



Métodos Algébricos da Física Teórica

Níckolas de Aguiar Alves ^{ID}

Versão preliminar de 29 de julho de 2021

RESUMO: Estas são notas de aula para um minicurso ministrado durante a III Escola Jayme Tiomno de Física Teórica, organizada entre 02 e 06 de agosto de 2021 pela Dead Physicists Society de modo online, embora voltada aos alunos de graduação do Instituto de Física da Universidade de São Paulo. O minicurso aborda uma introdução à teoria de grupos e representações para físicos com alguns exemplos e motivações vindos da Relatividade Geral e Física de Partículas.

PALAVRAS-CHAVE: Grupos, Tensores, Representações, Relatividade Restrita, Relatividade Geral, Física de Partículas, Teorias de Grande Unificação, Teoria Quântica de Campos em Espaços-Tempos Curvos.

Sumário

1	Grupos	1
1.1	Simetrias	1
1.2	Rotações	2
1.3	O Grupo de Galilei	7
1.4	Relatividade Restrita	11
2	Grupos e Álgebras de Lie	19
2.1	Noção de Grupo Contínuo	19
2.2	Álgebras de Lie	21
2.3	Grupo de Heisenberg	30
2.4	Grupo Especial Unitário em Duas Dimensões	32
3	Representações	37
3.1	Noções de Mecânica Quântica	37
3.2	Spin Eletrônico	39
3.3	Construções Básicas com Representações	41
3.4	Produtos Tensoriais e Tensores	46
3.5	Representações Projetivas e Spinors	51
3.6	Lema de Schur	52
4	Aplicações Físicas	54
4.1	O Caminho Óctuplo da Física de Partículas	54
4.2	Teorias de Grande Unificação	60
4.3	Efeitos Hawking e Unruh	60
	Referências	60

1 Grupos

Isn't it surprising that we can do physics at all?

K. SUNDERMEYER[30].

1.1 Simetrias

Suponha que desejemos descrever um certo sistema físico. Por exemplo, o movimento livre de uma partícula em Mecânica Clássica. O primeiro passo para estudar tal problema envolve fazer uma escolha: em qual referencial estamos trabalhando?

Há diversas escolhas necessárias para descrever o problema. Por exemplo, qual a posição inicial da partícula? Qual sua rapidez? Em qual direção ela se move? Em que instante o cronômetro foi iniciado?

A existência destas liberdades é reflexo das simetrias do sistema com o qual estamos lidando. Quando descrevemos o movimento de uma partícula livre, o Universo não se importa com qual ponto escolhemos chamar de origem. Ou com quando iniciamos o cronômetro. Ou com qual direção dizemos ser “para cima”. Na verdade, como podemos mudar de referencial ao considerar um observador se movendo com velocidade constante, o Universo não se importa sequer com qual a velocidade que atribuímos à partícula.

Embora a princípio isto pareça uma mera curiosidade, o estudo das simetrias de um sistema nos permitirá obter uma compreensão mais profunda sobre a Física de um sistema. De fato, pode-se dizer que as simetrias de um sistema determinam a sua dinâmica. As simetrias da Mecânica Clássica que mencionamos anteriormente são suficientes para obter de modo quase imediato a Lei da Inércia, por exemplo[2, 21].

Mesmo que não houvesse uma conexão tão profunda, como poderíamos estudar Física se não existissem tantas simetrias? Se as Leis de Newton valessem em apenas um único referencial privilegiado, quão complexo seria o movimento descrito em outros referenciais? E se não estivéssemos em um referencial tão privilegiado? Poderíamos ter que lidar com dinâmicas muito mais difíceis do que as que observamos de fato, e isso poderia tornar nosso trabalho completamente inviável.

O poder das simetrias é tão profundo e vasto que permeia a Física Fundamental. Ao estudar o Universo em grande escala, por exemplo, frequentemente assumimos que o espaço é homogêneo (não possui uma posição preferencial) e isotrópico (não possui uma direção preferencial)[31, Cap. 5]. Ao estudar modelos simples de buracos negros, assumimos que o espaço é esfericamente simétrico e que o espaço-tempo é estático no tempo[31, Cap. 6]. Podemos estudar alguns processos nucleares ao introduzir simetrias aproximadas em que prótons e nêutrons são tratadas como aspectos diferentes de uma só partícula[33, Sec. 4.6]. O Modelo Padrão da Física de Partículas — bem como tentativas de extensões conhecidas como Teorias de Grande Unificação — é profundamente baseado na noção de simetria[37, Partes IV e VII].

Em alguns destes casos, as simetrias nos permitem tornar o problema mais tratável. A existência de estrelas, planetas e galáxias nos dá certeza de que os arredores de um buraco negro não são esfericamente simétricos de maneira exata, mas apenas de forma aproximada. Em outros casos, como nos exemplos envolvendo Física de Partículas ou Gravitação Quântica, estas simetrias poderiam ser propriedades exatas do Universo.

Para motivarmos os conceitos matemáticos que utilizaremos em nosso estudo, comecemos considerando um exemplo familiar: a simetria esférica.

Links pros casos discutidos nas notas

1.2 Rotações

Diversos sistemas de interesse físico são modelados com o uso da simetria esférica. Por exemplo, o buraco negro de Schwarzschild ou o átomo de hidrogênio. Mas o que entendemos como sendo simetria esférica?

Um sistema físico é esfericamente simétrico quando é invariante por rotações. Ao rotacionarmos o sistema de coordenadas (ou todo o conteúdo do sistema, como preferir), a Física deve permanecer a mesma. Por exemplo, um buraco negro de Schwarzschild

possui a mesma aparência quando visto de cima, baixo, pela esquerda, pela direita, etc. Contudo, como podemos descrever matematicamente uma rotação?

Sabemos intuitivamente que rotações são transformações que preservam os ângulos entre vetores. Se temos dois vetores \mathbf{x} e \mathbf{y} e os rodamos para $R\mathbf{x}$ e $R\mathbf{y}$, os ângulos entre os dois vetores devem permanecer iguais. Além disso, a norma de \mathbf{x} e de $R\mathbf{x}$ devem coincidir: rotações não esticam nem encolhem vetores.

Ainda é razoável assumirmos que as rotações se comportam de modo que

$$R(\mathbf{x} + \mathbf{y}) = R\mathbf{x} + R\mathbf{y} \quad (1.1)$$

e

$$R(\alpha\mathbf{x}) = \alpha R\mathbf{x}, \quad (1.2)$$

onde α é um número real qualquer. A primeira propriedade pode ser entendida pela noção de que vetores que estão sendo somados são rotacionados juntos, de maneira “sólida”, e assim também é rotacionada a sua soma. Quanto à segunda propriedade, esperamos que rotações sejam “cegas” para o tamanho dos vetores: elas se importam apenas com a direção.

Em conjunto, as propriedades Eqs. (1.1) e (1.2) dizem que R é uma transformação linear. Pode-se mostrar que tais transformações podem ser representadas por matrizes [8, Sec. 3.4]. Assim, estamos interessados nas matrizes reais quadradas 3×3 que preservam ângulos e normas de vetores. Oras, ângulos e normas são calculados por meio do produto escalar. Assim, buscamos matrizes 3×3 com

$$R\mathbf{x} \cdot R\mathbf{y} = \mathbf{x} \cdot \mathbf{y}. \quad (1.3)$$

Exercício 1:

Seja R uma matriz quadrada com coeficientes reais. Mostre que $R^T R = \mathbb{1}$ se, e somente se,

$$R\mathbf{x} \cdot R\mathbf{y} = \mathbf{x} \cdot \mathbf{y} \quad (1.4)$$

para quaisquer dois vetores \mathbf{x} e \mathbf{y} .

Feito isso, mostra que $R^T R = \mathbb{1}$ se, e somente se,

$$\|R\mathbf{x}\| = \|\mathbf{x}\| \quad (1.5)$$

para todo vetor $\mathbf{x} \in \mathbb{R}^3$. Sugestão: prove a chamada identidade de polarização,

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 \right), \quad (1.6)$$

e a utilize para concluir a demonstração.

Acima, e no restante deste texto, R^T denota a matriz transposta de R e $\mathbb{1}$ denota a matriz identidade. \blacklozenge

Exercício 2:

Seja R uma matriz quadrada com coeficientes reais e tal que $R^T R = \mathbb{1}$. Mostre que se $\det R = -1$, então a transformação R mapeia um sistema de coordenadas de mão direita a um de mão esquerda e vice-versa. \blacklozenge

Deste modo, estamos interessados em estudar as propriedades do espaço

$$\mathrm{SO}(3) = \{R \in \mathrm{SL}_3(\mathbb{R}); R^\top R = \mathbb{1}\}. \quad (1.7)$$

A notação $\mathrm{SL}_3(\mathbb{R})$ pode parecer misteriosa, mas denota apenas as matrizes 3×3 com coeficientes reais e determinante 1. De modo geral, denotaremos por $\mathbb{M}_N(\mathbb{K})$ o espaço das matrizes $N \times N$ com coeficientes em algum conjunto \mathbb{K} . Além disso, definimos

$$\mathrm{GL}_{\mathbb{K}}(N) = \{M \in \mathbb{M}_N(\mathbb{K}); \det M \neq 0\}, \quad (1.8)$$

e

$$\mathrm{SL}_N(\mathbb{K}) = \{M \in \mathbb{M}_N(\mathbb{K}); \det M = 1\}. \quad (1.9)$$

“GL” abrevia o nome *general linear group*, ou grupo geral linear, que usamos para nos referir ao espaço das matrizes com determinante não-nulo. “SL”, de forma análoga, vem de *special linear group*, ou grupo linear especial, que diz respeito às matrizes com determinante 1.

Naturalmente, o “S” em $\mathrm{SO}(3)$ também significa *special*. O “O” por sua vez provém de *orthogonal*, ou ortogonal, que é o nome dado às matrizes quadradas com entradas reais e satisfazendo $R^\top R = \mathbb{1}$. Em particular, é comum definir

$$\mathrm{O}(3) = \{R \in \mathrm{GL}_{\mathbb{R}}(3); R^\top R = \mathbb{1}\}. \quad (1.10)$$

De todo modo, nosso interesse é em $\mathrm{SO}(3)$. Quais as propriedades básicas deste conjunto?

Dadas duas rotações quaisquer $R, S \in \mathrm{SO}(3)$, sua composição é também rotação. Dado um livro, por exemplo, podemos rodá-lo segundo R e imediatamente depois rodá-lo segundo S . Matematicamente, expressamos isso ao notar que

$$(RS)^\top RS = S^\top R^\top RS, \quad (1.11a)$$

$$= S^\top \mathbb{1} S, \quad (1.11b)$$

$$= S^\top S, \quad (1.11c)$$

$$= \mathbb{1}, \quad (1.11d)$$

onde usamos $R^\top R = \mathbb{1}$ e $S^\top S = \mathbb{1}$. Isto mostra que $RS \in \mathrm{O}(3)$. Porém, vemos ainda que

$$\det(RS) = \det(R) \cdot \det(S), \quad (1.12a)$$

$$= 1 \cdot 1, \quad (1.12b)$$

$$= 1, \quad (1.12c)$$

e assim $RS \in \mathrm{SO}(3)$.

Isso vale para rotações, mas é uma propriedade que esperaríamos de qualquer simetria. Começamos com o sistema em uma certa situação \mathcal{S} e aplicamos alguma transformação, como uma rotação ou translação, e obtemos a situação \mathcal{S}' , que não podemos distinguir de \mathcal{S} . Fazemos uma nova transformação sobre \mathcal{S}' e obtemos \mathcal{S}'' , que não conseguimos

distinguir de \mathcal{S}' . Ora, por que poderíamos distinguir \mathcal{S} de \mathcal{S}'' ? Assim, vemos que uma das propriedades que esperamos de simetrias certamente é que a composição de simetrias é também uma simetria.

Podemos a seguir considerar o que ocorre ao efetuar três rotações em sequência.

Exercício 3:

Mostre que, dadas três matrizes $n \times n$ com coeficientes complexos A , B e C vale que $(AB)C = A(BC)$. Ou seja, prove que o produto de matrizes com entradas complexas é associativo. \clubsuit

Como o produto de matrizes é associativo, o produto de elementos de $\text{SO}(3)$ também o é. De fato, chega a ser complexo pensarmos em como seria possível introduzir “parênteses” em rotações físicas. Sendo a associatividade uma propriedade algébrica particularmente simples, que nos permitirá descrever uma ampla gama de sistemas, exigiremos também que as simetrias satisfaçam uma propriedade associativa.

As rotações ainda possuem uma identidade: $\mathbb{1}$. De fato, $\mathbb{1}^\top \mathbb{1} = \mathbb{1}$ e $\det \mathbb{1} = 1$. Ademais, dada qualquer rotação R , $R\mathbb{1} = \mathbb{1}R = R$. Isto também é algo esperado de simetrias no geral: se o sistema está em uma configuração \mathcal{S} e não fazemos nada, ele segue na configuração \mathcal{S} , que é indistinguível de \mathcal{S} .

Rotações também são sempre inversíveis. Se giramos um corpo de uma certa forma, podemos reverter o processo. De modo matemático, isso significa que $R^{-1} \in \text{SO}(3)$ para todo $R \in \text{SO}(3)$. De fato sabemos que, dado $R \in \text{SO}(3)$,

$$\det(R^{-1}) = \det(R), \quad (1.13a)$$

$$= 1. \quad (1.13b)$$

Além disso,

$$R^\top R = \mathbb{1}, \quad (1.14)$$

$$R^\top = R^{-1}, \quad (1.15)$$

$$R = (R^{-1})^\top, \quad (1.16)$$

$$\mathbb{1} = (R^{-1})^\top R^{-1}, \quad (1.17)$$

o que prova que $R^{-1} \in \text{SO}(3)$. Perceba que inversibilidade também é uma propriedade desejada de simetrias: se podemos transformar \mathcal{S} em \mathcal{S}' , gostaríamos de poder desfazê-lo.

Assim, motivamos a definição do que é um grupo.

Definição 1 [Grupo]:

Seja G um conjunto e $\cdot : G \times G \rightarrow G$ uma função. Dizemos que (G, \cdot) é um grupo se valerem as propriedades

- i. $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$, para todos $g_1, g_2, g_3 \in G$;
- ii. existe um elemento $e \in G$, que chamaremos de identidade, tal que $e \cdot g = g \cdot e = g$, para todo $g \in G$;

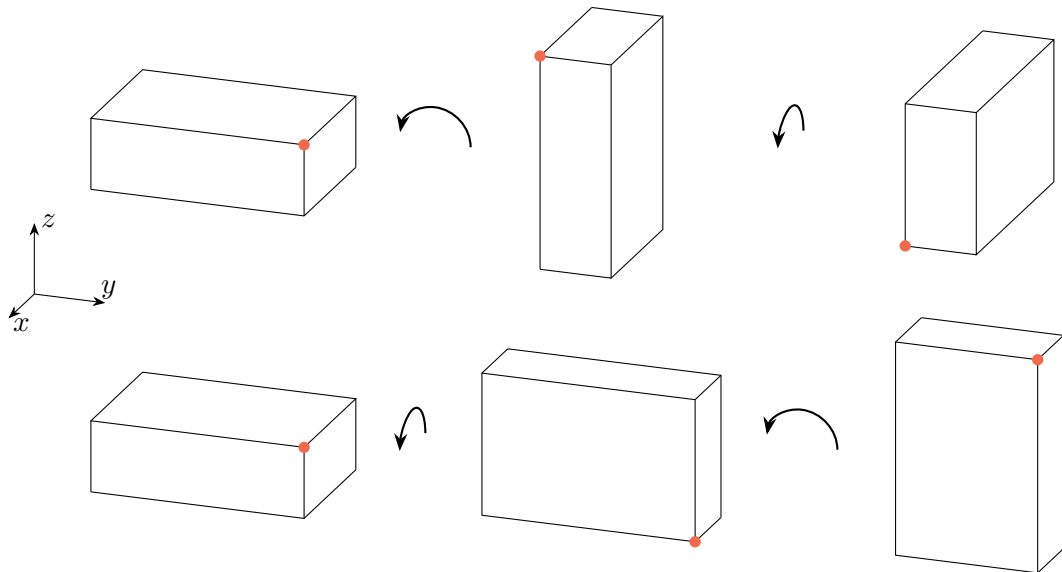


Figura 1: Pegue um livro. A partir de seu centro, escolha três eixos \hat{x} , \hat{y} e \hat{z} . Rotacionando o livro 90° em torno de \hat{x} e depois 90° em torno de \hat{y} não é a mesma coisa que rotacionar 90° em torno de \hat{y} e depois 90° em torno de \hat{x} .

- iii. para todo $g \in G$, existe $g^{-1} \in G$, que chamaremos de inverso de g , tal que $g \cdot g^{-1} = g^{-1} \cdot g = e$.

No caso particular em que $g_1 \cdot g_2 = g_2 \cdot g_1$ para todos os elementos de G , dizemos ainda que o grupo (G, \cdot) é comutativo ou Abelian. Em geral, dado um grupo (G, \cdot) , o denotaremos apenas por G sempre que o produto \cdot for claro pelo contexto. ♠

Note que o grupo de rotações em três dimensões não é Abelian. Pegue um livro. A partir de seu centro, escolha três eixos \hat{x} , \hat{y} e \hat{z} . Rotacionando o livro 90° em torno de \hat{x} e depois 90° em torno de \hat{y} não é a mesma coisa que rotacionar 90° em torno de \hat{y} e depois 90° em torno de \hat{x} . Veja, por exemplo, a Figura 1.

Exercício 4:

Convença-se, com base em argumentos físicos, que o grupo de rotações em duas dimensões,

$$\text{SO}(2) = \{R \in \text{SL}_2(\mathbb{R}); R^T R = \mathbb{1}\}, \quad (1.18)$$

é Abelian. Feito isso, mostre que todo elemento $R(\theta)$ de $\text{SO}(2)$ pode ser escrito na forma

$$R(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad (1.19)$$

onde $\theta \in [0, 2\pi)$. Use esta expressão geral para provar que $\text{SO}(2)$ é Abelian. ✖

Embora as rotações definam um grupo, pode ser pouco claro o motivo de ser interessante pensarmos nela desta forma. O que ganhamos com esse tratamento? Não estamos apenas introduzindo uma quantidade desnecessária de Matemática?

Diferentes pesquisadores têm diferentes respostas para esta pergunta, sendo que a resposta que daremos aqui é inspirada em [12, Cap. 1]. A Matemática frequentemente se preocupa em estudar de maneira concentrada e profunda quais são as propriedades de alguma estrutura. Por exemplo, os grupos. Enquanto físicos, é interessante tomarmos conhecimento desses estudos porque algumas das estruturas pelas quais os matemáticos se interessam (se não todas) ocorrem em Física, como as próprias rotações que acabamos de mencionar. Ao aprender sobre as propriedades dos grupos, aprendemos imediatamente sobre propriedades não apenas de rotações, mas de qualquer outro contexto físico em que a estrutura de grupo surja.

Como um exemplo exageradamente simples, temos a unicidade de identidade.

Proposição 2:

Seja (G, \cdot) um grupo. Existe um único elemento $e \in G$ tal que

$$e \cdot g = g \cdot e = g \quad (1.20)$$

para todo $g \in G$. □

Demonstração:

Suponha que $e, e^* \in G$ sejam elementos com esta propriedade. Então vale que

$$e = e \cdot e^* = e^* \quad (1.21)$$

e portanto ambos são iguais, provando a unicidade. ■

Este simples resultado vale imediatamente para qualquer estrutura que satisfaça as propriedades de grupo (na verdade, poderíamos ter assumido hipóteses ainda mais simples). Sejam as rotações ou qualquer outro grupo que considerarmos daqui em diante, sabemos de antemão que ele possui uma única identidade. Embora estas notas sejam voltadas a exemplos particulares na Física Fundamental, diversas das técnicas que estudaremos admitem este nível de generalidade.

Vamos então a outros exemplos de grupos.

1.3 O Grupo de Galilei

Como um primeiro exemplo, vamos considerar as simetrias presentes no espaço-tempo da Mecânica Clássica. Nossa discussão é inspirada por [4, 21, 23]. Notamos que toda esta discussão se concentra no caso particular da Mecânica Clássica Newtoniana.

Começamos por algumas observações experimentais e filosóficas. Começamos notando que vivemos em um Universo com três dimensões espaciais e uma dimensão temporal. Em uma grande porção de fenômenos, podemos tomá-lo como sendo geometricamente plano. Ou seja, podemos descrever o espaço-tempo utilizando geometria Euclideana.

A seguir, reconhecemos a existência de sistemas de referência “inerciais”. É possível tomar um referencial de modo que o espaço-tempo se apresente como homogêneo e espacialmente isotrópico.

“Homogêneo” remete à noção de que não há uma origem preferida no Universo. Não há um ponto com uma placa indicando “aqui começa o Universo”. Embora não tenhamos uma

prova disto, é um princípio que parece razoável com o que observamos experimentalmente* e filosoficamente aceitável. Assim, as translações espaciais e temporais são simetrias do espaço-tempo.

“Isotropia”, por sua vez, diz respeito à ausência de direção preferida. Ao fixarmos um instante de tempo e um ponto no espaço, não há nenhuma direção naturalmente privilegiada pela Física. O espaço é esfericamente simétrico ao redor deste ponto. Como o espaço-tempo é homogêneo, isso vale para qualquer ponto no espaço-tempo em particular. Assim, rotações são também simetrias do espaço-tempo.

Isto ainda não cobre todas as simetrias disponíveis. Suponha um referencial \mathcal{S}' se movendo em linha reta com velocidade constante \mathbf{V} em relação ao referencial inicial \mathcal{S} . Isto corresponde a transformar as posições das partículas por $\mathbf{r} \rightarrow \mathbf{r}' = \mathbf{r} - \mathbf{V}t$, como representado na Figura 2, enquanto mantemos o tempo inalterado $t \rightarrow t' = t$ — ou seja, usamos o mesmo cronômetro. Estas transformações são conhecidas como boosts de Galilei. As leis do movimento livre se mantêm inalteradas por esta transformação: partículas livres se movem com velocidade constante em ambos os referenciais (embora a velocidade em \mathcal{S} não seja a mesma que em \mathcal{S}').

