

# Teoria Clássica de Campos

Prof. Alysson F. Ferrari

<http://professor.ufabc.edu.br/~alysson.ferrari/>



Universidade Federal do ABC

12 de fevereiro de 2019



---

Estas notas são um resumo de algumas das aulas da disciplina e não constituem uma fonte de referência completa sobre os temas abordados, não substituindo assim a leitura da bibliografia recomendada. Em algum momento do futuro isso poderá se tornar um texto mais completo sobre o tema, mas por enquanto este momento continua envolto pelas brumas incertas da mera possibilidade.

*Esta é uma versão ainda muito preliminar destas notas, portanto não divulgue este material sem comunicar ao autor.*

Alysson Fábio Ferrari

alysson.ferrari (at) ufabc (pt) edu (pt) br

<http://professor.ufabc.edu.br/~alysson.ferrari/>

---



# Sumário

<b>1</b>	<b>Suplemento Matemático</b>	<b>7</b>
1.1	Grupos e Álgebras de Lie . . . . .	7
<b>2</b>	<b>Simetria Relativística</b>	<b>19</b>
2.1	O Grupo de Lorentz . . . . .	19
<b>3</b>	<b>O Campo Escalar</b>	<b>29</b>
3.1	Funcional ação e equações de movimento . . . . .	29
3.2	Soluções da Equação de Klein-Gordon livre . . . . .	31
3.3	Simetrias e o Teorema de Noether . . . . .	34
<b>A</b>	<b>Noções de Teoria das Distribuições</b>	<b>41</b>



# Suplemento Matemático

## 1.1 Grupos e Álgebras de Lie

A noção de grupo é uma das mais fundamentais em matemática, e encontra vasta aplicação na física teórica por ser uma ferramenta feita sob medida para se estudar *simetrias*. De fato, o estudo da teoria de grupos pode parecer abstrata demais ou árida ao estudante de física, mas o que se tem aprendido é que todas as poderosas ferramentas desenvolvidas pelos matemáticos para estudar a estrutura de grupos refletem-se num profundo conhecimento das consequências das simetrias da natureza nas teorias que descrevem os fenômenos naturais.

Matematicamente, um *grupo*  $\mathcal{G}$  é um conjunto de elementos em que está definida uma operação que chamamos de *produto*, que satisfaz os seguintes axiomas,

- se  $g_1, g_2 \in \mathcal{G}$ , então  $g_1 \cdot g_2 = g_3 \in \mathcal{G}$  (propriedade de fechamento)
- existe um elemento  $\mathbb{1} \in \mathcal{G}$  tal que, para qualquer  $g \in \mathcal{G}$ , vale que  $\mathbb{1} \cdot g = g \cdot \mathbb{1} = g$  (existência da unidade)
- para qualquer  $g \in \mathcal{G}$ , existe  $g^{-1} \in \mathcal{G}$  tal que  $g \cdot g^{-1} = g^{-1} \cdot g = \mathbb{1}$  (existência da inversa)
- para quaisquer  $g_1, g_2, g_3 \in \mathcal{G}$ , vale que  $(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$  (associatividade).

A sugestão de que grupos são estruturas que podem representar muito bem simetrias vem de se perceber que cada um desses axiomas é bastante natural se traduzido para transformações de simetria, desde que entendemos que o produto do grupo significa a aplicação sucessiva de transformações. Por exemplo, é natural que a aplicação sucessiva de duas transformações de simetria também é uma simetria do sistema.

Um grupo é dito *discreto* quando seus elementos podem ser enumerados,

$$\mathcal{G} = \{g_1, g_2, g_3, \dots\} ,$$

sendo que grupos discretos podem ser *finitos* ou *infinitos*. Um grupo é dito *contínuo* quando seus elementos são identificados por um certo número de parâmetros contínuos,

$$\mathcal{G} = \{g(\alpha_1, \alpha_2, \alpha_3, \dots)\}$$

onde  $\alpha_i$  são reais, e podem ser entendidos como coordenadas de um espaço chamado de *variedade de grupo* ("group manifold"). Cada ponto nesta variedade é associado a um elemento do grupo.

### Exemplo 1. Grupo de Permutações de Três Elementos

Considere uma operação de permutação, que pode ser representada por exemplo por

$$\begin{pmatrix} 123 \\ 312 \end{pmatrix},$$

onde a 1ª linha corresponde ao ordenamento original, e a 2ª ao resultado da permutação. Padronizando a 1ª linha como sendo sempre a sequência (123), podemos omiti-la, identificando cada permutação apenas pela 2ª linha. Desta forma, as  $3! = 6$  permutações diferentes podem ser escritas como

$$g_1 = (123) ; g_2 = (132) ; g_3 = (213) ; g_4 = (231) ; g_5 = (312) ; g_6 = (321) .$$

O conjunto  $\{g_i; i = 1, \dots, 6\}$  forma um grupo, chamado de  $S_3$ . De fato, o elemento  $g_1$  é a identidade. Aplicando sucessivamente permutações, podemos descobrir os produtos entre os diferentes elementos. Uma forma conveniente de organizar estas informações é montando uma tabela, chamada de *tabela de multiplicação do grupo*.

$S_3$	123	132	213	231	312	321
123	123	132	213	231	312	321
132	132	123	312	321	213	231
213	213	231	123	132	321	312
231	231	213	321	312	123	132
312	312	321	132	123	231	213
321	321	312	231	213	132	123

Trata-se de um grupo discreto e finito. Neste caso, a tabela de multiplicação resume toda a informação relevante sobre o grupo.

### Exemplo 2. Translações na reta.



Represente por  $T(a)$  a operação de translação  $\psi(x) \rightarrow \psi(x+a)$ . Então o conjunto dos  $T(a)$  para  $a \in \mathbb{R}$  forma grupo, com produto definido por

$$T(a) \cdot T(b) = T(a+b) .$$

Este produto claramente representa o resultado das translações sucessivas  $\psi(x) \rightarrow \psi(x+a) \rightarrow \psi(x+a+b)$ . A identidade é  $T(0)$  e  $T(-a)$  é a inversa de  $T(a)$ . A associatividade também é trivialmente verificada.

Trata-se de um grupo contínuo com um parâmetro. A variedade de grupo é a reta real.

### Exemplo 3. Rotações no plano.

Considere uma rotação por um ângulo  $\theta$  num espaço bidimensional. O efeito de tal transformação sobre um vetor  $\begin{bmatrix} x \\ y \end{bmatrix}$  corresponde à transformação linear

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = R(\theta) \begin{bmatrix} x \\ y \end{bmatrix} .$$

Então o conjunto das matrizes  $R(\theta)$ , com  $\theta \in [0, 2\pi]$ , forma um grupo, em que

$$R(\theta) \cdot R(\theta') = R(\theta + \theta') ,$$

$$R(-\theta) = [R(\theta)]^{-1} \quad ; \quad R(0) = \mathbb{1} .$$

Este grupo é chamado de  $SO(2)$ :  $O$  de “ortogonal” (porque  $R^T R = \mathbb{1}$ ),  $2$  por ser composto de matrizes bi-dimensionais, e  $S$  porque  $\det R = +1$ . Novamente, é um grupo contínuo, com um parâmetro, só que agora este parâmetro varia no intervalo  $[0, 2\pi]$ , que é a variedade de grupo neste caso.

Note que, se  $\varepsilon$  é muito pequeno, o elemento do grupo  $R(\varepsilon)$  resulta ser muito próximo à identidade,

$$R(\varepsilon) = \begin{bmatrix} 1 - \frac{\varepsilon^2}{2} & -\varepsilon \\ \varepsilon & 1 - \frac{\varepsilon^2}{2} \end{bmatrix} ,$$

ou seja, até a segunda ordem em  $\varepsilon$ ,

$$R(\varepsilon) \sim \mathbb{1}_{2 \times 2} + i\varepsilon \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} - \frac{\varepsilon^2}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

Note que

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

ou seja

$$R(\varepsilon) \sim \mathbb{1}_{2 \times 2} + i\varepsilon \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} + \frac{i^2 \varepsilon^2}{2} \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}^2.$$

Pode-se mostrar, por indução, que de fato

$$R(\varepsilon) = e^{i\varepsilon G},$$

onde  $G$  é a matriz  $\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}$ .

Esta matriz pode ser calculada diretamente da forma

$$G = \lim_{\varepsilon \rightarrow 0} \frac{R(\varepsilon) - \mathbb{1}}{i\varepsilon}.$$

Chamamos  $G$  de *gerador* do grupo. No caso, o grupo de rotações no plano tem um único gerador, associado ao seu único parâmetro.

Note que a definição de  $G$  depende unicamente dos elementos do grupo “próximos” da identidade. Contudo, neste caso é óbvio que podemos tomar estender o parâmetro  $\varepsilon$  para qualquer valor dentro do domínio  $[0, 2\pi]$ . Desta forma, podemos escrever qualquer elemento do grupo  $R(\theta)$  na forma

$$R(\theta) = e^{i\theta G}.$$

O crucial é observar que reconstruímos *todos os elementos do grupo* a partir de um gerador  $G$ , que no caso é uma única matriz constante, e que pode ser encontrado estudando-se apenas uma pequena parte do grupo, a vizinhança da identidade.

Note também que as relações que definem o grupo  $R(\theta)$  são idênticas ao do grupo  $T(a)$ , a única diferença é o domínio dos parâmetros de grupo:  $\theta$  varia num domínio compacto (que, para comparação, podemos tomar como  $[-\pi, \pi]$ ), enquanto que  $a$  pertence ao domínio não compacto que é a reta real  $\mathbb{R}$ . Os dois grupos são distintos, embora as regras de multiplicação do grupo sejam idênticas, e os dois grupos sejam os mesmos na proximidade da identidade, eles diferem pela *topologia* da variedade de grupo.

#### Exemplo 4. O grupo $U(1)$

Considere o conjunto de todas as fases complexas da forma  $\phi(\theta) = e^{i\theta}$ . Então tal conjunto

forma um grupo, com

$$\phi(\theta)\phi(\theta') = \phi(\theta + \theta') ; \quad [\phi(\theta)]^{-1} = \phi(-\theta) ; \quad \phi(0) = 1 .$$

A variedade de grupo é o intervalo  $[0, 2\pi]$ . O gerador pode ser calculado como

$$G = \lim_{\varepsilon \rightarrow 0} \frac{e^{i\varepsilon} - 1}{i\varepsilon} = 1 .$$

Este é o grupo  $U(1)$ , onde  $U$  significa “unitário”, já que  $\phi(\theta)\phi^\dagger(\theta) = 1$ . Note, contudo, que o produto do grupo é idêntico ao do exemplo anterior. De fato, podemos fazer a identificação

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \sim e^{i\theta} ,$$

que define um mapeamento um-para-um entre os elementos dos dois grupos, que preserva todas as propriedades do grupo (em particular, sua regra de multiplicação). Este mapeamento é um *isomorfismo de grupos*, e escrevemos que  $SO(2) \sim U(1)$ .

Num grupo contínuo, cada elemento do grupo está associado a um ponto na variedade de grupo. Esta variedade em princípio pode ter uma estrutura topológica muito complicada, mas se ela for suficientemente suave, então o grupo é dito um *Grupo de Lie*.

