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} T_a \psi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} T_a \partial_\mu \psi = \quad (IX.10)$$

$$\partial_\mu \left[\frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} T_a \psi \right] = -\frac{1}{i} \partial_\mu J_a^\mu,$$

gdzie przyjmuje się, że

$$J_a^\mu = -i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} T_a \psi . \quad (\text{IX.11})$$

Niech teraz wielkości $\epsilon^a = \epsilon^a(x)$. Wtedy

$$\delta \psi = -i \epsilon^a(x) T_a \psi(x), \quad (\text{IX.12})$$

$$\delta \partial_\mu \psi = -i \epsilon^a(x) T_a \partial_\mu \psi - i \partial_\mu \epsilon^a(x) T_a \psi ,$$

Wariacja lagrangianu w postaci niezmienniczej dla $\epsilon^a = \text{const.}$ ma teraz postać

$$\delta \mathcal{L} = -i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} T_a \psi \partial_\mu \epsilon^a(x) = J_a^\mu \partial_\mu \epsilon^a(x) , \quad (\text{IX.13})$$

i jest różna od zera. Widać, że złamanie niezmienniczości lagrangianu wynika z faktu, że pochodne $\partial_\mu \psi$ i sama funkcja ψ transformują się różnie. Niezmienniczość można odtworzyć, jeśli wprowadzić dodatkowe pola wektorowe $A_\mu^a(x)$, liczba których jest równa liczbie generatorów T_a , i za ich pomocą zdefiniować pochodne kowariantne pola $\psi(x)$:

$$D_\mu \psi(x) = (\partial_\mu - ig A_\mu^a(x) T_a) \psi(x) . \quad (\text{IX.14})$$

Prawo transformacji tej pochodnej kowariantnej jest takie samo, jak prawo transformacji funkcji pola

$$\delta D_\mu \psi(x) = -i \epsilon^a(x) T_a D_\mu \psi(x) . \quad (\text{IX.15})$$

Pola $A_\mu^a(x)$ nazywają się *polami cechowania*. Aby zachodził związek (IX.15) pola cechowania A_μ^a przy transformacji (IX.12) powinny podlegać następującej transformacji

$$\delta A_\mu^a(x) = \epsilon^c(x) f_{cb}^a A_\mu^b(x) - \frac{1}{g} \partial_\mu \epsilon^a(x) . \quad (\text{IX.16})$$

Rzeczywiście, z (IX.14) mamy, że

$$\begin{aligned} \delta D_\mu \psi(x) &= \partial_\mu \delta \psi(x) - ig \delta A_\mu^a(x) T_a \psi(x) - ig A_\mu^a(x) T_a \delta \psi(x) \\ &= -i \epsilon^a(x) T_a \partial_\mu \psi(x) - i \partial_\mu \epsilon^a(x) T_a \psi - ig \delta A_\mu^a T_a \psi \\ &\quad - ig A_\mu^a(x) T_a (-i \epsilon^b(x) T_a) \psi(x) \\ &= -i \epsilon^a T_a \partial_\mu \psi - i \partial_\mu \epsilon^a(x) T_a \psi - ig \delta A_\mu^a T_a \psi \\ &= -g \epsilon^a A_\mu^b [T_b, T_a] \psi - g \epsilon^a A_\mu^b T_a T_b \psi \end{aligned}$$

$$= -i \epsilon^a T_a (\partial_\mu - i g A_\mu^b T_b) \psi(x) - i g (\delta A_\mu^a + \frac{1}{g} \partial_\mu \epsilon^a + \epsilon^c A_\mu^b f_{bc}^a) T_a \psi,$$

skąd już bezpośrednio wynika (IX.16). Transformacje (IX.12) i (IX.16) nazywamy transformacjami cechowania.

Zwróćmy uwagę, że we wzorze (IX.16) w odróżnieniu od analogicznego wzoru dla transformacji pola elektromagnetycznego, występuje składnik zawierający stałe strukturalne f_{cb}^a . Związane jest to z faktem, że rozważana obecnie grupa transformacji jest grupa nieabelowa.

Z przeprowadzonych rozważań wynika, że zapewnimy sobie niezmienniczość lagrangianu jeśli zastąpimy pochodne $\partial_\mu \psi$ pochodnymi kowariantnymi $D_\mu \psi$. Tak więc lagrangian pola ψ oddziaływującego z polami cechowania A^a jest określony wzorem

$$\mathcal{L}(\psi, \partial_\mu \psi, A_\mu^a) = \mathcal{L}(\psi, D_\mu \psi), \quad (\text{IX.17})$$

gdzie $\mathcal{L}(\psi, \partial_\mu \psi)$ jest lagrangianem pola ψ . Wykazemy teraz niezmienniczość lagrangianu.

Ze względu na to, że wzór (IX.9) jest prawdziwy dla dowolnego ψ , to będzie i prawdziwy dla $D\psi$:

$$\frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a D_\mu \psi + \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial \psi} T_a \psi = 0,$$

skąd wykorzystując (IX.15) i (IX.12) mamy

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} \delta D_\mu \psi + \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial \psi} \delta \psi = 0,$$

a ten związek wykazuje niezmienniczość lagrangianu ze względu na transformacje (IX.12) i (IX.15).

Gęstość uogólnionych strumieni J_a^μ w obecności pól cechowania określamy następującym wzorem

$$J_a^\mu = -i \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a \psi. \quad (\text{IX.18})$$

Jeśli uwzględnimy wzór (IX.14), to otrzymamy

$$j_a^\mu = \frac{1}{g} \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial A_\mu^a} \quad (\text{IX.18})'$$

Rzeczywiście,

$$\frac{\partial \mathcal{L}(\psi, D\psi)}{\partial A_\mu^a} = \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} \frac{(D_\nu \psi)}{\partial A_\mu^a} \Rightarrow \frac{\partial \mathcal{L}}{\partial (D_\nu \psi)} = \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial A_\mu^a} \left[\frac{\partial (D_\nu \psi)}{\partial A_\mu^a} \right]^{-1},$$

$$\text{ale } D_\nu \psi = (\partial_\nu - ig A_\mu^a(x) T_a) \psi,$$

$$\text{czyli } \frac{\partial (D_\nu \psi)}{\partial A_\mu^a} = -ig T_a \psi \delta_\nu^\mu,$$

więc

$$j_a^\mu = -1 \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a \psi = -\frac{1}{ig T_a \psi} (-1) \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (A_\mu^a \psi)} T_a \psi = \frac{1}{g} \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (A_\mu^a \psi)}$$

Ponieważ lagrangian jest niezmienniczy względem transformacji (IX.12) i (IX.15), to możemy napisać

$$\mathcal{L}(\psi', D'\psi') = \mathcal{L}(\psi, D\psi),$$

gdzie

$$D'_\mu \psi' = \partial_\mu \psi' - ig A_\mu^a T_a \psi',$$

a ψ' i A' są przekształconymi polami ψ i A . Tak więc

$$\frac{\partial \mathcal{L}(\psi', D'\psi')}{\partial (D'_\mu \psi')} = \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\nu \psi)} \frac{\partial D_\nu \psi}{\partial (D'_\mu \psi')}$$

