

Capítulo 21

Grupos. Alguns Exemplos

Conteúdo

21.1	O Grupo de Permutações	1015
21.2	O Grupo de Permutações de n Elementos	1016
21.2.1	Ciclos, Transposições e Transposições Elementares	1016
21.2.1.1	O Sinal, ou Paridade, de uma Permutação. O Símbolo de Levi-Civita	1020
21.3	Alguns Grupos Matriciais	1022
21.3.1	Grupos Lineares e Grupos Lineares Especiais	1022
21.3.1.1	Grupos Lineares Projetivos	1025
21.3.2	O Grupo de Borel e o Grupo de Heisenberg	1027
21.3.2.1	O Grupo de Heisenberg	1027
21.3.3	Grupos Associados a Formas Bilineares e Sesquilineares	1034
21.3.3.1	Os Grupos Ortogonais	1038
21.3.3.2	Os Grupos Unitários	1040
21.3.3.3	Os Grupos Simpléticos	1041
21.4	Os Grupos $SO(2)$, $SO(3)$, $SU(2)$ e $SL(2, \mathbb{C})$	1049
21.4.1	Os Grupos $SO(2)$, $O(2)$, $SO(1, 1)$ e $O(1, 1)$	1049
21.4.2	O Grupo $SO(3)$	1053
21.4.2.1	Mais Propriedades das Matrizes de $SO(3)$	1061
21.4.2.2	$SO(3)$ e os Ângulos de Euler	1064
21.4.2.3	A Parametrização de Cayley de $SO(3)$ (e de $SO(n)$)	1069
21.4.3	O Grupo $O(3)$	1072
21.4.4	O Grupo $SU(2)$	1076
21.4.5	A Relação Entre $SO(3)$ e $SU(2)$	1081
21.4.6	O Grupo $SL(2, \mathbb{C})$	1085
21.5	Generalidades Sobre os Grupos $SU(n)$ e $SO(n)$	1087
21.5.1	Os Grupos $SU(n)$	1087
21.5.1.1	Um Pouco Sobre o Grupo $SU(3)$	1089
21.5.2	Os Grupos $SO(n)$	1091
21.6	O Grupo Afim e o Grupo Euclidiano	1095
21.7	O Grupo de Lorentz em $3 + 1$-Dimensões	1099
21.7.1	O Espaço-Tempo, a Noção de Intervalo e a Estrutura Causal	1099
21.7.2	A Invariância do Intervalo	1105
21.7.3	O Grupo de Lorentz	1108
21.7.4	Alguns Subgrupos do Grupo de Lorentz	1109
21.7.5	Alguns Fatos Sobre a Estrutura do Grupo de Lorentz	1112
21.7.6	Os Geradores Infinitesimais do Grupo de Lorentz	1116
21.7.6.1	Fórmula de Rodrigues para <i>Boosts</i> de Lorentz	1122
21.7.7	O Grupo de Galilei	1124
21.7.7.1	Comparação Entre os Grupos de Galilei e Lorentz. Contrações	1126
21.8	O Grupo de Poincaré	1129
21.9	Mais Sobre Grupos Simpléticos	1133
21.9.1	O Subgrupo Simplético Real Ortogonal	1134
21.9.2	Grupos Simpléticos. Álgebras de Lie e Parametrização de Cayley	1138
21.9.3	O Teorema de Williamson	1141
21.10	Operadores Diferenciais como Geradores Infinitesimais sobre Funções	1142
21.11	Exercícios Adicionais	1147

APÊNDICES 1149

21.A Extensão do Lema 21.3 e do Teorema 21.7 ao Caso Complexo 1149

21.B Prova do Teorema 21.11 1151



GRUPOS são objetos de suma importância na Física devido à sua relação com transformações de simetria¹. A noção abstrata de grupo foi introduzida na Seção 2.1.3, página 88. No presente capítulo apresentaremos alguns grupos de particular interesse na Física e na Matemática e estudaremos algumas de suas propriedades mais simples e importantes. Com particular detalhe trataremos do grupo de Lorentz na Seção 21.7, página 1099, grupo esse de fundamental importância na Teoria da Relatividade. No Capítulo 44, página 2373, apresentaremos diversos desenvolvimentos básicos da Mecânica Clássica, notadamente da teoria dos corpos rígidos, sob a luz da Teoria de Grupos, especialmente do grupo de rotações SO(3), o qual é estudado com certo detalhe na Seção 21.4.2, página 1053. O grupo de permutações de um conjunto finito, muito importante na Física Quântica, particularmente no estudo de sistemas atômicos e nucleares, é tratado na Seção 21.2, página 1016.

21.1 O Grupo de Permutações

Seja C um conjunto não vazio qualquer e seja $\text{Perm}(C)$ o conjunto de todas as funções bijetoras de C em C . $\text{Perm}(C)$ é naturalmente um grupo, onde o produto é a composição de funções (que é associativa) e o elemento neutro é a função identidade (que denotaremos doravante por id). O elemento inverso de uma função $f \in \text{Perm}(C)$ é a sua função inversa f^{-1} (que existe, por definição, pois todos os elementos de $\text{Perm}(C)$ são funções bijetoras). O grupo $\text{Perm}(C)$ é denominado *grupo de permutações* do conjunto C .

E. 21.1 *Exercício*. Mostre que $\text{Perm}(C)$ é um grupo Abelianiano se e somente se C possuir um ou dois elementos. ✦

De particular importância é o caso em que C é um conjunto finito. Tais grupos de permutação e suas representações também desempenham um papel de destaque na Física, particularmente na Mecânica Quântica, e por isso vamos nos deter um pouco nos mesmos logo adiante. Antes disso, mencionemos um resultado geral relevante, o chamado Teorema de Cayley.

• O Teorema de Cayley

Grupos de permutações desempenham um papel de destaque na Teoria de Grupos, em parte devido ao seguinte teorema estrutural, denominado *Teorema de Cayley*²:

Teorema 21.1 (Teorema de Cayley) *Seja G um grupo. Então, G é isomorfo a um subgrupo de um grupo de permutações, a saber, a um subgrupo de $\text{Perm}(G)$.* □

Prova. Para $g \in G$, defina-se a função $L_g : G \rightarrow G$ por $L_g(x) := gx, x \in G$. Afirmamos que $L_g : G \rightarrow G$ é bijetora. De fato, se $y \in G$, então $y = L_g(g^{-1}y)$, o que demonstra que L_g é sobrejetora. Fora isso, se $x, x' \in G$ são tais que $L_g(x) = L_g(x')$, então vale $gx = gx'$, o que implica $x = x'$, demonstrando que L_g é injetora e, portanto, bijetora.