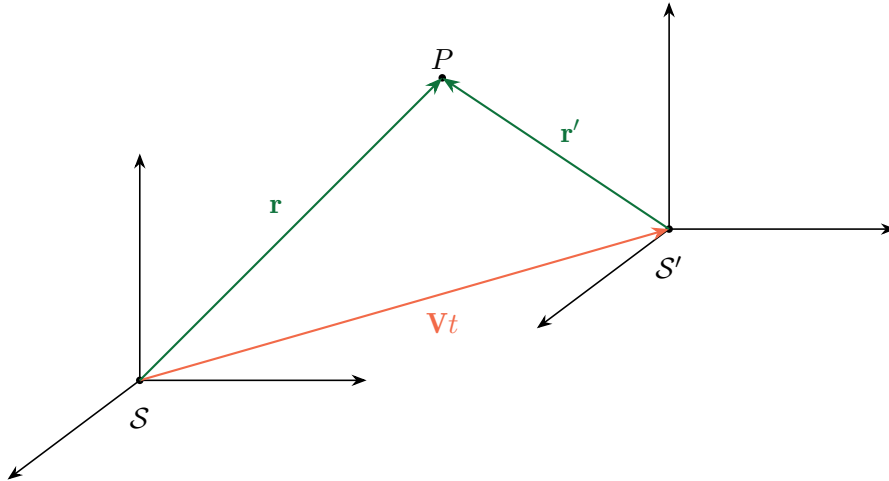


Figura 2: Um referencial \mathcal{S}' se movendo em linha reta com velocidade constante \mathbf{V} em relação ao referencial inicial \mathcal{S} .

Experimentalmente conclui-se que não apenas os movimentos de partículas livres seguem as mesmas leis tanto em \mathcal{S} quanto em \mathcal{S}' , mas na verdade todos os movimentos os fazem. Ao invés de apenas um referencial inercial \mathcal{S} , temos uma infinidade de referenciais inerciais relacionados pelas transformações

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_x & & & \\ -v_y & & R & \\ -v_z & & & \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_0 \\ x_0 \\ y_0 \\ z_0 \end{pmatrix}, \quad (1.22)$$

*Estamos considerando apenas situações cotidianas. Ao levar em conta efeitos cosmológicos, a hipótese de homogeneidade espacial frequentemente se mantém razoável, mas a de homogeneidade temporal não.

onde $R \in \text{SO}(3)$ e $t_0, x_0, y_0, z_0, v_x, v_y, v_z \in \mathbb{R}$. A matriz R representa uma rotação genérica, as constantes t_0, x_0, y_0, z_0 representam translações temporais e espaciais e as constantes v_x, v_y, v_z representam as velocidades associadas aos boosts de Galilei.

Denotaremos as transformações descritas na Eq. (1.22) por

$$G(t_0, \mathbf{r}_0, \mathbf{v}, R), \quad (1.23)$$

onde $\mathbf{v} = (v_x, v_y, v_z)$ e $\mathbf{r}_0 = (x_0, y_0, z_0)$. O conjunto de todas as transformações de referencial com a forma dada pela Eq. (1.22) na página precedente forma o chamado Grupo de Galilei próprio ortócrono*, denotado \mathcal{G}_+^\uparrow . Ou seja,

$$\mathcal{G}_+^\uparrow = \{G(t_0, \mathbf{r}_0, \mathbf{v}, R); t_0 \in \mathbb{R}, \mathbf{r}_0, \mathbf{v} \in \mathbb{R}^3, R \in \text{SO}(3)\}. \quad (1.24)$$

Exercício 5:

Suponha que a transformação de coordenadas de um referencial \mathcal{S} para um referencial \mathcal{S}' é dada por $G(t_0, \mathbf{r}_0, \mathbf{v}, R) \in \mathcal{G}_+^\uparrow$. Suponha que a transformação de coordenadas do referencial \mathcal{S}' para um referencial \mathcal{S}'' é dada por $G(t'_0, \mathbf{r}'_0, \mathbf{v}', R') \in \mathcal{G}_+^\uparrow$. Mostre que a transformação de coordenadas de \mathcal{S} para \mathcal{S}'' é dada por

$$G(t'_0 + t_0, \mathbf{r}'_0 + R'\mathbf{r}_0 + t_0\mathbf{v}', \mathbf{v}' + R'\mathbf{v}, R'R). \quad (1.25)$$

Podemos denotar isso como

$$G(t'_0, \mathbf{r}'_0, \mathbf{v}', R') \circ G(t_0, \mathbf{r}_0, \mathbf{v}, R) = G(t'_0 + t_0, \mathbf{r}'_0 + R'\mathbf{r}_0 - \mathbf{v}'t_0, \mathbf{v}' + R'\mathbf{v}, R'R). \quad (1.26)$$

Mostre que \mathcal{G}_+^\uparrow , quando equipado com a operação de composição de transformações, \circ , é um grupo. \blacktimes

Exercício 6:

Sejam $x = (t, \mathbf{r})$ e $x' = (t', \mathbf{r}')$ eventos no espaço-tempo. Vamos definir o intervalo temporal entre os dois eventos como

$$d_t(x, x') \equiv t - t' \quad (1.27)$$

e o intervalo espacial entre os dois eventos como

$$d_e(x, x') \equiv \|\mathbf{r} - \mathbf{r}'\|, \quad (1.28)$$

onde $\|\cdot\|$ denota a distância Euclideana em três dimensões.

Mostre que o grupo de Galilei próprio ortócrono preserva o intervalo temporal entre dois eventos quaisquer, isto é, mostre que

$$d_t(x, x') = d_t(Gx, Gx') \quad (1.29)$$

*“Próprio” e “ortócrono” se referem a estarmos nos limitando a transformações que não invertem o sentido do tempo, que seria implementado ao escrever -1 ao invés de $+1$ no canto superior esquerdo da matriz, e a rotações com $\det R = +1$ (proibindo o caso $\det R = -1$). Mais adiante justificaremos esta escolha.

para todo $G \in \mathcal{G}_+^\uparrow$. Ademais, mostre que o grupo de Galilei próprio ortócrono preserva o intervalo espacial entre eventos simultâneos, ou seja, vale que

$$d_e(x, x') = d_e(Gx, Gx') \quad (1.30)$$

para todo $G \in \mathcal{G}_+^\uparrow$ desde que $d_t(x, x') = 0$. A hipótese $d_t(x, x') = 0$ é necessária? O que ocorre nos casos em que $d_t(x, x') \neq 0$ e G é um boost de Galilei puro ($G = G(0, \mathbf{0}, \mathbf{v}, \mathbb{1})$) com $\mathbf{v} = \pm \frac{\mathbf{r}-\mathbf{r}'}{t-t'}$? Qual a razão física para isto? ✖

É interessante ainda contarmos quantos parâmetros são necessários para determinar completamente uma transformação de Galilei. Temos quatro translações, cada uma correspondendo a um parâmetro real. Três parâmetros determinam completamente os boosts. Restam as rotações.

Uma rotação em três dimensões pode ser descrita em termos de três ângulos, conhecidos como ângulos de Euler. Isto pode ser demonstrado formalmente*, mas também pode ser compreendido de maneira intuitiva: precisamos de dois ângulos, θ e φ , para especificar uma direção num espaço tridimensional, como fazemos no sistema de coordenadas esféricas. Isto nos permite determinar a nova direção do eixo $\hat{\mathbf{z}}$, mas ainda há uma infinidade de possíveis direções nas quais os eixos $\hat{\mathbf{x}}$ e $\hat{\mathbf{y}}$ podem apontar. Portanto, devemos especificar um terceiro ângulo, ψ , para determinar a direção dos últimos dois eixos. Como os três eixos são ortogonais e estamos assumindo sempre que a base é de mão direita (visto que estamos trabalhando com $\text{SO}(3)$), determinar a direção do eixo $\hat{\mathbf{z}}$ e a direção do eixo $\hat{\mathbf{x}}$ é suficiente para fixar o eixo $\hat{\mathbf{y}}$ de maneira única, e portanto bastam três ângulos.

Deste modo, vemos que uma transformação qualquer do grupo de Galilei próprio ortócrono é descrita em termos de dez parâmetros reais. Podemos considerar o que ocorre quando consideramos apenas alguns destes parâmetros por vez.

Exercício 7:

Mostre que os seguintes conjuntos são grupos sob a composição de transformações de Galilei:

- i. $\mathcal{G}_t = \{G(t_0, \mathbf{0}, \mathbf{0}, \mathbb{1}); t_0 \in \mathbb{R}\};$
- ii. $\mathcal{G}_e = \{G(0, \mathbf{r}_0, \mathbf{0}, \mathbb{1}); \mathbf{r}_0 \in \mathbb{R}^3\};$
- iii. $\mathcal{G}_b = \{G(0, \mathbf{0}, \mathbf{v}, \mathbb{1}); \mathbf{v} \in \mathbb{R}^3\};$
- iv. $\mathcal{G}_R = \{G(0, \mathbf{0}, \mathbf{0}, R); R \in \text{SO}(3)\}.$

Quais deles são Abelianos? ✖

Estes são exemplos de subgrupos do grupo de Galilei. Um subgrupo é, de modo intuitivo, um grupo menor dentro de um grupo maior, ambos possuindo estrutura semelhante.

Definição 3 [Subgrupo]:

Seja (G, \cdot) um grupo e seja $H \subseteq G$ um conjunto. Se (H, \cdot) for um grupo, dizemos que (H, \cdot) é subgrupo de (G, \cdot) . ♠

*Ver, por exemplo, [5]

Perceba que se (G, \cdot) é um grupo e $H \subseteq G$ um conjunto, (H, \cdot) é subgrupo de (G, \cdot) se, e somente se, valerem

- i. $h_1 \cdot h_2 \in H$, para todos h_1, h_2 em H ;
- ii. a identidade de G é elemento de H ;
- iii. o inverso de todo elemento de H está em H .

Notamos em particular que o grupo \mathcal{G}_R é, em certo sentido, semelhante ao grupo $\text{SO}(3)$.

Exercício 8:

Considere a função $\varphi: \text{SO}(3) \rightarrow \mathcal{G}_+^\uparrow$ dada por $\varphi(R) = G(0, \mathbf{0}, \mathbf{0}, R)$. Mostre que $\varphi(R \cdot R') = \varphi(R) \circ \varphi(R')$, ou seja, que φ preserva a estrutura algébrica do grupo. φ é injetora? Se sim, qual restrição deve ser feita ao contradomínio de φ para que ela seja tornada bijetora? ♣

Este é um exemplo de um homomorfismo: uma função entre grupos que preserva sua estrutura algébrica.

Definição 4 [Homomorfismos e Isomorfismos]:

Sejam (G, \cdot) e (H, \circ) grupos e $\varphi: G \rightarrow H$ uma função. Dizemos que φ é um homomorfismo de G em H se, e somente se, $\varphi(g_1 \cdot g_2) = \varphi(g_1) \circ \varphi(g_2)$, para todos g_1 e g_2 em G . Se φ for bijetora, dizemos que é um isomorfismo e que os grupos (G, \cdot) e (H, \circ) são isomorfos. ♠

Dois grupos isomorfos são indistinguíveis do ponto de vista algébrico. Embora \mathcal{G}_R e $\text{SO}(3)$ sejam definidos de maneiras completamente distintas, suas estruturas algébricas são exatamente iguais.

1.4 Relatividade Restrita

Um outro contexto em que a Teoria de Grupos é extremamente interessante é a Teoria da Relatividade Restrita, que se presta a descrever fenômenos macroscópicos (de modo que efeitos quânticos são desprezíveis), a altas velocidades (de modo que a descrição não-relativística não é mais suficiente) e em campos gravitacionais desprezíveis (de forma que efeitos da Relatividade Geral sejam desprezíveis). O tratamento que faremos aqui partirá de hipóteses mais fortes que o necessário e um tratamento mais tradicional pode ser encontrado, por exemplo, em [13, Cap. 12]. Nossa descrição é baseada na de [5].

Em geral, desejamos estudar Relatividade Restrita em $3 + 1$ dimensões (três para o espaço e uma para o tempo), mas por simplicidade analisaremos o caso simplificado em $1 + 1$ dimensões (uma dimensão espacial e uma dimensão temporal). Isto nos permitirá perceber efeitos gerais que ocorrem neste novo ambiente sem precisarmos nos preocupar com as sutilezas do espaço-tempo quadridimensional.

Ao dizer “espaço-tempo”, nos referimos ao conjunto de todos os pontos no espaço e no tempo. Como tanto o espaço quanto o tempo são números reais, vemos então que

estamos lidando com o conjunto \mathbb{R}^2 . Definimos um evento no espaço-tempo como sendo um elemento $x = (x_0, x_1)^\top \in \mathbb{R}^2$. Dado um tal evento, convencionamos que $x_0 = ct$, onde c é a velocidade da luz e t o instante de tempo em que ocorreu o evento, e que x_1 é a localização espacial do evento.

Dados dois eventos $x, y \in \mathbb{R}^2$ e escrevendo $\Delta x = y_1 - x_1$, $c\Delta t = y_0 - x_0$, há três possibilidades:

Tipo tempo

$$\frac{\Delta x^2}{\Delta t^2} < c^2, \quad (1.31a)$$

$$\Delta x^2 - c^2 \Delta t^2 < 0; \quad (1.31b)$$

Tipo luz

$$\frac{\Delta x^2}{\Delta t^2} = c^2, \quad (1.32a)$$

$$\Delta x^2 - c^2 \Delta t^2 = 0; \quad (1.32b)$$

Tipo espaço

$$\frac{\Delta x^2}{\Delta t^2} > c^2, \quad (1.33a)$$

$$\Delta x^2 - c^2 \Delta t^2 > 0. \quad (1.33b)$$

A Teoria da Relatividade Restrita proíbe a propagação de sinais com velocidade acima da velocidade da luz^{*}. Portanto, se x e y são do tipo espaço, nenhum sinal emitido em x pode alcançar y (pois isso necessitaria que o sinal viajasse acima da velocidade da luz). Por consequência, nada que tenha ocorrido em x pode ter influência sobre o que ocorrem em y , isto é, x e y não possuem ligação causal. Por outro lado, se fossem eventos do tipo tempo, seria possível a ligação causal dos dois eventos.

Perceba que estas questões de causalidade são traduzidas no estudo do sinal da quantidade

$$\Delta s^2 = \Delta x^2 - c^2 \Delta t^2. \quad (1.34)$$

Este estudo pode ser feito por meio da função

$$s(x, y) := (x_1 - y_1)^2 - (x_0 - y_0)^2, \quad (1.35)$$

ou, em notação matricial,

$$s(x, y) = \begin{pmatrix} x_0 - y_0 & x_1 - y_1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 - y_0 \\ x_1 - y_1 \end{pmatrix}. \quad (1.36)$$

^{*}Isto pode ser demonstrado (veja, por exemplo, [13, Cap. 12]), mas tomaremos como um postulado por simplicidade.

A matriz

$$\eta := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad (1.37)$$

é conhecida como métrica de Minkowski e, como veremos em breve, possui um papel central na teoria.

Voltamos agora nossa atenção ao problema de mudanças de referencial. Em Mecânica Clássica, dois referenciais distintos são conectados, em geral, por um sistema de equações da forma

$$\begin{cases} t' = t, \\ x' = x - vt, \end{cases} \quad (1.38)$$

que é o chamado boost de Galilei.

Exercício 9:

Mostre que os boosts de Galilei são precisamente os elementos do grupo \mathcal{G}_b definido no Exercício 7 na página 10. \blackbox

Esperamos então encontrar uma mudança de referencial da forma*

$$\begin{pmatrix} x'_0 \\ x'_1 \end{pmatrix} = \underbrace{\begin{pmatrix} \Lambda_{00} & \Lambda_{01} \\ \Lambda_{10} & \Lambda_{11} \end{pmatrix}}_{\Lambda} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}. \quad (1.39)$$

Como podemos restringir as possíveis opções de transformações de referencial?

O primeiro passo é considerarmos o Princípio da Relatividade de Galilei, que declara que a Física deve ser a mesma em todos os referenciais inerciais. Este princípio é uma necessidade para a consistência da Física: se a Física muda de referencial para referencial, faz algum sentido estudarmos Física em primeiro lugar? Como se poderia garantir estar em um referencial que satisfaz as leis que conhecemos?

Deste princípio, seguem duas consequências importantes para as nossas discussões: em primeiro lugar, certas constantes Físicas devem possuir os mesmos valores em todos os referenciais inerciais. Por exemplo, a velocidade da luz c , a constante de Newton G e a constante de Planck normalizada \hbar são partes essenciais da descrição de inúmeros fenômenos físicos e a alteração de seus valores alteraria a intensidade dos fenômenos associados. Por consequência, para que a Física seja a mesma, estas constantes precisam ser universais[†].

Além disto, temos a imposição de que a causalidade deve ser preservada entre diferentes referenciais. Se num dado referencial temos que o evento x não pode afetar o evento

*Estas não são todas as mudanças de referencial que podemos considerar. É possível admitir mudanças de referencial não-lineares, ou seja, que não são descritas por matrizes e isso é perfeitamente compatível com a Relatividade Restrita. Por exemplo, em $3 + 1$ dimensões, trocar de coordenadas cartesianas para esféricas é uma mudança de coordenadas não linear que claramente há de ser compatível com a teoria. Porém, não nos preocuparemos com esse nível de generalidade aqui, visto que nosso interesse é ligado à teoria de grupos.

[†]Na verdade, G pode depender das escalas de energia ou comprimento que estamos considerando, mas isso é assunto para outra hora.

y , então o mesmo deve ser satisfeito em todo outro referencial inercial. Do contrário, observadores em referenciais diferentes teriam realidades distintas, o que contradiz o Princípio da Relatividade.

Para que a causalidade seja preservada por uma mudança de referencial Λ , basta que os sinais de Δs sejam mantidos inalterados, isto é,

$$\text{sign}[s(\Lambda x, \Lambda y)] = \text{sign}[s(x, y)]. \quad (1.40)$$

A primeira proposta de transformações que podemos fazer são as dilatações: as transformações da forma $\Lambda = \lambda \mathbb{1}$, onde $\lambda \in \mathbb{R}_+^*$. No entanto, estas transformações não são físicas.

Isto pode parecer impactante a princípio. Uma dilatação apenas altera a escala em que estamos lidando, “esticando” o tempo e o espaço. Como isso pode ser problemático?

Comece por considerar o tamanho médio do átomo de hidrogênio. Embora nada nos garanta que este valor é o mesmo em todos os locais do Universo, a crença de que é segue de maneira natural da nossa crença no Princípio de Relatividade. Afinal, o tamanho do átomo de hidrogênio é definido pelas interações eletromagnéticas entre o próton e o elétron, e estamos impondo que estas interações se comportam da mesma forma.

De forma semelhante, considere a massa de repouso do elétron, m_e . Como \hbar tem dimensões de energia vezes tempo, a quantidade $\frac{\hbar}{m_e c}$ possui dimensões de comprimento (é o chamado comprimento de onda Compton do elétron). Se queremos acreditar que a massa de repouso do elétron é a mesma em todo o Universo (e não há motivação experimental para acreditar no contrário), segue como consequência um comprimento natural associado ao elétron. Uma dilatação alteraria esse comprimento natural e, por consequência, contraria a crença associada ao Princípio da Relatividade de que a Física precisa ser a mesma em todo referencial inercial.

Assim, saímos de nenhuma ideia sobre quais transformações são permitidas, obtivemos uma proposta e a descartamos. De volta à estaca zero, como podemos estudar as possíveis mudanças de referencial? A resposta para esta pergunta vem na forma do seguinte teorema.

Teorema 5:

Seja Λ uma matriz inversível 2×2 com coeficientes reais. Se Λ representa uma transformação entre sistemas de referenciais inerciais que preserva a estrutura causal do espaço-tempo e não envolve dilatações, então vale que

$$\Lambda^\top \eta \Lambda = \eta. \quad (1.41)$$

Por consequência, vale que

$$s(\Lambda x, \Lambda y) = s(x, y), \forall x, y \in \mathbb{R}^2. \quad (1.42)$$

□

Para uma demonstração, veja, por exemplo, [3, 5]. Nos concentraremos aqui em mostrar que isso é suficiente para que o sinal de $s(x, y)$ seja preservado.

Suponha que $\Lambda^\top \eta \Lambda = \eta$. Então note que

$$s(\Lambda x, \Lambda y) = [\Lambda(x - y)]^\top \eta [\Lambda(x - y)], \quad (1.43a)$$

$$= (x - y)^\top \Lambda^\top \eta \Lambda (x - y), \quad (1.43b)$$

$$= (x - y)^\top \eta (x - y), \quad (1.43c)$$

$$= s(x, y). \quad (1.43d)$$

Ou seja, as transformações Λ respeitando $\Lambda^\top \eta \Lambda = \eta$ mantém a função $s(x, y)$ invariante. O conteúdo do Teorema 5 na página precedente é que, se excluirmos as dilatações, a única forma de uma transformação de referenciais preservar o sinal de $s(x, y)$, ou seja, preservar a causalidade, é se $s(x, y)$ for invariante.

Exercício 10:

Mostre que

$$\mathcal{L} = \{\Lambda \in \text{GL}_{\mathbb{R}}(n); \Lambda^\top \eta \Lambda = \eta\} \quad (1.44)$$

é um grupo, onde $\eta = \text{diag}(-, +, \dots, +)$. \boxtimes

Exercício 11:

Mostre que se $\Lambda \in \mathcal{L}$, então $\det \Lambda = \pm 1$. \boxtimes

Agora que conhecemos o conjunto de todas as transformações conhecidas, queremos obter sua forma explícita.

Por enquanto, suporemos que $\det \Lambda = +1$, e trataremos o outro caso *a posteriori*.

Como $\eta^{-1} = \eta$, $\Lambda^\top \eta \Lambda = \eta \Rightarrow \eta \Lambda^\top \eta = \Lambda^{-1}$. Dada uma matriz 2×2 qualquer, temos que sua inversa é dada em termos de suas componentes por

$$\Lambda^{-1} = \frac{1}{\det \Lambda} \begin{pmatrix} \Lambda_{11} & -\Lambda_{01} \\ -\Lambda_{10} & \Lambda_{00} \end{pmatrix}. \quad (1.45)$$

Calculando o produto $\eta \Lambda^\top \eta$ em termos das componentes de Λ , temos que

$$\eta \Lambda^\top \eta = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda_{00} & \Lambda_{10} \\ \Lambda_{01} & \Lambda_{11} \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (1.46a)$$

$$= \begin{pmatrix} \Lambda_{00} & -\Lambda_{10} \\ -\Lambda_{01} & \Lambda_{11} \end{pmatrix}. \quad (1.46b)$$

Assim, a equação $\eta \Lambda^\top \eta = \Lambda^{-1}$ se torna (após impormos que $\det \Lambda = 1$)

$$\begin{pmatrix} \Lambda_{00} & -\Lambda_{10} \\ -\Lambda_{01} & \Lambda_{11} \end{pmatrix} = \begin{pmatrix} \Lambda_{11} & -\Lambda_{01} \\ -\Lambda_{10} & \Lambda_{00} \end{pmatrix}, \quad (1.47)$$

isto é,

$$\begin{cases} \Lambda_{00} = \Lambda_{11}, \\ \Lambda_{01} = \Lambda_{10}. \end{cases} \quad (1.48)$$

A condição $\det \Lambda = 1$ ainda nos ensina que

$$\Lambda_{00}\Lambda_{11} - \Lambda_{01}\Lambda_{10} = 1, \quad (1.49a)$$

$$\Lambda_{00}^2 = 1 + \Lambda_{01}^2, \quad (1.49b)$$

$$\Lambda_{00} = \pm \sqrt{1 + \Lambda_{01}^2}. \quad (1.49c)$$

Perceba que isto implica que $\Lambda_{00} \neq 0$. Por ora, assumiremos que $\Lambda_{00} > 0$. Lidaremos com o caso $\Lambda_{00} < 0$ *a posteriori*.