Dla nieskończenie małych transformacji, zgodnie z (IX.15)

$$D'_\mu \psi' = (1 - i\epsilon^a T_a) D_\mu \psi, \quad D_\mu \psi = (1 + i\epsilon^a T_a) D'_\mu \psi',$$

czyli

$$\frac{\partial \mathcal{L}(\psi', D'\psi')}{\partial (D'_\mu \psi')} = \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} (1 + i\epsilon^a T_a),$$

a więc

$$\delta \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\nu \psi)} = i \epsilon^a \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a \quad (IX.19)$$

Wariacja uogólnionego strumienia będzie więc miała postać

$$\begin{aligned} \delta J_a^\mu &= -i \delta \left[\frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a \psi \right] \\ &= -i \delta \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a \psi - i \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a \delta \psi \end{aligned} \quad (IX.20)$$

$$\begin{aligned} &= \epsilon^b \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_b T_a \psi - \epsilon^b \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} T_a T_b \psi \\ &= \epsilon^b \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} [T_b, T_a] \psi = -\epsilon^b f_{ba}^c J_c^\mu, \end{aligned}$$

gdzie skorzystaliśmy ze wzorów (IX.8), (IX.12), (IX.18) i (IX.19).

Ostatnio traktowaliśmy pole $\psi(x)$ jako pole rzeczywiste, co nie było istotnym ograniczeniem. W przypadku zespolonego pola Diraca, zamiast pola zespolonego możemy rozpatrzeć dwa pola rzeczywiste $\varphi_1(x)$ i $\varphi_2(x)$: $\psi(x) = \varphi_1(x) + i\varphi_2(x)$. Transformacji fazowej (IX.2) będzie w tym przypadku odpowiadać transformacja pola dwuskładnikowego

$$\varphi(x) = \begin{pmatrix} \varphi_1(\kappa) \\ \varphi_2(\kappa) \end{pmatrix}; \quad \delta\varphi(x) = -i\epsilon T\varphi(x), \quad (IX.21)$$

gdzie $\delta\varphi(x) = \begin{pmatrix} \delta\varphi_1(\kappa) \\ \delta\varphi_2(\kappa) \end{pmatrix}$, a macierz T rzędu 2 ma postać

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad (IX.22)$$

Ta transformacja pola rzeczywistego $\varphi(x)$ odpowiada ogólnemu prawu

transformacji (IX.7). Zdarza się jednakże dosyć często, że wygodniej jest posługiwać się lagrangianem zbudowanym z funkcji zespolonych. W tym przypadku wszystkie poprzednie wzory są słuszne, jeśli przez ψ rozumieć podwojona liczbę pól ψ, ψ^* , dla których prawo transformacji ma postać

$$\delta\psi(x) = -i\varepsilon^a T_a \psi(x), \quad \delta\psi^*(x) = i\varepsilon^a T_a^* \psi^*(x). \quad (\text{IX.23})$$

Widać, że polu $\psi^*(x)$ odpowiada macierz $-T_a^*$. Tak więc, w przypadku pola zespolonego strumień uogólniony będzie miał postać

$$J_a^\mu = -i \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} T_a \psi + i \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)} T_a^* \psi^*.$$

W przypadku pola zespolonego grupami wewnętrznego symetrii na ogół są grupy $SU(n)$. Generatorami tej grupy są macierze hermitowskie, a więc ostatnią równość można zapisać

$$J_a^\mu = -i \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} T_a \psi + i \psi^* T_a \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi^*)}.$$

Teraz będziemy chcieli zbudować lagrangian dla samych pól cechowania. Lagrangian ten będzie zależeć od pól A_μ^a i od ich pochodnych $\partial_\mu A_\mu^a$. Powinien on być niezmienniczy ze względu na transformacje (IX.16). Pojawia się więc zagadnienie znalezienia niezmienników tej transformacji. Problem ten rozwiążemy na przykładzie grupy $SU(n)$. Grupa $SU(n)$ to grupa 'szczególnie unitarnych' macierzy $n \times n$, o wyznaczniku równym jedności

$$U U^\dagger = U^\dagger U = 1.$$

Jest to grupa nieabelowa. Macierz unitarna z $\det U = 1$ można zapisać w postaci

$$U = e^{iH}, \quad (\text{IX.24})$$

gdzie H jest macierzą hermitowską ze śladem równym zeru: $\text{Tr } H = 0$. Niezmienniczość lagrangianu względem transformacji grupy $SU(n)$ nazywamy *niezmienniczością unitarną*.

Macierz hermitowska rzędu n jest zdefiniowana przez n^2 parametrów, ale elementy grupy $SU(n)$ są określone przez $n^2 - 1$ rzeczywistych parametrów, gdyż $\text{Tr } H = 0$. Dla infinitezymalnych

transformacji macierz \mathcal{U} można przedstawić w postaci

$$\mathcal{U} = 1 - i\varepsilon^a T_a, \quad (\text{IX.25})$$

gdzie ε^a są nieskończenie małymi parametrami ($a = 1, 2, \dots, n^2 - 1$), a T_a są macierzami hermitowskimi rzędu n ze śladem równym zeru (ponieważ $\det \mathcal{U} = 1$).

$$T_a^\dagger = T_a, \quad \text{Tr } T_a = 0.$$

Macierze T_a są generatorami grupy $SU(n)$ i spełniają związki komutacyjne

$$[T_a, T_b] = if_{ab}^c T_c. \quad (\text{IX.26})$$

Jeśli przy nieskończenie małych transformacjach grupy $SU(n)$ pewne wielkości H transformują się zgodnie ze wzorem

$$\delta H_a = -\varepsilon^b f_{ba}^c H_c, \quad (\text{IX.27})$$

to wielkości te nazywamy *składowymi kowariantnego $SU(n)$ -wektora*. Tak więc np. strumień J_a^μ (IX.20) jest kowariantnym $SU(n)$ -wektorem. Wielkości G^a nazywamy *składowymi kontrawariantnego $SU(n)$ -wektora* jeśli transformują się one zgodnie ze wzorem

$$\delta G^a = \varepsilon^b f_{bc}^a G^c. \quad (\text{IX.28})$$

Oprócz wektorów można wprowadzić *$SU(n)$ -tensory* drugiego i wyższych rzędów. Jako przykład można podać składowe mieszanego $SU(n)$ -tensora drugiego rzędu, które transformują się zgodnie ze wzorem

$$\delta Q_b^a = -\varepsilon^c f_{cb}^d Q_d^a + \varepsilon^c f_{cd}^a Q_b^d. \quad (\text{IX.29})$$

Tak więc tensorem mieszanym jest na przykład wielkość $G^a H_b$. Zgodnie ze wzorem (IX.29) wielkość $\delta Q_a^a = 0$, a więc Q_a^a jest *$SU(n)$ -skalarem*. Stałe strukturalne f_{ab}^c można rozpatrywać jako *$SU(n)$ -tensor trzeciego rzędu*, który transformuje się zgodnie ze wzorem

$$\delta f_{ab}^c = \varepsilon^d (f_{d1}^c f_{ab}^1 - f_{da}^1 f_{1b}^c - f_{db}^1 f_{a1}^c). \quad (\text{IX.30})$$

Wykorzystując tożsamość Jacobiego

$$f_{ab}^c f_{cd}^e + f_{bd}^c f_{ca}^e + f_{da}^c f_{cb}^e = 0, \quad (\text{IX.31})$$

otrzymujemy, że

$$\delta f_{ab}^c = 0. \quad (IX.32)$$

Ponieważ f_{ab}^c tworzą SU(n)-tensor, to wielkości

$$g_{ab} = \frac{1}{n} f_{ac}^d f_{db}^c, \quad (IX.33)$$

tworzą kowariantny symetryczny SU(n)-tensor rzędu drugiego, za pomocą którego można podnosić i opuszczać indeksy SU(n)-tensorów.