Afirmamos que o conjunto de tais funções $G^L \equiv \{L_g, g \in G\}$, que é um subconjunto de $\text{Perm}(G)$, é um grupo por composição e, portanto, um subgrupo de $\text{Perm}(G)$. De fato, para $g, h \in G$ e para cada $x \in G$ vale $L_g \circ L_h(x) = L_g(L_h(x)) = ghx = L_{gh}(x)$, o que estabelece que a composição é uma operação em G^L (que é naturalmente associativa, pois a composição de funções é associativa) e estabelece a regra

$$L_g \circ L_h = L_{gh}, \quad \text{para todos } g, h \in G. \tag{21.1}$$

Essa operação possui um elemento neutro, a saber L_e , com $e \in G$ sendo o elemento neutro de G . De fato, pela regra (21.1), vale $L_e \circ L_g = L_g \circ L_e = L_g$ para cada $g \in G$. Finalmente, L_g possui uma inversa por composição, a saber, $L_{g^{-1}}$, como é evidente também por (21.1).

¹Para um texto acessível sobre a história do desenvolvimento da noção de simetria na Física e na Matemática, recomendamos [368].

²Arthur Cayley (1821–1895).

Assim, estabelecemos que G^L é um subgrupo de $\text{Perm}(G)$. Resta estabelecer que G e G^L são isomorfos. O isomorfismo é dado simplesmente por $L : G \ni g \mapsto L_g \in G^L$. Essa aplicação é obviamente sobrejetora e a regra (21.1) mostra que se trata de um homomorfismo de grupos. Além disso, é injetora, pois se $L_g = L_h$ para $g, h \in G$, temos $gx = hx$ para todo $x \in G$, o que implica $g = h$. ■

E. 21.2 *Exercício.* Para cada $g \in G$, considere a função $R_g : G \rightarrow G$ definida por $R_g(x) := xg^{-1}$, $x \in G$. Mostre que cada R_g é bijetora e que $G^R \equiv \{R_g, g \in G\}$ é um subgrupo de $\text{Perm}(G)$ e que esse subgrupo também é isomorfo a G . ✦

21.2 O Grupo de Permutações de n Elementos

Seja $n \geq 1$, inteiro, e considere-se o conjunto $\{1, \dots, n\}$. O grupo $S_n = \text{Perm}(\{1, \dots, n\})$ é denominado *grupo de permutações de n elementos*.

E. 21.3 *Exercício.* Seja C um conjunto com n elementos. Mostre que $\text{Perm}(C)$ é isomorfo a S_n . ✦

E. 21.4 *Exercício.* Mostre que S_n tem exatamente $n!$ elementos. *Sugestão:* indução. ✦

Um elemento $\pi \in S_n$ é dito ser uma *permutação*. Como toda a permutação, π é uma função bijetora $\{1, \dots, n\} \rightarrow \{1, \dots, n\}$ e é costume representá-la na forma de um arranjo de tipo matricial:

$$\pi \equiv \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix},$$

onde na primeira linha ordenamos os elementos de $\{1, \dots, n\}$ e na segunda suas imagens por π .

Exemplos 21.1 Os elementos de S_2 são

$$\pi_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \quad \text{e} \quad \pi_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

π_1 é a identidade do grupo.

Os elementos de S_3 são

$$\begin{aligned} \pi_1 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, & \pi_2 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, & \pi_3 &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \\ \pi_4 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, & \pi_5 &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, & \pi_6 &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}. \end{aligned}$$

π_1 é a identidade do grupo. □

21.2.1 Ciclos, Transposições e Transposições Elementares

Vamos aqui estudar alguns fatos estruturais importantes sobre os grupos S_n .

- **Ciclos**

Precisamos da seguinte definição:

Definição. Uma permutação $\pi \in S_n$ é dita ser um *ciclo*, ou um *r-ciclo*, se existirem r elementos distintos i_1, \dots, i_r em $\{1, \dots, n\}$ tais que

$$\pi(j) = \begin{cases} j, & \text{se } j \notin \{i_1, \dots, i_r\}, \\ i_{a+1}, & \text{se } j = i_a, \text{ mas } a \neq r, \\ i_1, & \text{se } j = i_r. \end{cases}$$



E. 21.5 Exercício. Mostre que se π é um r -ciclo, então $\pi^r = \text{id}$.



A importância do conceito de ciclo manifesta-se no seguinte teorema:

Teorema 21.2 Toda permutação diferente da identidade é um produto de ciclos disjuntos dois a dois. □

Prova. Seja $\pi \in S_n, \pi \neq \text{id}$. Seja i_1 o menor elemento de $\{1, \dots, n\}$ para o qual $\pi(i) \neq i$. Vamos considerar a sequência (em princípio infinita)

$$i_1, \quad \pi(i_1), \quad \pi^2(i_1), \quad \pi^3(i_1), \dots$$

Os elementos dessa sequência são obviamente elementos de $\{1, \dots, n\}$ que é um conjunto finito. Consequentemente essa sequência tem, na verdade, elementos repetidos. Vamos supor que $\pi^p(i_1)$ e $\pi^q(i_1)$, $p < q$, sejam os primeiros elementos que se repetem: $\pi^p(i_1) = \pi^q(i_1)$. Essa igualdade implicaria $i_1 = \pi^{r_1}(i_1)$, onde $r_1 = q - p$. Assim, o primeiro par que se repete na sequência acima é, em verdade, o par i_1 e $\pi^{r_1}(i_1)$.

Isso nos diz que a sequência acima é uma repetição infinita da sequência finita

$$i_1, \quad \pi(i_1), \quad \pi^2(i_1), \quad \dots, \quad \pi^{r_1}(i_1),$$

sequência esta formada por r_1 elementos que, por construção, são distintos.