Concluimos então que $\Lambda \in \mathcal{L}$, se valermos $\det \Lambda = +1$ e $\Lambda_{00} > 0$, é da forma

$$\Lambda = \begin{pmatrix} \sqrt{1 + \Lambda_{01}^2} & \Lambda_{01} \\ \Lambda_{01} & \sqrt{1 + \Lambda_{01}^2} \end{pmatrix}. \quad (1.50)$$

Para obter Λ_{01} , precisaremos envolver um pouco mais de Física. Tome um referencial inercial, que chamaremos de \mathcal{S} . Considere uma partícula que se move com velocidade v em relação a este referencial \mathcal{S} e se encontrava na origem no instante $t = 0$. Queremos descobrir como as coordenadas de todos os eventos no referencial \mathcal{S} se apresentam no referencial \mathcal{S}' da partícula.

Em outras palavras, imagine um evento que ocorre no referencial \mathcal{S} . Por exemplo, um fóton é emitido pela partícula numa certa posição x e num certo instante de tempo t . Nosso objetivo é descobrir qual a posição x' e qual o instante de tempo t' em que um observador no referencial \mathcal{S}' observa o mesmo evento.

Como a partícula está sempre na origem do seu sistema de coordenadas, a emissão do fóton no referencial \mathcal{S}' se dá em algum instante de tempo t' , mas certamente na posição $x' = 0$. No referencial \mathcal{S} , sabemos que a partícula se move com velocidade v , então se o evento ocorreu no instante t , ele se deu na posição $x = vt$.

Temos então o seguinte sistema de equações, escrito em forma matricial:

$$\begin{pmatrix} ct' \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{1 + \Lambda_{01}^2} & \Lambda_{01} \\ \Lambda_{01} & \sqrt{1 + \Lambda_{01}^2} \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}. \quad (1.51)$$

Da segunda equação, temos que

$$0 = \Lambda_{01}ct + \sqrt{1 + \Lambda_{01}^2}vt, \quad (1.52a)$$

$$-\Lambda_{01}c = \sqrt{1 + \Lambda_{01}^2}v, \quad (1.52b)$$

$$\Lambda_{01}^2c^2 = v^2 + \Lambda_{01}^2v^2, \quad (1.52c)$$

$$\Lambda_{01}^2(c^2 - v^2) = v^2, \quad (1.52d)$$

$$\Lambda_{01}^2 = \frac{v^2}{c^2} \frac{1}{1 - \frac{v^2}{c^2}}, \quad (1.52e)$$

$$\Lambda_{01} = -\beta\gamma, \quad (1.52f)$$

onde definimos

$$\beta := \frac{v}{c}, \quad \gamma := \frac{1}{\sqrt{1 - \beta^2}}. \quad (1.53)$$

Note que o sinal negativo escolhido na Eq. (1.52f) na página anterior é devido ao fato de que a Eq. (1.52b) na página precedente implica que $\Lambda_{01} < 0$.

Como $\Lambda_{00} = +\sqrt{1 + \Lambda_{01}^2}$, temos agora que

$$\Lambda_{00} = +\sqrt{\frac{c^2 - v^2}{c^2 - v^2} + \frac{v^2}{c^2 - v^2}}, \quad (1.54a)$$

$$= +\sqrt{\frac{c^2}{c^2 - v^2}}, \quad (1.54b)$$

$$= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (1.54c)$$

$$= \gamma. \quad (1.54d)$$

Logo, se $\Lambda \in \mathcal{L}$, $\det \Lambda = +1$, $\Lambda_{00} > 0$, temos que

$$\Lambda = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix}. \quad (1.55)$$

Observe que as equações de transformação se tornam

$$\begin{cases} ct' = \gamma(ct - \beta x), \\ x' = \gamma(x - \beta ct). \end{cases} \quad (1.56)$$

Estas são as chamadas transformações de Lorentz. A simetria entre x e ct é a razão de termos escolhido trabalhar com ct ao invés de simplesmente t .

Perceba que, se tomarmos $v \ll c$, temos que $\frac{v}{c} \ll 1$ e $\gamma \approx 1$. Temos então o limite não-relativístico:

$$\begin{cases} ct' = ct, \\ x' = x - vt, \end{cases} \quad (1.57)$$

que são as transformações de Galilei.

Para $v = 0$, temos

$$\begin{cases} ct' = ct, \\ x' = x. \end{cases} \quad (1.58)$$

Estes limites indicam que estamos no caminho certo. Ao mudar para um referencial em repouso em relação ao referencial original, nada é alterado. De forma um pouco mais forte, a relatividade aparenta recuperar os resultados previamente conhecido sobre a Mecânica a baixas velocidades. Como nós sabemos que estes resultados fornecem uma boa descrição do Universo, é importante que nossa nova teoria os recupere, pois do contrário obteríamos uma descrição falha no que já conhecemos ao tentar estender o nosso conhecimento.

Em analogia com a nossa definição anterior de \mathcal{G}_+^\uparrow introduzimos as notações

$$\mathcal{L}_+^\uparrow = \{\Lambda \in \mathcal{L}; \det \Lambda = +1, \Lambda_{00} > 0\}, \quad (1.59a)$$

$$\mathcal{L}_+^\downarrow = \{\Lambda \in \mathcal{L}; \det \Lambda = +1, \Lambda_{00} < 0\}, \quad (1.59b)$$

$$\mathcal{L}_-^\uparrow = \{\Lambda \in \mathcal{L}; \det \Lambda = -1, \Lambda_{00} > 0\}, \quad (1.59c)$$

$$\mathcal{L}_-^\downarrow = \{\Lambda \in \mathcal{L}; \det \Lambda = -1, \Lambda_{00} < 0\}. \quad (1.59d)$$

$\mathcal{L}_+ = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow$ é chamado de grupo de Lorentz próprio, $\mathcal{L}^\uparrow = \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$ de grupo de Lorentz ortócrono e \mathcal{L}_+^\uparrow de grupo de Lorentz próprio ortócrono, ou ainda grupo de Lorentz restrito.

Exercício 12:

Sejam as matrizes P e T dadas por

$$P = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & 0 \\ 0 & +1 \end{pmatrix}. \quad (1.60)$$

Então mostre que vale que

$$\mathcal{L}_+^\downarrow = \{TP\Lambda; \Lambda \in \mathcal{L}_+^\uparrow\}, \quad (1.61a)$$

$$\mathcal{L}_-^\uparrow = \{P\Lambda; \Lambda \in \mathcal{L}_+^\uparrow\}, \quad (1.61b)$$

$$\mathcal{L}_-^\downarrow = \{T\Lambda; \Lambda \in \mathcal{L}_+^\uparrow\}. \quad (1.61c)$$

✚

Note que, embora estranhos, os elementos de \mathcal{L}_+^\downarrow , \mathcal{L}_-^\uparrow e \mathcal{L}_-^\downarrow ainda possuem significado físico. Eles são apenas os elementos de \mathcal{L}_+^\uparrow compostos com as operações de troca de paridade (reversão da orientações dos eixos de coordenadas espaciais), denotada por P , ou de reversão temporal, denotada por T . Caso queiramos descrever transformações de caráter geral, as escolhas de orientações eixos também precisam poder ser alteradas.

Então por que nós definimos anteriormente o grupo \mathcal{G}_+^\uparrow ao invés de uma versão mais geral que permitisse a inclusão de transformações de paridade e reversões temporais? Porque descobriu-se, em meados do século XX, que transformações de paridade e reversão temporal não são simetrias do Universo. Veja, por exemplo, [27, pp. 278–279], [33, p. 153–154].

Exercício 13:

Considere o conjunto \mathcal{P} de todas as transformações no espaço-tempo (1+3)-dimensional da forma

$$x \rightarrow \Lambda x + a, \quad (1.62)$$

onde x é um evento no espaço-tempo, $a \in \mathbb{R}^4$ e $\Lambda \in \mathcal{L}$. Mostre que \mathcal{P} é um grupo sob o produto usual de matrizes. Este grupo, que nada mais é que a versão relativística do grupo de Galilei, é o chamado grupo de Poincaré. Defina o grupo de Poincaré próprio ortócrono, \mathcal{P}_+^\uparrow , e mostre que é de fato um grupo. ✚

Exercício 14:

A definição que fizemos para um grupo pode ser reescrita em termos de axiomas um pouco mais simples. Seja G um conjunto dotado de uma operação $\cdot : G \times G \rightarrow G$ tal que valem

- i. $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$, para todos $g_1, g_2, g_3 \in G$,
- ii. existe um elemento $e \in G$ tal que $e \cdot g = g$ para todo $g \in G$,
- iii. para todo $g \in G$, existe um elemento $g^{-1} \in G$ tal que $g^{-1} \cdot g = e$.

Prove que G é um grupo. Note que todo grupo satisfaz os axiomas acima, e portanto podemos tomá-los como a definição de um grupo. Esta formulação nos permite concluir que uma certa estrutura é um grupo sem precisar checar que as expressões para identidade e inverso valem dos dois lados: basta conferir se valem à esquerda. \blacktimes

2 Grupos e Álgebras de Lie

2.1 Noção de Grupo Contínuo

Há uma forma conveniente de escrever as transformações de Lorentz da Eq. (1.55) na página 17 em termos da chamada rapidez, que parametriza o grupo de Lorentz em $(1+1)$ -dimensões através do uso de funções hiperbólicas.

Exercício 15 [Funções Hiperbólicas]:

Uma identidade frequentemente útil em Física é a chamada fórmula de Euler, $e^{i\theta} = \cos \theta + i \sin \theta$, onde i é a unidade imaginária (satisfazendo $i^2 = -1$) e θ é um ângulo real qualquer. Demonstre-a formalmente* expandindo cada uma das funções em sua série de Taylor.

A partir da fórmula de Euler, mostre que valem as expressões

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad \text{e} \quad \cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}. \quad (2.1)$$

Estas identidades podem ser entendidas como motivação para a definição das chamadas funções hiperbólicas, que são

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \text{e} \quad \cosh x = \frac{e^x + e^{-x}}{2}. \quad (2.2)$$

Mostre que $\cosh^2 x - \sinh^2 x = 1$, $\frac{d}{dx} \cosh x = \sinh x$ e $\frac{d}{dx} \sinh x = \cosh x$.

As demais funções hiperbólicas — como \tanh , sech , csch , coth , arsinh , arcosh , etc. — são definidas de maneira análoga às suas correspondentes na trigonometria esférica. \blacktimes

*Formalmente, não necessariamente rigorosamente (que envolveria preocupação com a manipulação correta de séries e boa definição de expressões envolvendo cálculo complexo). É comum utilizar o termo “formal” no sentido de cálculos feitos de maneira pouco rigorosa.

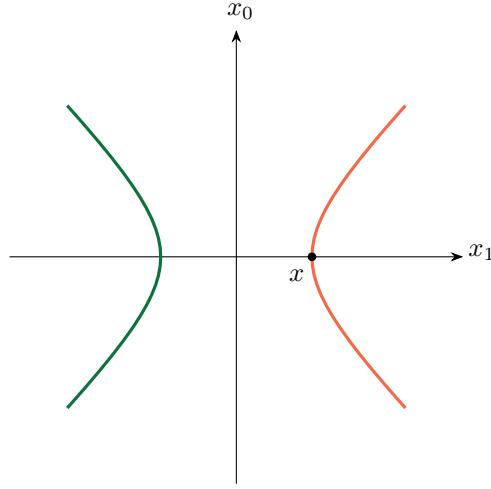


Figura 3: Algumas das transformações de Lorentz do evento $x = (0, 1)^\top$. Em laranja (direita), as transformações de \mathcal{L}_+^\uparrow . Em verde (esquerda), as de \mathcal{L}_-^\uparrow .

Exercício 16:

Defina a rapidez z por $z = \text{artanh } \beta$, onde $\beta = \frac{v}{c}$. Mostre que a transformação de Lorentz da Eq. (1.55) na página 17 pode ser escrita como

$$\Lambda = \begin{pmatrix} \cosh z & -\sinh z \\ -\sinh z & \cosh z \end{pmatrix}. \quad (2.3)$$

Desta forma, podemos identificar cada elemento de \mathcal{L}_+^\uparrow com um número real $z \in \mathbb{R}$. \blackcross

Vamos tomar agora um evento particular no espaço-tempo. Por simplicidade, escolhamos $x_0 = 0$ e $x_1 = 1$. Veja que os elementos de \mathcal{L}_+^\uparrow aplicados neste evento x definem agora uma curva no espaço tempo por meio da função

$$l_+^\uparrow(z) = \begin{pmatrix} \cosh z & -\sinh z \\ -\sinh z & \cosh z \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sinh z \\ \cosh z \end{pmatrix}. \quad (2.4)$$

Exercício 17:

De modo análogo, mostre que os elementos de \mathcal{L}_-^\uparrow definem uma curva no espaço-tempo por meio da função

$$l_-^\uparrow(z) = \begin{pmatrix} -\sinh z \\ \cosh z \end{pmatrix}. \quad (2.5)$$

\blackcross

Para visualizar o porquê disto ser interessante, podemos graficar estas curvas, como feito na Figura 3.

Vemos explicitamente que \mathcal{L}_+^\uparrow contém a identidade ($x \rightarrow x$), mas \mathcal{L}_-^\uparrow não. Logo, \mathcal{L}_-^\uparrow não é um grupo. Porém, além disso, cada ponto na curva laranja representa um elemento

de \mathcal{L}_+^\uparrow . Podemos identificar a curva $l_+^\uparrow(z)$ com o grupo \mathcal{L}_+^\uparrow , onde associamos aos diferentes pontos da curva o produto

$$l_+^\uparrow(z_1)l_+^\uparrow(z_2) = l_+^\uparrow(z_1 + z_2). \quad (2.6)$$

Exercício 18:

Mostre que \mathcal{L}_+^\uparrow é isomorfo ao grupo $\{l_+^\uparrow(z); z \in \mathbb{R}\}$ dotado do produto dado na Eq. (2.6). \clubsuit

Assim, vemos que \mathcal{L}_+^\uparrow não apenas é um grupo, mas também é uma curva no \mathbb{R}^2 . Em mais dimensões, poderíamos ter grupos que correspondem a outras curvas, a superfícies, a hiperfícies, e estruturas multi-dimensionais, conhecidas como variedades*, em geral.

Estes tipos de grupos, conhecidos como grupos de Lie, possuem uma estrutura contínua, no sentido de que podemos “acessar” diferentes elementos do grupo ao variar continuamente um parâmetro. Em particular, ao menos uma parte do grupo é acessível como variações contínuas a partir da identidade.

Em geral, nem todos os elementos do grupo são variações contínuas da identidade. \mathcal{L}^\uparrow , por exemplo, não é, como ilustrado na Figura 3 na página anterior. Para acessar o trecho em verde do grupo a partir da identidade é preciso que façamos uma transformação discreta: a reversão de paridade. No entanto, ainda é fascinante estudarmos o que ocorre nos casos em que o grupo é formado por deformações contínuas da identidade (em linguagem técnica, o que ocorre quando o grupo é conexo).

Exercício 19:

$\text{SO}(N)$, \mathcal{G}_+^\uparrow e \mathcal{P}_+^\uparrow são grupos de Lie? A qual curva no plano corresponde $\text{SO}(2)$? \clubsuit

2.2 Álgebras de Lie

Esta seção é parcialmente inspirada nos tratamentos de [10, 14, 36]. Vamos assumir que estamos lidando com um grupo de Lie G cujos elementos são descritos em termos de N parâmetros contínuos reais que denotaremos por α_a , $1 \leq a \leq N$, e coletivamente por α . A identidade do grupo é de particular interesse, visto que é o único elemento e com $e \cdot g = g \cdot e = g$ para todo $g \in G$. Por isso, tomaremos a parametrização de modo que enquanto $g(\alpha)$ é um elemento genérico do grupo, $g(0) = e$. Por fim, assumiremos também, por simplicidade, que o grupo em questão é formado por matrizes.

Assim, elementos suficientemente próximos da identidade devem poder ser escritos na forma

$$g(\epsilon) = \mathbb{1} + i\epsilon_a T_a + \mathcal{O}(\epsilon^2), \quad (2.7)$$

onde uma soma sobre índices a repetidos é subentendida, escrevemos ϵ ao invés de α para reforçar que estamos lidando com parâmetros pequenos e T^a são matrizes associadas a cada um dos parâmetros ϵ_a . A unidade imaginária na Eq. (2.7) é uma convenção comum em Física que será justificada posteriormente.

Justificar

*Para mais informações sobre o que são variedades de um ponto de vista físico, veja por exemplo [1, 4, 7, 19, 31].

Tomemos um exemplo a partir do grupo de Lorentz próprio ortócrono. Sabemos que os elementos de \mathcal{L}_+^\uparrow são dados em termos da rapidez por

$$\Lambda(z) = \begin{pmatrix} \cosh z & -\sinh z \\ -\sinh z & \cosh z \end{pmatrix}. \quad (2.8)$$

Para $z = \epsilon \ll 1$, podemos expandir cada uma dessas funções em série de Taylor.

Exercício 20:

Mostre que

$$\cosh z = \sum_{k=0}^{+\infty} \frac{z^{2k}}{(2k)!}, \quad \sinh z = \sum_{k=0}^{+\infty} \frac{z^{2k+1}}{(2k+1)!}. \quad (2.9)$$

✂

Assim, temos que

$$\Lambda(\epsilon) = \begin{pmatrix} 1 + \mathcal{O}(\epsilon^2) & -\epsilon + \mathcal{O}(\epsilon^3) \\ 1 + \mathcal{O}(\epsilon^2) & -\epsilon + \mathcal{O}(\epsilon^3) \end{pmatrix}, \quad (2.10a)$$

$$= \mathbb{1} + \epsilon \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad (2.10b)$$

$$= \mathbb{1} + i\epsilon \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad (2.10c)$$

$$(2.10d)$$

A matriz $T = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$ é dita ser o gerador do grupo uniparamétrico \mathcal{L}_+^\uparrow . O nome uniparamétrico se refere ao fato de que \mathcal{L}_+^\uparrow é descrito em termos de um único parâmetro. Note que T é dado em termos de $\Lambda(\epsilon)$ por

$$T \equiv -i \left. \frac{d}{d\epsilon} \Lambda(\epsilon) \right|_{\epsilon=0}. \quad (2.11)$$

Isto descreve apenas elementos infinitesimalmente próximos da identidade. De que é útil para descrever elementos gerais do grupo?

Tomemos então algum $z \in \mathbb{R}$ finito. Notemos que, para qualquer n natural,

$$\left[\Lambda\left(\frac{z}{n}\right) \right]^n = \Lambda\left(\frac{z}{n} + \cdots + \frac{z}{n}\right) = \Lambda(z). \quad (2.12)$$

Assim, formalmente temos que

$$\Lambda(z) = \lim_{n \rightarrow +\infty} \left[\Lambda\left(\frac{z}{n}\right) \right]^n, \quad (2.13a)$$

$$= \lim_{n \rightarrow +\infty} \left[\mathbb{1} + i \frac{z}{n} T \right]^n, \quad (2.13b)$$

$$= e^{izT}, \quad (2.13c)$$

onde usamos o resultado $e^x = \lim_{n \rightarrow +\infty} \left(1 + \frac{x}{n}\right)^n$.

Naturalmente, este cálculo a princípio pode soar estranho: o que seria a exponencial de uma matriz?

Por um lado, podemos a ver como sendo definida pelo procedimento que fizemos: dada qualquer matriz M , definimos

$$e^M = \lim_{n \rightarrow +\infty} \left(1 + \frac{M}{n}\right)^n. \quad (2.14)$$

Alternativamente, definimos com base na série de Taylor da exponencial,

$$e^M = \sum_{n=0}^{+\infty} \frac{M^n}{n!}, \quad (2.15)$$

onde convencionamos que $M^0 = \mathbb{1}$.

Esta ideia faz sentido? Vamos verificar que $\Lambda(z) = e^{izT}$ de fato recupera a expressão para $\Lambda(z)$ que possuíamos anteriormente*.

Queremos calcular

$$e^{izT} = \exp \begin{pmatrix} 0 & -z \\ -z & 0 \end{pmatrix}. \quad (2.16)$$

Para isso, vamos expandir a exponencial em série de Taylor. Será conveniente utilizar-mos um resultado auxiliar.

Exercício 21:

Dada $T = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$, mostre que $(iT)^{2n} = \mathbb{1}$ e $(iT)^{2n+1} = iT$ para todo n inteiro não-negativo. \blacklozenge

Temos então que

$$e^{izT} = \sum_{n=0}^{+\infty} \frac{i^n z^n T^n}{n!}, \quad (2.17a)$$

$$= \sum_{n=0}^{+\infty} \frac{z^{2n} (iT)^{2n}}{(2n!)} + \sum_{n=0}^{+\infty} \frac{z^{2n+1} (iT)^{2n+1}}{(2n+1)!}, \quad (2.17b)$$

$$= \left(\sum_{n=0}^{+\infty} \frac{z^{2n}}{(2n!)} \right) \mathbb{1} + \left(\sum_{n=0}^{+\infty} \frac{z^{2n+1}}{(2n+1)!} \right) (iT), \quad (2.17c)$$

$$= \cosh(z) \mathbb{1} + \sinh(z) (iT), \quad (2.17d)$$

$$= \begin{pmatrix} \cosh z & -\sinh z \\ -\sinh z & \cosh z \end{pmatrix}, \quad (2.17e)$$

$$= \Lambda(z), \quad (2.17f)$$

como esperado. Note que usamos também o resultado do Exercício 20 na página anterior.