O essencial da teoria de grupos de Lie é que: para cada parâmetro  $\alpha_i$  podemos associar um gerador definido por

$$G_i = \lim_{\varepsilon \rightarrow 0} \frac{g(\alpha_i = \varepsilon) - \mathbb{1}}{i\varepsilon} ,$$

em que  $g(\alpha_i = \varepsilon)$  é o elemento do grupo com  $\alpha_i = \varepsilon$  e todos os demais  $\alpha$ 's nulos. Os geradores representam a “ação linearizada” do grupo, ou seja, correspondem à ação dos elementos que são muito próximos da identidade. Definidos os geradores, pode-se mostrar que os elementos do grupo são representados pelo *mapa exponencial*

$$g(\alpha) = e^{i \sum_i \alpha_i G_i} ,$$

sendo que as condições para que o mapa exponencial exista, e respeite os axiomas de fechamento e associatividade do grupo, é que os geradores  $G_i$  satisfaçam uma *álgebra de Lie*,

$$[G_i, G_j] = c_{ij}^k G_k ,$$

bem como a identidade de Jacobi

$$[G_i, [G_j, G_k]] + \text{permutações cíclicas} = 0.$$

Esta descrição é essencialmente correta ao menos numa proximidade da identidade. Para muitos grupos, ela vale para todos os elementos do grupo que podem ser continuamente conectados à identidade. Existem casos mais sutis, envolvendo variedades de grupo com topologias não triviais: o caso do Grupo de Lorentz será estudado em detalhe adiante.

A vantagem de se lidar com grupos de Lie é que ao invés de estudar o grupo em si, que inclui um infinito não enumerável de elementos  $g(\alpha)$ , essencialmente toda a estrutura do grupo pode ser investigada a partir da sua versão linearizada, ou seja, sua álgebra de Lie, que é o conjunto finito de geradores  $G_i$ . Na verdade, existe uma poderosa teoria de *representações de álgebras de Lie* em que aprendemos que, na verdade, salvo os detalhes topológicos citados acima, *toda a estrutura essencial do grupo está contida nas constantes de estrutura  $c_{ij}^k$ , um conjunto finito de números reais!*

Consideremos, por pura simplicidade de notações, um grupo discreto  $\mathcal{G} = \{g_1, g_2, \dots\}$ . O grupo é considerado como o conjunto abstrato de elementos  $g_i$  que satisfazem uma determinada lei de produto  $g_i \cdot g_j = g_{p(i,j)}$ , chamada de *tabela de multiplicação do grupo*. Um mesmo grupo pode ter várias *representações* na forma de matrizes quadradas de uma certa dimensionalidade. Sempre que construímos um conjunto de matrizes  $[g_i]$ , uma associada a cada elemento abstrato do grupo  $g_i$ , dizemos que este conjunto de matrizes define uma *representação* do grupo se a multiplicação matricial reproduz a lei de produto do grupo, ou seja,

$$[g_i] \cdot [g_j] = [g_{p(i,j)}]$$

para qualquer  $g_i, g_j \in \mathcal{G}$ . Tipicamente, as matrizes  $[g_i]$  são construídas como transformações lineares num certo espaço vetorial, que representam os estados do sistema físico que sofrem a ação do grupo considerado.

O fato é que um mesmo grupo  $\mathcal{G}$  pode ter muitas representações de diferentes dimensionalidades. Um exemplo muito familiar é o do grupo  $SU(2)$ , muito estudado em Mecânica Quântica quando se discute o spin.

### Exemplo 5. O grupo $SU(2)$

O grupo  $SU(2)$  possui uma álgebra de Lie chamada de  $su(2)$ , composta por três geradores  $T_1, T_2, T_3$  que satisfazem as relações de comutação

$$[T_i, T_j] = i\epsilon_{ijk}T_k.$$

Em cursos de Mecânica Quântica, tradicionalmente se aprende um método que permite construir representações de  $su(2)$  da seguinte forma: define-se

$$T^\pm = T_1 \pm iT_2 \quad ; \quad T^2 = (T_1)^2 + (T_2)^2 + (T_3)^2 .$$

Prova-se então que  $[T_i, T^2] = 0$ , de forma que a matriz  $T^2$  pode ser diagonalizada junto com  $T_3$ , e o espaço de estados será composto por vetores da forma  $|j; m\rangle$ . Após uma série de análises, se descobre que, para que as matrizes que representam os  $T_i$  tenham dimensionalidade finita,  $j$  tem que assumir valores semi-inteiros,  $j = 0, \frac{1}{2}, 1, \dots$ , e para cada  $j$ ,  $m$  varia de unidade em unidade dentro do limite  $|m| \leq j$ . Isso significa que, para cada  $j$ , a álgebra será representada por matrizes com dimensão  $2j + 1$ .

Desta forma, pode-se fazer um catálogo de apresentações de  $su(2)$ :

- $j = 0$ : a representação trivial, em que  $[T_i] = 0$  e portanto o grupo associado só tem o elemento identidade.
- $j = \frac{1}{2}$ : as matrizes  $T_i$  são representadas pelas matrizes de Pauli,

$$T_i = \frac{\sigma_i}{2} ,$$

que são matrizes  $2 \times 2$  associadas ao *spin*.

- $j = 1$ : as matrizes  $T_i$  tem representação

$$T_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{bmatrix} \quad ; \quad T_2 = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} \quad ; \quad T_3 = \begin{bmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

agindo sobre um espaço de vetores tridimensionais  $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ . É fácil ver que, por exemplo,

$$e^{i\theta T_1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} ,$$

donde se vê que as transformações geradas por  $T_i$  correspondem a rotações em torno dos três eixos coordenados, o que significa que  $su(2) = so(3)$ , onde  $so(3)$  é a álgebra de Lie do grupo de rotações em três dimensões<sup>1</sup>.

<sup>1</sup>Os grupos  $SU(2)$  e  $SO(3)$  não são idênticos, contudo, devido à diferente topologia da variedade de grupo. O fato de que ambos possuem a mesma álgebra de Lie, contudo, significa que numa vizinhança da identidade os dois grupos são de fato idênticos.

Do exemplo anterior fica evidente que um mesmo grupo pode ter diferentes representações de diferentes dimensionalidades,

$$g_i \rightarrow [g_i]_{n \times n} \quad ; \quad g_i \rightarrow [g_i]_{m \times m} \quad ; \quad \dots$$

Em geral, toda representação pode ser, por uma transformação unitária, levada a uma forma diagonal por blocos, ou seja

$$[g_i] = \begin{bmatrix} [g_i]_{n_1 \times n_1} & 0 & \dots & \dots \\ 0 & [g_i]_{n_2 \times n_2} & 0 & \dots \\ 0 & 0 & \ddots & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

em que cada bloco  $[g_i]_{n_i \times n_i}$  não pode, por sua vez, ser reduzido da mesma forma. As representações  $[g_i]_{n_i \times n_i}$  são ditas *representações irredutíveis* do grupo, e elas são essenciais pois prova-se que *qualquer* representação pode ser construída combinando-se, da forma acima, representações irredutíveis.

No caso de  $SU(2)$  discutido anteriormente, o catálogo de representações citadas é o catálogo completo de todas as representações irredutíveis. O método aprendido em cursos de MQ, portanto, exaure todas as possíveis representações de  $SU(2)$ .

Lembrando-se que cada representação  $[g_i]_{n_i \times n_i}$  pode ser entendida como um grupo de matrizes que operam num dado espaço vetorial  $\begin{bmatrix} v \end{bmatrix}_{n_i}$ , então vemos que a ação de uma representação *redutível* é da forma

$$\begin{bmatrix} [g_i]_{n_1 \times n_1} & 0 \\ 0 & [g_i]_{n_2 \times n_2} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} v \end{bmatrix}_{n_1} \\ \begin{bmatrix} v \end{bmatrix}_{n_2} \end{bmatrix} = \begin{bmatrix} [g_i]_{n_1 \times n_1} & \cdot \\ [g_i]_{n_2 \times n_2} & \cdot \end{bmatrix} \begin{bmatrix} \begin{bmatrix} v \end{bmatrix}_{n_1} \\ \begin{bmatrix} v \end{bmatrix}_{n_2} \end{bmatrix},$$

ou seja, o subespaço  $\begin{bmatrix} v \end{bmatrix}_{n_1}$  não se mistura com o subespaço  $\begin{bmatrix} v \end{bmatrix}_{n_2}$  frente a uma transformação

do grupo. Fisicamente, isto significa que possivelmente o espaço de matrizes

$$\left[ \begin{array}{c} \left[ \begin{array}{c} v \\ \vdots \\ v \end{array} \right]_{n_1} \\ \left[ \begin{array}{c} v \\ \vdots \\ v \end{array} \right]_{n_2} \end{array} \right]$$

representa um *sistema não elementar*, composto por subsistemas menores, estes sim representados pelos subespaços irredutíveis  $\left[ \begin{array}{c} v \\ \vdots \\ v \end{array} \right]_{n_1}$  e  $\left[ \begin{array}{c} v \\ \vdots \\ v \end{array} \right]_{n_2}$ .

Na Teoria de Campos, um subespaço irredutível  $\left[ \begin{array}{c} v \\ \vdots \\ v \end{array} \right]_{n_i}$  das simetrias físicas das interações elementares é associado aos estados de uma *partícula elementar*: a irreducibilidade da representação  $[g_i]_{n_i \times n_i}$  é a contrapartida matemática ao fato de que a partícula elementar não pode ser “quebrada” em subsistemas menores, mais elementares. O catálogo de todas as representações irredutíveis do grupo que define a simetria da natureza, portanto, é também o catálogo de todos os tipos de *partículas elementares* que podem existir na natureza.

---

### Exemplo 6. O Campo Eletromagnético

Os vetores  $\vec{E}$  e  $\vec{B}$  se transformam, frente ao grupo de rotações tridimensionais  $SO(3)$  de forma

$$\begin{bmatrix} E'_x \\ E'_y \\ E'_z \end{bmatrix} = e^{i\theta_i T_i} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix},$$

e similarmente para  $\vec{B}$ , onde  $T_i$  é a representação irredutível de 3 dimensões da álgebra  $su(2)$  apresentada em anteriormente. Frente ao grupo  $SO(3)$ , portanto, os vetores  $\vec{E}$  e  $\vec{B}$  representam entidades fisicamente distintas.

Contudo, frente a uma transformação de Lorentz, que é representada matematicamente pelo grupo  $SO(2,1)$ , que será estudado em detalhes mais adiante, as componentes de  $\vec{E}$  e  $\vec{B}$  “se misturam” entre si. Isto significa que, frente ao grupo de Lorentz,  $\vec{E}$  e  $\vec{B}$  constituem uma única representação irredutível, e portanto são fisicamente entendidos como componentes de uma mesma entidade: o campo eletromagnético  $F_{\mu\nu}$ .

---

Tipicamente, é de grande interesse conseguir construir todas as representações irredutíveis de um dado grupo de Lie. Existe uma tecnologia bastante bem desenvolvida para este fim. Tipicamente, isso é feito construindo representações dos geradores  $G_i$  do grupo como matrizes

atuando sobre um certo espaço de estados representados por vetores

$$|\alpha_i\rangle ,$$

onde  $\alpha_i$  denota coletivamente um conjunto apropriado de rótulos contínuos ou discretos. Como usual na Mecânica Quântica, estes rótulos representam autovalores de um determinado conjunto completo de operadores que comutam entre si, e que portanto podem ser simultaneamente diagonalizados.

Os chamados *operadores de Casimir* da álgebra são de suma importância para identificar as representações irredutíveis. Um operador  $\mathcal{C}_i$  (que geralmente é uma combinação não linear dos geradores  $G_i$ ) é chamado de Casimir da álgebra se ele comuta com todos os geradores:

$$[\mathcal{C}_i, G_j] = 0.$$

A relevância dos operadores de Casimir é dada pelo por um teorema devido a Schuur: *numa representação irredutível, todo operador de Casimir é representado por uma matriz múltipla da identidade*. Isso permite dividir os rótulos dos estados  $|\alpha_i\rangle$  em dois tipos,

$$|\alpha_i; \beta_i\rangle ,$$

em que agora os  $\alpha_i$  são os autovalores associados aos operadores de Casimir. Como estes comutam com todos os  $G_i$ , os autovalores  $\alpha_i$  ficam “congelados” em qualquer transformação do grupo, e portanto servem para *rotular as diferentes representações irredutíveis*.