Vamos denominar

$$i_1, \quad i_2 := \pi(i_1), \quad i_3 = \pi^2(i_1), \quad \dots, \quad i_{r_1} = \pi^{r_1}(i_1)$$

e definir $\pi_1 \in S_n$ por

$$\pi_1(j) = \begin{cases} j, & \text{se } j \notin \{i_1, \dots, i_{r_1}\}, \\ i_{a+1} = \pi^a(i_1), & \text{se } j = i_a, \text{ mas } a \neq r_1, \\ i_1, & \text{se } j = i_{r_1}. \end{cases}$$

É evidente que π_1 é um ciclo e que π_1 e π coincidem no conjunto $\{i_1, \dots, i_{r_1}\}$. Podemos então escrever

$$\pi = \pi_1 \pi' = \pi' \pi_1,$$

onde $\pi' \in S_n$ é a identidade em $\{i_1, \dots, i_{r_1}\}$ e coincide com π no complemento:

$$\pi'(j) = \begin{cases} j, & \text{se } j \in \{i_1, \dots, i_{r_1}\}, \\ \pi(j), & \text{de outra forma.} \end{cases}$$

O que fazemos em seguida é repetir o procedimento, mas agora para a permutação π' . Obteremos $\pi' = \pi_2 \pi'' = \pi'' \pi_2$, onde π_2 é novamente um ciclo (disjunto de π_1 , por construção). Como $\{1, \dots, n\}$ é um conjunto finito, a repetição desse procedimento deve ter um fim, e obtemos

$$\pi = \pi_1 \pi_2 \dots \pi_k$$

para k ciclos π_1, \dots, π_k disjuntos dois a dois. Isso completa a prova. ■

• **Transposições**

2-ciclos são denominados *transposições*. Sejam p e q dois elementos distintos de $\{1, \dots, n\}$. A transposição de p e q , denotada por $t_{p,q}$ é a permutação definida por

$$t_{p,q}(j) = \begin{cases} j, & \text{se } j \neq p \text{ e } j \neq q, \\ q, & \text{se } j = p, \\ p, & \text{se } j = q. \end{cases}$$

Observe-se que $t_{p,q} = t_{q,p}$.

Transposições são importantes pela seguinte razão:

Teorema 21.3 *Todo ciclo pode ser escrito como um produto de transposições.* □

Prova. Seja π o ciclo associado ao conjunto $\{i_1, \dots, i_r\} \subset \{1, \dots, n\}$:

$$\pi(j) = \begin{cases} j, & \text{se } j \notin \{i_1, \dots, i_r\}, \\ i_{a+1}, & \text{se } j = i_a, \text{ mas } a \neq r, \\ i_1, & \text{se } j = i_r. \end{cases}$$

A prova resume-se na constatação que

$$\pi = t_{i_{r-1}, i_r} t_{i_{r-2}, i_{r-1}} \dots t_{i_2, i_3} t_{i_1, i_2}.$$
■

E. 21.6 *Exercício.* Complete os detalhes e/ou faça alguns casos particulares para convencer-se. ✦

O seguinte teorema é um corolário imediato dos Teoremas 21.2 e 21.3:

Teorema 21.4 *Toda permutação diferente da identidade é um produto de transposições.* □

• **Transposições elementares**

De particular importância são as transposições de vizinhos $t_i = t_{i, i+1}$ com $i = 1, \dots, n - 1$:

$$t_i(j) = \begin{cases} j, & \text{se } j \neq i \text{ e } j \neq i + 1, \\ i + 1, & \text{se } j = i, \\ i, & \text{se } j = i + 1. \end{cases}$$

e que são chamadas *transposições elementares*. A importância das mesmas reside nos dois teoremas abaixo.

Teorema 21.5 *Toda transposição é um produto de transposições elementares.* □

Prova. Seja $t_{p,q}$ uma transposição com $p < q$. A prova resume-se na constatação que

$$t_{p,q} = t_{q-1,q} \cdots t_{p+1,p+2} t_{p,p+1} t_{p+1,p+2} \cdots t_{q-1,q} = t_{q-1} \cdots t_{p+1} t_p t_{p+1} \cdots t_{q-1}.$$

■

E. 21.7 *Exercício.* Complete os detalhes e/ou faça alguns casos particulares para convencer-se. ✦

O seguinte teorema é um corolário imediato dos Teoremas 21.2, 21.3, 21.4 e 21.5:

Teorema 21.6 *Toda permutação diferente da identidade é um produto de transposições elementares.* □

E. 21.8 *Exercício.* Determine quais dos elementos π_1, \dots, π_6 do grupo S_3 (página 1016) são transposições elementares e escreva os demais como produtos de tais transposições elementares. ✦

O Teorema 21.6 afirma que S_n é um grupo gerado por transposições elementares, ou seja, todo $\pi \in S_n$ (distinto da identidade) é da forma

$$\pi = t_{i_1} \cdots t_{i_k}, \tag{21.2}$$

para certas transposições elementares t_{i_1}, \dots, t_{i_k} . Podemos nos perguntar, essa forma de escrever π é única? A resposta é não, pelas razões que agora expomos (Proposição 21.1, logo abaixo).

• Transposições elementares e suas relações

Proposição 21.1 *Seja $n \in \mathbb{N}$ com $n \geq 3$. Em S_n as transposições elementares $t_i, i = 1, \dots, n - 1$ satisfazem as seguintes relações:*

$$(t_i)^2 = \text{id}, \tag{21.3}$$

$$t_i t_j = t_j t_i, \quad \text{se } |i - j| \geq 2, \tag{21.4}$$

$$t_i t_{i+1} t_i = t_{i+1} t_i t_{i+1}, \quad \text{se } i = 1, \dots, n - 2. \tag{21.5}$$

□

Prova. \longleftrightarrow *Exercício.* Basta verificar as igualdades seguindo as definições. ■

Essa proposição explica por que a representação (21.2) não é geralmente única: o lado direito de (21.2) pode eventualmente ser reescrito se aplicarmos quaisquer das relações (21.3)-(21.5). Estas, porém, são as únicas relações que as transposições elementares t_i satisfazem. Desses fatos extraímos a seguinte conclusão:

Proposição 21.2 *Todo grupo gerado por $n - 1$ elementos t_1, \dots, t_{n-1} que satisfaçam as relações (21.3)-(21.5) (e somente elas) é isomorfo a S_n .* □

Prova. \longleftrightarrow *Exercício.*

O seguinte resultado será usado adiante.