*Caso deseje uma demonstração de que a série que usamos para definir a exponencial de matriz de fato converge, veja [14, Proposição 2.1]

Exercício 22:

Considere a expressão

$$e^{i\theta M} = e^{\theta a} \begin{pmatrix} \cos(\theta b) & \sin(\theta b) \\ -\sin(\theta b) & \cos(\theta b) \end{pmatrix}, \quad (2.18)$$

onde θ , a e b são parâmetros reais. Obtenha M .

Refleta se dada e^X , é sempre possível obter X e se tal X é única. Caso o assunto lhe interesse, veja [14, Sec. 2.3] para uma discussão sobre o logaritmo de matriz. \blackbox

Exercício 23:

Considere o grupo $\text{SO}(2)$. Qual é o seu gerador? Como podemos escrever um elemento genérico de $\text{SO}(2)$ utilizando uma exponencial de matriz? \blackbox

Consideremos agora um caso mais complexo, em que o grupo possui mais que um parâmetro: o grupo de Lorentz próprio ortócrono em $(3+1)$ -dimensões.

Queremos escrever $\Lambda(\epsilon) = \mathbb{1} + i\epsilon_a T^a$ para parâmetros ϵ pequenos. Sabemos que os elementos do grupo devem satisfazer a condição $\Lambda^\top \eta \Lambda = \eta$. Em termos dos geradores, esta condição se torna

$$(\mathbb{1} + i\epsilon_a T^{a\top}) \eta (\mathbb{1} + i\epsilon_a T^a) = \eta, \quad (2.19a)$$

$$\eta + i\epsilon_a T^{a\top} \eta + i\epsilon_a \eta T^a + \mathcal{O}(\epsilon^2) = \eta, \quad (2.19b)$$

$$\epsilon_a \eta T^a = -\epsilon_a T^{a\top} \eta + \mathcal{O}(\epsilon^2). \quad (2.19c)$$

Como os parâmetros ϵ são arbitrários, temos que, para cada a ,

$$\eta T^a = -T^{a\top} \eta, \quad (2.20)$$

onde já descartamos os termos desprezíveis de ordem ϵ^2 . Como $\eta^2 = \mathbb{1}$, temos que

$$T^a = -\eta T^{a\top} \eta \quad (2.21)$$

para cada a . Estes são os possíveis geradores do grupo de Lorentz. Abrindo esta expressão em componentes obtemos

$$\begin{pmatrix} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & T_{11} & T_{12} & T_{13} \\ T_{20} & T_{21} & T_{22} & T_{23} \\ T_{30} & T_{31} & T_{32} & T_{33} \end{pmatrix} = \begin{pmatrix} -T_{00} & T_{10} & T_{20} & T_{30} \\ T_{01} & -T_{11} & -T_{21} & -T_{31} \\ T_{02} & -T_{12} & -T_{22} & -T_{32} \\ T_{03} & -T_{13} & -T_{23} & -T_{33} \end{pmatrix}. \quad (2.22)$$

Assim, definimos as matrizes

$$-iK_1 = \begin{pmatrix} 0 & +1 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad -iK_2 = \begin{pmatrix} 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad -iK_3 = \begin{pmatrix} 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 \end{pmatrix}, \quad (2.23a)$$

$$-iJ_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & +1 & 0 \end{pmatrix}, \quad -iJ_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & +1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad -iJ_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (2.23b)$$

onde os fatores de $-i$ provém da convenção que fizemos de que escrevemos os elementos infinitesimais de um grupo com um fator de i , tal qual na Eq. (2.7) na página 21.

Cross reference

Desta maneira podemos escrever um elemento qualquer do grupo de Lorentz próprio ortócrono em $(3+1)$ -dimensões na forma

$$\Lambda = \exp(i\mathbf{z} \cdot \mathbf{K} + i\boldsymbol{\theta} \cdot \mathbf{J}), \quad (2.24)$$

onde $\mathbf{z} \cdot \mathbf{K} \equiv z_1 K_1 + z_2 K_2 + z_3 K_3$ e $\boldsymbol{\theta} \cdot \mathbf{J} \equiv \theta_1 J_1 + \theta_2 J_2 + \theta_3 J_3$.

Porém, ainda não sabemos o significado desta decomposição.

Exercício 24:

Mostre que

$$e^{izK_1} = \begin{pmatrix} \cosh z & -\sinh z & 0 & 0 \\ -\sinh z & \cosh z & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.25)$$

e

$$e^{i\theta J_1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \cos \theta & \sin \theta \\ 0 & 0 & -\sin \theta & \cos \theta \end{pmatrix}. \quad (2.26)$$

Conclua que os geradores K_i estão associados a boosts em direções arbitrárias e os geradores J_i estão associados a rotações em torno de eixos espaciais arbitrários. ✖

Isto nos permite concluir algo de interessante sobre as transformações de referencial relativísticas? Na verdade, sim!

Dadas duas matrizes A e B , é conveniente definir seu comutador $[A, B] = AB - BA$. Perceba que o comutador “mede o quanto as matrizes falham em comutar”.

Exercício 25:

Mostre, manualmente ou por meios computacionais, que

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \quad [J_i, K_j] = i\epsilon_{ijk} K_k, \quad [K_i, K_j] = -i\epsilon_{ijk} J_k, \quad (2.27)$$

onde ϵ_{ijk} denota o símbolo completamente antissimétrico de Levi-Civita com $\epsilon_{123} = +1$. Ou seja, permutações ijk cíclicas de 123 fornecem $\epsilon_{ijk} = +1$ ($\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1$), permutações não-cíclicas ijk fornecem $\epsilon_{ijk} = -1$ ($\epsilon_{132} = \epsilon_{321} = \epsilon_{213} = -1$) e repetição de quaisquer dois índices fornece 0. ✖

Isso reflete o fato de que boosts de Lorentz em direções diferentes não comutam, tal qual rotações em torno de eixos diferentes não comutam, tal qual boosts de Lorentz e rotações não comutam.

Isso pode parecer um pouco estranho. Não deveria valer, por exemplo, que $e^{i\theta J_1} e^{izK_1} = e^{i\theta J_1 + izK_1} = e^{izK_1 + i\theta J_1} = e^{izK_1} e^{i\theta J_1}$? Na verdade, esta propriedade das exponenciais não vale para matrizes. Perceba que

$$e^{i\theta J_1} e^{izK_1} = \left(\mathbb{1} + i\theta J_1 - \frac{\theta^2}{2} J_1^2 + \mathcal{O}(\theta^3) \right) \left(\mathbb{1} + izK_1 - \frac{z^2}{2} K_1^2 + \mathcal{O}(z^3) \right), \quad (2.28a)$$

$$= \mathbb{1} + i(\theta J_1 + zK_1) - \frac{1}{2}(\theta^2 J_1^2 + 2\theta z J_1 K_1 + z^2 K_1^2) + \dots, \quad (2.28b)$$

$$= \mathbb{1} + i(\theta J_1 + zK_1) - \frac{1}{2}(\theta^2 J_1^2 + \theta z J_1 K_1 + \theta z K_1 J_1 + z^2 K_1^2) - \frac{\theta z}{2} [J_1, K_1] + \dots, \quad (2.28c)$$

$$= \mathbb{1} + i(\theta J_1 + zK_1) - \frac{1}{2}(\theta J_1 + zK_1)^2 - \frac{\theta z}{2} [J_1, K_1] + \dots, \quad (2.28d)$$

onde os pontos denotam termos com $\mathcal{O}(\theta^3, \theta^2 z, \theta z^2, z^3)$.

Vemos então que um comutador $[J_1, K_1]$ ocorre em segunda ordem. Ele não poderá ser cancelado por outros termos, visto que são de ordens superiores, e assim não podemos escrever $e^{i\theta J_1} e^{izK_1} = e^{i\theta J_1 + izK_1}$. Desta forma, concluímos que o Exercício 25 na página anterior de fato implica a não-comutatividade de diferentes “setores” do grupo de Lorentz.

Estes exemplos nos ensinam que podemos aprender muito sobre grupos de Lie ao estudar os seus geradores. Isto motiva a definição do que é uma álgebra de Lie.

Definição 6 [Álgebra de Lie de um Grupo de Lie Matricial]:

Seja G um grupo de Lie matricial. Definimos sua álgebra de Lie, \mathfrak{g} , como o conjunto das matrizes X tais que $e^{itX} \in G$ para todo $t \in \mathbb{R}$. ♠

Exercício 26:

Seja G um grupo de Lie matricial e \mathfrak{g} sua álgebra de Lie. Mostre que, dados $g \in G$ e $T \in \mathfrak{g}$, vale que $gTg^{-1} \in \mathfrak{g}$. ✚

Estudemos então algumas propriedades das álgebras de Lie.

Proposição 7:

Seja G um grupo de Lie matricial e \mathfrak{g} sua álgebra de Lie. Mostre que, dados $X, Y \in \mathfrak{g}$ e $s \in \mathbb{R}$,

i. $X + Y \in \mathfrak{g}$;

ii. $sX \in \mathfrak{g}$;

iii. $-i[X, Y] \in \mathfrak{g}$. □

Demonstração:

Nossa demonstração será feita no estilo dos físicos. Para uma demonstração ~~correta~~

rigorosa, veja [14, Teorema 2.18]. [10, Sec. 2.2] toma uma abordagem também pouco rigorosa, mas um tanto distinta.

Como $X, Y \in \mathfrak{g}$, $e^{itX}, e^{itY} \in G$ para todo $t \in \mathbb{R}$. Sabemos que $e^{itX}e^{itY}$ precisa ser elemento de G pela própria definição de grupo. Tomando $t = \epsilon$ infinitesimal, vemos que

$$e^{i\epsilon X}e^{i\epsilon Y}, \quad (2.29a)$$

$$= [\mathbb{1} + i\epsilon X + \mathcal{O}(\epsilon^2)][\mathbb{1} + i\epsilon Y + \mathcal{O}(\epsilon^2)], \quad (2.29b)$$

$$= \mathbb{1} + i\epsilon X + i\epsilon Y + \mathcal{O}(\epsilon^2), \quad (2.29c)$$

$$= \mathbb{1} + i\epsilon(X + Y) + \mathcal{O}(\epsilon^2). \quad (2.29d)$$

Assim, vemos que $\mathbb{1} + i\epsilon(X + Y)$ é um elemento infinitesimal de G . Para obter um elemento finito, podemos apenas compor este elemento consigo mesmo infinitas vezes, no espírito da Eq. (2.14) na página 23, de forma que obtemos no limite $e^{it(X+Y)} \in G$ para todo $t \in \mathbb{R}$. Assim, $X + Y \in \mathfrak{g}$.

A seguir, queremos provar que $sX \in \mathfrak{g}$. Como e^{itX} para todo $t \in \mathbb{R}$ e $s \in \mathbb{R}$, há de valer que e^{itsX} para todo $t \in \mathbb{R}$, visto que $ts \in \mathbb{R}$. Logo, $sX \in \mathfrak{g}$.

Por fim, devemos lidar com o comutador. Sabemos do Exercício 26 na página anterior que $e^{itX}Ye^{-itX} \in \mathfrak{g}$. Tomando $t = \epsilon$ infinitesimal, temos que

$$e^{i\epsilon X}Ye^{-i\epsilon X} = [\mathbb{1} + i\epsilon X + \mathcal{O}(\epsilon^2)]Y[\mathbb{1} - i\epsilon X + \mathcal{O}(\epsilon^2)], \quad (2.30a)$$

$$= [Y + i\epsilon XY + \mathcal{O}(\epsilon^2)][\mathbb{1} - i\epsilon X + \mathcal{O}(\epsilon^2)], \quad (2.30b)$$

$$= Y + i\epsilon XY - i\epsilon YX + \mathcal{O}(\epsilon^2), \quad (2.30c)$$

$$= Y + i\epsilon[X, Y] + \mathcal{O}(\epsilon), \quad (2.30d)$$

e assim $Y + i\epsilon[X, Y]$ é elemento da álgebra de Lie. Subtraindo Y , a primeira propriedade que provamos implica que $i\epsilon[X, Y]$ é elemento da álgebra de Lie. Multiplicando por $-\frac{1}{\epsilon}$ concluímos da segunda propriedade que provamos que $-i[X, Y] \in \mathfrak{g}$. ■

Destas propriedades e dos nossos conhecimentos gerais sobre matrizes, podemos concluir que a álgebra de Lie constitui um espaço vetorial real.

Definição 8 [Espaço Vetorial]:

Seja V um conjunto, \mathbb{K} um corpo* e $+: V \times V \rightarrow V$ e $\cdot: \mathbb{K} \times V \rightarrow V$ funções. Dizemos que $(V, \mathbb{K}, +, \cdot)$ é um espaço vetorial sobre \mathbb{K} (ou \mathbb{K} -espaço vetorial) se, e somente se, valerem as propriedades

- i. $(V, +)$ é um grupo Abelian;
- ii. $(\alpha\beta) \cdot u = \alpha \cdot (\beta \cdot u)$ para quaisquer $\alpha, \beta \in \mathbb{K}$ e $u \in V$;
- iii. $1 \cdot u = u$, para qualquer $u \in V$ (1 denota a unidade de \mathbb{K});
- iv. $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$, para quaisquer $\alpha \in \mathbb{K}$ e $u, v \in V$;

*A definição formal de corpo (em inglês, *field*) pode ser encontrada, por exemplo, em [8, p. 20], mas para nossos propósitos nos interessamos apenas pelos casos $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

v. $(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$, para quaisquer $\alpha, \beta \in \mathbb{K}$, $u \in V$. ♠

Exercício 27:

Mostre que \mathbb{K}^n é um espaço vetorial sobre \mathbb{K} para $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$. Mostre que \mathbb{C}^n é um espaço vetorial sobre \mathbb{R} . ✚

Exercício 28:

Mostre que $\mathbb{M}_N(\mathbb{K})$ é um espaço vetorial sobre \mathbb{K} para $\mathbb{K} = \mathbb{R}$ e $\mathbb{K} = \mathbb{C}$. Mostre também que $\mathbb{M}_N(\mathbb{C})$ é um espaço vetorial real. ✚

Exercício 29:

Anteriormente definimos as noções de subgrupo, homomorfismo e isomorfismos de grupos ao pensar em estruturas que preservavam a estrutura de grupos. Como podemos definir subespaços vetoriais e homomorfismos e isomorfismos de espaços vetoriais?

Prove que se $(V, \mathbb{K}, +, \cdot)$ é um espaço vetorial sobre \mathbb{K} , então $U \subseteq V$ constitui um espaço vetorial sobre \mathbb{K} se, e somente se, as condições

- i. $0 \in U$;
- ii. $u + v \in U$ para quaisquer elementos $u, v \in U$;
- iii. $\alpha \cdot v \in U$ para quaisquer elementos $\alpha \in \mathbb{K}$, $u \in U$

valerem. Utilize este resultado e o Exercício 28 para provar que toda álgebra de Lie de um grupo de Lie matricial é um espaço vetorial *real*. ✚

Não é comum nos referirmos a “homomorfismos entre espaços vetoriais” por este nome. Geralmente os chamamos de transformações lineares.

Dizemos que um espaço vetorial V possui dimensão n se bastarem n números de \mathbb{K} para caracterizar qualquer elemento de V completamente. Por exemplo, tomemos \mathbb{R}^3 . Todo ponto de \mathbb{R}^3 é da forma (a, b, c) com $a, b, c \in \mathbb{R}$, e assim podemos caracterizar qualquer elemento de \mathbb{R}^3 completamente com 3 números reais. Assim, \mathbb{R}^3 enquanto espaço vetorial real tem dimensão 3.

Consideremos a seguir o espaço $\mathbb{M}_2(\mathbb{C})$, das matrizes complexas 2×2 . Estas matrizes possuem quatro entradas complexas, então podemos descrevê-las completamente usando 4 números complexos (e portanto é um espaço vetorial complexo de dimensão 4) ou 8 números reais (e portanto é um espaço vetorial real de dimensão 8).

De modo geral, vemos que $\mathbb{M}_N(\mathbb{K})$ tem dimensão N^2 sobre \mathbb{K} : poderíamos rearranjar as entradas de cada matriz em um vetor longo com N^2 componentes. Logo, todos os espaços vetoriais que nos interessam aqui são da forma \mathbb{K}^n .

Exercício 30:

Convença-se de que $\mathbb{M}_N(\mathbb{C})$ é um espaço vetorial real de dimensão $2N^2$. ✚

No caso de álgebras de Lie, sempre podemos escolher um conjunto mínimo de geradores a partir do qual escreveremos os demais. Como as álgebras de Lie do nosso interesse sempre são subespaços de $\mathbb{M}_N(\mathbb{K})$, sua dimensão enquanto espaço vetorial real será não

mais do que $2N^2$. Em geral, poderemos escolher um conjunto de geradores T_a que nos permitirá escrever todos os outros elementos $X \in \mathfrak{g}$ na forma

$$X = \alpha_a T_a, \quad (2.31)$$

onde a soma sobre os índices repetidos está subentendida e $\alpha_a \in \mathbb{R}$. Dizemos que X é uma combinação linear dos T_a . Queremos que esse conjunto seja mínimo no sentido de que não é possível escrever nenhum dos T_a em termos dos outros. Por exemplo, os vetores $(1, 0, 0)^\top$, $(0, 1, 0)^\top$ e $(0, 0, 1)^\top$ nos permitem escrever qualquer elemento de \mathbb{R}^3 na forma

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (2.32)$$

Porém, não podemos escrever $(1, 0, 0)^\top$ como combinação linear de $(0, 1, 0)^\top$ e $(0, 0, 1)^\top$. Se um conjunto de vetores é tal que nenhum de seus elementos pode ser escrito como combinação linear dos demais, dizemos que ele é linearmente independente. Do contrário, o dizemos linearmente dependente.

É possível mostrar, embora não o faremos aqui, que a menos que queiramos trabalhar em dimensão^{*} infinita, os \mathbb{K} -espaços vetoriais \mathbb{K}^n são os únicos \mathbb{K} -espaços vetoriais que existem.

Teorema 9:

Seja $n \in \mathbb{N}$ e \mathbb{K} um corpo. Todo \mathbb{K} -espaço vetorial V de dimensão n é isomorfo a \mathbb{K}^n . \square

Assim, dada uma álgebra de Lie, especificar sua dimensão a caracteriza quase por completo. Se uma álgebra de Lie tem n geradores independentes, ela tem dimensão n (cada um dos geradores faz o papel que um vetor da forma $(0, \dots, 0, 1, 0, \dots, 0)$ faria no \mathbb{R}^n) e isso a torna única entre outros espaços vetoriais a menos de isomorfismo. Porém, ainda é preciso considerarmos mais um fator para determiná-la unicamente: o comutador.

Álgebras de Lie não são meros espaços vetoriais, elas também possuem um produto. Se \mathfrak{g} é uma álgebra de Lie e $X, Y \in \mathfrak{g}$, então $-i[X, Y] \in \mathfrak{g}$. Este tipo de estrutura não acompanha a definição de espaço vetorial, e portanto poderíamos ter duas álgebras de Lie com a mesma dimensão e comutadores diferentes.

Exercício 31:

Demonstre, manualmente ou por meios computacionais, as seguintes propriedades do comutador de matrizes. Sendo $X, Y, Z \in \mathbb{M}_N(\mathbb{K})$, $\alpha \in \mathbb{K}$,

- i. $[X + \alpha \cdot Y, Z] = [X, Z] + \alpha \cdot [Y, Z]$ (linearidade na primeira entrada);
- ii. $[X, Y + \alpha \cdot Z] = [X, Y] + \alpha \cdot [X, Z]$ (linearidade na segunda entrada);

^{*}Não definimos apropriadamente o que significa “dimensão”, confiando que o leitor possui uma noção intuitiva a partir de sua experiência com vetores bi e tridimensionais. Em caso de incômodo, veja, por exemplo, [8, Sec. 2.2–2.3].

- iii. $[X, X] = 0$ (alternatividade);
- iv. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (identidade de Jacobi);
- v. $[X, Y] = -[Y, X]$ (anticomutatividade).

✂

Para especificar o comutador, vamos ver qual a forma mais geral que ele pode tomar.

Suponha que a álgebra de Lie tenha dimensão n , ou seja, que podemos especificar qualquer elemento em termos de n geradores independentes. Dados $X, Y \in \mathfrak{g}$, podemos escrevê-los como

$$X = x_a T_a \quad \text{e} \quad Y = y_a T_a, \quad (2.33)$$

onde T_a são os geradores independentes. Usando as propriedades de linearidade do Exercício 31 na página anterior, vemos que

$$[X, Y] = -i[x_a T_a, y_b T_b], \quad (2.34a)$$

$$= x_a y_b [T_a, T_b]. \quad (2.34b)$$

Assim, se especificarmos $[T_a, T_b]$, os comutadores dos geradores, podemos obter os comutadores de quaisquer outros elementos da álgebra de Lie, bastando escrevê-los em termos dos geradores.

Sabemos ainda que $-i[T_a, T_b]$ é elemento da álgebra de Lie. Logo, também pode ser escrito em termos dos geradores. Sua expressão mais genérica será

$$-i[T_a, T_b] = f_{abc} T_c, \quad (2.35)$$

onde as constantes f_{abc} , chamadas de constantes de estrutura, podem depender dos geradores T_a, T_b entrando no comutador e possuem o índice c para formar uma combinação linear com os T_c do lado direito. Assim, para caracterizar uma álgebra de Lie completamente, devemos determinar sua dimensão e as constantes de estrutura que fornecem o comutador por meio de

$$[T_a, T_b] = i f_{abc} T_c. \quad (2.36)$$

É comum que textos de Física se refiram à Eq. (2.36) como a álgebra do grupo sendo considerado.

Sendo este assunto consideravelmente abstrato, vamos a alguns exemplos.

2.3 Grupo de Heisenberg

Consideremos a função $h: \mathbb{R}^3 \rightarrow \mathbb{M}_3(\mathbb{R})$ dada por

$$h(a, b, c) = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}. \quad (2.37)$$

Vamos definir o grupo de Heisenberg em uma dimensão, H , por

$$H = \{h(a, b, c); a, b, c \in \mathbb{R}^3\}. \quad (2.38)$$

Exercício 32:

Mostre que H é, de fato, um grupo sob o produto usual de matrizes. Em particular, mostre que o produto é dado por

$$h(a, b, c)h(a', b', c') = h(a + a', b + b', c + c' + ab'), \quad (2.39)$$

e que a identidade é dada por $h(0, 0, 0)$. Qual o inverso de $h(a, b, c)$? \boxtimes

O grupo de Heisenberg é definido em termos de três parâmetros, e portanto sua álgebra de Lie possuirá três geradores independentes. Os obteremos calculando

$$p = -i \frac{d}{da} h(a, 0, 0), \quad q = -i \frac{d}{db} h(0, b, 0), \quad \hbar = -i \frac{d}{dc} h(0, 0, c), \quad (2.40)$$

onde a notação q, p, \hbar será justificada posteriormente devido à conexão deste grupo com a Mecânica Quântica.