### Exemplo 7. As representações irredutíveis de $SU(2)$

Os geradores da álgebra são os operadores de Momento Angular  $J_i$ . Existe apenas um operador de Casimir:  $\vec{J}^2 = \sum_i J_i^2$ . O conjunto completo de operadores comutantes tradicionalmente utilizado é  $\{\vec{J}^2, J_3\}$ .

Os geradores são representados por matrizes que operam em estados  $|j; m\rangle$ . O autovalor  $j = 0, \frac{1}{2}, 1, \dots$  identifica cada representação irredutível. O operador de Casimir  $\vec{J}^2$  é representado pela matriz

$$\vec{J}^2 = j(j+1) \mathbb{1}$$



e o operador  $J_3$  pela matriz diagonal]

$$J^3 = \begin{bmatrix} j & 0 & \cdots & \cdots & \cdots \\ 0 & j-1 & 0 & \cdots & \cdots \\ \vdots & 0 & \ddots & 0 & \cdots \\ \vdots & \vdots & 0 & -j+1 & 0 \\ \vdots & \vdots & \vdots & 0 & -j \end{bmatrix} .$$

Os demais geradores  $J_1$  e  $J_2$  são representados por matrizes não diagonais:

$$J_1 |j; m\rangle = \sum_{m'} [J_1]_{mm'} |j; m'\rangle \quad ; \quad J_2 |j; m\rangle = \sum_{m'} [J_2]_{mm'} |j; m'\rangle .$$



## Simetria Relativística

### 2.1 O Grupo de Lorentz

O princípio fundamental da Relatividade Restrita pode ser enunciado da seguinte maneira:

*Uma mudança de um referencial inercial para outro implica na mudança das coordenadas associadas a qualquer evento do espaço tempo  $x^\mu \rightarrow x'^\mu$ , de tal forma que  $x^\mu$  e  $x'^\mu$  sejam relacionados por uma transformação linear que deixa o intervalo relativístico entre dois eventos*

$$\Delta s^2 = (x - y)^\mu (x - y)_\mu = (x^0 - y^0)^2 - (\vec{x} - \vec{y})^2 \quad (2.1)$$

*constante.*

A restrição da *linearidade* da transformação  $x^\mu \rightarrow x'^\mu$  é retirada na Relatividade Geral, que é invariante por reparametrizações arbitrárias  $x^\mu \rightarrow x'^\mu(x)$ . Ao tratar da Relatividade Restrita, contudo, basta supor transformações da forma

$$x'^\mu = \Lambda^\mu_\nu x^\nu + b^\mu. \quad (2.2)$$

O parâmetro  $b^\mu$  representa uma *translação* da origem das coordenadas. Tais transformações obviamente deixam (2.1) invariante. Menos trivial é a parte homogênea da transformação,

$$x'^\mu = \Lambda^\mu_\nu x^\nu, \quad (2.3)$$

que iremos inicialmente considerar. Para respeitar (2.1), temos que ter

$$x'^\mu x'_\mu = \eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} \Lambda^\mu_\rho x^\rho \Lambda^\nu_\sigma x^\sigma = \eta_{\mu\nu} x^\mu x^\nu; \quad (2.4)$$

como isto vale para  $x^\mu$  arbitrário, tem que valer

$$\eta_{\rho\sigma}\Lambda^\rho_\mu\Lambda^\sigma_\nu = \eta_{\mu\nu}. \quad (2.5)$$

Aqui, obviamente,  $\eta_{\mu\nu}$  é a métrica plana do espaço de Minkowski,  $\eta_{\mu\nu} = \text{diag}(+1, -1, -1, -1)$ . Escrevendo  $\eta$  como uma matriz  $4 \times 4$ , e também entendendo  $\Lambda^\mu_\nu$  como elementos de uma matriz  $4 \times 4$ , podemos reescrever (2.5) da forma matricial da forma

$$\Lambda^T \eta \Lambda = \eta. \quad (2.6)$$

A equação acima define um conjunto de matrizes  $\Lambda$  que forma um grupo conhecido pelos matemáticos de  $O(3,1)$ <sup>1</sup>, e que será por nós chamado de *grupo de Lorentz*, simbolizado por  $\mathcal{L}$ .

Tomando-se o determinante de ambos os membros de (2.6) concluímos que

$$\det^2 \Lambda = 1 \Rightarrow \det \Lambda = \pm 1.$$

De forma que  $\mathcal{L} = \mathcal{L}_+ \cup \mathcal{L}_-$ , onde  $\mathcal{L}_+$  é o subconjunto *próprio* (matrizes com determinante +1) e  $\mathcal{L}_-$  é o subconjunto *impróprio* (matrizes com determinante -1). Note que são dois subconjuntos disjuntos, e não há nenhuma mudança contínua de parâmetros que leva uma matriz de um para o outro. Além disso, tomando a componente  $\mu = \nu = 0$  de (2.5) temos

$$\eta_{\rho\sigma}\Lambda^\rho_0\Lambda^\sigma_0 = \left(\Lambda^0_0\right)^2 - \sum_i \left(\Lambda^i_0\right)^2 = \eta_{00} = 1$$

logo

$$\left|\Lambda^0_0\right| \geq 1$$

para qualquer  $\Lambda \in \mathcal{L}$ . Isto significa que podemos ter matrizes  $\Lambda$  com  $\Lambda^0_0 \geq 1$  (chamadas *ortócronas*) e outras com  $\Lambda^0_0 \leq -1$  (chamadas *não-ortócronas*), novamente definindo uma divisão de  $\mathcal{L}$  em subconjuntos que não podem ser continuamente conectados. Adota-se a seguinte simbologia e nomenclatura para estes subconjuntos:

	$\det \Lambda$	$\text{sgn} \Lambda^0_0$	
$\mathcal{L}^\uparrow_+$	+1	+	próprio ortócrono (restrito)
$\mathcal{L}^\uparrow_-$	-1	+	impróprio ortócrono
$\mathcal{L}^\downarrow_+$	+1	-	próprio não-ortócrono
$\mathcal{L}^\downarrow_-$	-1	-	impróprio não-ortócrono

As diferentes componentes de  $\mathcal{L}$  são conectadas por transformações de Lorentz especiais chamadas de *transformações discretas*, que são:

<sup>1</sup>Leia-se: conjunto de transformações que deixam invariante uma métrica indefinida com 3 autovalores de um sinal e 1 autovalor com outro sinal.

- *inversão espacial* (também chamada de *paridade*),  $\vec{x} \rightarrow -\vec{x}$ , representada pela matriz  $P = \text{diag}(1, -1, -1, -1)$
- *inversão temporal*,  $t \rightarrow -t$ , representada pela matriz  $T = \text{diag}(-1, 1, 1, 1)$
- *inversão espaço-temporal*, representada por  $PT = -\mathbb{1}$ .

Claramente,  $P \in \mathcal{L}_+^\uparrow$  e  $T \in \mathcal{L}_-^\downarrow$ . Além disso, seja  $\Lambda \in \mathcal{L}$ : é fácil ver que o produto  $P\Lambda$  tem determinante de sinal oposto ao de  $\Lambda$ , enquanto que  $T\Lambda$  tem determinante e a componente  $\Lambda_0^0$  de sinais opostos ao de  $\Lambda$ . Isto significa que

$$P\mathcal{L}_+^\uparrow = \mathcal{L}_-^\uparrow \quad ; \quad T\mathcal{L}_+^\uparrow = \mathcal{L}_-^\downarrow \quad ; \quad PT\mathcal{L}_+^\uparrow = \mathcal{L}_+^\downarrow. \quad (2.7)$$

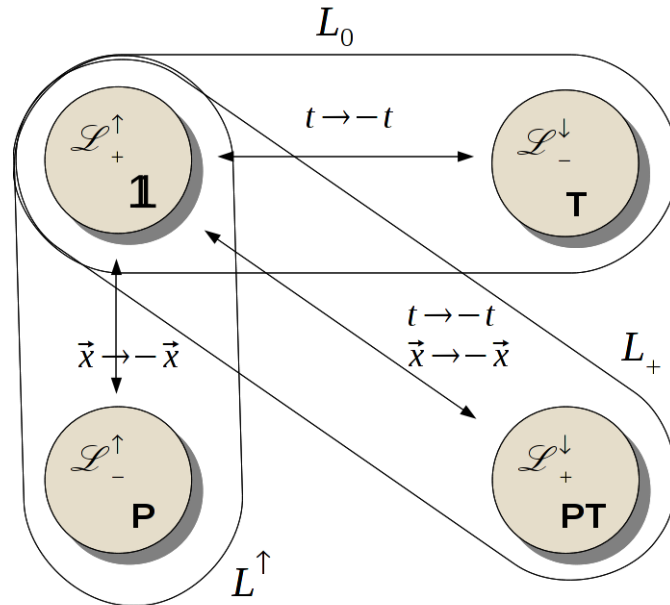
As transformações discretas  $P$ ,  $T$  e  $PT$ , portanto, comunicam as várias componentes de  $\mathcal{L}$ .

Note que  $\mathbb{1} \in \mathcal{L}_+^\uparrow$ , de forma que apenas o subconjunto de transformações de Lorentz *restritas* é um subgrupo. Podemos, contudo, definir outros subgrupos de  $\mathcal{L}$  da seguinte forma:

$$L_+ = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow \quad ; \quad L_0 = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\downarrow \quad ; \quad L^\uparrow = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$$

Como o subconjunto  $\mathcal{L}_+$  contém apenas matrizes com determinante  $+1$ , trata-se do grupo que os matemáticos chamam de  $SO(3, 1)$ .

Toda esta informação sobre a estrutura do grupo de Lorentz é representada no diagrama abaixo.



Na prática, vamos sempre considerar apenas o subgrupo restrito  $\mathcal{L}_+^\uparrow$ , e trataremos das transformações discretas separadamente. Na verdade, quando dissermos que uma teoria é *invariante de Lorentz*, estaremos nos referindo apenas ao subgrupo restrito  $\mathcal{L}_+^\uparrow$ . Existem teorias que são

invariantes por  $\mathcal{L}_+^\uparrow$  mas não pelo grupo  $\mathcal{L}$  completo: por exemplo, a teoria eletrofraca, responsável pela descrição unificada das interações eletromagnéticas e nucleares fracas, viola a simetria de paridade (Prêmio Nobel a T. D. Lee e C. N. Yang em 1957), e também a transformação de inversão temporal  $T$  é violada em interações elementares (Nobel de 1964 e 2008).

A interpretação física das transformações descritas por matrizes  $\Lambda \in \mathcal{L}_+^\uparrow$  vem do seguinte teorema:

**Teorema 1.** *Seja  $\Lambda \in \mathcal{L}_+^\uparrow$ . Então  $\Lambda$  pode ser escrito da forma*

$$\Lambda = \Lambda_{R_2} \Lambda_\beta \Lambda_{R_1}$$

onde  $\Lambda_{R_{1,2}} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{R}_{1,2} \end{pmatrix}$ , sendo que  $\mathbf{R}_{1,2}$  são matrizes do grupo de rotação tridimensional  $SO(3)^2$ , enquanto que

$$\Lambda_\beta = \begin{pmatrix} \cosh \beta & \sinh \beta & 0 \\ -\sinh \beta & \cosh \beta & 0 \\ 0 & 0 & \mathbb{1}_{2 \times 2} \end{pmatrix},$$

representa uma transformação de Lorentz usual no plano  $x_0 x_1$ .