Lema 21.1 *Para cada $i = 1, \dots, n - 1$ existe ao menos um $\pi_i \in S_n$ (dependente de i) tal que $t_i = \pi_i t_1 \pi_i^{-1}$. Segue disso que para quaisquer $i, j \in \{1, \dots, n - 1\}$ existe ao menos um $\pi_{ij} \in S_n$ (dependente de i e de j) tal que $t_i = \pi_{ij} t_j \pi_{ij}^{-1}$. □*

Prova. Se $i = 1$ podemos tomar $\pi_1 = \text{id}$ e não há o que se demonstrar. Consideremos $i > 1$. Escolhamos π_i de sorte que $\pi_i(1) = i$ e $\pi_i(2) = i + 1$. Teremos, por um lado,

$$t_i(\pi_i(j)) = \begin{cases} t_i(i) = i + 1, & j = 1, \\ t_i(i + 1) = i, & j = 2, \\ \pi_i(j), & \text{demais casos.} \end{cases} \quad \text{Por outro lado,} \quad \pi_i(t_1(j)) = \begin{cases} \pi_i(2) = i + 1, & j = 1, \\ \pi_i(1) = i, & j = 2, \\ \pi_i(j), & \text{demais casos.} \end{cases}$$

Isso exhibe o fato que $t_i\pi_i = \pi_it_1$, provando a primeira afirmação.

Segue disso que para quaisquer $i, j \in \{1, \dots, n - 1\}$ temos $t_i = \pi_it_1\pi_i^{-1}$ e $t_j = \pi_jt_1\pi_j^{-1}$. Assim, $t_i = \pi_i\pi_j^{-1}t_j\pi_j\pi_i^{-1}$, ou seja, $t_i = \pi_{ij}t_j\pi_{ij}^{-1}$ com $\pi_{ij} := \pi_i\pi_j^{-1}$. ■

• **O grupo de tranças**

As relações (21.4)–(21.5) inspiram a definição de um outro grupo, aparentado ao grupo S_n , e que é denominado *grupo de n tranças* e denotado por B_n (do inglês *braid* = trança). Este é, por definição, o grupo gerado por $n - 1$ elementos b_1, \dots, b_{n-1} que satisfaçam as relações

$$b_ib_j = b_jb_i, \quad \text{se } |i - j| \geq 2, \tag{21.6}$$

$$b_ib_{i+1}b_i = b_{i+1}b_ib_{i+1}, \quad \text{se } i = 1, \dots, n - 2, \tag{21.7}$$

de tal forma que para todo $\beta \in B_n$ existem $\{b_{i_1}, \dots, b_{i_k}\} \subset \{b_1, \dots, b_{n-1}\}$ e números inteiros $n_1, \dots, n_k \in \mathbb{Z}$ tais que

$$\beta = (b_{i_1})^{n_1} \dots (b_{i_k})^{n_k}.$$

Para $n > 1$ os grupos B_n são não Abelianos.

Note-se que as relações (21.6)–(21.7) são formalmente idênticas às relações (21.4)–(21.5), mas a relação (21.3) não tem análogo em B_n , ou seja, ao contrário do que ocorre em S_n , os elementos b_i não têm a si mesmos como inversa. Por essa razão, elementos como $(b_i)^n$ para n 's diferentes são todos distintos entre si. Assim, ao contrário de S_n , o grupo B_n tem um número infinito de elementos, apesar de ter um número finito de geradores.

E. 21.9 Exercício. Seja $p : \mathbb{Z} \rightarrow \{0, 1\}$ definida por $p(n) = 0$ se n for par e $p(n) = 1$ se n for ímpar. Mostre que $\phi : B_n \rightarrow S_n$ definido por $\phi((b_{i_1})^{n_1} \dots (b_{i_k})^{n_k}) = t_{i_1}^{p(n_1)} \dots t_{i_k}^{p(n_k)}$ é um homomorfismo. ✦

O grupo de tranças foi inventado pelo matemático E. Artin³ em 1925 e desempenha um papel importante na chamada *Teoria dos Nós*, um rico capítulo do estudo das propriedades topológicas do espaço tridimensional. Nesse contexto os elementos b_i têm uma interpretação interessante em termos de transposições elementares de tranças (barbantes) no espaço tridimensional. Por falta de espaço e habilidade em apresentar as figuras correspondentes, não entraremos em mais detalhes aqui e remetemos o estudante à leitura de [212], por exemplo. No final dos anos 80 e nos anos 90 do Século XX encontrou-se aplicações dos grupos de tranças na Física, no contexto das Teorias Quânticas de Campos em dimensões 2 e 3, assim como na Física dos Materiais (problema da Supercondutividade a altas temperaturas).

21.2.1.1 O Sinal, ou Paridade, de uma Permutação. O Símbolo de Levi-Civita

Vamos agora apresentar a noção de sinal, ou paridade, de uma permutação. Essa noção é empregada em vários lugares e é relevante na definição dos chamados *símbolos de Levi-Civita*, que apresentaremos adiante, e que são empregados no Cálculo Vetorial e na Geometria Diferencial.

Se $m \in \mathbb{Z}$ é um número inteiro, definimos a paridade de m por $(-1)^m$. Assim, dois números $m_1, m_2 \in \mathbb{Z}$ têm a mesma paridade se ambos forem pares ou se ambos forem ímpares.

³Emil Artin (1889–1962).

Definimos o *polinômio de Vandermonde*⁴ de n variáveis reais x_1, \dots, x_n , sendo $n \geq 2$, por

$$V_n(x_1, \dots, x_n) := \prod_{1 \leq i < j \leq n} (x_i - x_j) = \prod_{i=1}^{n-1} \prod_{j=1+1}^n (x_i - x_j). \tag{21.8}$$

Assim, tem-se, por exemplo,

$$V_2(x_1, x_2) = (x_1 - x_2),$$

$$V_3(x_1, x_2, x_3) = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3),$$

$$V_4(x_1, x_2, x_3, x_4) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)(x_2 - x_3)(x_2 - x_4)(x_3 - x_4) \text{ etc.}$$

Observe-se que para cada par $i, j \in \{1, \dots, n\}$ com $i < j$ o fator $x_i - x_j$ ou $x_j - x_i$ ocorre precisamente uma vez em $V_n(x_1, \dots, x_n)$.