Rozpatrzmy przypadek gdy nieskończenie małe parametry ε^a we wzorze (IX.25) są funkcjami x : $\varepsilon^a = \varepsilon^a(x)$. W tym przypadku pochodne $\partial_{\mu} H_a$, $\partial_{\mu} G^a$, $\partial_{\mu} Q_b^a$ już nie będą SU(n)-tensorami, gdyż transformują się one zgodnie z nowymi wzorami, innymi niż (IX.27), (IX.28) i (IX.29), np.:

$$\delta \partial_{\mu} H_a = -\varepsilon^b f_{ba}^c \partial_{\mu} H_c - \partial_{\mu} \varepsilon^b f_{ba}^c H_c, \quad (IX.34)$$

$$\delta \partial_{\mu} G^a = \varepsilon^b f_{bc}^a \partial_{\mu} G^c + \partial_{\mu} \varepsilon^b f_{bc}^a G^c.$$

Wprowadźmy pochodne kowariantne D_{μ} SU(n)-wektorów

$$D_{\mu} H_a = \partial_{\mu} H_a + gf_{ab}^c A_{\mu}^b H_c, \quad (IX.35)$$

$$D_{\mu} G^a = \partial_{\mu} G^a + gf_{bc}^a A_{\mu}^b G^c.$$

Zbadajmy ich prawo transformacji, np dla $D_{\mu} H_a$:

$$\begin{aligned} \delta D_{\mu} H_a &= \delta \partial_{\mu} H_a + gf_{ab}^c \delta A_{\mu}^a H_c + gf_{ab}^c A_{\mu}^b \delta H_c \\ &= -\varepsilon^b f_{ba}^c \partial_{\mu} H_c - \partial_{\mu} \varepsilon^b f_{ba}^c H_c \\ &\quad + gf_{ab}^c (\varepsilon^d f_{dh}^b A_{\mu}^h - \frac{1}{g} \partial_{\mu} \varepsilon^b) H_c + gf_{ab}^c A_{\mu}^b (-\varepsilon^d f_{dc}^h) H_h \end{aligned}$$

$$\begin{aligned}
&= -\epsilon^b f_{ba}^c \partial_\mu H_c - \partial_\mu \epsilon^b f_{ba}^c H_c + g f_{ab}^c \epsilon^d f_{dh}^b A_\mu^h H_c + \\
&\quad - \partial_\mu \epsilon^b f_{ab}^c H_c - g \epsilon^d f_{ab}^c f_{dc}^h A_\mu^b H_h \\
&= -\epsilon^b f_{ba}^c \partial_\mu H_c - g \epsilon^d A_\mu^l H_c (f_{dl}^b f_{ba}^c + f_{la}^b f_{bd}^c) .
\end{aligned}$$

Wykorzystując (IX.31) otrzymujemy

$$\delta D_\mu H_a = -\epsilon^b f_{ba}^c D_\mu H_c \quad (\text{IX.36})$$

a więc $D_\mu H_a$ transformuje się jak $SU(n)$ -wektor. Podobnie jest z D_μ^a i z pochodnymi kowariantnymi tensorów wyższych rzędów, np. pochodna kowariantna $SU(n)$ -tensora drugiego rzędu Q_a^b ma postać

$$D_\mu Q_a^b = \partial_\mu Q_a^b + g f_{ac}^d A_\mu^c Q_d^b + g f_{cd}^b A_\mu^c Q_a^d \quad (\text{IX.37})$$

Pochodna kowariantna $SU(n)$ -tensora f_{ab}^c jest równa zero

$$D_\mu f_{ab}^c = 0 \quad (\text{IX.38})$$

Pole cechowania A_μ^a przy zmiennym $\epsilon^a(x)$ nie jest $SU(n)$ -wektorem. Pokażemy teraz jak się buduje $SU(n)$ -wektory. Jeśli G^a jest pewnym kowariantnym $SU(n)$ -wektorem, to

$$D_\nu (D_\mu G^a) = \partial_\nu D_\mu G^a + g f_{d1}^a A_\nu^d D_\mu G^1$$

Podstawiając jeszcze raz (IX.35) otrzymujemy

$$\begin{aligned}
D_\nu D_\mu G^a &= \partial_\nu \partial_\mu G^a + g f_{bc}^a \partial_\nu A_\mu^b Q^c + g f_{bc}^a A_\mu^b \partial_\nu Q^c \\
&\quad + g f_{d1}^a A_\nu^d \partial_\mu Q^1 + g^2 f_{d1}^a A_\nu^d f_{bc}^1 A_\mu^b Q^c .
\end{aligned}$$

Wynika stąd, że

$$\begin{aligned}
&(D_\nu D_\mu - D_\mu D_\nu) G^a = \\
&= g f_{bc}^a G^c (\partial_\nu A_\mu^b - \partial_\mu A_\nu^b) + g^2 (f_{bc}^1 f_{d1}^a + f_{cd}^1 f_{b1}^a) A_\nu^d A_\mu^b Q^c .
\end{aligned}$$

Jeśli wykorzystamy (IX.31) to

$$(D_\nu D_\mu - D_\mu D_\nu)G^a = -g F_{\mu\nu}^b f_{bc}^a G^c, \quad (\text{IX.39})$$

gdzie

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{bc}^a A_\mu^b A_\nu^c. \quad (\text{IX.40})$$

Lewa strona (IX.39) jest SU(n)-wektorem, wielkość $f_{bc}^a G^c$ jest mieszanym SU(n)-tensorem rzędu drugiego, więc wielkość $F_{\mu\nu}^a$ jest kowariantnym SU(n)-wektorem. Tak więc, jeśli $F_{\mu\nu}^a \neq 0$, to różniczkowanie kowariantne jest nieprzemienne, ale spełnia tożsamość

$$D_\rho F_{\mu\nu}^a + D_\mu F_{\nu\rho}^a + D_\nu F_{\rho\mu}^a = 0. \quad (\text{IX.41})$$

Różniczkowanie kowariantne funkcji pola jest też nieprzemienne.