Não vamos provar o teorema aqui, mas sua interpretação física é clara: uma transformação de Lorentz genérica deve representar a passagem de um referencial  $S$  para outro referencial  $S'$  que se move com relação a  $S$  com a velocidade constante  $\vec{v}$ . Tal transformação pode ser decomposta, pelo teorema, como o produto de três transformações consecutivas:

- uma rotação que alinha os eixos  $x_1$  de  $S$  com a velocidade  $\vec{v}$ ,
- um *boost*, ou seja, uma transformação de Lorentz ao longo de  $x_1$ , da forma que conhecemos de cursos elementares
- uma rotação que alinha o eixo  $x_1$  do referencial após o *boost* ao do referencial  $S'$ .

Vamos agora nos restringir a matrizes  $\Lambda \in \mathcal{L}_+^\uparrow$ , e vamos estudar os *geradores* do grupo de Lorentz restrito. Para tanto, vamos considerar uma transformação infinitesimal

$$\Lambda^\mu_\nu = \delta^\mu_\nu + \delta\lambda^\mu_\nu; \quad |\delta\lambda^\mu_\nu| \ll 1. \quad (2.8)$$

Inserindo em (2.5), pode-se ver que  $\delta\lambda$  tem que satisfazer  $\delta\lambda_{\mu\nu} = -\delta\lambda_{\nu\mu}$ , ou seja,  $\delta\lambda_{\mu\nu}$  é *antisimétrico*. Isto significa que a matriz de componentes  $\delta\lambda_{\mu\nu}$  tem *seis* componentes independentes, e portanto teremos *seis* geradores na álgebra de Lie de  $\mathcal{L}_+^\uparrow$ .

<sup>2</sup>Ou seja, matrizes  $3 \times 3$  ortogonais com determinante unitário satisfazendo  $\mathbf{R}^T \mathbf{R} = \mathbb{1}$

Tais geradores serão representados por matrizes antisimétricas  $M_{\alpha\beta}$ . Note que, a esta altura, os índices  $\alpha, \beta$  variam de 0 a 3 mas não são necessariamente índices espaço-temporais (discutiremos isto mais adiante). Pode-se encontrar explicitamente a forma das matrizes  $L_{\alpha\beta}$  e citamos aqui o resultado:

$$(M_{\alpha\beta})^\mu_\nu = \frac{i}{2} \left( \eta_{\alpha\nu} \delta^\mu_\beta - \eta_{\beta\nu} \delta^\mu_\alpha \right). \quad (2.9)$$

Desta equação, pode-se por cálculo direto encontrar as relações de comutação entre os  $M_{\alpha\beta}$ . Obtemos assim a *álgebra de Lie do grupo de Lorentz*:

$$[M^{\mu\nu}, M^{\alpha\beta}] = i \left( \eta^{\mu\alpha} M^{\nu\beta} - \eta^{\nu\alpha} M^{\mu\beta} - \eta^{\nu\beta} M^{\mu\alpha} + \eta^{\mu\beta} M^{\nu\alpha} \right). \quad (2.10)$$

Para facilitar a interpretação física desta relação, convém “decompor” os geradores  $M^{\alpha\beta}$  em componentes espaciais e temporais da seguinte forma,

$$J_i = -\frac{1}{2} \varepsilon_{ijk} M_{jk} \quad ; \quad K_i = M_{i0}. \quad (2.11)$$

Pode-se obter das relações:

$$[J_i, J_j] = i\varepsilon_{ijk} J_k \quad ; \quad [J_i, K_j] = i\varepsilon_{ijk} K_k \quad ; \quad [K_i, K_j] = -i\varepsilon_{ijk} J_k, \quad (2.12)$$

Estas relações significam que

- $J_i$  obedecem à álgebra do momento angular, logo correspondem a rotações espaciais
- $K_i$ , por exclusão, são os geradores responsáveis por *boosts* na direção dos três eixos coordenados.

Outra conclusão interessante é que, escrevendo-se a forma matricial explícita de  $J_i$  e  $K_i$ , percebemos que

$$J_i^\dagger = J_i \quad ; \quad K_i^\dagger = -K_i.$$

A consequência deste fato é que, ao representar uma transformação de Lorentz da forma exponencial

$$\Lambda = e^{i(\omega_i J_i + \nu_i K_i)}, \quad (2.13)$$

podemos perceber que os parâmetros  $\omega_i$  aparecem dentro de *senos e cossenos*, ou seja, são definidos num intervalo compacto:  $\omega_i \in [0, 2\pi)$ . Por outro lado, os parâmetros  $\nu_i$  aparecem dentro de *senos e cossenos hiperbólicos*, logo são definidos num intervalo não-compacto:  $\nu_i \in [0, \infty)$ . Ou seja, a variedade de grupo de  $\mathcal{L}_+^\uparrow$  é não compacta, o que caracteriza por sua vez  $\mathcal{L}_+^\uparrow$  como um grupo de Lie não-compacto. Grupos não-compactos são consideravelmente mais complicados de se estudar matematicamente, mas felizmente isso não será um obstáculo para o nível de profundidade que pretendemos alcançar nestas notas.

Vamos agora resolver o seguinte problema: *determinar a lista completa das representações irreduzíveis do grupo de Lorentz restrito  $\mathcal{L}_+^\uparrow$* . Nosso ponto de partida é a álgebra (2.10). A resolução deste problema é do nosso interesse pelo seguinte: o objeto do estudo da teoria clássica de campos serão campos da forma

$$\phi_A(x) ,$$

onde  $x = x^\mu$  são coordenadas do espaço-tempo, e  $A$  representa coletivamente outros índices que podem existir na teoria. Se queremos que nossa teoria respeite a invariância relativística, temos que nos perguntar o que acontece com os índices  $A$  frente a uma mudança de coordenadas. Supomos que translações não interferem nos índices  $A$ , ou seja, frente a uma translação  $x^\mu \rightarrow x^\mu + a^\mu$ , temos que

$$\phi_A(x) \rightarrow \phi'_A(x') = \phi_A(x - a) .$$

Contudo, frente a uma rotação de Lorentz, os índices  $A$  podem ser modificados. Se queremos ter esperança de construir Lagrangeanas que sejam invariantes frente a  $\mathcal{L}_+^\uparrow$  envolvendo os campos  $\phi_A(x)$ , vamos exigir que, frente a uma transformação de Lorentz

$$x'^\mu = \Lambda^\mu_\nu x^\nu ,$$

os campos se transformam segundo

$$\phi_A(x) \rightarrow \phi'_A(x') = \sum_B D_{AB}(\Lambda) \phi_B(x) ,$$

onde  $D_{AB}(\Lambda)$  é, em princípio, uma matriz que forma uma representação irreduzível do grupo  $\mathcal{L}_+^\uparrow$ . Desta forma, garantimos que os campos possuem transformação bem definida frente a uma mudança de coordenadas e, a partir disso, seremos capazes de construir invariantes, contraindo adequadamente os índices  $A, B$ .

Para resolver este problema, o truque será definir as matrizes abaixo

$$\begin{cases} N_i = \frac{1}{2} (J_i + iK_i) \\ M_i = \frac{1}{2} (J_i - iK_i) \end{cases} , \quad (2.14)$$

que possuem as seguintes relações de comutação:

$$\begin{cases} [N_i, M_j] = 0 \\ [N_i, N_j] = i\varepsilon_{ijk} N_k \\ [M_i, M_j] = i\varepsilon_{ijk} M_k \end{cases} . \quad (2.15)$$

Ou seja:  $N$  e  $M$  obedecem duas álgebras desacopladas de  $su(2)$ .



Conhecemos, de cursos elementares de Mecânica Quântica, o catálogo completo de *uma* álgebra  $su(2)$ : as representações irredutíveis (*irreps*) são catalogadas pelo “momento angular total”  $j = 0, \frac{1}{2}, 1, \dots$ . Para cada  $j$ , a representação se baseia numa base de  $(2j + 1)$  estados  $|jm\rangle$ , onde  $|m| < j$ , logo a representação é  $(2j + 1)$  dimensional, cuja forma explícita  $[J_i]_{(2j+1) \times (2j+1)}$  pode ser encontrada por fórmulas bem conhecidas.

**Exemplo 8.** Para  $j = \frac{1}{2}$ , a base de estados é  $\{|+\rangle, |-\rangle\}$ , e as matrizes  $J_i$  são representadas pelas matrizes de Pauli:  $[J_i]_{2 \times 2} = \frac{1}{2}\sigma_i$ .

Para  $j = 1$ , a base de estados é  $\{|1\rangle, |0\rangle, |-1\rangle\}$ , e as matrizes  $J_i$  são agora representadas por

$$[J_1]_{3 \times 3} = i \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; [J_2]_{3 \times 3} = i \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}; [J_3]_{3 \times 3} = i \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

que são os conhecidos geradores de rotações em três dimensões (lembre-se:  $so(3) \sim su(2)$ ).

O que queremos fazer é representar *duas* álgebras  $su(2)$  mas num mesmo espaço de estados, para que possamos “inverter” as matrizes  $N, M$  em termos de  $J, K$ , encontrando assim a representação para  $J, K$  neste espaço. A solução é usar o conceito de produto tensorial de espaços lineares.

Para cada par de números  $(j, j')$ , teremos uma base de  $(2j + 1)(2j' + 1)$  dimensões dada por  $|jm\rangle \otimes |j'm'\rangle$ . A matriz  $N_i$  opera apenas no espaço dos  $|jm\rangle$ , e sua ação no espaço dos  $|j'm'\rangle$  é trivial (identidade), o reverso acontecendo para as  $M_i$ . No espaço definido pelos vetores  $|jm\rangle \otimes |j'm'\rangle$ , as matrizes  $N$  e  $M$  serão  $(2j + 1)(2j' + 1) \times (2j + 1)(2j' + 1)$  dimensionais. Elas podem ser obtidas pelo produto externo de matrizes como segue

$$\begin{cases} [N_i] = [N_i]_{(2j+1)^2} \times \mathbb{1}_{(2j'+1)^2} \\ [M_i] = \mathbb{1}_{(2j+1)^2} \times [M_i]_{(2j'+1)^2} \end{cases}. \quad (2.16)$$

Conhecendo assim as matrizes  $[N_i]$  e  $[M_i]$ , podemos voltar a (2.14) e obter as matrizes  $[J_i]$  e  $K_i$ , que são nosso objetivo final.

O resumo é que possuímos assim um catálogo completo das *irreps* da álgebra de Lorentz (e, correspondentemente, do grupo  $\mathcal{L}_+^\uparrow$ ). Elas são identificadas por dois índices semi-inteiros, que denotaremos como  $(j, j')$ . Uma informação adicional, que não iremos justificar aqui, é que a soma  $j + j'$  representa o *spin* da partícula que eventualmente poderá ser descrita por um campo que se transforma segundo a representação  $(j, j')$  do grupo de Lorentz<sup>3</sup>. Abaixo, vamos cons-

<sup>3</sup>A justificativa desta afirmação vem do estudo da teoria de Wigner sobre representações do grupo de Poincaré, que podem ser encontrados em bons livros de Teoria Quântica de Campos.

truir as representações que, até o momento, possuem importância fundamental para a física teórica.