O polinômio V_n é um polinômio de grau $n - 1$ em cada uma de suas variáveis.

Afirmamos que para $\pi \in S_n$, o polinômio $V_n(x_{\pi(1)}, \dots, x_{\pi(n)})$, definido nas mesmas variáveis x_1, \dots, x_n , difere do polinômio $V_n(x_1, \dots, x_n)$ apenas por um sinal. De fato, como π é uma bijeção, ocorrem também em $V_n(x_{\pi(1)}, \dots, x_{\pi(n)})$ todos os fatores $x_i - x_j$, mas alguns deles podem estar em ordem invertida.

O sinal com que $V_n(x_{\pi(1)}, \dots, x_{\pi(n)})$ e $V_n(x_1, \dots, x_n)$ diferem depende apenas de π , sendo denominado *sinal*, ou *paridade*, da permutação π e denotado por $(-1)^\pi$ ou por $\text{sinal}(\pi)$. Em símbolos,

$$V_n(x_{\pi(1)}, \dots, x_{\pi(n)}) = \text{sinal}(\pi)V_n(x_1, \dots, x_n),$$

com $\text{sinal}(\pi) = \pm 1$. É evidente que $\text{sinal}(\text{id}) = 1$.

O seguinte fato é relevante:

Lema 21.2 Para todos $\pi, \pi' \in S_n$ vale

$$\text{sinal}(\pi \circ \pi') = \text{sinal}(\pi)\text{sinal}(\pi'), \tag{21.9}$$

o que mostra que a função *sinal* é um homomorfismo de S_n em \mathbb{Z}_2 .

Segue de (21.9) que $S_n^+ := \{\pi \in S_n \mid \text{sinal}(\pi) = +1\}$ é um subgrupo normal de S_n . □

Prova. Vamos assumir que as variáveis x_1, \dots, x_n sejam diferentes entre si. Com isso, podemos escrever

$$\text{sinal}(\pi) = \frac{V_n(x_{\pi(1)}, \dots, x_{\pi(n)})}{V_n(x_1, \dots, x_n)}$$

para todo $\pi \in S_n$. Logo,

$$\text{sinal}(\pi \circ \pi') = \frac{V_n(x_{\pi(\pi'(1))}, \dots, x_{\pi(\pi'(n))})}{V_n(x_1, \dots, x_n)} = \underbrace{\frac{V_n(x_{\pi(\pi'(1))}, \dots, x_{\pi(\pi'(n))})}{V_n(x_{\pi'(1)}, \dots, x_{\pi'(n)})}}_{=\text{sinal}(\pi)} \underbrace{\frac{V_n(x_{\pi'(1)}, \dots, x_{\pi'(n)})}{V_n(x_1, \dots, x_n)}}_{=\text{sinal}(\pi')},$$

provando (21.9).

De (21.9) segue imediatamente que $\text{sinal}(\pi) = \text{sinal}(\pi^{-1})$. Com isso, fica claro que S_n^+ é um subgrupo de S_n . Que é um subgrupo normal segue da observação, também por (21.9), que

$$\text{sinal}(\pi\pi'\pi^{-1}) = \text{sinal}(\pi)\text{sinal}(\pi')\text{sinal}(\pi^{-1}) = \text{sinal}(\pi'),$$

o que implica que se $\pi' \in S_n^+$, então $\pi\pi'\pi^{-1}$ também é elemento de S_n^+ para todo $\pi \in S_n$, provando que S_n^+ é normal. ■

O subgrupo S_n^+ é também denominado *subgrupo alternante de grau n* .

⁴Alexandre-Théophile Vandermonde (1735–1796).

E. 21.10 *Exercício*. Já mencionamos que S_n tem $n!$ elementos. Mostre que o grupo S_n^+ tem $n!/2$ elementos. ✦

Se $t_k, k = 1, \dots, n - 1$, é uma transposição elementar de S_n , afirmamos que

$$\text{sinal}(t_k) = -1. \tag{21.10}$$

Pelo Lema 21.1, página 1019, é suficiente provar que $\text{sinal}(t_1) = -1$. Agora,

$$V_n(x_{t_1(1)}, x_{t_1(2)}, x_{t_1(3)}, \dots, x_{t_1(n)}) = V_n(x_2, x_1, x_3, \dots, x_n)$$

que difere de $V_n(x_1, x_2, x_3, \dots, x_n)$ apenas na substituição do fator $x_1 - x_2$ por $x_2 - x_1$ (justifique!), mostrando que $\text{sinal}(t_1) = -1$.

Vamos supor que $\pi \in S_n$ possa ser representado com um produto de k transposições elementares: $\pi = t_{i_1} \cdots t_{i_k}$. Segue de (21.9) e de (21.10) que

$$\text{sinal}(\pi) = \begin{cases} 1, & \text{se } \pi = \text{id} , \\ (-1)^k, & \text{se } \pi \neq \text{id} , \pi = t_{i_1} \cdots t_{i_k} . \end{cases}$$

Um corolário evidente dessa relação é que, ainda que uma permutação π possa ser representada de diversas formas como produto de transposições elementares (usando-se, por exemplo, as relações (21.3)-(21.5)) a paridade do número de transposições elementares necessárias para se obter π é sempre a mesma.

E. 21.11 *Exercício*. Determine o sinal das permutações π_1, \dots, π_6 do grupo S_3 dadas nos Exemplos 21.1, página 1016. ✦

• **O símbolo de Levi-Civita**

Seja π uma função de $\{1, \dots, n\}$ em si mesmo. O *símbolo de Levi-Civita*⁵ associado à função π , denotado por ε_π ou por $\varepsilon_{\pi(1), \dots, \pi(n)}$, é definido por

$$\varepsilon_{\pi(1), \dots, \pi(n)} := \begin{cases} \text{sinal}(\pi) , & \text{se } \pi \text{ for uma bijeção,} \\ 0 , & \text{se } \pi \text{ não for uma bijeção.} \end{cases} \tag{21.11}$$

Como π é suposto ser uma função de $\{1, \dots, n\}$ em si mesmo, ela não será uma bijeção se não for injetora, ou seja, se existirem i, j distintos em $\{1, \dots, n\}$ tais que $\pi(i) = \pi(j)$.