Korzystając z (IX.14) mamy

$$\begin{aligned} D_\nu D_\mu \psi &= \partial_\nu D_\mu \psi - ig A_\nu^b T_b D_\mu \psi \\ &= \partial_\nu \partial_\mu \psi - ig \partial_\nu (A_\mu^a T_a \psi) - ig A_\nu^b T_b \partial_\mu \psi - g^2 A_\nu^b T_b A_\mu^a T_a \psi. \end{aligned}$$

Otrzymujemy stąd, że

$$(D_\nu D_\mu - D_\mu D_\nu)\psi = -ig(\partial_\nu A_\mu^a - \partial_\mu A_\nu^a)T_a \psi - g^2(A_\nu^b A_\mu^a)[T_b, T_a]\psi.$$

Jeśli wykorzystamy (IX.26), to

$$(D_\mu D_\nu - D_\nu D_\mu)\psi = ig F_{\mu\nu}^a T_a \psi. \quad (\text{IX.42})$$

Oznaczmy przez \mathcal{L}_0 lagrangian zbudowany tylko z pól cechowania $A_\mu^a(x)$ i z ich pierwszych pochodnych. Z warunku niezmienniczości lagrangianu ze względu na transformacje cechowania i transformacje Lorentza wynika, że powinien być on pewną funkcją wielkości

$$-\frac{1}{4} g_{ab} F_{\mu\nu}^a F^{b\mu\nu} \equiv -\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu},$$

k która w związku z przedstawionymi powyżej rozważaniami jest niezmiennikiem ze względu na transformacje z grupy SU(n) i ze względu na transformacje Lorentza. Najprostszy lagrangian spełniający zadane warunki będzie miał postać

$$\mathcal{L}_0(A_\mu^a, \partial_\mu A_\mu^a) = -\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu}. \quad (\text{IX.43})$$

Zgodnie z (IX.17) pełny lagrangian jest postaci

$$\mathcal{L} = \mathcal{L}(\psi, D\psi) + \mathcal{L}_0(A, \partial A) . \quad (\text{IX. 44})$$

Równanie ruchu pola ψ jest następujące

$$\partial_\mu \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} - \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial \psi} = -1 g \frac{\partial \mathcal{L}(\psi, D\psi)}{\partial (D_\mu \psi)} A_\mu^a T_a , \quad (\text{IX. 45})$$

ponieważ

$$\begin{aligned} \left. \frac{\partial \mathcal{L}}{\partial \psi} \right|_{\partial \psi = \text{const.}} &= \frac{\partial \mathcal{L}}{\partial \psi} + \frac{\partial \mathcal{L}}{\partial (D_\mu \psi)} \frac{\partial (D_\mu \psi)}{\partial \psi} \\ &= \frac{\partial \mathcal{L}}{\partial \psi} - 1 g \frac{\partial \mathcal{L}}{\partial (D_\mu \psi)} A_\mu^a T_a \end{aligned}$$

$$\left. \frac{\partial \mathcal{L}}{\partial (\partial_\mu \psi)} \right|_{\psi = \text{const.}} = \frac{\partial \mathcal{L}}{\partial (D_\mu \psi)} .$$

Równanie ruchu pola A_μ^a przyjmuje następującą postać

$$\partial_\mu \frac{\partial \mathcal{L}_0(A, \partial A)}{\partial (\partial_\mu A_\nu^a)} - \frac{\partial \mathcal{L}_0(A, \partial A)}{\partial A_\nu^a} = g J_a^\nu . \quad (\text{IX. 46})$$

Policzmy teraz wariacje lagrangianu związane z wariacjami δA_μ^a i A_μ^a :

$$\begin{aligned} \delta_{\partial A_\nu^a} \mathcal{L}_0 &= - F_a^{\mu\nu} \delta \partial_\mu A_\nu^a , \\ \delta_{A_\nu^a} \mathcal{L}_0 &= - \frac{1}{2} g F_a^{\mu\nu} f_{bc}^a (\delta A_\nu^b A_\nu^c + A_\mu^b \delta A_\nu^c) \\ &= - g F_a^{\mu\nu} f_{bc}^a A_\nu^c \delta A_\mu^b . \end{aligned}$$

Wynika stąd, że

$$\frac{\partial \mathcal{L}_0(A, \partial A)}{\partial (\partial_\mu A_\nu^a)} = F_a^{\mu\nu} ,$$

$$\frac{\delta \mathcal{L}_o(A, \partial A)}{\delta A_\nu^a} = -g F_b^{\nu\mu} f_{ac}^b A_\nu^c .$$

Wstawiając te wyrażenia do (IX.46) otrzymujemy

$$\partial_\mu F_a^{\mu\nu} + g F_b^{\nu\mu} f_{ac}^b A_\nu^c = g J_a^\nu . \quad (\text{IX.47})$$

Wzór ten można zapisać w postaci bardziej zwartej, jeśli oznaczyć,

$$I_a^\nu = J_a^\nu - f_{ac}^b A_\nu^c F_b^{\nu\mu} . \quad (\text{IX.48})$$

W tym przypadku otrzymujemy, że

$$\partial_\mu F_a^{\mu\nu} = g I_a^\nu . \quad (\text{IX.49})$$

Tensor $F_a^{\mu\nu}$ jest antysymetryczny ze względu na permutację wskaźników μ i ν , więc

$$\partial_\nu I_a^\nu = 0 . \quad (\text{IX.50})$$

W ten sposób otrzymaliśmy prawo zachowania uogólnionego strumienia i odpowiadającemu mu ładunkowi

$$Q_a = \int d^3x I_a^0(x) . \quad (\text{IX.51})$$

DODATEK A

Podamy teraz kilka podstawowych definicji i faktów, które mogą być przydatne czytelnikowi zaznajomionemu z teorią grup tylko pobieżnie.

Zbiór G nazywamy grupą, jeśli jego elementy spełniają następujące aksjomaty:

- I. jeśli $a \in G$ i $b \in G$ i $a b = c$, to $c \in G$,
(prawo kompozycji);
- II. $(a b)c = a(b c)$, (prawo łączności);
- III. $a e = a$, (istnienie prawego elementu jednostkowego e);
- IV. $a x = e$, (istnienie prawego elementu odwrotnego x do każdego elementu $a \in G$).

Przestrzeń topologiczna $T = (A, \{U\})$, to dowolny zbiór A z określoną topologią $\{U\}$, czyli taką rodziną podzbiorów zbioru A , ze spełnione są następujące aksjomaty:

- I. \emptyset i A należą do $\{U\}$;
- II. Jeśli $U_1 \in \{U\}$ i $U_2 \in \{U\}$, to $U_1 \cap U_2 \in \{U\}$;
- III. $\bigcup_1 U_1 \in \{U\}$, $1 = 1, 2, \dots, M$.

Elementy rodziny $\{U\}$ nazywamy zbiorami otwartymi przestrzeni T .

Odwzorowaniem homeomorficznym (topologicznym) dwu przestrzeni metrycznych $M_1 = (A_1, d_1)$ i $M_2 = (A_2, d_2)$ (d_1 jest odległością określoną na zbiorze A_1), nazywamy takie odwzorowanie $f: M_1 \rightarrow M_2$, ze:

- I. f jest wzajemnie jednoznaczne;
- II. f jest ciągle;
- III. f^{-1} jest ciągle.

Mówimy wtedy, że przestrzeń M_1 jest homeomorficzna z M_2 , lub że M_1 i M_2 są topologicznie ekwiwalentne.