- Representação  $(0,0)$  – spin 0

A base do espaço de representação possui um único vetor,  $|0\rangle \times |0\rangle$ , e as matrizes  $N$  e  $M$  são representadas por matrizes nulas. Desta forma,  $[J] = [K] = 0$  e, daí,  $\Lambda = e^{i(\omega_i J_i + v_i K_i)} = 1$ . Esta representação descreve um campo que não possui índices que se modificaram frente a uma transformação de Lorentz, e por isso mesmo, são chamados de *campos escalares*:

$$\phi_A(x) \rightarrow \phi'_A(x') = \phi_A(x)$$

- Representação  $\left(\frac{1}{2}, 0\right)$  – spin  $\frac{1}{2}$

A base do espaço de representação possui dois vetores:  $\{|+\rangle \otimes |0\rangle, |-\rangle \otimes |0\rangle\}$ , logo trata-se de uma representação bi-dimensional. Em cada subespaço, as matrizes envolvidas são  $[N]_{2 \times 2} = \frac{1}{2}\sigma_i$  e  $[M]_{1 \times 1} = 0$ . As matrizes que representam  $N$  e  $M$  no espaço total são portanto

$$[N_i]_{2 \times 2} = [N]_{2 \times 2} \otimes \mathbb{1}_{1 \times 1} = \frac{1}{2}\sigma_i,$$

$$[M_i]_{2 \times 2} = \mathbb{1}_{2 \times 2} \otimes 0_{1 \times 1} = 0_{2 \times 2}.$$

Voltando a (2.14), concluímos que

$$\begin{cases} [J_i] = \frac{1}{2}\sigma_i \\ [K_i] = -\frac{i}{2}\sigma_i \end{cases}.$$

Como esperávamos,  $[J]$  é uma matriz hermitiana, enquanto  $[K]$  é anti-hermiteana. A matriz que representa uma transformação de Lorentz nesta representação é portanto, a partir de (2.13),

$$\Lambda_L = e^{\frac{i}{2}\sigma_i(\omega_i - iv_i)},$$

e os campos que se transformam segundo esta representação serão chamados de *espinores de Weyl de mão-esquerda*,

$$\psi_L(x) \rightarrow \psi'_L(x') = \Lambda_L \psi_L(x).$$

O motivo de chamarmos esta representação de “mão esquerda” não ficará claro antes de discutirmos com mais detalhe o campo de Dirac.

- Representação  $\left(0, \frac{1}{2}\right)$  – spin  $\frac{1}{2}$

Como a anterior, a base do espaço de representação possui dois vetores:  $\{|0\rangle \otimes |+\rangle, |0\rangle \otimes |-\rangle\}$ , logo esta representação também é bi-dimensional. Em cada subespaço, as matrizes envolvidas são  $[N]_{1 \times 1} = 0$  e  $[M]_{2 \times 2} = \frac{1}{2}\sigma_i$ . As matrizes que representam  $N$  e  $M$  no espaço total

são portanto

$$[N_i]_{2 \times 2} = 0_{1 \times 1} \otimes \mathbb{1}_{2 \times 2} = 0_{2 \times 2},$$

$$[M_i]_{2 \times 2} = \mathbb{1}_{1 \times 1} \otimes [M]_{2 \times 2} = \frac{1}{2} \sigma_i,$$

logo

$$\begin{cases} [J_i] = \frac{1}{2} \sigma_i \\ [K_i] = \frac{i}{2} \sigma_i \end{cases}.$$

Chegamos assim a

$$\Lambda_R = e^{\frac{i}{2} \sigma_i (\omega_i + i v_i)},$$

e os campos que se transformam segundo esta representação serão chamados de *espinores de Weyl de mão-direita*,

$$\psi_R(x) \rightarrow \psi'_R(x') = \Lambda_R \psi_R(x).$$

Note que as duas representações,  $(\frac{1}{2}, 0)$  e  $(0, \frac{1}{2})$  correspondem a partículas com spin  $\frac{1}{2}$ , e no entanto são duas representações *inequivalentes* do grupo de Lorentz. Entenderemos um pouco melhor este fato quando estudarmos o Campo de Dirac, quando também entenderemos a necessidade de, em teorias quânticas do elétron, considerar um espinor de *quatro-componentes*, que na verdade não consta da lista das *irreps* da álgebra de  $\mathcal{L}_+^\uparrow$ !

- Representação  $(\frac{1}{2}, \frac{1}{2})$  – spin 1

Não discutiremos em detalhe esta representação, mas pode-se mostrar que esta representação da álgebra de Lorentz por matrizes  $4 \times 4$  mostra-se ser igual a representação de definição utilizada para definir o grupo  $\mathcal{L}_+^\uparrow$  a partir da invariância do intervalo relativístico. Em outras palavras, esta é a representação *vetorial* do grupo de Lorentz, que descreve a transformação de campos vetoriais como o potencial de Maxwell:

$$A^\mu(x) \rightarrow A^{\mu'}(x') = \Lambda^\mu_{\nu} A^\nu(x).$$

Reencontramos, assim, a lei de transformação de tetra-vetores como uma das *irreps* da álgebra de  $\mathcal{L}_+^\uparrow$ . Como é bem conhecido, a partir desta representação podemos generalizar o conceito de tetra-vetor para o de tensores em geral, desenvolvendo assim toda a notação covariante para teorias relativísticas.



# O Campo Escalar

## 3.1 Funcional ação e equações de movimento

Um *campo* é a associação de uma certa quantidade a cada ponto do espaço e do tempo. O caso mais simples é o de uma função real  $\phi(\vec{x}, t)$ , que define um *campo escalar*. Exemplos que devem ser familiares não faltam: a temperatura numa determinada sala, a pressão em um fluido, etc. Também podemos associar um vetor a cada ponto do espaço e do tempo: os campos eletromagnéticos  $\vec{E}(\vec{x}, t)$  e  $\vec{B}(\vec{x}, t)$  são exemplos canônicos. Em princípio, um campo deve variar no espaço e no tempo, obedecendo a certas *Equações de Movimento*. Tais equações podem ser postuladas a partir de considerações físicas, mas do ponto de vista formal, vamos sempre supor que podem ser obtidas por meio de um *Princípio Variacional* aplicado a uma *Funcional Ação* que depende do campo e de suas derivadas.

Queremos estudar a dinâmica de teorias que envolvem um certo número de campos, que podem incluir campo escalares, vetoriais, etc... A simetria relativística dita exatamente os tipos de campos que podemos ter numa teoria relativística, bem a forma como estes se transformam frente a uma mudança de referencial inercial: a saber, se o campo  $\phi(x)$  possui índices adicionais, ou estes se transformam segundo alguma representação do grupo  $\mathcal{L}_+^\uparrow$ , ou não se transformam frente a mudanças de referenciais. Este requerimento garante que saibamos construir invariantes a partir dos campos e de suas derivadas.

Para obter os resultados que estamos interessados neste capítulo, basta considerar um único campo escalar  $\phi(x)$ . Num caso mais geral, todos os resultados podem ser obtidos para cada componente de cada campo considerado, de forma que esta simplificação não reduz em nada a generalidade destes resultados. Vamos considerar portanto a dinâmica de um campo escalar, que deve ser obtida a partir de uma *Funcional Ação*,

$$S[\phi] = \int_{\Omega} d^4x \mathcal{L}(\phi, \partial_{\mu}\phi) , \quad (3.1)$$

onde a *densidade Lagrangeana*  $\mathcal{L}(\phi, \partial_\mu \phi)$  é uma função dos campos e de suas derivadas primeiras. Nesta notação,  $d^4x$  é um elemento de volume,  $\Omega$  é uma região do espaço-tempo que, se for limitada, terá uma borda  $\partial\Omega$ . Em muitos casos, consideraremos que a integral se estende por todo o espaço-tempo, o que simbolizaremos formalmente por  $\partial\Omega \rightarrow \infty$  e/ou pela omissão do volume  $\Omega$  na integral. Usaremos sempre unidades naturais, o que significa que  $\hbar = c = 1$ , e  $S[\phi]$  não tem dimensões, de forma que  $\dim \mathcal{L} = \text{comprimento}^{-4}$ .

A dinâmica dos campos vem do requerimento de que a ação (3.1) seja estacionária frente a uma variação infinitesimal arbitrária do campo,

$$\phi(x) \rightarrow \phi(x) + \delta_0 \phi(x), \quad (3.2)$$

ou seja,

$$\delta S[\phi] \equiv S[\phi + \delta\phi] - S[\phi] = 0, \quad (3.3)$$

onde  $|\delta_0 \phi(x)| \ll 1$  para  $x \in \Omega$ , e  $\delta_0 \phi$  se anula na borda do volume de integração<sup>1</sup>. Estamos usando a notação  $\delta_0 \phi$  para a variação no campo porque mais adiante, teremos que definir uma forma mais geral de variação.

Até a primeira ordem em  $\delta\phi$ , temos

$$\delta S[\phi] = \int_{\Omega} d^4x \left[ \frac{\partial \mathcal{L}}{\partial \phi} \delta_0 \phi + \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} \partial_\mu \delta_0(\phi) \right]. \quad (3.4)$$

Os índices espaço-temporais estão apropriadamente contraídos, como deve ser já que a ação é um escalar de Lorentz. Integrando por partes:

$$\delta S[\phi] = \int_{\Omega} d^4x L_\phi(x) \delta_0 \phi + \oint_{\partial\Omega} \delta_0 \phi \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} d\sigma_\mu, \quad (3.5)$$

onde

$$L_\phi(x) = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]}. \quad (3.6)$$

Como  $\delta_0 \phi(\partial\Omega) = 0$ , o termo de superfície desaparece.

Generalizando o Princípio de Hamilton, vamos dizer que se  $\phi(x)$  é solução das equações de movimento, então  $\delta S[\phi] = 0$  para  $\delta_0 \phi$  arbitrário, o que implica na seguinte forma das *Equações de Euler Lagrange*.

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} = 0. \quad (3.7)$$

---

<sup>1</sup>Esta condição é similar à condição de que a posição inicial e final estão fixas no Princípio de Hamilton da Mecânica Clássica.

Por outro lado, podemos lembrar que a propriedade fundamental da derivada funcional,

$$\delta S[\phi] = \int d^4x \frac{\delta S[\phi]}{\delta \phi(x)} \delta_0 \phi(x), \quad (3.8)$$

de forma que podemos enunciar as Equações de Euler-Lagrange simplesmente como

$$\frac{\delta S[\phi]}{\delta \phi(x)}, \quad (3.9)$$

que é uma forma mais direta de enunciar o Princípio da Ação Estacionária.

---

### Exemplo 9. O Campo Escalar Real

Considere a densidade Lagrangiana envolvendo um campo escalar real  $\phi$ ,

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2} (\partial^\mu \phi \partial_\mu \phi - m^2 \phi^2) - V(\phi), \quad (3.10)$$

onde  $V(\phi)$  é, tipicamente, uma soma de termos com potências de  $\phi$  maiores que dois. Note que usamos, aqui, o sistema de unidades naturais, em que  $c = \hbar = 1$ .

Calculando-se as derivadas

$$\frac{\partial \mathcal{L}}{\partial \phi} = -m^2 \phi - V'(\phi)$$

e

$$\frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} = \frac{\partial}{\partial [\partial_\mu \phi]} \left( \frac{1}{2} \eta^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi \right) = \partial^\mu \phi$$

reescrevemos a equação (3.7) na forma

$$-m^2 \phi - V'(\phi) - \underbrace{\partial_\mu \partial^\mu \phi}_{\square} = 0,$$

ou seja

$$(\square + m^2) \phi = -V'(\phi). \quad (3.11)$$

## 3.2 Soluções da Equação de Klein-Gordon livre

Na mecânica clássica, resolver um determinado problema significa encontrar uma solução das equações diferenciais que determinam o problema, satisfazendo determinadas condições iniciais. Da mesma forma, numa teoria de campos, podemos procurar pela função  $\phi(\vec{x}, t)$  que satisfaz a equação de Klein-Gordon

$$(\square + m^2) \phi(x) = -V'(\phi), \quad (3.12)$$

e também satisfaz uma determinada condição inicial, tipicamente na forma de funções conhecidas  $\phi(\vec{x}, t_0)$  e  $\partial_0 \phi(\vec{x}, t_0)$ , onde  $t_0$  é fixo. No caso em que  $V = 0$ , tais soluções podem ser encontradas explicitamente, e é o que faremos nesta seção.