Os símbolos de Levi-Civita para o caso $n = 3$ são empregados no Cálculo Vetorial, por exemplo, na definição do chamado produto vetorial em \mathbb{R}^3 , e são usados para expressar operadores diferenciais, como o rotacional. Os símbolos de Levi-Civita do caso $n = 3$ e suas propriedades são estudados no Capítulo 4, página 258. Os símbolos de Levi-Civita também são empregados na Geometria Diferencial, no estudo das chamadas *formas diferenciais*. Naquele contexto dedicamos a eles o Apêndice 35.A, página 1762.

21.3 Alguns Grupos Matriciais

21.3.1 Grupos Lineares e Grupos Lineares Especiais

Vamos denotar por $\text{Mat}(n, \mathbb{R})$ ou $\text{Mat}(\mathbb{R}, n)$ o conjunto de todas as matrizes reais $n \times n$ e por $\text{Mat}(n, \mathbb{C})$ ou $\text{Mat}(\mathbb{C}, n)$ o conjunto de todas as matrizes complexas $n \times n$.

$\text{Mat}(n, \mathbb{R})$ e $\text{Mat}(n, \mathbb{C})$ são naturalmente dois grupos (Abelianos) em relação à operação de soma de matrizes. Não, porém, em relação à operação de produto, pois é bem sabido que nem toda matriz possui uma inversa.

⁵Tullio Levi-Civita (1873–1941).

O conjunto de todas as matrizes de $\text{Mat}(n, \mathbb{R})$ que são inversíveis forma naturalmente um grupo não Abelian⁶ em relação ao produto usual de matrizes. Esse grupo, denominado *grupo linear real*, é denotado por $\text{GL}(n, \mathbb{R})$. Analogamente, o conjunto de todas as matrizes de $\text{Mat}(n, \mathbb{C})$ inversíveis forma um grupo não Abelian⁷ que é denominado *grupo linear complexo* e denotado por $\text{GL}(n, \mathbb{C})$. Em símbolos,

$$\text{GL}(n, \mathbb{R}) := \{A \in \text{Mat}(n, \mathbb{R}), \det(A) \neq 0\} \quad \text{e} \quad \text{GL}(n, \mathbb{C}) := \{A \in \text{Mat}(n, \mathbb{C}), \det(A) \neq 0\} .$$

Devido à propriedade bem conhecida $\det(AB) = \det(A)\det(B)$, o produto de duas matrizes com determinante igual a 1 é novamente uma matriz com determinante igual a 1. Assim,

$$\text{SL}(n, \mathbb{R}) := \{A \in \text{Mat}(n, \mathbb{R}), \det(A) = 1\} \quad \text{e} \quad \text{SL}(n, \mathbb{C}) := \{A \in \text{Mat}(n, \mathbb{C}), \det(A) = 1\}$$

são subgrupos de $\text{GL}(n, \mathbb{R})$ e $\text{GL}(n, \mathbb{C})$, respectivamente.

E. 21.12 *Exercício.* Para qualquer matriz $n \times n$ real ou complexa e inversível A vale que $(A^T)^{-1} = (A^{-1})^T$. Além disso, para qualquer matriz $n \times n$ complexa A vale que $(A^*)^{-1} = (A^{-1})^*$. Usando esses fatos, mostre que se $A \in \text{GL}(n, \mathbb{R})$, então $A^T \in \text{GL}(n, \mathbb{R})$. Analogamente, mostre que se $A \in \text{GL}(n, \mathbb{C})$, então A^* e $A^T \in \text{GL}(n, \mathbb{C})$. *

E. 21.13 *Exercício.* Para qualquer matriz $n \times n$ real ou complexa A vale que $\det(A) = \det(A^T)$. Fora isso, para qualquer matriz $n \times n$ complexa A vale que $\det(A) = \det(A^*)$. Usando esses fatos, mostre que se $A \in \text{SL}(n, \mathbb{R})$ então $A^T \in \text{SL}(n, \mathbb{R})$. Analogamente, mostre que se $A \in \text{SL}(n, \mathbb{C})$ então A^* e $A^T \in \text{SL}(n, \mathbb{C})$. *

*

Os grupos $\text{GL}(n, \mathbb{R})$, $\text{GL}(n, \mathbb{C})$, $\text{SL}(n, \mathbb{R})$ e $\text{SL}(n, \mathbb{C})$ possuem vários outros subgrupos de interesse. Discutiremos alguns adiante, como os grupos de Borel, os grupos ortogonais, unitários e simpléticos.

• Os grupos $\text{GL}(n, \mathbb{Q})$, $\text{SL}(n, \mathbb{Q})$ e $\text{SL}(n, \mathbb{Z})$

Vamos denotar por $\text{Mat}(n, \mathbb{Z})$ ou $\text{Mat}(\mathbb{Z}, n)$ o conjunto de todas as matrizes $n \times n$ cujos elementos de matriz são números inteiros e por $\text{Mat}(n, \mathbb{Q})$ ou $\text{Mat}(n, \mathbb{Q})$ o conjunto de todas as matrizes $n \times n$ cujos elementos de matriz são números racionais. Analogamente, defina-se

$$\text{GL}(n, \mathbb{Z}) := \{A \in \text{Mat}(n, \mathbb{Z}), \det(A) \neq 0\} \quad \text{e} \quad \text{GL}(n, \mathbb{Q}) := \{A \in \text{Mat}(n, \mathbb{Q}), \det(A) \neq 0\}$$

e

$$\text{SL}(n, \mathbb{Z}) := \{A \in \text{Mat}(n, \mathbb{Z}), \det(A) = 1\} \quad \text{e} \quad \text{SL}(n, \mathbb{Q}) := \{A \in \text{Mat}(n, \mathbb{Q}), \det(A) = 1\} .$$

Então, valem as seguintes afirmações:

1. $\text{GL}(n, \mathbb{Q})$ é um grupo em relação à operação de produto usual de matrizes.
2. $\text{SL}(n, \mathbb{Q})$ é um grupo em relação à operação de produto usual de matrizes.
3. $\text{GL}(n, \mathbb{Z})$ **não** é um grupo em relação à operação de produto usual de matrizes, mas sim um monóide.
4. $\text{SL}(n, \mathbb{Z})$ é um grupo em relação à operação de produto usual de matrizes.