Przestrzeń topologiczna T nazywamy przestrzenią topologiczną jednorodną, jeśli dla dowolnej pary punktów a, b ($a \in T, b \in T$) istnieje taki homeomorfizm

$h: T \rightarrow T$, że $h(a) = b$.

Zbiór G nazywamy grupą topologiczną ciągłą, jeśli:

- I. G jest grupa;
- II. G jest przestrzenią topologiczną;
- III. odwzorowanie $h: G \times G \rightarrow G$, określone jako $h(x, y) = x \cdot y^{-1}$ jest ciągle.

Grupa topologiczna jest przestrzenią jednorodną.

Grupa Lie, to grupa topologiczna w której otoczenie elementu jednostkowego jest homeomorficzne z r -wymiarową przestrzenią euklidesową \mathbb{R}^r .

Z jednorodności grupy topologicznej wynika, że dowolny element grupy Lie ma otoczenie homeomorficzne z \mathbb{R}^r . Liczba r nazywa się rzędem lub wymiarem grupy Lie. Oznacza to, że dowolny element grupy $g \in G$

można określić przez r parametrów.

Grupę G nazywamy *grupą ciągłą*, jeśli zbiór jej elementów tworzy przestrzeń topologiczną.

Grupę ciągłą nazywamy *grupą zwartą*, jeśli każda funkcja $f(g)$, ciągła na wszystkich elementach grupy G jest jednocześnie ograniczona.

Funkcja $f(g)$ na grupie G jest ciągła na elemencie $g_1 \in G$, jeśli dla każdej liczby dodatniej δ istnieje takie otoczenie F elementu g_1 , że dla wszystkich elementów g z otoczenia F spełniona jest nierówność

$$|f(g) - f(g_1)| < \delta, \quad (g \in F).$$

Będziemy mówili, że zadana jest *reprezentacja* T grupy G w pewnej przestrzeni liniowej L , jeśli każdemu elementowi g grupy G odpowiada operator $T(g)$ w przestrzeni L tak, że iloczynowi elementów grupy odpowiada iloczyn operatorów, czyli

$$T(g_1) T(g_2) = T(g_1 g_2).$$

Wymiar przestrzeni L nazywa się *wymiarem reprezentacji*. Grupa może mieć zarówno skończone, jak i nieskończone reprezentacje.

Podamy teraz kilka ważnych własności reprezentacji.

Jakiegokolwiek dwie reprezentacje grupy G związane zależnością

$$T_A(g) = A T(g) A^{-1},$$

nazywają się *reprezentacjami równoważnymi*. We wzorze tym T jest reprezentacją grupy G w przestrzeni L , a operator A jest dowolnym nieosobliwym operatorem liniowym przeprowadzającym wektory z L w wektory przestrzeni L_1 tego samego wymiaru (w szczególności L_1 może pokrywać się z L). Operator $T_A(g)$ działa w przestrzeni L_1 .

Wszystkie reprezentacje danej grupy G rozpadają się na klasy równoważnych reprezentacji, a tym samym zadanie odnalezienia wszystkich reprezentacji sprowadza się do trudniejszego problemu odnalezienia wszystkich nierównoważnych reprezentacji. Każda klasa równoważnych reprezentacji grupy skończonej zawiera reprezentacje

unitarne.

Podprzestrzeń L_1 przestrzeni liniowej L nazywamy *przestrzenią niezmienniczą* względem operatora A , jeżeli dla każdego $x \in L_1$ zachodzi $Ax \in L_1$.

Podprzestrzeniami nietrywialnymi przestrzeni L są wszystkie podprzestrzenie przestrzeni L oprócz niej samej i przestrzeni zerowej (złożonej z jednego elementu - zera przestrzeni).

Reprezentacja T grupy G w przestrzeni L nazywa się *reprezentacją przywiedlną*, jeżeli w L istnieje choćby jedna nietrywialna podprzestrzeń L_1 , niezmiennicza względem wszystkich operatorów $T(g)$, $g \in G$.

Reprezentacja T grupy G w przestrzeni L nazywa się *reprezentacją nieprzywiedlną*, jeżeli w L nie istnieje ani jedna nietrywialna podprzestrzeń L_1 , niezmiennicza względem wszystkich operatorów $T(g)$, $g \in G$. Tak więc wszystkie reprezentacje jednowymiarowe są nieprzywiedlne. Każda przywiedlna, skończenie wymiarowa reprezentacja unitarna, rozkłada się na nieprzywiedlne reprezentacje unitarne i jest ich sumą. Równocześnie przestrzeń L rozkłada się na sumę wzajemnie ortogonalnych podprzestrzeni niezmienniczych.

Charakterem $\chi(g)$ reprezentacji T ($g \rightarrow T(g)$) nazywa się sumę diagonalnych elementów macierzy odpowiadającej w dowolnej bazie operatorowi $T(g)$.

$$\chi(g) = \text{tr } T(g)$$

Wartość charakteru nie zależy od wyboru bazy. Charakter reprezentacji jest bardzo ważnym pojęciem, ponieważ stanowi jej nietrywialną charakterystykę. Tak więc np. charaktery reprezentacji równoważnych są sobie równe; charaktery nierównoważnych reprezentacji nieprzywiedlnych są ortogonalne (ich iloczyn skalarny znika); kwadrat skalarny charakteru reprezentacji nieprzywiedlnej jest równy jedności, a reprezentacji przywiedlnej jest większy od jedności; charakter reprezentacji przywiedlnej jest równy sumie charakterów wszystkich

reprezentacji nieprzywiedlnych w niej zawartych.

Będziemy interesowali się tylko takimi reprezentacjami T grupy G , których operatory są różniczkowalnymi funkcjami parametrów $\lambda_1, \dots, \lambda_m$. Ponieważ zadanie parametrów jednoznacznie określa element grupy, operator $T(g)$ można rozpatrywać jako funkcję parametrów

$$T(g) = T(\lambda_1, \dots, \lambda_m) \equiv T(\lambda).$$

Wyberzemy w sposób dowolny bazę w przestrzeni L (zadajemy układ współrzędnych), w której działają operatory $T(g)$ ($g \in G$) i za jej pomocą przyporządkujemy operatorowi $T(g)$ macierz $T_{ik}(g)$. Każdy element macierzy jest oczywiście też funkcją parametrów.

Reprezentację T będziemy nazywali różniczkowalną, jeśli wszystkie elementy macierzyowe

$$T_{ik}(\lambda_1, \dots, \lambda_m), \quad i, k = 1, \dots, s$$

są funkcjami różniczkowalnymi, zależnymi od parametrów grupy $\lambda_1, \dots, \lambda_m$.

Operator, któremu odpowiada macierz

$$\frac{\partial T_{ik}(\lambda)}{\partial \lambda_j},$$

nazywamy pochodną operatora $T(\lambda)$ ze względu na parametr λ_j i oznaczamy symbolem

$$\frac{\partial T(\lambda)}{\partial \lambda_j}.$$

Operatorem inifinitezymalnym X_j odpowiadającym parametrowi λ_j nazywa się pochodną operatora $T(g)$ względem λ_j wzięta przy $g = e$, czyli w punkcie $\lambda_1 = \dots = \lambda_n = 0$.