Focalizaremos nossa atenção, portanto, à equação de Klein-Gordon livre,

$$\left(\square + m^2\right) \phi(x) = 0. \quad (3.13)$$

Iremos resolvê-la usando o método de Fourier: tentamos uma solução da forma

$$\phi(x) = \int \frac{d^4 k}{(2\pi)^2} e^{ikx} \tilde{\phi}(k), \quad (3.14)$$

onde  $kx \equiv k_\mu x^\mu = (k^0 x^0) - \vec{k} \cdot \vec{x}$ . Inserindo em (3.13), usando a independência linear das exponenciais  $e^{ikx}$ , vemos que a equação de Klein-Gordon será satisfeita se

$$\left(k^2 - m^2\right) \tilde{\phi}(k) = 0.$$

Como precisamos ter  $\tilde{\phi}(k) \neq 0$  para não ter uma solução trivial, temos que impor  $k^2 - m^2 = 0$ . Fazemos isso escrevendo

$$\tilde{\phi}(k) = \delta(k^2 - m^2) \hat{\phi}(k),$$

onde  $\hat{\phi}(k)$  é, em princípio, arbitrária (oportunamente, será relacionada às condições iniciais do campo). Inserindo esta expressão em (3.14),

$$\phi(x) = \int \frac{d^4 k}{(2\pi)^2} e^{ikx} \delta(k^2 - m^2) \hat{\phi}(k). \quad (3.15)$$

Vamos refletir sobre o resultado desta expressão: chegamos à conclusão que a amplitude de Fourier  $\hat{\phi}(k)$  pode ser arbitrária, desde que as integrais sejam sobre todos os momentos  $k^\mu$  tais que  $k^2 - m^2 = 0$ ; diz-se que deve-se somar os momentos *na concha de massa*, ou seja, na hipersuperfície contida no espaço dos momentos definida pela condição  $k^2 - m^2 = 0$ .

Esta última condição, fisicamente, é importante porque pode ser reescrita como

$$\left(k^0\right)^2 = \vec{k}^2 + m^2, \quad (3.16)$$

ou seja, a componente  $k^0$  do momento  $k^\mu$  deve corresponder à energia física de uma partícula de momento  $\vec{k}$  e massa  $m$ . Tornamos isto mais claro, escrevendo a função  $\delta$  em (3.15) da forma

$$\delta(k^2 - m^2) = \delta\left(\left(k^0\right)^2 - \omega_{\vec{k}}^2\right), \quad (3.17)$$

onde



$$\omega_{\vec{k}} \equiv +\sqrt{\vec{k}^2 + m^2}. \quad (3.18)$$

Note que, de (3.16) segue que  $k^0 = \pm\omega_{\vec{k}}$ , ou seja, na expressão (3.15) teremos tanto contribuições com frequência positiva quanto com frequência negativa. Para levar isto em conta, consideremos o seguinte truque: a identidade

$$1 = \theta(x^0) + \theta(-x^0), \quad (3.19)$$

em que  $\theta(x)$  é a função salto (ou de Heavyside),

$$\theta(x) = \begin{cases} 0, & x < 0 \\ 1/2, & x = 0 \\ 1, & x \geq 0 \end{cases}, \quad (3.20)$$

é inserida em (3.15),

$$\phi(x) = \int \frac{d^4k}{(2\pi)^2} e^{i(k^0x^0 - \vec{k} \cdot \vec{x})} \delta\left((k^0)^2 - \omega_{\vec{k}}^2\right) \left[\theta(k^0) + \theta(-k^0)\right] \hat{\phi}(k). \quad (3.21)$$

Fazemos a substituição  $k \rightarrow -k$  no termo proporcional a  $\theta(-k^0)$ ,

$$\phi(x) = \int \frac{d^4k}{(2\pi)^2} \delta\left((k^0)^2 - \omega_{\vec{k}}^2\right) \theta(k^0) \left[e^{i(k^0x^0 - \vec{k} \cdot \vec{x})} \hat{\phi}(k) + e^{-i(k^0x^0 - \vec{k} \cdot \vec{x})} \hat{\phi}(-k)\right]. \quad (3.22)$$

Agora efetuamos a integral em  $k^0$ , usando a identidade

$$\delta\left((k^0)^2 - \omega_{\vec{k}}^2\right) = \frac{1}{2\omega_{\vec{k}}} \left[\delta(k^0 - \omega_{\vec{k}}) + \delta(k^0 + \omega_{\vec{k}})\right]. \quad (3.23)$$

Como  $\delta(k^0 + \omega_{\vec{k}}) \theta(k^0) = 0$ , chegamos a

$$\phi(x) = \int \frac{d^3k}{(2\pi)^2} \left[ \frac{e^{i(\omega_{\vec{k}}x^0 - \vec{k} \cdot \vec{x})}}{2\omega_{\vec{k}}} \hat{\phi}(\omega_{\vec{k}}, \vec{k}) + \frac{e^{-i(\omega_{\vec{k}}x^0 - \vec{k} \cdot \vec{x})}}{2\omega_{\vec{k}}} \hat{\phi}(-\omega_{\vec{k}}, -\vec{k}) \right]. \quad (3.24)$$

Agora algumas definições,

$$\phi^\pm(\vec{k}) = \sqrt{2\pi} \hat{\phi}(\pm\omega_{\vec{k}}, \pm\vec{k}), \quad (3.25)$$

$$f_{\vec{k}}(x) = \frac{e^{i(\omega_{\vec{k}}x^0 - \vec{k} \cdot \vec{x})}}{(2\pi)^{3/2}}, \quad (3.26)$$

nos permitem reescrever (3.24) da forma

$$\phi(x) = \int \frac{d^3k}{2\omega_{\vec{k}}} \left[ f_{\vec{k}}(x) \phi^+(\vec{k}) + f_{\vec{k}}^*(x) \phi^-(\vec{k}) \right]. \quad (3.27)$$

O fato de que  $\phi$  deve ser real significa que  $\left(\phi^+(\vec{k})\right)^* = \phi^-(\vec{k})$ .

Pausa para reflexão: observe que, afora algumas normalizações, essencialmente a solução (3.27) corresponde a construir a solução geral  $\phi(x)$  como uma superposição linear, com coeficientes arbitrários  $\phi^\pm(\vec{k})$ , de soluções da equação de Klein-Gordon livre  $f_{\vec{k}}(x) \sim e^{ik^\mu x_\mu}$ .

As funções  $f_{\vec{k}}$  definidas em (3.26) obedecem a relações de ortogonalidade,

$$\begin{cases} i \int d^3x f_{\vec{k}}^*(x) \overleftrightarrow{\partial}_0 f_{\vec{q}}(x) = 2\omega_{\vec{k}} \delta^3(\vec{k} - \vec{q}) \\ i \int d^3x f_{\vec{k}}(x) \overleftrightarrow{\partial}_0 f_{\vec{q}}(x) = i \int d^3x f_{\vec{k}}^*(x) \overleftrightarrow{\partial}_0 f_{\vec{q}}^*(x) = 0 \end{cases}, \quad (3.28)$$

cujas provas ficam como exercício para o leitor. Aqui, emprega-se a notação

$$A \overleftrightarrow{\partial}_0 B \equiv A \partial_0 B - (\partial_0 A) B. \quad (3.29)$$

Estas relações de ortogonalidade permitem, finalmente, relacionar as amplitudes de Fourier  $\phi^\pm(\vec{k})$  com as condições iniciais do campo, num dado tempo fixo  $t_0$ . Fica também a cargo do leitor mostrar que

$$\begin{cases} \phi^+(\vec{k}) = i \int d^3x f_{\vec{k}}^*(x) \overleftrightarrow{\partial}_0 \phi(\vec{x}, t) \\ \phi^-(\vec{k}) = i \int d^3x \phi(\vec{x}, t) \overleftrightarrow{\partial}_0 f_{\vec{k}}(x) \end{cases}, \quad (3.30)$$

o que pode ser feito facilmente usando-se as relações de ortogonalidade (3.28). Note que o membro da esquerda destas relações é explicitamente independente de  $t$ , portanto, o membro da direita *também deve ser*. Podemos assim fixar  $t = t_0$ , e desta forma encontrar as amplitudes de Fourier  $\phi^\pm(\vec{k})$  em termos dos valores iniciais do campo,  $\phi(\vec{x}, t_0)$  e  $\partial_0 \phi(\vec{x}, t_0)$ . O campo  $\phi(x)$  fica, assim, completamente determinado por (3.30) e (3.27) a partir de suas condições iniciais.

### 3.3 Simetrias e o Teorema de Noether

Emmy Noether foi uma matemática alemã que demonstrou um dos resultados mais fundamentais da física teórica, o chamado *Teorema de Noether*, pois demonstra da forma mais geral a profunda relação que existe entre leis de conservação e simetrias. Iremos demonstrar o teorema no contexto de teorias de campos descritas por uma Funcional Ação, como acabamos de definir.

Entendemos uma *transformação de simetria* como uma mudança que, aplicada a um sistema físico, resulta numa nova configuração do sistema que é idêntica à inicial. No contexto de

teoria de campos, uma transformação de simetria em geral engloba tanto uma mudança de coordenadas quanto uma mudança no valor do campo em cada ponto do espaçotempo, ou seja,

$$\begin{cases} x^\mu \rightarrow x'^\mu \\ \phi(x) \rightarrow \phi'(x') \end{cases} . \quad (3.31)$$

A interpretação desta última equação é: focalizando a atenção a um determinado ponto do espaçotempo, localizado pelas coordenadas  $x^\mu$ , inicialmente o valor do campo neste ponto é  $\phi(x)$ ; após a transformação, as coordenadas deste ponto mudam para  $x'^\mu$  e o valor do campo neste ponto também muda para  $\phi'(x')$ . Ou seja: fisicamente, a mudança no valor do campo num determinado ponto fixo do espaço-tempo é dado pela diferença

$$\delta\phi = \phi'(x') - \phi(x) ,$$

que chamaremos de *variação total* do campo.