Cross reference

Da Eq. (2.37) na página anterior vemos que

$$p = \begin{pmatrix} 0 & -i & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix}, \quad \hbar = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (2.41)$$

Assim, temos

$$[q, p] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -i & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -i & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.42a)$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.42b)$$

$$= i \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (2.42c)$$

$$= i\hbar. \quad (2.42d)$$

Exercício 33:

Mostre que para os demais casos, teremos

$$[q, \hbar] = 0, \quad [p, \hbar] = 0. \quad (2.43)$$

\boxtimes

Assim, temos que as constantes de estrutura não-nulas são

$$f_{qp\hbar} = -f_{pq\hbar} = 1. \quad (2.44)$$

Todas as outras 25 constantes se anulam. Com as constantes de estrutura (que já contém o fato de que a álgebra possui três dimensões), caracterizamos completamente a álgebra de Lie do grupo de Heisenberg, \mathfrak{h} .

2.4 Grupo Especial Unitário em Duas Dimensões

No começo destas notas nós definimos os grupos especiais ortogonais $SO(N)$, que caracterizam rotações em várias dimensões. Podemos também considerar agora o seu análogo para os números complexos.

Definimos o grupo unitário em dimensão N , $U(N)$, por

$$U(N) = \left\{ U \in GL_{\mathbb{C}}(N); U^\dagger U = \mathbb{1} \right\} \quad (2.45)$$

e o grupo especial unitário em dimensão N , $SU(N)$, por

$$SU(N) = \left\{ U \in SL_N(\mathbb{C}); U^\dagger U = \mathbb{1} \right\}. \quad (2.46)$$

A notação U^\dagger representa uma versão “complexificada” da transposição. Ela representa transposição, como no caso real, seguida da conjugação complexa das entradas da matriz. Por exemplo,

$$\begin{pmatrix} a + ib & c + id \\ e + if & g + ih \end{pmatrix}^\dagger = \begin{pmatrix} a - ib & e - if \\ c - id & g - ih \end{pmatrix}. \quad (2.47)$$

Da mesma forma que matrizes com $R^\top R = \mathbb{1}$ são ditas ortogonais, matrizes com $U^\dagger U = \mathbb{1}$ são ditas unitárias. Enquanto matrizes com $S^\top = S$ são ditas simétricas, matrizes com $H^\dagger = H$ são ditas hermitianas.

Note que, nos números complexos, números com $u^* u = 1$ são da forma $u = e^{i\theta}$ e números com $h^* = h$ são reais. Assim, matrizes unitárias e hermitianas generalizam a ideia de número complexo com módulo um e de número real para as matrizes complexas.

Exercício 34:

Seja $U \in U(N)$. Mostre que $\det U = e^{i\theta}$ para algum ângulo θ . ✂

Por enquanto iremos nos debruçar sobre o grupo $SU(2)$, que encontra aplicações na Física do spin eletrônico, na Física nuclear (ver [33, Cap. 4] para estes dois exemplos) e na descrição das interações fracas da Física de Partículas (por exemplo, [26, 28, 37]). Veremos explicitamente exemplos de aplicações na Física nuclear mais adiante.

Queremos estudar as matrizes com a forma

$$U = e^{i\alpha_a T_a} \quad (2.48)$$

e tais que

$$U^\dagger U = \mathbb{1}, \quad \det U = 1. \quad (2.49)$$

Tomemos os parâmetros $\alpha_a = \epsilon_a$ como sendo infinitesimais para podermos impor estas condições sobre os elementos da álgebra de Lie. Temos

$$U^\dagger U = \mathbb{1}, \quad (2.50a)$$

$$[\mathbb{1} + i\epsilon_a T_a + \mathcal{O}(\epsilon^2)]^\dagger [\mathbb{1} + i\epsilon_a T_a + \mathcal{O}(\epsilon^2)] = \mathbb{1}, \quad (2.50b)$$

$$[\mathbb{1} - i\epsilon_a T_a^\dagger + \mathcal{O}(\epsilon^2)] [\mathbb{1} + i\epsilon_a T_a + \mathcal{O}(\epsilon^2)] = \mathbb{1}, \quad (2.50c)$$

$$\mathbb{1} - i\epsilon_a T_a^\dagger + i\epsilon_a T_a + \mathcal{O}(\epsilon^2) = \mathbb{1}, \quad (2.50d)$$

$$T_a = T_a^\dagger. \quad (2.50e)$$

Ou seja, para que U seja unitária, é preciso que seus geradores sejam hermitianos. É por isso que introduzimos a convenção com a unidade imaginária anteriormente: do contrário, os geradores seriam anti-hermitianos ($T_a^\dagger = -T_a$), e matrizes hermitianas são um tanto mais naturais de se lidar em Física, visto que experimentos em geral medem números reais, não imaginários.

Para prosseguir, enunciamos sem demonstração um resultado útil acerca das exponenciais de matriz.

Teorema 10:

Seja M uma matriz. A expressão

$$\det(e^M) = e^{\text{Tr}[M]} \quad (2.51)$$

é válida. □

Demonstração:

Veja [14, Teorema 2.11]. Daremos aqui um exemplo para o caso de uma matriz diagonal. Perceba que

$$\exp \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} e^a & 0 \\ 0 & e^b \end{pmatrix}. \quad (2.52)$$

Ademais,

$$\det \left(\exp \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \right) = e^a e^b = e^{a+b}. \quad (2.53)$$

Porém,

$$\exp \left(\text{Tr} \left[\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \right] \right) = e^{a+b}, \quad (2.54)$$

e assim vemos que, de fato, o teorema vale nesse caso simplificado. ■

Com isso em mente, vemos que

$$\det(e^{i\alpha_a T_a}) = 1, \quad (2.55a)$$

$$e^{i\alpha_a \text{Tr}[T_a]} = 1, \quad (2.55b)$$

$$\text{Tr}[T_a] = 0. \quad (2.55c)$$

Ou seja, para que a matriz U tenha determinante 1, seus geradores devem ter traço nulo.

Assim, buscamos geradores hermitianos de traço nulo. Como não há outras condições sobre as matrizes especiais unitárias além da dimensão, devemos pegar o máximo possível de geradores hermitianos no número escolhido de dimensões satisfazendo essa propriedade.

No caso $N = 2$, teremos três matrizes independentes. Vejamos o porquê. Explicitamente, a matriz 2×2 complexa mais geral possível é

$$\begin{pmatrix} a + ib & c + id \\ e + if & g + ih \end{pmatrix}, \quad (2.56)$$

onde todos os parâmetros são reais. Impor que ela seja hermitiana implica que

$$\begin{pmatrix} a - ib & e - if \\ c - id & g - ih \end{pmatrix} = \begin{pmatrix} a + ib & c + id \\ e + if & g + ih \end{pmatrix}, \quad (2.57)$$

o que impõe restrições sobre diversos parâmetros. Vemos então que a matriz 2×2 hermitiana mais geral possível é

$$\begin{pmatrix} a & c - if \\ c + if & g \end{pmatrix}. \quad (2.58)$$

A seguir impomos que a matriz deve ter traço nulo, ou seja,

$$a + g = 0. \quad (2.59)$$

Assim, a matriz 2×2 hermitiana de traço nulo mais geral possível é

$$\begin{pmatrix} a & c - if \\ c + if & -a \end{pmatrix}. \quad (2.60)$$

Exercício 35:

Convença-se de que em N dimensões há $N^2 - 1$ matrizes hermitianas com traço nulo linearmente independentes. \blackcross

É então comum definirmos as chamadas matrizes de Pauli como

$$\sigma_1 = \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ +i & 0 \end{pmatrix} \quad \text{e} \quad \sigma_3 = \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.61)$$

Exercício 36:

Mostre que valem as relações

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \quad (2.62)$$

e

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\mathbb{1}, \quad (2.63)$$

onde o anti-commutador $\{A, B\}$ é definido por $\{A, B\} = AB + BA$. \blackcross

Em geral, convencionase tomar os geradores de $SU(2)$ como sendo as matrizes $\tau_i = \frac{\sigma_i}{2}$, para que tenhamos a normalização $\text{Tr}[\tau_i\tau_j] = \frac{1}{2}\delta_{ij}$. Além disso, vemos que sob essa normalização a álgebra de $SU(2)$, denotada $\mathfrak{su}(2)$, é caracterizada por

$$[\tau_i, \tau_j] = i\epsilon_{ijk}\tau_k. \quad (2.64)$$

Exercício 37:

Mostre que a álgebra de $SO(3)$, $\mathfrak{so}(3)$, é dada por

$$[J_i, J_j] = i\epsilon_{ijk}J_k. \quad (2.65)$$

Sugestão: utilize os resultados dos Exercícios 24 e 25 na página 25. \blackcross

$\mathfrak{so}(3)$ e $\mathfrak{su}(2)$ possuem a mesma dimensão e as mesmas constantes de estrutura. Por consequência, elas são isomorfas no sentido de álgebras de Lie, e escrevemos $\mathfrak{so}(3) \cong \mathfrak{su}(2)$. Também podemos dizer que $SU(2)$ é localmente isomorfo a $SO(3)$, visto que a álgebra de Lie diz respeito ao comportamento do grupo nas vizinhanças da identidade. Em outras palavras, se olhássemos apenas os elementos infinitesimais de $SU(2)$ e $SO(3)$, não seríamos capazes de distinguir os dois grupos. Contudo, os dois grupos não são globalmente isomorfos.

Para compreender melhor este fato, consideremos um exemplo de [36]. Seja $\mathbf{x} \in \mathbb{R}^3$ e considere a matriz X dada por

$$X = \mathbf{x} \cdot \boldsymbol{\tau} = x \cdot \tau_1 + y \cdot \tau_2 + z \cdot \tau_3 = \frac{1}{2} \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}. \quad (2.66)$$

Perceba que, por ser uma combinação linear de matrizes de Pauli, X é hermitiana e tem traço nulo.

Tome $U \in SU(2)$ e vamos considerar a transformação $X \rightarrow X' = U^\dagger X U$.

Exercício 38:

Mostre que $\det X = -\frac{\|\mathbf{x}\|^2}{4}$. Mostre que X' é hermitiana e tem traço nulo, e por consequência pode ser escrita na forma $X' = \mathbf{x}' \cdot \boldsymbol{\tau}$. Mostre que os vetores \mathbf{x} e \mathbf{x}' tem a mesma norma e que são linearmente relacionados, isto é, existe uma matriz R tal que $\mathbf{x}' = R\mathbf{x}$ com $\|R\mathbf{x}\| = \|\mathbf{x}\|$. Conclua que $R \in SO(3)$. \blacktimes

Assim, a cada elemento U de $SU(2)$ está associada uma rotação $R \in SO(3)$. Isto é consequência do isomorfismo local. Contudo, esta associação não é bijetora: cada R está associado a duas matrizes $U \in SU(2)$, pois

$$X' = U^\dagger X U = (-U)^\dagger X (-U), \quad (2.67)$$

ou seja, tanto U quanto $-U$ são mapeados para R . Embora localmente os dois grupos sejam indistinguíveis, globalmente $SU(2)$ tem o dobro do tamanho de $SO(3)$.

Não iremos adentrar temas mais profundos em topologia, mas é possível mostrar que $SU(2)$ é simplesmente conexo, mas $SO(3)$ não. Intuitivamente falando, isso significa que $SO(3)$ possui um “buraco”, mas $SU(2)$ não.

Se um grupo G é conexo*, um grupo H simplesmente conexo com um homomorfismo $\Phi: H \rightarrow G$ cujo homomorfismo de álgebra de Lie associado, $\varphi: \mathfrak{h} \rightarrow \mathfrak{g}$, é um isomorfismo é dito ser o recobrimento universal de G .

Em outras palavras, o recobrimento universal de um grupo é um grupo maior, mas sem buracos. Para que a situação faça mais sentido, vamos a um exemplo com um grupo unidimensional, que facilita a visualização.

A Figura 4 na próxima página ilustra o grupo $SO(2)$ como uma curva. Perceba que a curva cerca um buraco. De maneira mais técnica, existem vários caminhos distintos que saem da identidade (denotada por um ponto preto) e a ela retornam de maneiras não

*Ou seja, é formado por um único “pedaço”, como \mathcal{L}_+^\uparrow . \mathcal{L}^\uparrow , como vimos na Figura 3 na página 20, é formado por dois “pedaços”, e portanto não é conexo

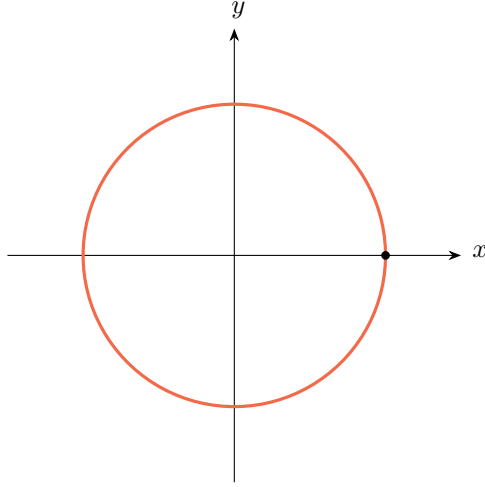


Figura 4: Visualização do grupo $SO(2)$ análoga à da Figura 3 na página 20 para \mathcal{L}^\dagger . Note que $SO(2)$ possui um “buraco”: existe mais de uma forma de fazer um caminho saindo do ponto preto e a ele retornando e elas não podem ser deformadas umas nas outras.

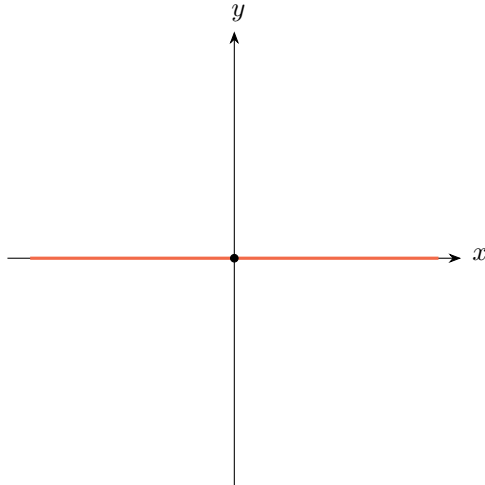


Figura 5: Visualização do grupo $(\mathbb{R}, +)$ análoga à da Figura 3 na página 20 para \mathcal{L}^\dagger . Note que $(\mathbb{R}, +)$ não possui um “buraco”.

equivalentes. Por exemplo, pode-se sair da identidade, mover-se 30° e retornar. Porém, também pode-se dar a volta completa. Ambos os caminhos representam curvas ao redor de $SO(2)$, mas não se pode distorcer um no outro continuamente.

Comparemos agora com o caso do grupo $(\mathbb{R}, +)$, formado pelos números reais com a operação de adição e representado na Figura 5. Perceba que qualquer caminho que saia da identidade e a ela retorne pode ser deformado a qualquer outro, sem a necessidade de cortes.

Exercício 39:

Mostre que a função $\Phi: \mathbb{R} \rightarrow \text{SO}(2)$ definida por

$$\Phi(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (2.68)$$

é um homomorfismo. ✚

$(\mathbb{R}, +)$ não é um grupo matricial, então não podemos usar livremente todas as ferramentas que desenvolvemos. Para perceber que ele de fato é localmente isomorfo a $\text{SO}(2)$, apenas note que ambos são Abelianos e uniparamétricos (e portanto tem dimensão 1). Assim, suas álgebras de Lie precisam ser isomorfas porque tem a mesma dimensão e todas as constantes de estrutura são nulas.

Desta forma, vemos que $(\mathbb{R}, +)$ é o recobrimento universal de $\text{SO}(2)$. Observando-os apenas nas proximidades da identidade, não podemos distingui-los. Porém, globalmente nós percebemos que são grupos distintos.

Embora isso possa parecer tolo agora, será um resultado útil em breve, quando o fato de certos grupos não serem simplesmente conexos nos atormentará.

3 Representações

3.1 Noções de Mecânica Quântica

Para ter uma compreensão mais adequada das discussões que faremos daqui em diante, é conveniente conhecer algumas ideias básicas da Mecânica Quântica. Esta seção serve apenas como uma introdução simplificada aos aspectos que iremos utilizar, e em particular irá se restringir ao caso de dimensão finita. Para uma apresentação mais completa, veja, por exemplo, [9, 22, 27, 29, 32, 33].

Uma forma de caracterizar de maneira genérica o objetivo de uma teoria física é: dado o estado de um sistema, obter os valores numéricos de seus observáveis.

Isto é uma forma ligeiramente técnica de descrever o que fazemos com frequência, por exemplo, em Mecânica Clássica. Se conhecemos a posição e o momento de todas as partículas do sistema, somos capazes de obter qualquer informação que possa ser medida sobre ele naquele instante de tempo. Por exemplo, podemos calcular a energia, o momento angular, as próprias posições e momentos, etc. De modo dinâmico, dado o estado do sistema em um instante de tempo queremos poder prever qual será seu estado em um instante futuro, para que possamos então calcular os observáveis em um novo instante.

Nossa formulação simplificada de Mecânica Quântica seguirá o mesmo preceito: queremos, a partir do estado do sistema, obter os valores dos observáveis, como energia, momento, etc. Não nos preocuparemos com evolução temporal e analisaremos apenas sistemas quânticos simples, que não envolvem partículas andando pelo espaço tridimensional.

Em Mecânica Quântica, faremos os seguintes postulados*:

*Estamos assumindo sistemas um tanto simplificados e apresentando postulados que não são suficientes

- i. os estados de um sistema físico são vetores $|\psi\rangle$ em um espaço vetorial complexo \mathcal{H} com produto escalar denotado $\langle\varphi|\psi\rangle$;
- ii. os observáveis de um sistema físico são matrizes hermitianas agindo em \mathcal{H} ;
- iii. um estado $|a\rangle$ que não seja representado pelo vetor nulo tem valor definido a para o observável \hat{A} se, e somente se, $\hat{A}|a\rangle = a|a\rangle$
- iv. o valor esperado de um observável \hat{A} no estado ψ é dado por $\langle\hat{A}\rangle_\psi = \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle}$.

Frequentemente nos referiremos ao espaço vetorial complexo com produto escalar como um espaço de Hilbert. Por definição, um espaço de Hilbert é um espaço vetorial complexo com produto escalar que seja completo, o que significa que ele não deve possuir “buracos”. Embora este requisito seja extremamente relevante para que a teoria quântica faça sentido, o deixaremos de lado sob a escusa de estarmos lidando em dimensão finita. É também comum referir-se ao que chamamos de “produto escalar”, em analogia com o caso dos vetores tridimensionais, como produto interno.

Por “valor esperado” queremos dizer que ao preparar N sistemas no mesmo estado $|\psi\rangle$ e medir o valor de \hat{A} em cada um deles, a média das medições se aproxima de $\langle\hat{A}\rangle_\psi$ conforme $N \rightarrow +\infty$.

A notação com $|\psi\rangle$ (lê-se “ket ψ ”), $\langle\varphi|$ (“bra φ ”) e afins, conhecida como notação de Dirac, é popular em Mecânica Quântica por facilitar a manipulação de expressões comuns na teoria. Notamos que, por estarmos em um espaço vetorial complexo, o produto escalar satisfaz

$$\langle\psi|\varphi\rangle = \overline{\langle\varphi|\psi\rangle}, \quad (3.1)$$

onde a barra denota conjugação complexa. Podemos escrever $\langle\psi| = |\psi\rangle^\dagger$.

Exercício 40:

Considere um estado $|\psi\rangle$ e um outro estado $|\varphi\rangle = e^{i\alpha}|\psi\rangle$, onde α é uma constante real. Mostre que todos os valores esperados de observáveis no estado $|\psi\rangle$ coincidem com os do estado $|\varphi\rangle$. Logo, dois estados que distinguem por uma fase constante são fisicamente idênticos.

Note que assumimos que a fase é constante. Fases que dependem da posição ou do tempo podem ter consequências físicas observáveis. \blacktimes

Este formalismo é consideravelmente diferente do que utilizamos na Mecânica Clássica, mas perceba que sua essência é idêntica: dado um estado, obtemos os resultados possíveis das medições de diferentes observáveis. Em Mecânica Clássica os resultados eram determinísticos e agora são apenas probabilísticos e agora utilizamos uma matematização que aparenta ser completamente diferente, mas a ideia essencial do que queremos fazer

para descrever sistemas gerais. Esta escolha é feita por este não ser um curso de Mecânica Quântica, mas notamos que esse tipo de abordagem é extremamente comum. Como um comentário para aqueles que já estudaram Mecânica Quântica, perceba que raramente se vê a literatura postular que estados são funcionais lineares agindo sobre o espaço de observáveis, embora essa seja uma definição mais geral que engloba tanto estados puros quanto mistos desde o início.

enquanto físicos permanece igual. Para discussões mais aprofundadas sobre o porquê dos postulados e afins, recomenda-se a leitura dos textos citados anteriormente, em particular o primeiro capítulo do clássico [9], que fornece uma motivação excelente para a nossa formulação da Mecânica Quântica.

Vamos a um exemplo que nos permita compreender um pouco como utilizar este formalismo matemático.

3.2 Spin Eletrônico

No início do século XIX, Joseph Fraunhofer utilizou um telescópio para focalizar a luz do Sol em uma fenda e então passá-la por um prisma. O resultado foi uma faixa de cores indo do vermelho ao violeta, mas povoada por algumas linhas escuras no meio, que hoje conhecemos como linhas espectrais. A ausência das cores das linhas espectrais foi explicada qualitativamente com a percepção de que elas eram absorvidas pela parte fria da atmosfera solar após terem sido emitidas da superfície quente. Experimentos semelhantes foram feitos com outros astros e com outros objetos na própria Terra, que permitiram concluir que a cada elemento estão associadas algumas linhas espectrais. O sódio está associado a algumas cores, o magnésio a outras, etc. Em particular, isso possibilitou a descoberta de alguns elementos no Sol antes deles serem observados na Terra, como foi o caso do hélio[32, p. 77].

Isto eventualmente foi explicado com a teoria atômica. Descobriu-se que os elétrons possuem níveis de energia discretos em torno dos átomos e as frequências espectrais correspondem às frequências associadas aos fótons emitidos ou absorvidos quando um elétron transita de um nível de energia para outro. Para um processo de emissão de um fóton, por exemplo, tem-se

$$E_\gamma = E_i - E_f, \quad (3.2)$$

onde E_γ é a energia do fóton, E_i a energia inicial do elétron e E_f sua energia final.