Liczba operatorów inifinitezymalnych jest równa wymiarowi grupy.

Twierdzenie 1.

Jeśli T i T_1 są dwiema reprezentacjami grupy G działającymi w jednej i tej samej przestrzeni L i mającymi ten sam zespół operatorów inifinitezymalnych, to reprezentacje T i T_1 pokrywają się.

Aby dowieść to twierdzenie wykazemy, że operatory infinytezymalne X_j w sposób jednoznaczny określają operatory $T(g)$. Wybierzmy dowolnie wektor $x \in L$ i oznaczmy

$$y(g) = T(g)x. \quad (A.1)$$

Jasne, że $y(g)$ jest funkcją parametrów λ_j . Jeśli określimy tę funkcję, to wobec dowolności wektora x , będą określone również wszystkie operatory $T(g)$. Zastąpmy element g elementem g^{-1} i podziałajmy na obie strony operatorem $T(h)$, $h \in G$

$$T(h) y(g^{-1}) = T(h) y(g^{-1})x = T(hg^{-1})x = y(hg^{-1}).$$

Oznaczmy $hg^{-1} = f$, a wtedy

$$T(fg) y(g^{-1}) = y(f). \quad (A.2)$$

Oznaczmy przez $\lambda_1(f)$, $\lambda_1(g)$ i $\lambda_1(fg)$ parametry charakteryzujące odpowiednio elementy f , g i fg , przy czym parametry $\lambda_1(fg)$ są funkcjami $2m$ parametrów $\lambda_1(f)$ i $\lambda_1(g)$:

$$\lambda_1(fg) = \varphi_1 \left[\lambda_1(f), \dots, \lambda_m(f); \lambda_1(g), \dots, \lambda_m(g) \right].$$

Jasne jest, że postać funkcji φ_1 ($i=1, \dots, m$) jest określona wyłącznie przez sposób mnożenia w grupie G i wybór parametrów λ_1 charakteryzujących elementy grupy i nie zależy od rozważanej reprezentacji T . Zróżniczkujmy $y(f)$ względem $\lambda_1(f)$ przy ustalonym g

$$\frac{\partial y(f)}{\partial \lambda_j(f)} = \frac{\partial T(fg)}{\partial \lambda_1(fg)} \frac{\partial \varphi_1(fg)}{\partial \lambda_j(f)} y(g^{-1}),$$

i połóżmy $g = f^{-1}$. W tym przypadku pochodne $\frac{\partial T(fg)}{\partial \lambda_1(fg)}$ przejdą w odpowiadające im operatory infinytezymalne

$$\frac{\partial y(f)}{\partial \lambda_j(f)} = S_j^1(f) X_1 y(f), \quad (A.3)$$

gdzie

$$S_j^1(f) = \left. \frac{\partial \varphi_1(fg)}{\partial \lambda_j(f)} \right|_{g=f^{-1}}. \quad (A.4)$$

Jako warunek początkowy do układu równań różniczkowych pierwszego rzędu (A.3) dołączmy

$$y(e) = x, \quad (y(0, \dots, 0) = x),$$

i wtedy funkcja $y(g)$ będzie wyznaczona jednoznacznie. Aby się o tym przekonać, należy rozwiązać ciąg równań, z warunkami początkowymi wynikającymi z poprzedniego rozwiązania. Tak więc najpierw rozwiązujemy równanie (A.3) dla $\lambda_1 \neq 0$, $\lambda_2, \dots, \lambda_m = 0$ z warunkiem początkowym $y(0, \dots, 0) = x$, następnie rozwiązujemy równanie (A.3) dla λ_1 i $\lambda_2 \neq 0$, a $\lambda_3, \dots, \lambda_m = 0$ z warunkiem początkowym $y(\lambda_1, 0, \dots, 0) = y_1(\lambda_1)$ (tu λ_1 traktujemy jako parametr swobodny) itd. aż do sytuacji, gdy wszystkie $\lambda_i \neq 0$, a warunek początkowy ma postać

$$y(\lambda_1, \lambda_2, \dots, \lambda_{m-1}, 0) = y_{m-1}(\lambda_1, \dots, \lambda_{m-1}) = 0.$$

Wykazaliśmy więc, że operatory infinytezymalne X_i , $i = 1, \dots, m$ utworzone za pomocą reprezentacji T , jednoznacznie tę reprezentację wyznaczają. Pokazaliśmy ponadto, jak można za pomocą operatorów infinytezymalnych utworzyć reprezentację T . Tym samym, zadanie znalezienia reprezentacji T sprowadza się do znalezienia odpowiadających jej operatorów infinytezymalnych. To ostatnie zadanie znacznie ułatwiają następujące dwa twierdzenia.

Twierdzenie 2.

Operatory infinytezymalne X_j ($j=1, \dots, m$) odpowiadające dowolnej reprezentacji T grupy G , spełniają następujące relacje komutacyjne

$$X_j X_k - X_k X_j = C_{jk}^i X_i, \quad (\text{A.5})$$

gdzie stałe C_{jk}^i , zwane stałymi strukturalnymi grupy G , nie zależą od wyboru reprezentacji T .

Dowód wynika natychmiast z rozpatrzenia układu równań (A.3), którego rozwiązaniem jest funkcja $y(g)$ (A.1). Wiemy, że rozwiązanie istnieje, tym samym musi być również spełniony warunek konieczny całkowalności układu równań, czyli musi zachodzić równość drugich pochodnych

$$\frac{\partial^2 y}{\partial \lambda_j \partial \lambda_k} = \frac{\partial^2 y}{\partial \lambda_k \partial \lambda_j}. \quad (\text{A.6})$$

Sprawdźmy ten warunek

$$\begin{aligned} \frac{\partial^2 y}{\partial \lambda_k \partial \lambda_j} &= \frac{\partial S_j^1(f)}{\partial \lambda_k} X_1 y(f) + S_j^1(f) X_1 \frac{\partial y(f)}{\partial \lambda_k} = \\ &= \frac{\partial S_j^1(f)}{\partial \lambda_k} X_1 y(f) + S_j^1(f) X_1 S_k^1(f) X_1 y(f). \end{aligned}$$

Podobnie

$$\frac{\partial^2 y}{\partial \lambda_j \partial \lambda_k} = \frac{\partial S_k^1(f)}{\partial \lambda_j} X_1 y(f) + S_k^1(f) X_1 S_j^1(f) X_1 y(f).$$

Otrzymane wyrażenia podstawiamy do (A.6) i otrzymujemy

$$\begin{aligned} \left(\frac{\partial S_k^1}{\partial \lambda_j} - \frac{\partial S_j^1}{\partial \lambda_k} \right) X_1 y(f) &= S_j^1 S_k^1 X_1 X_1 y(f) - S_k^1 S_j^1 X_1 X_1 y(f) \\ &= S_j^1 S_k^1 X_1 X_1 y(f) - S_k^1 S_j^1 X_1 X_1 y(f) = S_j^1 S_k^1 (X_1 X_1 - X_1 X_1) y(f). \end{aligned}$$