Iremos oportunamente particularizar para uma *transformação infinitesimal*, considerando que

$$\begin{cases} x'^\mu = x^\mu + \delta x^\mu \\ \phi'(x') = \phi(x) + \delta\phi(x) \end{cases} , \quad (3.32)$$

onde  $\delta x^\mu$  e  $\delta\phi$  são muito pequenos. Neste ponto, vale lembrar a variação de forma de um campo, já definida em (3.2),

$$\delta_0\phi(x) = \phi'(x) - \phi(x) . \quad (3.33)$$

Note a diferença: a variação total  $\delta\phi$  corresponde à variação do campo num mesmo ponto físico. A variação de forma  $\delta_0\phi$  é um artifício matemático, que usamos para derivar as equações de movimento (caso em que  $\delta_0\phi$  era arbitrária, exceto pela condição  $\delta_0\phi(\partial\Omega) = 0$ ). Abaixo, precisaremos relacionar a variação de forma com a variação total, o que pode ser feito para uma transformação infinitesimal da seguinte forma:

$$\delta\phi(x) = \phi'(x') - \phi(x) = \phi'(x) + \delta x^\mu \partial_\mu \phi'(x) - \phi(x) .$$

Note que podemos fazer a simplificação  $\delta x^\mu \partial_\mu \phi'(x) = \delta x^\mu \partial_\mu \phi(x)$ , já que a diferença entre  $\phi'$  e  $\phi$  já é de primeira ordem. Portanto:

$$\delta\phi(x) = \delta_0\phi(x) + \delta x^\mu \partial_\mu \phi(x) . \quad (3.34)$$

Aplicando a transformação (3.31) a um determinado sistema físico cuja dinâmica é descrito por uma ação  $S[\phi]$ , definimos uma ação transformada

$$S'[\phi'] = \int_{\Omega'} d^4x' \mathcal{L}(\phi'(x'), \partial'_\mu \phi'(x')) , \quad (3.35)$$

onde  $\partial'_\mu = \frac{\partial}{\partial x'^\mu}$  e  $\Omega'$  é a imagem de  $\Omega$  frente à transformação de coordenadas considerada. Usando o teorema de transformação de coordenadas<sup>2</sup> na integral acima, podemos escrever

$$S'[\phi'_A] = \int_{\Omega} d^4x \mathcal{J}\left(\frac{x'^\mu}{x^\nu}\right) \mathcal{L}(\phi'(x'), \partial'_\mu \phi'(x')) , \quad (3.36)$$

onde agora  $x' = x'(x)$  é entendido como função de  $x$ , e

$$\mathcal{J}\left(\frac{x'^\mu}{x^\nu}\right) = \det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right) \quad (3.37)$$

é o determinante Jacobiano da transformação  $x^\mu \rightarrow x'^\mu$ . A diferença entre a ação original e a transformada pode ser escrita, portanto, como

$$\delta S[\phi] = \int_{\Omega} d^4x \left[ \mathcal{J}\left(\frac{x'^\mu}{x^\nu}\right) \mathcal{L}(\phi'(x'), \partial'_\mu \phi'(x')) - \mathcal{L}(\phi, \partial_\mu \phi) \right] . \quad (3.38)$$

Agora vamos particularizar para uma transformação infinitesimal, conforme (3.32), e fazer todos os cálculos apenas em primeira ordem de  $\delta x^\mu$  e  $\delta\phi$ . O determinante Jacobiano, em primeira ordem em  $\delta x^\mu$  é dado por

$$\det\left(\frac{\partial x'^\mu}{\partial x^\nu}\right) \sim 1 + \partial_\mu \delta x^\mu , \quad (3.39)$$

de forma que a variação infinitesimal da ação vale

$$\delta S[\phi] = \int_{\Omega} d^4x [\delta \mathcal{L}(\phi, \partial_\mu \phi) + \partial_\mu \delta x^\mu \mathcal{L}(\phi, \partial_\mu \phi)] \quad (3.40)$$

onde

$$\begin{aligned} \delta \mathcal{L} &= \mathcal{L}(\phi'(x'), \partial'_\mu \phi'(x')) - \mathcal{L}(\phi(x), \partial_\mu \phi(x)) \\ &\sim \delta x^\mu \partial_\mu \mathcal{L}(\phi(x), \partial_\mu \phi(x)) + \mathcal{L}(\phi'(x), \partial_\mu \phi'(x)) - \mathcal{L}(\phi(x), \partial_\mu \phi(x)) , \end{aligned}$$

sempre desconsiderando termos de ordem mais alta. Note que a diferença envolvendo os dois últimos termos é justamente o que já calculamos na derivação das equações de movimento

<sup>2</sup>Veja, por exemplo, [http://en.wikipedia.org/wiki/Change\\_of\\_variable](http://en.wikipedia.org/wiki/Change_of_variable).

podemos portanto copiar o resultado da Eq. (3.5),

$$\delta \mathcal{L} = \delta x^\mu \partial_\mu \mathcal{L}(\phi, \partial_\mu \phi) + L_\phi(x) \delta_0 \phi + \partial_\mu \left( \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} \delta_0 \phi \right), \quad (3.41)$$

onde aparece novamente a variação de forma  $\delta_0 \phi$  já definida em (3.33), e  $L_\phi$  definido em (3.6). Juntando (3.40) e (3.41),

$$\delta S[\phi] = \int_\Omega d^4x \left[ L_\phi \delta_0 \phi + \partial_\mu \left( \delta_0 \phi \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} + \delta x^\mu \mathcal{L} \right) \right]. \quad (3.42)$$

Será conveniente, contudo, aplicar (3.34) no parêntesis mais interno da última expressão, reescrevendo  $\delta_0 \phi$  em termos de  $\delta \phi$ , pois é este último que efetivamente corresponde à transformação física que estamos considerando, conforme discutido anteriormente. Desta forma, chegamos finalmente a

$$\delta S[\phi] = \int_\Omega d^4x \{ L_\phi \delta_0 \phi - \partial^\mu C_\mu \}, \quad (3.43)$$

onde

$$C^\mu = -\delta x^\nu \left( \delta^\mu_\nu \mathcal{L}(\phi, \partial_\mu \phi) - \partial_\nu \phi \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} \right) - \delta \phi \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]}. \quad (3.44)$$

Estamos finalmente prontos a definir o que entendemos como uma *simetria da teoria*. Dize-mos que uma transformação da forma (3.32) é uma *simetria* se o valor da ação  $S[\phi]$  é invariante frente a esta transformação, ou seja, se  $\delta S[\phi] = 0$ .

Logo, se a transformação (3.32) é uma simetria, para um *campo arbitrário*  $\phi(x)$  temos que

$$L_\phi \delta_0 \phi_A - \partial_\mu C^\mu = 0.$$

Em particular, se  $\phi(x)$  é solução das equações de movimento (3.7),  $L_\phi = 0$  e portanto

$$\partial_\mu C^\mu = 0, \quad (3.45)$$

de forma que  $C^\mu$  é uma *corrente conservada*.

Tipicamente, as simetrias da natureza são descritas por *grupos contínuos (ou de Lie)*<sup>3</sup>. Um resultado fundamental da teoria de grupos de Lie é que qualquer transformação pode ser obtida por um número finito de *geradores* que fazem parte de uma *álgebra de Lie*. Suponha, portanto, que consideramos uma simetria descrita por um grupo  $\mathcal{G}$  contendo  $N$  geradores. Isso significa que a transformação infinitesimal (3.32) depende dos  $N$  geradores do grupo e de  $N$  coeficientes

<sup>3</sup>O leitor encontrará uma brevíssima revisão dos conceitos mais fundamentais no Suplemento Matemático 1.1.

infinitesimais  $\varepsilon_\ell$ , ou seja, é possível escrever

$$\begin{cases} \delta x^\mu = \left( \frac{\delta x^\mu}{\delta \varepsilon^\ell} \right) \varepsilon_\ell \\ \delta \phi = \left( \frac{\delta \phi}{\delta \varepsilon^\ell} \right) \varepsilon_\ell \end{cases}, \quad \ell = 1, \dots, N, \quad (3.46)$$

onde  $\left( \frac{\delta x^\mu}{\delta \varepsilon^\ell} \right)$  e  $\left( \frac{\delta \phi}{\delta \varepsilon^\ell} \right)$  representam simbolicamente uma representação adequada de cada um dos geradores do grupo  $\mathcal{G}$ .

Inserindo (3.46) em (3.45), vemos que existem  $N$  correntes conservadas  $C_\ell^\mu$ , dadas por

$$C_\ell^\mu \equiv \left( \frac{\delta x^\mu}{\delta \varepsilon^\ell} \right) \left( \partial_\nu \phi \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} - \delta_\nu^\mu \mathcal{L}(\phi, \partial_\mu \phi) \right) - \left( \frac{\delta \phi}{\delta \varepsilon^\ell} \right) \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]}. \quad (3.47)$$

Este é o enunciado do principal resultado de Noether: *a cada gerador de um grupo de simetria da teoria considerada, existe uma corrente conservada associada, dada pela Eq. (3.47).*

Vamos discutir em mais detalhes em que sentido a equação (3.45) define uma lei de conservação. Note que  $C^\mu$  pode ser, em princípio, uma função complicada de  $\phi$  e  $\partial_\mu \phi$  que, por sua vez, dependem de  $x^\mu$ . Logo, podemos definir a partir de  $C^\mu$

$$\begin{cases} \rho(\vec{x}, t) = C^0(\vec{x}, t) \\ j^i(\vec{x}, t) = C^i(\vec{x}, t) \end{cases}. \quad (3.48)$$

De (3.45), obtemos

$$\partial_0 C^0 + \partial_i C^i = 0 \Rightarrow \frac{\partial}{\partial t} \rho(\vec{x}, t) + \nabla \cdot \vec{j}(\vec{x}, t) = 0, \quad (3.49)$$

que é uma *equação da continuidade*. Considere um volume tridimensional  $V$  arbitrário, e integre os dois membros da equação anterior neste volume: de um lado, teremos

$$\int_V d^3x \frac{\partial}{\partial t} \rho(\vec{x}, t) = \frac{d}{dt} \int_V d^3x \rho(\vec{x}, t),$$

por outro,

$$\int_V d^3x \nabla \cdot \vec{j}(\vec{x}, t) = \oint_{\partial V} d^3x \vec{j}(\vec{x}, t) \cdot d\vec{\sigma}.$$

Ou seja, se entendemos  $Q(t) = \int_V d^3x \rho(\vec{x}, t)$  como a quantidade de uma determinada grandeza contida em  $V$ , mostramos que

$$\frac{d}{dt} Q(t) = - \oint_{\partial V} d^3x \vec{j}(\vec{x}, t) \cdot d\vec{\sigma}, \quad (3.50)$$

onde o membro da direita corresponde ao *fluxo* desta grandeza que atravessa a fronteira de  $V$ .

Ou seja: a grandeza que consideramos é necessariamente conservada.

Por outro lado, tomando o volume  $V$  como todo o espaço, e supondo que  $\vec{j}$  tende a zero rapidamente no infinito, temos que

$$\frac{d}{dt}Q(t) = 0. \quad (3.51)$$

Desta forma a grandeza  $Q(t)$  é na verdade independente do tempo. Ou seja, embora a integral  $\int d^3x \rho(\vec{x}, t)$  formalmente depende do tempo, ao efetuarmos a integral por todo o espaço definiremos uma grandeza *independente do tempo*,

$$Q = \int d^3x \rho(\vec{x}, t). \quad (3.52)$$

Diz-se que  $Q$  é a *carga* associada à simetria considerada.

---

#### Exemplo 10. Simetria de Translação

As leis da física devem ser invariantes por uma mudança na escolha da origem do referencial de espaço e tempo. Isso reflete-se no fato de que uma *translação*,

$$\delta x^\mu = \varepsilon^\mu \quad (3.53)$$

deve ser uma transformação de simetria de qualquer teoria física. Pelo teorema de Noether, deve haver uma carga conservada associada à simetria por translação. Reescrevendo a equação (3.46) para reproduzir (3.53), lembrando que por definição o campo escalar tem que satisfazer

$$\delta\phi = \phi'(x') - \phi(x) = 0,$$

vemos que

$$\begin{cases} \delta x^\mu = (\delta^\mu_\nu) \varepsilon^\nu \\ \delta\phi = 0 \end{cases},$$

de onde podemos identificar  $\left(\frac{\delta x^\mu}{\delta \varepsilon^\alpha}\right)$  e  $\left(\frac{\delta\phi}{\delta \varepsilon^\ell}\right)$ , que inserido em (3.47) fornece

$$C^\mu_\nu = \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} \partial_\nu \phi - \delta^\mu_\nu \mathcal{L},$$

ou, equivalentemente

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial [\partial_\mu \phi]} \partial^\nu \phi - \eta^{\mu\nu} \mathcal{L}. \quad (3.54)$$

O tensor  $T^{\mu\nu}$  é chamado de *tensor energia-momento*, e terá fundamental importância para a definição da teoria quântica do campo escalar. Conforme o teorema de Noether, a componente

$\mu = 0$  do tensor de energia momento define a densidade de uma grandeza conservada,

$$P^\nu = \int d^3x T^{0\nu}, \quad (3.55)$$

grandeza que oportunamente identificaremos como o *momento linear*.