Utilizando os postulados da Mecânica Quântica numa versão um pouco mais elaborada que a que apresentamos, pode-se obter uma descrição detalhada de quais são os níveis de energia e, por consequência, quais são as linhas espectrais de cada átomo. Porém, os resultados falharam em explicar o que era observado experimentalmente: para cada linha prevista pela teoria, o experimento detectava duas linhas próximas.

Wolfgang Pauli sugeriu que existia um grau de liberdade até então desconhecido e que assumiria apenas dois valores. Assim, todas as linhas previstas apareceriam duas vezes.

Por enquanto não vamos discutir o significado físico deste grau de liberdade, que hoje em dia é conhecido como spin. Apenas o usaremos como exemplo para ilustrar os axiomas da Mecânica Quântica. Usaremos a discussão sobre spin mais tarde, mas por enquanto apenas precisamos do fato de que é um sistema quântico de dois níveis.

Denotaremos por \hat{H} a matriz hermitiana associada à energia do sistema, conhecida como operador Hamiltoniano (também o chamaremos de Hamiltoniana). Como temos apenas dois valores possíveis, o espaço de Hilbert que escolhemos é \mathbb{C}^2 . Assim, a forma geral da Hamiltoniana é

$$\hat{H} = \begin{pmatrix} a & b + ic \\ b - ic & d \end{pmatrix}, \quad (3.3)$$

onde $a, b, c, d \in \mathbb{R}$. Note que essa restrição é devida à hipótese de que \hat{H} é hermitiano. Por simplicidade, vamos assumir que $c = 0$ e $d = a$. Definimos os estados

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{e} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (3.4)$$

que formam uma base de \mathbb{C}^2 .

Os estados do sistema são todos os elementos de \mathbb{C}^2 . Assim, os estados possíveis são

$$|\psi\rangle = \begin{pmatrix} x \\ y \end{pmatrix}, \quad (3.5)$$

com $x, y \in \mathbb{C}$. Como a expressão para o valor esperado de um observável envolve uma divisão por $\langle\psi|\psi\rangle$, é conveniente normalizarmos os estados para $\langle\psi|\psi\rangle = 1$. Isso corresponde a

$$\langle\psi|\psi\rangle = 1, \quad (3.6a)$$

$$\begin{pmatrix} x^* & y^* \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1, \quad (3.6b)$$

$$|x|^2 + |y|^2 = 1. \quad (3.6c)$$

O valor esperado de \hat{H} no estado $|0\rangle$ (que tem $\langle 0|0\rangle = 1$) é

$$\langle\hat{H}\rangle_0 = \langle 0|\hat{H}|0\rangle, \quad (3.7a)$$

$$= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (3.7b)$$

$$= \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \quad (3.7c)$$

$$= a. \quad (3.7d)$$

Note que como

$$\hat{H}|0\rangle = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}, \quad (3.8)$$

não vale que $\hat{H}|0\rangle = \alpha|0\rangle$ para algum α , e portanto não é certo que uma medição de \hat{H} retornará o valor a . Porém, se vários sistemas forem preparados no estado $|0\rangle$, a média das medições tenderá para a conforme o número de experimentos tende a infinito.

Exercício 41:

Mostre que $\langle\hat{H}\rangle_1 = a$.

✂

Exercício 42:

A incerteza de um observável \hat{A} em um certo estado é dada por

$$\Delta A = \sqrt{\langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2}. \quad (3.9)$$

Esta incerteza não tem caráter experimental, ela diz respeito ao fato de que \hat{A} simplesmente pode não possuir um valor definido no estado considerado. Mostre que no estado $|0\rangle$ a incerteza de \hat{H} é $\Delta H = b$. Qual a incerteza no estado $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$? \clubsuit

Eventualmente, foi sugerido que o spin teria origem a partir de uma noção de momento angular intrínseco do elétron. O que isso significa?

A primeira interpretação que temos seria que o elétron é como uma esfera rotacionando em torno de seu próprio eixo. Vamos analisar essa possibilidade.

Exercício 43:

Devido ao fato de que o spin tem apenas dois graus de liberdade, é possível mostrar que seu momento angular precisaria ser sempre $\pm \frac{\hbar}{2}$. Assumindo que o elétron é uma esfera sólida e que a velocidade na superfície do elétron não pode ultrapassar a velocidade da luz, obtenha o raio mínimo do elétron como sendo da ordem de $R \gtrsim 10^{-13}$ m. \clubsuit

Acreditava-se desde então que isso seria impossível, visto que um elétron tão grande dificilmente passaria despercebido[33, pp. 104–105]. Logo, a proposta que temos é que o elétron possui momento angular, apesar de não ser como uma esfera em rotação. É uma forma distinta de momento angular do que a que estamos acostumados em situações menos extremas.

Este aparente incômodo pode ser simplificado, e pretendemos fazer isso, sob a luz da teoria de representações.

3.3 Construções Básicas com Representações

Nossa motivação para a introdução da noção de representação é inspirada por [31, Chap. 13]. Além disso, boa parte das nossas discussões é baseada em [5].

Dado um sistema físico qualquer, é preciso que sejamos capazes de implementar, de alguma forma, os efeitos de transformações de Poincaré sobre o sistema.

De fato, suponha que você decida girar a mesa em que realiza um certo experimento. É preciso que haja uma forma de codificar essa alteração no sistema, e o mesmo valeria se fizéssemos um boost, uma translação espacial, começássemos a cronometrar num instante diferente, etc.

Assim, queremos ter uma forma de agir sobre os estados e observáveis do sistema de uma maneira que represente a ação do grupo. Esta é a ideia por trás de uma representação.

Como podemos fazer isso? Sabemos as noções básicas do que um grupo deve satisfazer: ser associativo, possuir identidade e inverso. Sabemos também como podemos agir em estados: por meio de matrizes. Logo, gostaríamos de representar um grupo qualquer por meio de matrizes agindo em um espaço vetorial V . Como estas matrizes deveriam ser inversíveis, faz sentido escolhermos o grupo $GL(V)$ das transformações lineares inversíveis em V .

Exercício 44:

Mostre que $GL(V)$, o grupo das transformações lineares inversíveis em um espaço vetorial V , é de fato um grupo sob a operação de composição. Recorde que a composição de funções é sempre associativa. \clubsuit

Definição 11 [Representação]:

Seja G um grupo. Uma representação de G é uma dupla (ρ, V) onde V é um espaço vetorial e $\rho: G \rightarrow \text{GL}(V)$ é um homomorfismo. Caso ρ seja injetivo, ou seja, caso $g_1, g_2 \in G, g_1 \neq g_2 \Rightarrow \rho(g_1) \neq \rho(g_2)$, dizemos que a representação é fiel. ♠

Um grupo familiar que serve um bom propósito como primeiro exemplo é $(\mathbb{R}, +)$, o grupo dos números reais dotado da operação de adição. Queremos representar sua estrutura por meio de multiplicação de matrizes. Como isso pode ser feito?

Há várias representações possíveis. Vamos começar exibindo uma representação unidimensional. Tome o espaço vetorial como sendo \mathbb{R} e o homomorfismo como sendo $\rho_1: \mathbb{R} \rightarrow \mathbb{R}$ tal que $\rho_1(g) = e^g$. De fato,

$$\rho_1(g_1 + g_2) = e^{g_1 + g_2} = e^{g_1} e^{g_2} = \rho_1(g_1) \rho_1(g_2). \quad (3.10)$$

Note também que esta é uma representação fiel: se $g_1 \neq g_2$, então $e^{g_1} \neq e^{g_2}$.

Podemos também escolher uma representação bidimensional. Tome o espaço vetorial como sendo \mathbb{R}^2 e escolha o homomorfismo

$$\rho_2(g) = \begin{pmatrix} 1 & g \\ 0 & 1 \end{pmatrix}. \quad (3.11)$$

Exercício 45:

Mostre que a Eq. (3.11) define um homomorfismo de $(\mathbb{R}, +)$ em $\text{GL}_{\mathbb{R}}(2)$. ✖

Há outras opções de representações bidimensionais. Tome, por exemplo,

$$\rho_3(g) = \begin{pmatrix} 1 & 0 \\ g & 1 \end{pmatrix} \quad (3.12)$$

ou

$$\rho_4(g) = \begin{pmatrix} \cos g & \sin g \\ -\sin g & \cos g \end{pmatrix}. \quad (3.13)$$

Exercício 46:

Mostre que as Eqs. (3.12) e (3.13) definem homomorfismos de $(\mathbb{R}, +)$ em $\text{GL}_{\mathbb{R}}(2)$. ✖

Exercício 47:

Quais das representações definidas pelas Eqs. (3.11) a (3.13) são fiéis? ✖

Podemos então nos perguntar se estas representações são de fato fundamentalmente distintas ou apenas formas distintas de escrever a mesma coisa.

Exercício 48:

Mostre que existe uma matriz unitária e ortogonal M tal que

$$M \rho_2(g) M^{-1} = \rho_3(g) \quad (3.14)$$

vale. Sugestão: a solução precisa ser uma matriz simétrica. ✖

Note que a existência de uma matriz como a que ocorre no Exercício 48 na página anterior significa que as duas representações diferem apenas por uma mudança de base. Logo, no fundo, são como duas formas de escrever a mesma representação. Porém, é possível que existam representações distintas numa mesma dimensão.

Exercício 49:

Mostre que não existe nenhuma matriz M tal que

$$M\rho_2(g)M^{-1} = \rho_4(g) \quad (3.15)$$

vale para todo $g \in \mathbb{R}$. Ademais, mostre que a única matriz M que satisfaz a condição mais fraca

$$M\rho_2(g) = \rho_4(g)M \quad (3.16)$$

é a matriz nula $M = 0$. ✖

É interessante que tenhamos uma forma de classificar se duas representações são, no fundo, a mesma.

Definição 12 [Intertwiner]:

Seja G um grupo e sejam (V_1, ρ_1) e (V_2, ρ_2) representações de G . Um intertwiner de ρ_1 e ρ_2 é um operador linear $U: V_1 \rightarrow V_2$ tal que $U\rho_1(g) = \rho_2(g)U$ para todo $g \in G$. ♠

Por ora, a noção de intertwiner só nos é interessante para definir representações equivalentes, mas a visitaremos novamente mais tarde. Lema de Schur

Definição 13 [Representações Equivalentes]:

Duas representações (V_1, ρ_1) e (V_2, ρ_2) de um mesmo grupo G são ditas equivalentes quando possuem um intertwiner inversível, isto é, quando há um operador linear $U: V_1 \rightarrow V_2$ tal que

$$U\rho_1(g)U^{-1} = \rho_2(g) \quad (3.17)$$

para todo $g \in G$. ♠

Vamos considerar uma outra representação, desta vez em \mathbb{R}^3 . Tome o homomorfismo como sendo

$$\rho_5(g) = \begin{pmatrix} e^g & 0 & 0 \\ 0 & \cos g & \sin g \\ 0 & -\sin g & \cos g \end{pmatrix}. \quad (3.18)$$

Exercício 50:

Mostre que Eq. (3.18) define um homomorfismo. ✖

Apesar de Eq. (3.18) definir uma representação, percebemos que ela se assemelha com representações que estudamos anteriormente. Em outras palavras, podemos quebrá-la em pedaços menores.

Por exemplo, considere os vetores de \mathbb{R}^3 que pertencem ao subespaço $V_1 = \text{span}\{(1, 0, 0)\}$, ou seja, os vetores da forma $\lambda \cdot (1, 0, 0)$. Para estes vetores, temos

$$\rho(g)v = \begin{pmatrix} e^g & 0 & 0 \\ 0 & \cos g & \sin g \\ 0 & -\sin g & \cos g \end{pmatrix} \begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix} = e^g \begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix} \in V_1. \quad (3.19)$$

Ou seja, o subespaço V_1 é mantido invariante pela representação: após agir com a representação num elemento de V_1 , continuamos em V_1 .

Note, contudo, que

$$\rho(g) \begin{pmatrix} 0 \\ \lambda \\ 0 \end{pmatrix} = \begin{pmatrix} e^g & 0 & 0 \\ 0 & \cos g & \sin g \\ 0 & -\sin g & \cos g \end{pmatrix} \begin{pmatrix} 0 \\ \lambda \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \cos g \lambda \\ -\sin g \lambda \end{pmatrix} \quad (3.20)$$

e

$$\rho(g) \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix} = \begin{pmatrix} e^g & 0 & 0 \\ 0 & \cos g & \sin g \\ 0 & -\sin g & \cos g \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ \sin g \lambda \\ \cos g \lambda \end{pmatrix}. \quad (3.21)$$

Logo, $\text{span}\{(0, 1, 0)\}$ e $\text{span}\{(0, 0, 1)\}$ não são invariantes, mas $V_2 = \text{span}\{(0, 1, 0), (0, 0, 1)\}$, composto dos vetores da forma $(0, \lambda, \mu)$, é mantido invariante pela representação. Se $v \in V_2$, $\rho(g)v \in V_2$ para todo $g \in G$.

Definição 14 [Subespaços Invariantes por uma Representação]:

Seja G um grupo e (V, ρ) uma representação de G . Um subespaço W de V é dito ser invariante por ρ se, e somente se, $w \in W$ implica que $\rho(g)w \in W$ para todos $g \in G$. ♠

Note que $W = V$ e $W = \{0\}$ sempre são invariantes sob qualquer representação.

Definição 15 [Representações Redutíveis e Irredutíveis]:

Seja G um grupo e (V, ρ) uma representação de G . ρ é dita ser irredutível se os únicos subespaços invariantes forem os subespaços triviais $W = V$ e $W = \{0\}$. Do contrário, ρ é dita ser redutível. ♠

Também é comum nos referirmos a representações irredutíveis como “irreps” (*irreducible representations*).

Eq. (3.18) na página precedente define uma representação redutível, pois vimos que V_1 e V_2 são subespaços invariantes. Na verdade, Eq. (3.18) na página anterior define um tipo especial de representação redutível: uma representação maximalmente redutível, pois podemos de fato expressá-la como uma matriz com blocos diagonais compostos por representações irredutíveis. Vamos rephrasing isso de maneira mais precisa.

Definição 16 [Soma Direta Algébrica]:

Sejam V e W espaços vetoriais de dimensão finita sobre um mesmo corpo \mathbb{F} . Definimos sua soma direta, $V \oplus W$, como o espaço vetorial $(V \times W, \mathbb{F}, +, \cdot)$ onde a soma e produto são dados por

$$(v_1, w_1) + \alpha \cdot (v_2, w_2) = (v_1 + \alpha v_2, w_1 + \alpha w_2) \quad (3.22)$$

para todos $v_1, v_2 \in V, w_1, w_2 \in W$ e $\alpha \in K$. ♠

O adjetivo “algébrica” se refere ao fato de que como estamos em dimensão finita, não é preciso se preocupar com questões envolvendo convergência. Note que se V e W são espaços de Hilbert, $V \oplus W$ também é um espaço de Hilbert se utilizarmos o produto interno dado por

$$\langle (v_1, w_1) | (v_2, w_2) \rangle_{V \oplus W} = \langle v_1 | v_2 \rangle_V \langle w_1 | w_2 \rangle_W. \quad (3.23)$$

Como um exemplo imediato, veja que

$$\mathbb{R}^3 = \mathbb{R} \oplus \mathbb{R}^2. \quad (3.24)$$

Porém, há uma outra forma de caracterizarmos a soma direta. Suponha que W_1 e W_2 são subespaços vetoriais de V tais que $W_1 \cap W_2 = \{0\}$ e tais que todo elemento v de V é da forma $v = w_1 + w_2$ para $w_1 \in W_1$ e $w_2 \in W_2$. Então vale que $W_1 \oplus W_2$ é isomorfo a V . Em particular, podemos reescrever Eq. (3.24) como

$$\mathbb{R}^3 = \text{span}\{(1, 0, 0)\} \oplus \text{span}\{(0, 1, 0), (0, 0, 1)\}. \quad (3.25)$$

Por vezes é conveniente denotarmos o elemento (v, w) como $v \oplus w$. Além disso, se temos uma matriz $A: V \rightarrow V$ e outra $B: W \rightarrow W$, então podemos definir $A \oplus B$ por $(A \oplus B)(v \oplus w) = Av \oplus Bw$. Note que, desta forma, matricialmente temos (digamos que V tem duas dimensões e W uma, por simplicidade)

$$(A \oplus B)(v \oplus w) = \begin{pmatrix} A_1 1 & A_1 2 & 0 \\ A_2 1 & A_2 2 & 0 \\ 0 & 0 & B \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ w \end{pmatrix} = \begin{pmatrix} A_1 1 v_1 + A_1 2 v_2 \\ A_2 1 v_1 + A_2 2 v_2 \\ B w \end{pmatrix}, \quad (3.26)$$

de forma que a matriz $A \oplus B$ é dada por uma diagonal em blocos.

Definição 17 [Soma Direta de Representações]:

Seja G um grupo e (V_1, ρ_1) e (V_2, ρ_2) duas de suas representações. Definimos a soma direta destas duas representações como a representação (V, ρ) com espaço vetorial $V = V_1 \oplus V_2$ e homomorfismo $\rho(g) = \rho_1(g) \oplus \rho_2(g)$ para todo $g \in G$. ♠

Com relação aos exemplos que demos anteriormente, vemos agora que

$$\rho_5 = \rho_1 \oplus \rho_4, \quad (3.27)$$

sendo que ρ_1 e ρ_4 são irredutíveis.

Em espaços de Hilbert de dimensão finita, esta decomposição é sempre possível.

Teorema 18:

Seja \mathcal{H} um espaço de Hilbert de dimensão finita e seja ρ uma representação de um grupo G em \mathcal{H} por operadores unitários, isto é, tal que $\rho(g)^ = \rho(g)^{-1} = \rho(g^{-1})$ para todo $g \in G$. Então ou ρ é irredutível ou existem representações irredutíveis (V_i, ρ_i) , $1 \leq i \leq n$ para algum n natural tais que $\mathcal{H} = V_1 \oplus \cdots \oplus V_n$ e $\rho = \rho_1 \oplus \cdots \oplus \rho_n$. □*

Demonstração:

Se ρ é irredutível, está encerrada a demonstração. Do contrário, existe algum subespaço não-trivial W que é invariante por ρ .

Denotemos $W^\perp = \{v \in \mathcal{H}; \langle v|w \rangle = 0, \forall w \in W\}$. É possível mostrar que $V = W \oplus W^\perp$. Note que, se $v \in W^\perp$ e $w \in W$,

$$\langle \rho(g)v|w \rangle = \langle v|\rho(g)^*w \rangle, \quad (3.28a)$$

$$= \langle v|\rho(g^{-1})w \rangle, \quad = 0, \quad (3.28b)$$

onde a última linha se deve ao fato de que $\rho(g^{-1})w \in W$ por W ser invariante. Contudo, assim provamos que $\langle \rho(g)v|w \rangle = 0$ para todo $g \in G$ e todo $w \in W$. Logo $v \in W^\perp \Rightarrow \rho(g)v \in W^\perp$ para todo $g \in G$, o que significa que W^\perp é invariante sob ρ .

Defina $V_1 = W$ e $V_2 = W^\perp$. Defina $\rho_i(g)v = \rho(g)v$ para todos $g \in G, v \in V_i$. Vale então que $\rho = \rho_1 \oplus \rho_2$. Se ρ_1 e ρ_2 forem irredutíveis, a demonstração está concluída. Do contrário, basta repetir o processo. Como a dimensão de W é sempre estritamente menor que a de \mathcal{H} , que é finita, o processo acaba. ■

Exercício 51:

Mostre que o Teorema 18 na página anterior também vale para o caso de representações anti-unitárias, ou seja, tais que $\langle v|\rho(g)w \rangle = \langle \rho(g)^*w|v \rangle$ e $\rho(g)^* = \rho(g)^{-1} = \rho(g^{-1})$. ✖

Em geral, nosso interesse se concentrará nestes tipos de representações por serem exatamente as que possuem interesse físico em Mecânica Quântica. Isso é garantido pelo chamado Teorema de Wigner.

Teorema 19 [Wigner]:

Toda transformação de simetria pode ser representada no espaço de Hilbert de estados físicos por operadores unitários ou anti-unitários fixos a menos de uma fase. □

Demonstração:

Ver [34, Sec. 2.A]. ■

Por transformação de simetria, entendemos que todas as probabilidades e valores esperados devem ser preservados. O teorema de Wigner garante que a única forma de fazer isso é por meio de operadores unitários ou anti-unitários. A observação acerca de uma fase se deve ao fato de que não precisamos de algo tão forte quanto uma representação para estudar um sistema quântico. Voltaremos a esse assunto na Seção 3.5 na página 51.

3.4 Produtos Tensoriais e Tensores

A esta altura, estamos vendo que as representações que realmente nos interessam são as irredutíveis. Na maioria das situações de nosso interesse, seremos capazes de escrever outras representações como somas diretas de irreps. Porém, como podemos obter novas irreps de um grupo?

Uma das maneiras de o fazer é utilizar as representações que já conhecemos para construir novas representações maiores e então as decompor em componentes irredutíveis.

Nossas referências principais para esta discussão são [11, 36]. Nosso exemplo principal será o grupo $\text{SO}(3)$.

O grupo $\text{SO}(3)$, assim como todo grupo matricial, admite uma representação particularmente simples devido a já ser um subgrupo de algum $\text{GL}(V)$: basta tomar $\text{id}: \text{SO}(3) \rightarrow \text{GL}_3(\mathbb{R})$, ou seja, cada matriz de $\text{SO}(3)$ é representada por si mesma. Esta representação, conhecida como representação fundamental, atua como

$$v^i \rightarrow v'^i = R^i_j v^j, \quad (3.29)$$

onde estamos denotando os índices referentes a linhas em cima (v^i são as coordenadas de um vetor coluna) e os referentes a colunas embaixo (v_i são as coordenadas de um vetor linha). Por exemplo, o objeto

$$v = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad (3.30)$$

é descrito nesta notação por $v^1 = a$, $v^2 = b$ e $v^3 = c$. Convencionamos também que índices repetidos estão sendo somados. Em notação matricial, a Eq. (3.29) seria escrita como $\mathbf{v} \rightarrow \mathbf{v}' = R\mathbf{v}$.

A representação fundamental também é conhecida pelos físicos como representação vetorial. Enquanto no jargão matemático um vetor é um elemento de um espaço vetorial, o jargão físico é mais restritivo: vetores são definidos pela forma como se transformam sob certos grupos. Em geral, ao dizer que v^i é um “vetor”, isso significa que v^i é um elemento do espaço vetorial associado à representação fundamental de $\text{SO}(3)$. Isso é uma forma técnica de dizer que vale a Eq. (3.29). Para evitar usar o mesmo termo para mais de um conceito, tentarei usar a palavra “trivetor” para vetores de $\text{SO}(3)$, mas provavelmente cometerei alguns deslizes.