Para provar 1, notemos que o produto de matrizes $n \times n$ com entradas racionais é também uma matriz $n \times n$ com entradas racionais (por quê?). Assim, a operação de produto é uma operação binária em $\text{GL}(n, \mathbb{Q})$. O elemento neutro é a matriz identidade, que é elemento de $\text{GL}(n, \mathbb{Q})$ (pois os números 0 e 1 são racionais). Por fim, resta mostrar que a inversa de uma matriz inversível com entradas racionais também tem entradas racionais.

Para mostrar isso, notemos primeiramente que o determinante de uma matriz com entradas racionais é também um número racional, pois o cálculo do determinante de uma matriz M envolve apenas operações de soma e produto dos

⁶Exceto no caso $n = 1$, onde o grupo é Abelian, trivialmente.

⁷Idem.

elementos de matriz de M . Além disso, lembremos a chamada “regra de Laplace⁸”), expressão (10.20), página 453, que para qualquer matriz A o elemento ij da sua matriz inversa (se houver) é dado por

$$(A^{-1})_{ij} = \frac{(-1)^{i+j}}{\det(A)} \text{Men}(A)_{ji}, \tag{21.12}$$

onde $\text{Men}(A)_{ij}$ é o determinante da matriz $(n - 1) \times (n - 1)$ obtida eliminando-se a i -ésima linha e a j -ésima coluna da matriz A . (A matriz $\text{Men}(A)$ é por vezes denominada matriz dos menores de A). Vê-se claramente daí que se A é uma matriz com entradas racionais então os números $\text{Men}(A)_{ji}$ são também racionais, assim como $\det(A)$. Logo, $(A^{-1})_{ij}$ é um número racional e, portanto, se $A \in \text{GL}(n, \mathbb{Q})$ então $A^{-1} \in \text{GL}(n, \mathbb{Q})$.

O item 2 se prova da mesma maneira.

No caso do item 3, notemos que o produto de matrizes $n \times n$ com entradas inteiras é também uma matriz $n \times n$ com entradas inteiras (por quê?). Assim, a operação de produto é uma operação binária em $\text{GL}(n, \mathbb{Z})$. O elemento neutro é a matriz identidade, que é elemento de $\text{GL}(n, \mathbb{Z})$ (pois os números 0 e 1 são inteiros). Com isso, $\text{GL}(n, \mathbb{Z})$ é um monóide. O problema que faz com que $\text{GL}(n, \mathbb{Z})$ não seja um grupo reside no fato de que a inversa de uma matriz com entradas inteiras nem sempre é uma matriz com entradas inteiras. Isso se vê claramente no exemplo da matriz $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ cuja inversa é $\begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}$. No entanto, se uma matriz A , inversível com entradas inteiras, tiver determinante igual a 1, segue imediatamente de (21.12) que A^{-1} tem também entradas inteiras. Daí, prova-se facilmente a afirmativa 4.

E. 21.14 *Exercício.* Complete os detalhes das afirmações feitas acima. *

E. 21.15 *Exercício.* Verifique que $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \in \text{SL}(n, \mathbb{Z})$ e que $A^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \in \text{SL}(n, \mathbb{Z})$. Mais genericamente, se

a, b, c e d são números inteiros tais que $ad - bc = 1$, então $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(n, \mathbb{Z})$ e $A^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \in \text{SL}(n, \mathbb{Z})$. *

E. 21.16 *Exercício.* Verifique que todas as matrizes da forma $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ com $b \in \mathbb{Z}$ são elementos de $\text{SL}(n, \mathbb{Z})$. Verifique que

todas as matrizes da forma $\begin{pmatrix} 1 & 1 \\ c & c+1 \end{pmatrix}$ com $c \in \mathbb{Z}$ são elementos de $\text{SL}(n, \mathbb{Z})$. *

• **Outros subgrupos de $\text{GL}(n, \mathbb{R})$ e de $\text{GL}(n, \mathbb{C})$**

Há vários outros subgrupos de $\text{GL}(n, \mathbb{R})$ e $\text{GL}(n, \mathbb{C})$ aos quais eventualmente faremos referência. Deixamos ao estudante provar em cada caso que se trata realmente de grupos. Dois deles são os grupos de matrizes com determinante positivo:

$$\text{GL}(n, \mathbb{R})^+ := \{A \in \text{Mat}(n, \mathbb{R}), \det(A) > 0\}, \tag{21.13}$$

$$\text{GL}(n, \mathbb{C})^+ := \{A \in \text{Mat}(n, \mathbb{C}), \det(A) > 0\}. \tag{21.14}$$

Outro grupo relevante é o chamado *grupo de Weyl*⁹ de $\text{GL}(n, \mathbb{C})$:

$$W_n := \left\{ A \in \text{GL}(n, \mathbb{R}), A_{ij} \in \{0, 1\} \forall i, j, \text{ com } \sum_{i=1}^n A_{ij} = 1 = \sum_{j=1}^n A_{ij} \right\}.$$

⁸Pierre-Simon Laplace (1749–1827).

⁹Hermann Klaus Hugo Weyl (1885–1955).

Em palavras, as matrizes de W_n são matrizes $n \times n$ cujas entradas valem 0 ou 1, sendo que exatamente um elemento “1” ocorre em cada linha e em cada coluna.

E. 21.17 *Exercício.* Mostre que W_n é realmente um grupo pelo produto usual de matrizes. Mostre também que as matrizes de W_n têm determinante ± 1 . *

E. 21.18 *Exercício.* Mostre que W_2 contém apenas dois elementos, a saber, as matrizes $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ e $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. *

E. 21.19 *Exercício.* Determine os (seis) elementos de W_3 . *

E. 21.20 *Exercício.* Prove que W_n é isomorfo ao grupo de permutações de n elementos S_n , definido na Seção 21.2, página 1016. *

21.3.1.1 Grupos Lineares Projetivos

Vimos na Seção 2.2.2.2, página 138 que se G é um grupo e $\mathbf{Z}(G)$ o seu centro, o grupo quociente $P(G) := G/\mathbf{Z}(G)$ é denominado *grupo projetivo associado a G* . Vamos aqui estudar alguns desses grupos no caso de grupos matriciais.