Considere por exemplo a Lagrangiana do campo escalar real (3.10),

$$\mathcal{L}(\phi, \partial_\mu \phi) = \frac{1}{2} \left( \partial^\mu \phi \partial_\mu \phi - m^2 \phi^2 \right) - V(\phi).$$

Calculando explicitamente  $T^{\mu\nu}$  neste caso, encontraremos

$$T^{\mu\nu} = \partial^\mu \phi \partial^\nu \phi - \frac{1}{2} \eta^{\mu\nu} \left( \partial^\rho \phi \partial_\rho \phi - m^2 \phi^2 \right) + \eta^{\mu\nu} V(\phi). \quad (3.56)$$

Em particular:

$$\begin{aligned} T^{00} &= \left( \partial^0 \phi \right)^2 - \frac{1}{2} \eta^{00} \left( \left( \partial^0 \phi \right)^2 - (\nabla \phi)^2 - m^2 \phi^2 \right) + \eta^{00} V(\phi) \\ &= \frac{1}{2} \left( \partial^0 \phi \right)^2 + \frac{1}{2} (\nabla \phi)^2 + \frac{1}{2} m^2 \phi^2 + V(\phi), \end{aligned} \quad (3.57)$$

e

$$T^{0i} = \partial^0 \phi \partial^i \phi = -\dot{\phi} \nabla \phi. \quad (3.58)$$

Note que  $T^{00}$  é uma grandeza estritamente positiva: este tinha que ser justamente o caso pois, como veremos adiante,  $T^{00}$  será interpretado como a densidade de energia associado ao campo.



## Noções de Teoria das Distribuições

Considere uma carga puntiforme  $q$  localizada na origem. Podemos nos perguntar qual a densidade de carga associada a esta carga: uma função  $\rho(\mathbf{r})$  tal que

$$\int_{\text{todo o espaço}} \rho(\mathbf{r}) d^3V = q$$

mas, ao mesmo tempo,

$$\rho(\mathbf{r}) = 0 \text{ se } \mathbf{r} \neq \mathbf{0}.$$

O problema é que não existe nenhuma função que satisfaça, ao mesmo tempo, as duas condições acima. O físico Paul A. M. Dirac percebeu que podia-se trabalhar com “funções generalizadas”, que satisfizessem estas propriedades. O uso de tais “funções” está cuidadosamente justificado por um ramo da matemática chamada de *teoria das distribuições* ou *das funções generalizadas*. Vamos dar uma noção intuitiva desta teoria, e mostrar as regras práticas que permitem o uso de distribuições para a modelagem e solução de problemas físicos.

Suponha que tenhamos uma certa quantidade de carga, digamos  $q = 1C$ , numa certa região de largura  $L$  próxima à origem. Uma possível densidade de carga seria a função

$$\rho_L(x) = \begin{cases} \frac{1}{L}, & \text{se } -\frac{L}{2} < x < \frac{L}{2} \\ 0, & \text{se } x < -\frac{L}{2} \text{ ou } x > \frac{L}{2} \end{cases}$$

Note que  $\int_{-\infty}^{\infty} \rho_L(x) dx = 1$  o que justamente nos diz que  $\rho_L(x)$  descreve uma distribuição de carga total  $1C$ .

Pensando em modelar uma partícula pontual, tomamos o limite  $L \rightarrow 0$  da função  $\rho_L(x)$ , obtendo uma função  $\delta(x)$  “infinitamente concentrada” em  $x = 0$ . O problema é que esta função não está bem definida em  $x = 0$ . Seríamos tentados a escrever

$$\delta(x) = \lim_{L \rightarrow 0} \rho_L(x) = \begin{cases} 0, & \text{se } x \neq 0 \\ \infty, & \text{se } x = 0 \end{cases}.$$

Contudo, o valor de uma integral

$$\int f(x) dx$$

não muda se mudamos o valor da função  $f(x)$  num único ponto  $x_0$ . Ou seja, mesmo que eliminemos a divergência em  $x = 0$  associando um valor arbitrário qualquer para  $\delta(0)$ , a integral  $\int \delta(x) dx$  não vai diferir da integral de uma função  $f(x)$  que é identicamente nula (pois  $\delta$  e  $f$  só diferem num único ponto) – ou seja, teremos um resultado nulo.

Isto significa que não existe nenhuma função que faz o papel de uma carga puntiforme na origem. Contudo, a “função”  $\delta(x)$  definida acima intuitivamente representa a idéia de uma carga totalmente concentrada na origem, se entendermos sua integral da forma

$$\int_{-\infty}^{\infty} \delta(x) dx = \lim_{L \rightarrow 0} \int_{-\infty}^{\infty} \rho_L(x) dx = 1.$$

Podemos multiplicar a “função”  $\delta(x)$  por uma função contínua  $f(x)$ , e obter uma expressão bem definida, se integrada conforme

$$\int_{-\infty}^{\infty} f(x) \delta(x) dx = \lim_{L \rightarrow 0} \int_{-\infty}^{\infty} f(x) \rho_L(x) dx.$$

Para calcular o limite, note que  $\rho_L(x) = 0$  se  $|x| > L/2$ , logo

$$\int_{-\infty}^{\infty} f(x) \rho_L(x) dx = \int_{-L/2}^{L/2} f(x) \rho_L(x) dx$$

e, para  $L$  suficientemente pequeno, podemos tomar  $f$  aproximadamente constante, igual a  $f(0)$ , na região de integração,

$$\int_{-\infty}^{\infty} f(x) \rho_L(x) dx \approx f(0) \underbrace{\int_{-L/2}^{L/2} \rho_L(x) dx}_1 = f(0).$$

Portanto, tomando o limite  $L \rightarrow 0$ ,

$$\int_{-\infty}^{\infty} f(x) \delta(x) dx = f(0),$$

o que nos permite escrever, no sentido de distribuição,

$$f(x) \delta(x) = f(0) \delta(x).$$

Esta igualdade é razoável já que  $\delta(x) = 0$  se  $x \neq 0$ , portanto, o produto  $f(x) \delta(x)$  não pode depender do valor de  $f$  em qualquer  $x \neq 0$ .

Obviamente, o ponto  $x = 0$  não tem nada de especial, e poderíamos ter escolhido um  $x = a$  qualquer. Temos assim, mais geralmente, a propriedade fundamental da “função”  $\delta(x)$  como sendo

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) dx = f(a) .$$

A “função”  $\delta(x)$  é chamada de *função delta de Dirac*. A resumo da história é que  $\delta(x)$  não faz sentido se tentamos defini-la como uma função usual, definida para todo  $x$ , mas sua integração dá um resultado bem definido, obtido por um procedimento “honesto” de limite de integrais de funções regulares.

Um pouco mais formalmente: vamos considerar um espaço de “funções teste”  $\mathcal{F}$ , contendo funções  $\xi(x) \in \mathcal{F}$  que são contínuas, infinitamente diferenciáveis, e que se anulam fora de um intervalo finito. Este último requerimento garante que podemos integrar por partes qualquer integral em que apareça uma função de teste, sem gerar termos de superfície.

Dada qualquer função  $f(x)$  razoavelmente suave<sup>1</sup>, claramente podemos definir uma operação que associa a cada função teste  $\xi(x)$  o valor  $f[\xi]$  dado pela integral

$$f[\xi] = \int_{-\infty}^{\infty} f(x) \xi(x) dx . \quad (\text{A.1})$$

Definimos, assim, uma *funcional* no espaço de funções teste, associada à função  $f$ . Estas funcionais serão chamadas de *distribuições*.

Considere agora uma distribuição que chamaremos de  $\delta_{x_0}$ , que satisfaz

$$\delta_{x_0}[\xi] \equiv \xi(x_0) , \quad (\text{A.2})$$

para qualquer funções de teste  $\xi(x)$ . Aqui,  $x_0$  é uma parâmetro real arbitrário. Como vimos, não existe nenhuma função  $\delta(x - x_0)$  que satisfaça

$$\int_{-\infty}^{\infty} \delta(x - x_0) \xi(x) dx = \xi(x_0) \quad (\text{A.3})$$

e no entanto a *funcional*  $\delta_{x_0}$  está bem definida. Como vimos, podemos construir sequencias de funções  $\delta_L(x - x_0)$  que, num limite apropriado, produzem o resultado desejado. Ou, como tradicionalmente é feito, omitimos a necessidade deste processo de limite e formalmente consideramos uma “função”  $\delta(x - x_0)$  satisfazendo (A.3). Este procedimento, aparentemente pouco formal, nunca vai nos levar a nenhuma inconsistência desde que entendemos que a “função”  $\delta(x - x_0)$  só faz sentido se *aparecer sob o sinal de integral, e multiplicada por uma função teste*. Desta

<sup>1</sup>As condições de suavidade sobre  $f$  são muito mais fracas do que sobre as funções teste  $\xi$ : a continuidade de  $f$  já é suficiente para garantir a convergência da integral.

forma, podemos mesmo aplicar à “função”  $\delta(x - x_0)$  operações como derivação, por exemplo.

Por exemplo, se  $g(x)$  é uma função contínua, que significado podemos dar a um produto da forma  $g(x)\delta(x - x_0)$ ? A resposta é obtida integrando-se com uma função teste,

$$\int g(x)\delta(x - x_0)\xi(x)dx = g(x_0)\xi(x_0) .$$

Note que, se  $\xi$  é uma função de teste, o produto  $g\xi$  também o será, desde que  $g$  seja uma função suficientemente bem comportada. Isto significa que podemos fazer a identificação

$$g(x)\delta(x - x_0) = g(x_0)\delta(x - x_0) ,$$

desde que os dois membros da equação produzem o mesmo resultado quando integrado com uma função teste.

Por outro lado, podemos considerar a derivada  $\partial_x\delta(x - x_0)$ . O resultado de tal expressão, lembrando que  $\xi$  se anula no infinito e portanto podemos integrar por partes sem produzir qualquer termo de superfície, é

$$\int \partial_x\delta(x - x_0)\xi(x)dx = \int \delta(x - x_0)(-\partial_x\xi(x))dx = -\partial_x\xi(x_0) .$$

Ou seja: mesmo que  $\delta(x - x_0)$  não seja uma função, no sentido de distribuição a derivada  $\partial_x\delta(x - x_0)$  está bem definida! Por outro lado, podemos também calcular a derivada de  $\delta(x - x_0)$  com relação ao parâmetro  $x_0$ . O resultado:

$$\int (\partial_{x_0}\delta(x - x_0))\xi(x)dx = \partial_{x_0} \int \delta(x - x_0)\xi(x)dx = \partial_{x_0}\xi(x_0) .$$

Considere agora a expressão

$$g(x)\partial_x\delta(x - x_0) ,$$

em que  $g$  é uma função contínua e derivável. Integrando com uma função teste, podemos escrever

$$\begin{aligned} \int g(x)\partial_x\delta(x - x_0)\xi(x)dx &= - \int \delta(x - x_0)\partial_x(g(x)\xi(x))dx \\ &= - \int \delta(x - x_0)\partial_xg(x)\xi(x)dx - \int \delta(x - x_0)g(x)\partial_x\xi(x)dx \\ &= -\partial_xg(x_0) \int \delta(x - x_0)\xi(x)dx - g(x_0) \int \delta(x - x_0)\partial_x\xi(x)dx \\ &= \int [-\partial_xg(x_0)\delta(x - x_0) + g(x_0)\partial_x\delta(x - x_0)]\xi(x)dx \end{aligned}$$

ou seja,

$$g(x) \partial_x \delta(x - x_0) = g(x_0) \partial_x \delta(x - x_0) - \partial_x g(x_0) \delta(x - x_0) .$$

Este último exemplo deve ser estudado com cuidado pelo leitor, já que mostra a importância de considerar a presença da função teste ao manipular uma expressão envolvendo “funções”  $\delta$ .