A partir desta noção, podemos começar a pensar em novos objetos que se transformam de outras formas. No final do dia, devemos conseguir novos espaços vetoriais, que de maneira simplista podemos enxergar como formados por conjuntos de números. Afinal, um vetor coluna é um conjunto de N números, por exemplo. Porém, não há regras sobre como vamos arranjar estes números, desde que eles ainda se combinem de maneira linear. Por exemplo, o objeto na Eq. (3.30) poderia ser reescrito na forma de uma matriz 3×3 antissimétrica,

$$v = \begin{pmatrix} 0 & a & -b \\ -a & 0 & c \\ b & -c & 0 \end{pmatrix}, \quad (3.31)$$

que contém a mesma informação em um arranjo diferente. Outra opção seria usar matrizes 2×2 hermitianas e escrever

$$v = v^i \sigma_i = \begin{pmatrix} c & a - ib \\ a + ib & -c \end{pmatrix}. \quad (3.32)$$

*Isso varia com a área de atuação, mas físicos também frequentemente se referem às componentes de um vetor como vetor, sem distinguir o objeto de suas componentes.

Todas estas são maneiras de armazenar informação sobre um espaço vetorial tridimensional, e podemos explorar essa liberdade para escrever novas representações com mais facilidade.

Começemos pensando em objetos com dois índices, T^{ij} . Note que isto não é uma matriz: matrizes tem um índice linha e um índice coluna, M^i_j , mas o objeto que estamos considerando possui dois índices de linhas. Logo, é um objeto que pode ser multiplicado com dois vetores linha (da forma v_i) para obter um número real. Como T^{ij} possui dois índices coluna, esperamos que se transforme como dois trivetores, ou seja, como

$$T^{ij} \rightarrow T'^{ij} = R^i_k R^j_l T^{kl}. \quad (3.33)$$

Cada índice de T^{ij} é transformado como se pertencesse a um trivetor. Em outras palavras, o objeto T^{ij} se transforma da mesma maneira que um objeto $u^i v^j$, onde u^i e v^i são trivetores.

Para que Eq. (3.33) defina uma representação, é preciso que os objetos T^{ij} formem um espaço vetorial. A construção desse espaço vetorial, que é conhecido como produto tensorial, é mais técnica do que é cabido para este curso, e recomendamos a leitura de textos mais profundos, como [5, 11, 18], para estes detalhes. Uma forma de perceber esta estrutura é enxergando os vetores linha como funções que “comem” um vetor coluna e “cospem” um número. De fato, dados u_i e v^j , $u_i v^i$ é um número real. A mesma construção pode ser feita ao contrário: vetores coluna “comem” vetores linha e “cospem” números reais. Nesta linha de raciocínio, os objetos T^{ij} formam o espaço vetorial de todos as funções $T: V^* \times V^* \rightarrow \mathbb{R}$, onde V^* denota o espaço dos vetores linha. [17, 24, 31] discutem estas construções em situações de interesse físico.

Exercício 52:

Convença-se que a transformação da Eq. (3.33) de fato define uma representação. Por simplicidade, tome como hipótese que os objetos T^{ij} formam um espaço vetorial. ✂

Os objetos T^{ij} são conhecidos como tensores. Mais especificamente, tensores de tipo $(2, 0)$. De forma genérica, um tensor de tipo (p, q) possui p índices em cima e q índices embaixo, ou seja, é da forma $T^{a_1 \dots a_p}_{b_1 \dots b_q}$. O rank de um tensor do tipo (p, q) é, por definição, $p + q$.

É interessante nos perguntarmos se existe alguma diferença real entre índices em cima e embaixo. No caso geral, há. Em diversos grupos esta distinção é fundamental. Porém, no caso do grupo $SO(3)$ a distinção é supérflua.

Exercício 53:

Mostre que v^i e v_i se transformam da mesma maneira sob $SO(3)$, onde convencionamos que o vetor linha associado a um vetor coluna é obtido por $v_i = \delta_{ij} v^j$ e $v^i = \delta^{ij} v_j$, onde δ^{ij} e δ_{ij} valem 1 se $i = j$ e 0 caso contrário. Em outras palavras, mostre que as representações definidas por v^i e v_i são equivalentes. ✂

O próximo passo é verificarmos se a representação fornecida pelos tensores $(2, 0)$ é irredutível ou não. Em geral, tomar produtos tensoriais para formar novas representações nos leva a representações redutíveis, e uma forma de as decompor em representações redutíveis é utilizando informações sobre combinações simétricas e antissimétricas.

Por exemplo, tomemos T^{ij} . Em geral, $T^{ij} \neq T^{ji}$ e $T^{ij} \neq -T^{ji}$, o que significa que T^{ij} não é um tensor simétrico nem um tensor antissimétrico. Porém, perceba que podemos escrevê-lo como

$$T^{ij} = \frac{T^{ij} + T^{ji}}{2} + \frac{T^{ij} - T^{ji}}{2}, \quad (3.34)$$

e assim decompô-lo em uma parte simétrica, dada por

$$T^{(ij)} = \frac{1}{2}(T^{ij} + T^{ji}), \quad (3.35)$$

e uma parte antissimétrica,

$$T^{[ij]} = \frac{1}{2}(T^{ij} - T^{ji}). \quad (3.36)$$

As notações $T^{[ij]}$ e $T^{(ij)}$ são particularmente úteis e comuns em geometria diferencial. É útil as definirmos também para os casos em que mais índices ocorrem. Para um tensor do tipo $(l, 0)$ define-se

$$T^{(i_1 \cdots i_l)} = \frac{1}{l!} \sum_{\pi} T^{(i_{\pi(1)} \cdots i_{\pi(l)})} \quad (3.37)$$

e

$$T^{[i_1 \cdots i_l]} = \frac{1}{l!} \sum_{\pi} \delta_{\pi} T^{(i_{\pi(1)} \cdots i_{\pi(l)})}, \quad (3.38)$$

onde π são as permutações possíveis e δ_{π} vale $+1$ para permutações pares e -1 para permutações ímpares. Por exemplo, $T^{[ij]} = \frac{1}{2}(T^{ij} - T^{ji})$, pois trocar (i, j) por (j, i) é uma permutação par (0 trocas são necessárias), mas trocar (i, j) por (j, i) é uma permutação ímpar (1 troca necessária). As definições para tensores mais gerais são análogas

Exercício 54:

Prove que as identidades

$$T^{(ij)}_k = \frac{1}{2}(T^{ij}_k + T^{ji}_k), \quad (3.39)$$

$$T^{[ijk]} = \frac{1}{6}(T^{ijk} + T^{jki} + T^{kij} - T^{jik} - T^{ikj} - T^{kji}), \quad (3.40)$$

e

$$T^{(ij)}_{[kl]} = \frac{1}{4}(T^{ij}_{kl} - T^{ij}_{lk} + T^{ji}_{kl} - T^{ji}_{lk}) \quad (3.41)$$

valem.

✂

Exercício 55:

Mostre que a expressão

$$T^{[ij]} = \frac{1}{2} \epsilon^{ijk} \epsilon_{klm} T^{lm} \quad (3.42)$$

vale.

✂

Exercício 56:

Mostre que $S^{(ij)}A_{[ij]} = 0$ para quaisquer dois tensores S^{ij} e A_{ij} . \clubsuit

Exercício 57:

Suponha que $S^{ij} = S^{ji}$ e $A^{ij} = -A^{ji}$ são tensores de $\text{SO}(3)$. Mostre que suas propriedades de simetria ou antissimetria são preservadas por transformações de $\text{SO}(3)$. \clubsuit

Logo, o subespaço de tensores antissimétricos é invariante sob a representação que escolhemos, e o mesmo vale para o subespaço de tensores simétricos. Podemos então decompor a representação em uma soma direta da representação simétrica e da representação antissimétrica.

Façamos essa análise de maneira mais explícita, seguindo os métodos descritos, por exemplo, em [10, Cap. 10]. Como T^{ij} se transforma como $u^i v^j$, podemos fazer a análise de representações com $u^i v^j$ para deixar mais explícito que estamos considerando o produto tensorial da representação fundamental consigo mesma. Note que podemos começar escrevendo

$$u^i v^j = u^{(i} v^{j)} + u^{[i} v^{j]}. \quad (3.43)$$

Para um tensor simétrico, podemos ainda separar sua parte com traço do restante. O traço de um tensor S^{ij} é dado por $S = \delta_{ij} S^{ij}$. Note que $\frac{1}{3} S \delta^{ij}$ e S^{ij} tem o mesmo traço (o fator de 3 é necessário porque, em três dimensões, $\delta^{ij} \delta_{ij} = 3$). Logo, podemos escrever

$$S^{ij} = \frac{1}{3} S \delta^{ij} + \left(S^{ij} - \frac{1}{3} S \delta^{ij} \right). \quad (3.44)$$

Exercício 58:

Mostre que transformações de $\text{SO}(3)$ preservam o traço de tensores de rank 2. \clubsuit

Assim, podemos escrever

$$u^i v^j = \frac{1}{3} u_k v^k \delta^{ij} + u^{[i} v^{j]} + \left(u^{(i} v^{j)} - \frac{1}{3} u_k v^k \delta^{ij} \right). \quad (3.45)$$

Calculando de forma um pouco mais explícita como as transformações de cada setor ocorrem, pode-se mostrar que essas representações são de fato irredutíveis (veja [36, Chap. IV.1]).

Exercício 59:

Mostre que a representação de $\text{SO}(3)$ descrita pelos tensores que se transformam como $\frac{1}{3} u_k v^k \delta^{ij}$ tem dimensão 1. Mostre ainda que as representações de $u^{[i} v^{j]}$ e $(u^{(i} v^{j)} - \frac{1}{3} u_k v^k \delta^{ij})$ tem dimensões 3 e 5, respectivamente.

Sugestão: quantos números são necessários para escrever a matriz mais geral com entradas dadas por $\frac{1}{3} u_k v^k \delta^{ij}$? \clubsuit

Note em particular que a representação de $u^{[i} v^{j]}$ é equivalente à representação fundamental, visto que a podemos escrever como $\epsilon_{ijk} u^{[j} v^{k]}$ e $\text{SO}(3)$ não distingue índices em cima e embaixo.

Exercício 60:

Mostre que $\epsilon_{ijk}u^{[j}v^{k]}$ se transforma como um trivetor. Sugestão: se R é uma matriz 3×3 , então $\epsilon_{ijk}R^i_l R^j_m R^k_n = \det R \epsilon_{lmn}$. \blacktimes

Assim, concluímos que o produto tensorial da representação fundamental consigo mesma é uma soma direta de três representações. É comum denotarmos as representações em termos de sua dimensão e por isso escrevemos

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{5} \oplus \mathbf{3} \oplus \mathbf{1}. \quad (3.46)$$

Esta notação pode ser preocupante caso tenhamos mais de uma representação irredutível com a mesma dimensão, mas $\text{SO}(3)$ é um grupo particularmente bem comportado. Sempre que temos dois índices antissimétricos, podemos usar o símbolo de Levi-Civita e escrever $\epsilon_{ijk}u^{[j}v^{k]}$. Na prática, para este grupo, precisamos lidar apenas com tensores totalmente simétricos.

3.5 Representações Projetivas e Spinors

Nossa definição de representação exigia a presença de um homomorfismo, ou seja, uma função $U: G \rightarrow \text{GL}(\mathcal{H})$, onde \mathcal{H} é agora o espaço de Hilbert da teoria quântica, tal que $U(g_1 g_2) = U(g_1)U(g_2)$. Contudo, as probabilidades quânticas não seriam alteradas se tivéssemos apenas $U(g_1, g_2) = e^{i\alpha(g_1, g_2)}U(g_1)U(g_2)$, pois as expressões diferem apenas por uma fase que independe do espaço e do tempo. Nossa discussão é inspirada por [16, 31, 34].

Por conta desta liberdade de fase, a teoria quântica nos obriga a considerar também representações um tanto mais gerais, conhecidas como representações projetivas, que são precisamente estas noções generalizadas de representações com uma fase adicional nos produtos (veja [16, 31, 34] para abordagens mais detalhadas). Contudo, uma forma de contornar isto é admitindo um grupo maior. Como mencionamos anteriormente, o grupo $\text{SU}(2)$ é o recobrimento universal de $\text{SO}(3)$ e é possível mostrar que as representações unitárias de $\text{SU}(2)$ correspondem às representações unitárias projetivas ou não de $\text{SO}(3)$. Porém, todas as representações unitárias de $\text{SU}(2)$ são representações genuínas, não projetivas. Assim, trabalhando com $\text{SU}(2)$ temos uma maneira mais simples de lidar com o problema de encontrar as representações de $\text{SO}(3)$.

Consideremos então a representação fundamental de $\text{SU}(2)$. Note que podemos entender esta representação como uma representação de $\text{SO}(3)$, dado que podemos construir um homomorfismo injetor de $\text{SO}(3)$ em $\text{SU}(2)$ pelo procedimento da Seção 2.4 na página 32. Isto nos explica então a forma como rotações podem ser descritas de maneira não-trivial num espaço de Hilbert bidimensional, como o que utilizamos mais cedo ao pensar sobre spin eletrônico.

Isso explica também em partes o porquê da noção de spin ter sido tão curiosa à primeira vista. Ela ocorre de uma maneira pouco convencional dentro do grupo $\text{SO}(3)$.

Um exercício comum em cursos de Mecânica Quântica e que é interessante como um novo exemplo no nosso contexto é o estudo de um sistema com dois elétrons. Cada um dos elétrons possui um spinor — o nome que damos à representação $\mathbf{2}$ de $\text{SO}(3)$,

embora o termo também ocorra em outros contextos — que descreve o seu spin. Os dois spinors são independentes, mas ocorrem simultaneamente na descrição, o que nos motiva a descrever o sistema completo em termos de um produto tensorial. Assim, o vetor que descreve o sistema se transforma na representação $\mathbf{2} \otimes \mathbf{2}$ de $\text{SO}(3)$. Podemos escrever essa representação em termos de representações irredutíveis?

Faremos a análise por métodos tensoriais, mas desta vez trabalhando com tensores de $\text{SU}(2)$, visto que isto deixa a análise em termos de representações usuais, não projetivas. Há uma diferença desta vez, no entanto: nós precisaremos distinguir índices em cima e embaixo, pois para abaixar índices em grupos formados por matrizes complexas é preciso fazer uma conjugação complexa, ou seja, definimos $u_i = \delta_{ij} \bar{u}_j$, onde a barra denota conjugação complexa. A razão para isto é que matrizes unitárias preservam $\sum_i \bar{u}^i v^i$, mas não $\sum_i u^i v^i$.

Exercício 61:

Mostre que a combinação $\delta_{ij} u^i v^j$ não é invariante por $\text{SU}(2)$, mas se transforma de acordo com $\delta_{ij} u^i v^j \rightarrow u^k v^l (U^\dagger U)^{kl}$, onde deixamos todos os índices em cima para denotar que representam os elementos de matriz usuais. \boxtimes

Exercício 62:

Mostre que $\bar{u}^i \rightarrow \bar{u}^j (U^\dagger)^{ji}$. \boxtimes

Da mesma forma que denotamos $u_i = \bar{u}^j$, também denotaremos $(U^\dagger)_j^i = (U^\dagger)^{ji}$. Assim, a lei de transformação dos vetores conjugados u_i se torna $u^i \rightarrow u^j (U^\dagger)_j^i$.

Com isto em mente, começamos como com $\mathbf{3} \otimes \mathbf{3}$, separando as partes simétrica e antissimétrica. Temos

$$u^i v^j = u^{(i} v^{j)} + u^{[i} v^{j]}. \quad (3.47)$$

A seguir, tentaríamos separar a parte simétrica em parte com traço e parte sem traço, mas isso não é necessário: nos grupos unitários, o traço é tomado apenas entre índices diferentes (um em cima e um embaixo), pois a quantidade preservada pelas matrizes unitárias é $u_i v^i$, não $\delta_{ij} u^i v^j$. Assim, não podemos separar traço algum. Percebendo que a parte simétrica tem 3 dimensões e a antissimétrica tem 1, concluímos que

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3}_S \oplus \mathbf{1}_A, \quad (3.48)$$

onde os subscritos são apenas lembretes das propriedades de simetria. Em Mecânica Quântica, isso reflete o fato de que há um tripleto de estados em que a parte espinorial do estado é simétrica e um singlete em que a parte espinorial é antissimétrica.

3.6 Lema de Schur

Antes de partirmos para aplicações mais gerais em Física, vamos a um último resultado geral sobre teoria de representações: o Lema de Schur. Dadas algumas condições, seu resultado (conforme enunciado pelos físicos em geral) é simples: se uma matriz A comuta com $\rho(g)$ para todo g pertencente a um grupo G e alguma representação irredutível

ρ , então $A = \lambda \mathbb{1}$ para alguma constante λ . A forma que demonstraremos é um tanto diferente, mas este é um de seus corolários e o foco de nosso interesse.

Por ora, este parece um resultado meramente técnico, mas veremos em breve que ele possui consequências físicas profundas. Nossa demonstração seguirá [5].

Teorema 20 [Lema de Schur]:

Seja G um grupo. Tome duas representações irredutíveis de G , (ρ_1, V_1) e (ρ_2, V_2) , sendo que V_1 e V_2 não são triviais (ou seja, não são compostos apenas pelo vetor nulo). Nestas condições, se a transformação linear $A: V_1 \rightarrow V_2$ for um intertwiner de ρ_1 e ρ_2 , isto é, se $A\rho_1(g) = \rho_2(g)A$ para todo $g \in G$, então ou A é inversível ou $A = 0$. \square

Faremos a demonstração por meio de exercícios guiados.

Exercício 63:

Defina o núcleo $\text{Ker } A$ e a imagem $\text{Ran } A$ de uma transformação linear $A: V_1 \rightarrow V_2$ por

$$\text{Ker } A = \{v \in V_1; Av = 0\} \quad (3.49)$$

e

$$\text{Ran } A = \{Av; v \in V_1\}. \quad (3.50)$$

Mostre que, nas condições do Lema de Schur, $\text{Ker } A$ é invariante por ρ_1 e $\text{Ran } A$ é invariante por ρ_2 . \blackcross

Como assumimos que as representações são irredutíveis, isso implica que $\text{Ker } A = \{0\}$ ou $\text{Ker } A = V_1$ e, de modo análogo, $\text{Ran } A = \{0\}$ ou $\text{Ran } A = V_2$. Assim, temos quatro casos:

- i. $\text{Ker } A = \{0\}$ e $\text{Ran } A = \{0\}$;
- ii. $\text{Ker } A = \{0\}$ e $\text{Ran } A = V_2$;
- iii. $\text{Ker } A = V_1$ e $\text{Ran } A = \{0\}$;
- iv. $\text{Ker } A = V_1$ e $\text{Ran } A = V_2$.

Exercício 64:

Seja $A: V_1 \rightarrow V_2$ uma transformação linear. Mostre que $\text{Ker } A = \{0\}$ e $\text{Ran } A = \{0\}$ não pode valer sob as condições do Lema de Schur. Faça o mesmo para $\text{Ker } A = V_1$ e $\text{Ran } A = V_2$. \blackcross

Assim, restam duas opções apenas.

Exercício 65:

Seja $A: V_1 \rightarrow V_2$ uma transformação linear. Mostre que $\text{Ker } A = V_1$ e $\text{Ran } A = \{0\}$ implica $A = 0$. \blackcross

Se este não for o caso, então resta uma única opção.

Exercício 66:

Seja $A: V_1 \rightarrow V_2$ uma transformação linear. Mostre que $\text{Ker } A = \{0\}$ e $\text{Ran } A = V_2$ implica que A é injetor e sobrejetor. Por consequência, A é bijetor, ou seja, inversível. \blackbox

Com o Lema de Schur em mãos, podemos agora estudar um de seus corolários, que nos será de grande interesse (e frequentemente nos referiremos a ele como Lema de Schur).

Corolário 21:

Seja G um grupo e (ρ, \mathcal{H}) uma representação irredutível e unitária de G num espaço de Hilbert. Suponha que $A: \mathcal{H} \rightarrow \mathcal{H}$ é um operador linear limitado* tal que $A\rho(g) = \rho(g)A$ para todo $g \in G$. Nestas condições, existe $\lambda \in \mathbb{C}$ tal que $A = \lambda\mathbb{1}$. \square

Demonstração:

Faremos a demonstração no caso de dimensão finita. A demonstração geral se aplica a qualquer dimensão, mas por utilizar conceitos mais delicados de Análise Funcional que generalizam as noções que utilizaremos abaixo, é menos pedagógica dado o escopo dessas notas. Contudo, ela pode ser encontrada em [5].

Como A é um intertwiner de ρ consigo mesmo, sabemos imediatamente que é inversível ou nulo. Porém, o mesmo vale para $A - \lambda\mathbb{1}$, onde λ é qualquer número complexo. Se provarmos que existe λ tal que $A - \lambda\mathbb{1}$ é nulo, então concluímos que $A = \lambda\mathbb{1}$ e o resultado está provado. Pelo Lema de Schur, basta mostrarmos que $A - \lambda\mathbb{1}$ não é inversível.

Para saber se $A - \lambda\mathbb{1}$ é ou não inversível, podemos observar seu determinante. Para que não seja inversível, é preciso que $\det(A - \lambda\mathbb{1}) = 0$. Contudo, esta equação determina um polinômio de grau N em λ , onde N é a dimensão do espaço de Hilbert em consideração. Pelo Teorema Fundamental da Álgebra, todo polinômio com coeficientes em \mathbb{C} possui raízes em \mathbb{C} , e portando existe algum $\lambda \in \mathbb{C}$ tal que $\det(A - \lambda\mathbb{1}) = 0$. O Lema de Schur então garante que $A - \lambda\mathbb{1} = 0$ e, por consequência, vale que $A = \lambda\mathbb{1}$. \blacksquare

4 Aplicações Físicas

Agora que temos em mãos um bom conhecimento sobre como funciona a teoria de grupos e representações, é instrutivo (e divertido) explorarmos algumas formas em que estes conceitos ocorrem em Física Fundamental.