Kładziemy $f = e$, wtedy $S_j^1(e) = \delta_j^1$, $y(e) = x$,

$$(X_j X_k - X_k X_j) x = \left. \left(\frac{\partial S_k^1}{\partial \lambda_j} - \frac{\partial S_j^1}{\partial \lambda_k} \right) \right|_{f=e} X_1 x.$$

Ponieważ wektor x jest dowolny, więc otrzymujemy równość operatorowa

$$(X_j X_k - X_k X_j) x = C_{jk}^1 X_1 x,$$

gdzie

$$C_{jk}^1 = \left. \left(\frac{\partial S_k^1(f)}{\partial \lambda_j} - \frac{\partial S_j^1(f)}{\partial \lambda_k} \right) \right|_{f=e}$$

Z ostatniego wzoru widać, że współczynniki C_{jk}^1 są antysymetryczne ze względu na dolne wskaźniki

$$C_{jk}^1 = -C_{kj}^1.$$

Operatory X_1 spełniają tożsamość Jacobiego

$$\left[[X_1, X_j], X_k \right] + \left[[X_j, X_k], X_1 \right] + \left[[X_k, X_1], X_j \right] = 0,$$

skąd otrzymujemy tożsamość dla stałych strukturalnych

$$C_{jk}^1 C_{im}^1 + C_{km}^1 C_{ij}^1 + C_{mj}^1 C_{ik}^1 = 0.$$

Pożyteczny jest również następujący fakt z teorii grup

Twierdzenie 3.

Jeżeli dowolne operatory liniowe

$$A_1, A_2, \dots, A_m$$

określone w pewnej przestrzeni L spełniają te same relacje komutacyjne

$$A_j A_k - A_k A_j = C_{jk}^1 A_1, \quad (\text{A.7})$$

co i operatory nieskończone X_1, \dots, X_m , to operatory A_j są nieskończonymi operatorami pewnej działającej w L reprezentacji T grupy G .

Zanim wykazemy prawdziwość powyższego twierdzenia, wprowadzimy pewną pomocniczą tożsamość.

Rozpatrzmy iloczyn fgg_1 ($f, g, g_1 \in G$) i oznaczmy $fg=h$ i $gg_1=h_1$. Wtedy $hg_1 = fh_1$ i

$$\varphi_1[\lambda(h), \lambda(g_1)] = \varphi_1[\lambda(f), \lambda(h_1)], \quad (\text{A.8})$$

$$(\lambda(h)) = (\lambda_1(h), \dots, \lambda_m(h))$$

Ponadto

$$\lambda_j(h) = \varphi_j[\lambda(f), \lambda(g)],$$

$$\lambda_j(h_1) = \varphi_j[\lambda(g), \lambda(g_1)].$$

Zrózniczkujmy teraz wzór (A.8) względem $\lambda_k(f)$ przy ustalonych g i g_1 .

$$\frac{\partial \varphi_1(hg_1)}{\partial \lambda_j(h)} \frac{\partial \lambda_j(fg)}{\partial \lambda_k(f)} = \frac{\partial \varphi_1(fh_1)}{\partial \lambda_k(f)}.$$

Położmy teraz w tym wyrażeniu $g_1 = (fg)^{-1} = h^{-1}$ i odpowiednio $h_1 = gg_1 = g(fg)^{-1} = gg^{-1}f^{-1} = f^{-1}$. Uwzględniając definicje (A.4) macierzy $S_{.j}^1(f)$

$$S_{,j}^1(f) = \left. \frac{\partial \varphi_1(fg)}{\partial \lambda_j(f)} \right|_{g=f^{-1}}$$

otrzymamy poszukiwaną tożsamość

$$S_{,j}^1(h) = \frac{\partial \varphi_1(fg)}{\partial \lambda_j(f)} = S_{,j}^1(f),$$

lub

$$S_{,j}^1(fg) = \frac{\partial \varphi_1(fg)}{\partial \lambda_j(f)} = S_{,j}^1(f). \quad (\text{A.9})$$

Dowód ostatniego twierdzenia, opiera się na fakcie, że warunkiem koniecznym i wystarczającym całkowalności układu równań różniczkowych

$$\frac{\partial y(g)}{\partial \lambda_j} = S_{,j}^1 A_{,1} y(g), \quad (\text{A.10})$$

jest równość drugich pochodnych

$$\frac{\partial^2 y}{\partial \lambda_k \partial \lambda_j} = \frac{\partial^2 y}{\partial \lambda_j \partial \lambda_k}.$$

Warunek ten ma postać (A.7)

$$A_{,j} A_{,k} - A_{,k} A_{,j} = C_{,jk}^1 A_{,1},$$

która jest jednoznaczna i spełniająca warunek początkowy $y(e) = x$ dla dowolnego wektora $x \in L$. Wobec liniowości układu (A.10), operator $T_A(g)$ przyporządkowujący wektorowi x wektor $y(g)$

$$T_A(g) x = y(g)$$

jest operatorem liniowym. Aby odwzorowanie $g \rightarrow T_A(g)$ było reprezentacją grupy G , to musi zachodzić związek $T_A(gh) = T_A(g)T_A(h)$ dla dowolnej pary elementów g, h grupy G . Jeśli zastosujemy ten związek do wektora x , to przyjmie on następującą postać

$$y(gh) = T_A(gh) x = T_A(g) T_A(h) x = T_A(g) y(h)$$

co oznacza, że wektor $y(gh)$ rozpatrywany jako funkcja g przy ustalonym h , również jest szczególnym rozwiązaniem równań różniczkowych (A.10), ale przy warunku początkowym

$$y(gh) \Big|_{g=e} = y(h) .$$

Musimy więc jeszcze tylko sprawdzić, czy $y(gh)$ rzeczywiście spełnia równanie (A.9). Policzmy pochodną

$$\frac{\partial y(gh)}{\partial \lambda_1(g)} = \frac{\partial y(gh)}{\partial \lambda_k(gh)} \frac{\partial \varphi_k(gh)}{\partial \lambda_1(g)} = S^i_k(g) A_j y(gh) \frac{\partial \varphi_k(gh)}{\partial \lambda_1(g)} .$$

Korzystając z tożsamości (A.9) otrzymujemy

$$\frac{\partial y(gh)}{\partial \lambda_1(g)} = S^j_{i1} A_j y(gh) .$$

Tak więc funkcja $y(gh)$ spełnia równanie (A.10), a zatem operatory $T_A(g)$ stanowią reprezentację grupy G . Operatory infinitesimalne związane z tą reprezentacją znajdziemy korzystając z definicji operatorów infinitesimalnych

$$\frac{\partial T_A(g)}{\partial \lambda_j(g)} \Big|_{g=e} x = \frac{\partial y(g)}{\partial \lambda_j(g)} \Big|_{g=e} = S^i_j(g) A_i y(g) \Big|_{g=e} = A_j x .$$

Widzimy, że operatory infinitesimalne generowane przez reprezentację T_A pokrywają się z operatorami A .