4.1 O Caminho Óctuplo da Física de Partículas

Uma das consequências do Lema de Schur é a relação entre a existência de simetrias de um sistema quântico e a presença de degenerescências. Lembremos que se \hat{A} é um observável, então os valores possíveis que podem ser medidos para \hat{A} são os números a tais que $\hat{A}|\psi\rangle = a|\psi\rangle$ para algum vetor $|\psi\rangle$ no espaço de Hilbert. É possível que haja vários vetores distintos $|\psi_n\rangle$ para os quais isso vale para o mesmo número a , e assim temos

*“Limitado” significa que se $\|v\| = 1$, então $\|Av\| \leq C$ para alguma constante C . Essencialmente, A é limitado no sentido de não conseguir levar vetores unitários em vetores arbitrariamente grandes. Em dimensão finita, toda transformação linear é limitada, mas em dimensão infinita a distinção entre operadores limitados e ilimitados é relevante.

vários estados fisicamente distintos que retornariam a mesma medição. Nossa discussão relacionando esta noção com simetrias se inspira em [36].

Para ver como isso se relaciona com simetrias, suponha que temos alguma transformação U unitária que representa uma simetria, no sentido de que $U\hat{A}U^\dagger = \hat{A}$, ou seja, \hat{A} é mantido invariante por esta transformação.

Exercício 67:

Mostre que se $U\hat{A}U^\dagger = \hat{A}$ para U unitário, então $[U, \hat{A}] = 0$. \clubsuit

Exercício 68:

Mostre que, dada uma matriz $N \times N$ \hat{A} , $\{U \in U(N); U\hat{A}U^\dagger = \hat{A}\}$ é um grupo. \clubsuit

Exercício 69:

Mostre que se $U\hat{A}U^\dagger = \hat{A}$ para U unitário e $\hat{A}|\psi\rangle = a|\psi\rangle$, então $\hat{A}|\psi'\rangle = a|\psi'\rangle$, onde $|\psi'\rangle = U|\psi\rangle$. \clubsuit

Assim, tanto o estado $|\psi\rangle$ quanto o estado $|\psi'\rangle$ fornecem o mesmo valor quando medimos \hat{A} .

Note que a presença de d estados degenerados linearmente independentes (ou seja, que não podem ser escritos como combinações lineares dos demais) implica a existência de uma representação irredutível d -dimensional do grupo. Se nós soubéssemos qual o grupo G por trás das transformações U , poderíamos descobrir quais são os possíveis valores de d a partir do estudo das representações de G . Alternativamente, se observamos diversas formas de degenerescência num sistema quântico, podemos adivinhar qual o grupo de simetrias escondido com base num raciocínio reverso.

Esta noção possui uma relação íntima com o Lema de Schur: se $U(g)$ é uma representação irredutível unitária de algum grupo G e $U(g)\hat{A}U(g)^\dagger = \hat{A}$, então o Lema de Schur nos está garantindo que $\hat{A} = a\mathbb{1}$ para alguma constante a . Ou seja, a simetria impõe a estrutura da teoria.

Vamos exemplificar isto com alguns “estudos de caso”. Nos baseamos em [6, 15, 32, 33, 35, 36].

Em 1932, foi descoberto o nêutron. Um fato particularmente curioso acerca dele é o valor de sua massa em comparação com o próton. Tem-se[25]

$$m_p = 938.272\,081(6) \text{ MeV}, \quad (4.1)$$

e

$$m_n = 939.565\,413(6) \text{ MeV}, \quad (4.2)$$

onde usamos a unidade convencional para massas em Física de Altas Energias: MeV. Na verdade, esta é uma unidade de energia, mas é convencionalizado que o valor está sendo dividido por c^2 (onde c é a velocidade da luz) para obter uma dimensão de massa. O elétron-volt é definido como a variação na energia de um elétron ao atravessar uma diferença de potencial elétrico de 1 V. Mais especificamente, $1 \text{ eV} = 1.602\,176\,634 \times 10^{-19} \text{ J}$ (não há incerteza, pois as unidades do SI são definidas de modo a deixar algumas constantes fundamentais com valores exatos).

Como m_p e m_n são extremamente próximos, é razoável supor que existe uma simetria aproximada relacionando prótons e nêutrons. Sabemos que não é uma simetria exata, visto que as duas partículas tem cargas elétricas diferentes, e portanto não esperamos que a interação eletromagnética respeite essa simetria. Contudo, esperamos que em energias muito altas, em que a interação eletromagnética é essencialmente desprezível, essa simetria dite a Física do sistema. Em essência, é como um problema de bloquinhos em que sabemos que há atrito, mas escolhemos o desprezar: em algumas situações, a hipótese pode ser absurda, mas nas condições apropriadas podemos obter excelentes resultados com essa aproximação.

Temos uma degenerescência na massa de duas partículas. Podemos entender o próton e o nêutron como dois estados degenerados, o que sugere a existência de uma simetria. Estudos sobre a força de interação entre prótons e nêutrons e entre prótons e prótons reforçaram a ideia de que, a menos de interações eletromagnética, as duas partículas são essencialmente idênticas. Pelos argumentos que fizemos mais cedo, isso significa que o grupo de simetria subjacente deveria possuir uma representação bidimensional. Oras, já sabemos que o grupo $SU(2)$ tem uma representação **2**, então podemos tentar utilizá-lo neste contexto. Temos agora a proposta de que a teoria que descreve as interações entre prótons e nêutrons deveria ser simétrica sob transformações da forma

$$\begin{pmatrix} p \\ n \end{pmatrix} \rightarrow U \begin{pmatrix} p \\ n \end{pmatrix}, \quad (4.3)$$

onde $U \in SU(2)$. Enquanto as interações eletromagnéticas forem desprezíveis, poderíamos então “embaralhar” o que chamamos de próton e o que chamamos de nêutron sem alterar as consequências físicas.

Conhecemos essa simetria como isospin. Perceba que o seu significado é que, enquanto ignorarmos a carga elétrica, as leis físicas não apenas são invariantes sob a troca $p \leftrightarrow n$, mas também por misturas. Por exemplo, como toda matriz ortogonal é unitária, as leis físicas deveriam ser também invariantes pela substituição

$$p \rightarrow \cos \theta p + \sin \theta n, \quad n \rightarrow -\sin \theta p + \cos \theta n \quad (4.4)$$

para qualquer ângulo θ . E este é só um exemplo: qualquer transformação de $SU(2)$ deve preservar as leis físicas enquanto ignorarmos efeitos eletromagnéticos.

Podemos obter uma descrição da carga elétrica. Queremos uma matriz Q tal que $Qp = +p$ e $Qn = 0$, pois assim vale que p e n tem cargas elétricas bem definidas*. Ora, basta tomarmos

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (4.5)$$

Notemos que a carga elétrica de fato quebra a simetria de isospin. Caso Q fosse preservada por isospin, teríamos que $[Q, U] = 0$ para todo $U \in SU(2)$ pela argumentação do início desta seção. Contudo, este não é o caso.

*Estamos seguindo a convenção da Física de Partículas de medir as cargas em unidades da carga elétrica elementar $e = 1.602176634 \times 10^{-19} \text{ C}$.

Exercício 70:

Mostre que $[Q, \tau^i] \neq 0$ para $i = 1, 2$ e disso conclua que há elementos de $SU(2)$ que não comutam com Q . Perceba que $[Q, \tau^3] = 0$, e portanto parte da simetria é mantida ao final do processo. Mostre que essa simetria é, na verdade, uma representação redutível do grupo $U(1)$. Assim, dizemos que a simetria é quebrada de acordo com $SU(2) \rightarrow U(1)$.

Sugestão: note que $Q = \tau^3 + \frac{1}{2}$. \blacktimes

De forma um pouco mais geral, o operador de carga Q é dado por $Q = \tau^3 + \frac{B}{2}$, onde B é o chamado número bariônico, que vale $B = 1$ para prótons e nêutrons e essencialmente conta o número de núcleons. Núcleos atômicos podem ter valores maiores de B . Historicamente, uma das relevâncias disso é que eventualmente descobriu-se a existência dos píons π^+ e π^- , partículas que se transformavam como **1** sob $SO(3)$ e tinham carga ± 1 (para o pión π^\pm). Um dos processos em que os píons eram formados é $p + p \rightarrow p + n + \pi^+$. Se assumirmos que o número bariônico é conservado nesse processo, teríamos para os π^\pm que $B = 0$ e $\tau^3 \pi^\pm = \pm \pi^\pm$. Como isso é possível?

Como na representação fundamental τ^3 é dada por $\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, certamente os píons não podem estar na representação fundamental do grupo de isospin. Contudo, na representação **3** temos o que é preciso.

Exercício 71:

Mostre que as matrizes

$$J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad (4.6)$$

satisfazem as relações de comutação da álgebra $\mathfrak{su}(2)$, e portanto formam uma representação da álgebra. Assim, as exponenciando podemos obter uma representação tridimensional de $SU(2)$. \blacktimes

Perceba que J_3 , que corresponde a τ_3 nessa representação, satisfaz

$$J_3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad J_3 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad J_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \quad (4.7)$$

Assim, podemos acomodar os píons na representação **3** se houver um terceiro pión π^0 . Ele de fato existe. Por completude, comentamos ainda as massas dos píons, que são[25]

$$m_{\pi^\pm} = 139.570\,39(18) \text{ MeV} \quad (4.8)$$

e

$$m_{\pi^0} = 134.9768(5) \text{ MeV}. \quad (4.9)$$

Com o tempo, mais partículas foram descobertas e passaram a dar indícios de outras simetrias. Novos decaimentos sugeriram a existência de uma nova quantidade conservada

pelas interações fortes (como as que ocorrem entre núcleons) e eletromagnéticas, mas não pelas interações fracas (associadas a decaimentos). Esta quantidade foi batizada de estranheza (*strangeness*). É possível organizar mésons e bárions (a grosso modo, partículas “nem tão pesadas” e partículas “pesadas”) em termos de sua estranheza e carga elétrica, como feito nas Figuras 6 a 8 nas páginas 58–59 para parte das partículas conhecidas no início dos anos 1960.

As massas de cada um desses conjuntos é semelhante. Para os mésons pseudoescalares*, temos os píons com massas por volta de† 140 MeV, os káons com massas em torno de 490 MeV e o méson η em torno de 550 MeV.

Para os bárions de spin‡ $\frac{1}{2}$, temos o próton e o nêutron com cerca de 940 MeV, o sistema Σ com 1190 MeV, o sistema Ξ com 1320 MeV e o bárion Λ com 1115 MeV.

Por fim, os bárions de spin $\frac{3}{2}$ são compostos pelo sistema Δ com cerca de 1230 MeV, o sistema Σ^* com 1380 MeV e o sistema Ξ^* com 1530 MeV.

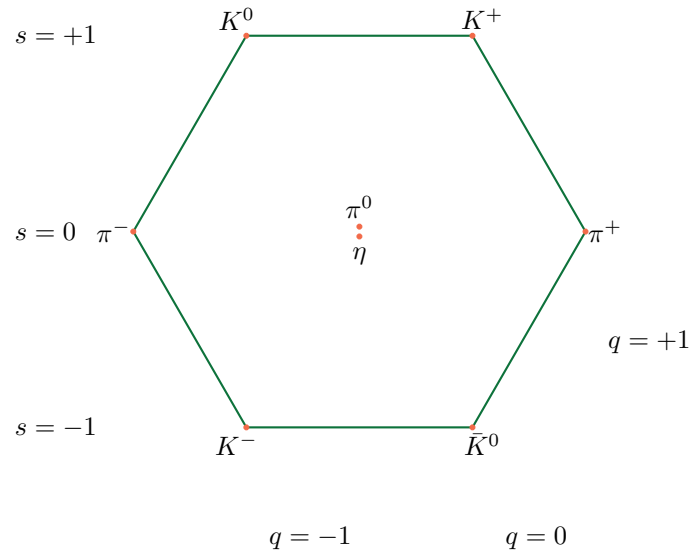


Figura 6: Classificação dos mésons pseudoescalares segundo estranheza e carga.

As massas dentro de cada um desses conjuntos são semelhantes. Embora os quase 500 MeV dos káons pareçam muito distintos dos modestos 140 MeV dos píons, repare que, comparados aos bárions, as massas são de fato parecidas. A grande diferença entre elas apenas indica que a possível simetria subjacente é uma aproximação pior que o isospin com $SU(2)$ que havíamos obtido anteriormente.

Gostariamos então de encontrar um grupo que possua uma representação com oito dimensões, para explicar as massas semelhantes dos mésons escalares. É possível que um

*Pseudoescalar significa que se transformam como **1** sob $SO(3)$ e trocam de sinal sob trocas de paridade.

†As massas que mencionamos para os mésons e bárions advém de [6, 25].

‡Spin $\frac{1}{2}$ significa que se transformam como **2** sob $SO(3)$. Adiante teremos ainda partículas de spin $\frac{3}{2}$, que corresponde à irrep **4** de $SO(3)$.

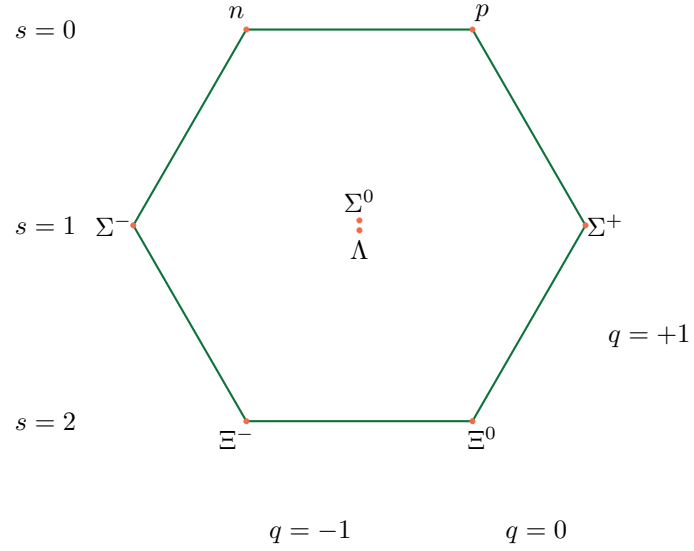


Figura 7: Classificação dos bárions de spin $\frac{1}{2}$ segundo estranheza e carga.

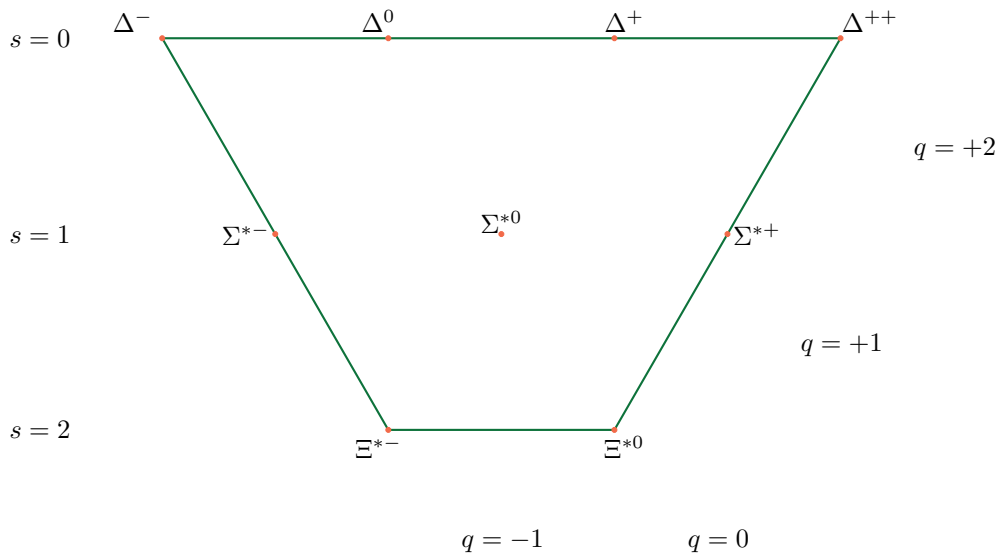


Figura 8: Classificação dos bárions de spin $\frac{3}{2}$ conhecidos no início dos anos 1960 segundo estranheza e carga.

tal grupo também admita outras representações que nos permitam entender as estruturas dos bárions também.

Assim, no início dos anos 1960 Gell-Mann e Ne'eman propuseram, de maneira independente, que o grupo desejado é o grupo $SU(3)$. Vamos considerar essa proposta analisando o produto $\mathbf{3} \otimes \bar{\mathbf{3}}$. Utilizando os métodos tensoriais, temos

$$u^i v_j = \frac{1}{3} \delta^i_j u^k v_k + \left(u^i v_j - \frac{1}{3} \delta^i_j u^k v_k \right), \quad (4.10)$$

e não podemos fazer outras decomposições, pois não há como simetrizar ou anti-simetrizar índices em posições opostas, dado que eles se transformam de modos distintos. Como o traço é descrito por um único número, vemos então que obtivemos a decomposição

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}, \quad (4.11)$$

que inclui uma representação em dimensão 8, exatamente como precisávamos — um octeto, como o da Figura 6 na página 58. Há ainda um singlete, que não parece se encaixar na nossa listagem de partículas, mas não há motivo para se preocupar: o méson η' foi detectado alguns anos após a proposta do caminho óctuplo[20] por Gell-Mann e Ne'eman e corresponde precisamente ao que se esperaria de um singlete de $SU(3)$.

Ainda há mais uma combinação interessante para explorarmos. A saber, $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$.

Exercício 72:

Notando que em $SU(3)$ podemos escrever $u^{[i} v^{j]} = \frac{1}{2} \epsilon^{ijk} \epsilon_{klm} u^l v^m$, mostre que $\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}}$. \boxtimes

4.2 Teorias de Grande Unificação

4.3 Efeitos Hawking e Unruh

Referências

1. Alves, N. A. *Differential Geometry* (Notas para estudo pessoal). https://alves-nickolas.github.io/pdf/Differential_Geometry.pdf.
2. Alves, N. A. *Mecânica Clássica* (notas de aula). https://alves-nickolas.github.io/pdf/Mec%C3%A2nica_Cl%C3%A1ssica.pdf.
3. Alves, N. A. *O Grupo de Lorentz: Relatividade Restrita aos olhos de um matemático* (notas de aula). https://alves-nickolas.github.io/pdf/O_Grupo_de_Lorentz.pdf.
4. Arnold, V. I. *Mathematical Methods of Classical Mechanics* (Springer, New York, 1989).
5. Barata, J. C. A. *Notas de Física-Matemática* http://denebola.if.usp.br/%20jbarata/Notas_de_aula/notas_de_aula.html.
6. Bertuzzo, E. *Lectures on Particle Physics* http://fmatrm.if.usp.br/~enrico/SM/Particle_Physics_Book.pdf (2021).

7. Carroll, S. *Spacetime and Geometry: An Introduction to General Relativity* (Pearson, Harlow, 2014).
8. Coelho, F. U. & Lourenço, M. L. *Um Curso de Álgebra Linear* (Editora da Universidade de São Paulo, São Paulo, 2013).
9. Dirac, P. A. M. *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1967).
10. Georgi, H. *Lie Algebras in Particle Physics* (CRC Press, Boca Raton, 2018).
11. Germano, G. R. *Representações Irredutíveis Unitárias Do Grupo de Poincaré* Master's thesis (University of São Paulo, São Paulo, 2016). DOI: [10.11606/D.43.2016.tde-08122016-160042](https://doi.org/10.11606/D.43.2016.tde-08122016-160042).
12. Geroch, R. *Mathematical Physics* (University of Chicago Press, Chicago, 1985).
13. Griffiths, D. J. *Introduction to Electrodynamics* (Cambridge University Press, Cambridge, 2017).
14. Hall, B. C. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction* (Springer, Cham, 2003).
15. Halzen, F. & Martin, A. D. *Quarks and Leptons: An Introductory Course in Modern Particle Physics* Includes index (Wiley, New York, 1984).
16. Hamermesh, M. *Group Theory and Its Application to Physical Problems* Originally published: Reading, Mass. : Addison-Wesley, 1962 (Dover Publications, New York, 1989).
17. Hawking, S. W. & Ellis, G. F. R. *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, 1973).
18. Hoffman, K. & Kunze, R. *Linear Algebra* (Prentice-Hall, Inc., Englewood Cliffs, 1971).
19. José, J. V. & Saletan, E. J. *Classical Dynamics: A Contemporary Approach* (Cambridge University Press, Cambridge, 1998).
20. Kupść, A. *What Is Interesting in η and η' Meson Decays?* em *AIP Conference Proceedings* MESON PHYSICS AT COSY-11 AND WASA-AT COSY. **950** (AIP, 2007), 165–179. DOI: [10.1063/1.2819029](https://doi.org/10.1063/1.2819029). arXiv: [0709.0603](https://arxiv.org/abs/0709.0603) [[nucl-ex](#)].
21. Landau, L. D. & Lifshitz, E. M. *Mechanics* (Butterworth-Heinemann, Oxford, 1976).
22. Le Bellac, M. *Quantum Physics* (Cambridge University Press, Cambridge, 2006).
23. Lévy-Leblond, J.-M. em *Group Theory and its Applications Vol. II* (ed. Loeb, E.) 221–299 (Academic Press, New York, 1971).
24. Misner, C. W., Thorne, K. S. & Wheeler, J. A. *Gravitation* (Freeman, San Francisco, 1973).
25. Particle Data Group *et al.* Review of Particle Physics. *Progress of Theoretical and Experimental Physics* **2020**. 083C01, 083C01 (14 de ago. de 2020). DOI: [10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104).

26. Peskin, M. E. & Schroeder, D. V. *An Introduction to Quantum Field Theory* (Westview Press, Boulder, 1995).
27. Sakurai, J. J. & Napolitano, J. *Modern Quantum Mechanics* (Cambridge University Press, Cambridge, 2017).
28. Schwartz, M. D. *Quantum Field Theory and the Standard Model* (Cambridge University Press, Cambridge, 2014).
29. Shankar, R. *Principles of Quantum Mechanics* (Springer, New York, 1994).
30. Sundermeyer, K. *Symmetries in Fundamental Physics* (Springer, Cham, 2014).
31. Wald, R. M. *General Relativity* (University of Chicago Press, Chicago, 1984).
32. Weinberg, S. *Foundations of Modern Physics* (Cambridge University Press, Cambridge, 2021).
33. Weinberg, S. *Lectures on Quantum Mechanics* (Cambridge University Press, Cambridge, 2015).
34. Weinberg, S. *The Quantum Theory of Fields Vol. I: Foundations* (Cambridge University Press, Cambridge, 1995).
35. Weinberg, S. *The Quantum Theory of Fields Vol. II: Modern Applications* (Cambridge University Press, Cambridge, 1996).
36. Zee, A. *Group Theory in a Nutshell for Physicists* (Princeton University Press, Princeton, 2016).
37. Zee, A. *Quantum Field Theory in a Nutshell* (Princeton University Press, Princeton, 2010).