Znaczenie powyższych twierdzeń tkwi w tym, że wykazują one równoważność zagadnienia znajdowania operatorów infinitesimalnych wszelkich możliwych reprezentacji i zagadnienia znajdowania wszelkich możliwych zbiorów operatorów

$$A_1, \dots, A_n$$

spełniających zadane z góry relacje komutacyjne (A.5). To ostatnie zagadnienie jest równoważne zagadnieniu odnalezienia wszelkich możliwych różniczkowalnych reprezentacji grupy G .

Z grupą Lie G związana jest jej algebra Lie \mathfrak{G} . Rzeczywista (zespolona) algebra Lie \mathfrak{G} m parametrowej grupy Lie G nazywamy m -wymiarową przestrzeń wektorową rozpiętą nad polem liczb rzeczywistych (zespolonych) przez operatory infinitesimalne X_i ($i = 1, \dots, m$) grupy G ; mnożenie elementów algebry \mathfrak{G} jest określone jako mnożenie komutatorowe

$$[X_i, X_j] = C_{ij}^k X_k,$$

gdzie C_{ij}^k są stałymi strukturalnymi grupy G . Operatory infinitesimalne spełniają rolę bazy algebry \mathfrak{G} .

Przykładem grupy Liego jest grupa obrotów $O(3)$, ponieważ każdy obrót można scharakteryzować trzema parametrami, np. kątami Eulera. Wymiar tej grupy wynosi 3.

Grupa Lorentza jest również grupą Lie. Jej wymiar wynosi 6, ponieważ do pełnego określenia wzajemnego ruchu dwu inercjalnych układów odniesienia potrzeba 6 parametrów: trzy rzuty prędkości początku układu współrzędnych drugiego układu odniesienia względem pierwszego i trzy kąty opisujące obrót osi drugiego układu współrzędnych względem pierwszego.

Rozpatrzmy reprezentację wektorową τ grupy $O(3)$. Dowolnemu obrotowi g przyporządkujemy wektor, skierowany w kierunku dodatniego zwrotu osi obrotu i równy co do wielkości kątowi obrotu. Jako parametry λ_1 przyjmujemy trzy jego rzuty na osie współrzędnych. Obrót g z parametrami $\lambda_1, 0, 0$ jest obrotem wokół osi OX o kąt λ_1 w kierunku dodatnim. W tym przypadku, w bazie kartezjańskiej, operatorowi $\tau(\lambda_1, 0, 0)$ odpowiada macierz

$$\tau(\lambda_1, 0, 0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \lambda_1 & -\sin \lambda_1 \\ 0 & \sin \lambda_1 & \cos \lambda_1 \end{pmatrix},$$

co należy rozumieć tak, że

$$y(g) = x' = \tau(\lambda_1, 0, 0)x, \quad \text{czyli}$$

$$x'_1 = x_1,$$

$$x'_2 = x_2 \cos \lambda_1 - x_3 \sin \lambda_1,$$

$$x'_3 = x_2 \sin \lambda_1 + x_3 \cos \lambda_1,$$

$$\frac{\partial y(f)}{\partial \lambda_j(f)} = S_j^i(f) X_i y(f),$$

$$S_j^1(f) = \left. \frac{\partial \varphi_1(gh)}{\partial \lambda_j(g)} \right|_{g=f^{-1}} .$$

Mamy stąd

$$\frac{\partial \tau(\lambda_1, 0, 0)}{\partial \lambda_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\sin \lambda_1 & -\cos \lambda_1 \\ 0 & \cos \lambda_1 & -\sin \lambda_1 \end{pmatrix} ,$$

a więc

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} .$$

Jeśli wprowadzimy oznaczenia,

$$\epsilon_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \quad \epsilon_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \quad \epsilon_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ itd.} ,$$

to widać, że

$$X_1 = \epsilon_{32} - \epsilon_{23} .$$

Stosując cykliczną permutację wskaźników otrzymamy

$$X_2 = \epsilon_{13} - \epsilon_{31} , \quad X_3 = \epsilon_{21} - \epsilon_{12} .$$

Wykorzystując fakt, że

$$[X_j, X_k] = C_{jk}^1 X_1 ,$$

otrzymujemy np. dla $j = 1$ i $k = 2$, że

$$X_1 X_2 - X_2 X_1 = c_{12}^1 X_1 + c_{12}^2 X_2 + c_{12}^3 X_3 ,$$

a więc

$$\begin{aligned} & (\epsilon_{32} - \epsilon_{23}) (\epsilon_{13} - \epsilon_{31}) - (\epsilon_{13} - \epsilon_{31}) (\epsilon_{32} - \epsilon_{23}) = \\ & = c_{12}^1 (\epsilon_{32} - \epsilon_{23}) + c_{12}^2 (\epsilon_{13} - \epsilon_{31}) + c_{12}^3 (\epsilon_{21} - \epsilon_{12}) . \end{aligned}$$

Po podstawieniu na miejsce ϵ_{1j} ich wartości otrzymujemy

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} =$$

$$= c_{12}^1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} + c_{12}^2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + c_{12}^3 \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Wynika stąd, że trzy pierwsze stałe strukturalne grupy mają następujące wartości:

$$c_{12}^1 = c_{12}^2 = 0, \quad c_{12}^3 = 1.$$

Z otrzymanych wyników wyciągamy wniosek, że

$$[X_1, X_2] = c_{12}^1 X_1 = \delta_3^1.$$

Powyższy przykład można wykorzystać przy znajdowaniu stałych strukturalnych innych grup.

LITERATURA

TEORIA GRUP

- [1] M. HAMERMESH, Teoria grup w zastosowaniu do zagadnień fizycznych, PWN, 1968.
- [2] JAN MOZRZYMAS, Zastosowania teorii grup w fizyce, PWN, 1976.
- [3] G. J. LUBARSKI, Teoria grup i jej zastosowania w fizyce, PWN, 1961.

TEORIA POLA

- [1] Н. Н. БОГОЛЮБОВ, Д. Б. ШИРКОВ, Введение в теорию квантованных полей, Наука, Москва, 1984.
- [2] А. І. АХЕЗЕР, С. В. ПЕЛЕТНИНСКИЙ, Поля и фундаментальные взаимодействия, Наукова Думка, Киев, 1986.
- [3] LEWIS H. RYDER, Quantum Field Theory, Cambridge University Press, 1985 (Л. Райдер, Квантовая теория поля, Мир, Москва, 1987).

- [4] PAUL H. FRAMPTON, Gauge Field Theories, Benjamin/Cummings Pub. Comp., Inc., Menlo Park, California, 1987.
- [5] А. А. СЛАВНОВ, Л. А. ФАДЕЕВ, Введение в квантовую теорию калибровочных полей, Наука, Москва, 1988.
- [6] PIERRE RAMOND, Field Theory. A Modern Primer, Benjamin/Cummings Publ. Comp., Inc., Reading, Massachusetts, 1981 (П. Рамон, Теория поля. Современный вводный курс, Мир, Москва, 1984).
- [7] ERHARD SEILER, Gauge Theories as a Problem of Constitutive Quantum Field Theory and Statistical Mechanics, Lecture Notes in Physics 159, Springer-Verlag, Berlin, 1982 (Э. Зайлер, Калибровочные теории, Мир, Москва, 1985).