• Grupos Lineares Projetivos

Grupos Lineares Projetivos são relevantes em certas aplicações à Física Quântica e à Geometria. Não se trata exatamente de grupos matriciais, mas eles são definidos pela tomada de quociente de grupos matriciais pelos seus centros. Vamos aqui apresentar sua definição sem detalharmos suas propriedades mais avançadas.

O grupo $GL(n, \mathbb{C})$ possui um centro $\mathbf{Z}(GL(n, \mathbb{C}))$, composto pelas matrizes que são múltiplos não nulos da matriz identidade: $\mathbf{Z}(GL(n, \mathbb{C})) := \{\lambda \mathbf{1}_n, \lambda \in \mathbb{C} \setminus \{0\}\}$. O grupo $\mathbf{Z}(GL(n, \mathbb{C}))$ (que é, obviamente, isomorfo ao grupo $(\mathbb{C} \setminus \{0\}, \cdot)$), é (trivialmente) um subgrupo normal¹⁰ de $GL(n, \mathbb{C})$, pois, evidentemente, $A(\lambda \mathbf{1}_n)A^{-1} = \lambda \mathbf{1}_n$ para todo $A \in GL(n, \mathbb{C})$. Podemos, assim, definir o grupo quociente de $GL(n, \mathbb{C})$ por $\mathbf{Z}(GL(n, \mathbb{C}))$:

$$PGL(n, \mathbb{C}) := GL(n, \mathbb{C})/\mathbf{Z}(GL(n, \mathbb{C})) .$$

Para cada $n \in \mathbb{N}$ esse grupo é denominado *grupo projetivo linear complexo* (de ordem $n \in \mathbb{N}$). Trata-se do grupo composto pelas matrizes $n \times n$ complexas inversíveis identificando-se todos os elementos $A \in GL(n, \mathbb{C})$ com seus múltiplos λA , $\lambda \in \mathbb{C} \setminus \{0\}$.

De forma totalmente análoga define-se o *grupo projetivo linear real* (de ordem $n \in \mathbb{N}$) por

$$PGL(n, \mathbb{R}) := GL(n, \mathbb{R})/\mathbf{Z}(GL(n, \mathbb{R})) ,$$

onde, aqui $\mathbf{Z}(GL(n, \mathbb{R})) := \{\lambda \mathbf{1}_n, \lambda \in \mathbb{R} \setminus \{0\}\}$. Trata-se do grupo composto pelas matrizes $n \times n$ reais inversíveis identificando-se todos os elementos $A \in GL(n, \mathbb{R})$ com seus múltiplos λA , $\lambda \in \mathbb{R} \setminus \{0\}$.

Os elementos de $PGL(n, \mathbb{C})$ ou $PGL(n, \mathbb{R})$ são denominados *transformações lineares projetivas*, ou simplesmente *transformações projetivas* (o termo *homografias* é também usado).

Diversas generalizações e especializações dessas definições são possíveis. Se considerarmos o grupo $SL(n, \mathbb{C})$ das matrizes especiais complexas, temos o *grupo projetivo especial complexo*

$$PSL(n, \mathbb{C}) := SL(n, \mathbb{C})/\mathbf{Z}(SL(n, \mathbb{C})) ,$$

onde $\mathbf{Z}(SL(n, \mathbb{C}))$ é o subgrupo de $SL(n, \mathbb{C})$ composto por matrizes múltiplas da identidade e com determinante 1:

$$\mathbf{Z}(SL(n, \mathbb{C})) := \{\lambda \mathbf{1}_n, \lambda \in \mathbb{C} \text{ tal que } \lambda^n = 1\} .$$

¹⁰As noções de subgrupo normal e grupo quociente foram desenvolvidas na Seção 2.2.2, página 132.

Os elementos de $\mathbf{Z}(\mathrm{SL}(n, \mathbb{C}))$ são da forma $\exp\left(\frac{2\pi ik}{n}\right)\mathbb{1}_n$, com $k = 0, 1, \dots, n - 1$. Trata-se do grupo composto pelas matrizes $n \times n$ complexas de determinante 1 identificando-se todos os elementos $A \in \mathrm{GL}(n, \mathbb{C})$ com seus múltiplos λA , com λ da forma $\lambda = \exp\left(\frac{2\pi ik}{n}\right)\mathbb{1}_n$, com $k = 0, 1, \dots, n - 1$, as raízes n -ésimas da unidade.

Os grupos projetivos especiais reais são definidos analogamente:

$$\mathrm{PSL}(n, \mathbb{R}) := \mathrm{SL}(n, \mathbb{R})/\mathbf{Z}(\mathrm{SL}(n, \mathbb{R})),$$

também com

$$\mathbf{Z}(\mathrm{SL}(n, \mathbb{R})) := \{\lambda\mathbb{1}_n, \lambda \in \mathbb{R} \text{ tal que } \lambda^n = 1\}.$$

Note-se, aqui, porém, que $\mathbf{Z}(\mathrm{SL}(n, \mathbb{R})) = \{-\mathbb{1}_n, \mathbb{1}_n\}$ caso n seja par e $\mathbf{Z}(\mathrm{SL}(n, \mathbb{R})) = \{\mathbb{1}_n\}$, caso n seja ímpar. Assim, o quociente no caso n ímpar nada fornece de novo: $\mathrm{PSL}(n, \mathbb{R}) \simeq \mathrm{SL}(n, \mathbb{R})$, n ímpar. No caso n par, $\mathrm{PSL}(n, \mathbb{R})$ pode ser descrito como o grupo das matrizes de entradas reais e determinante 1 para o qual identificamos os elementos A e $-A$ de $\mathrm{SL}(n, \mathbb{R})$.

• Alguns isomorfismos entre grupos lineares projetivos

É importante dizer que, os grupos $\mathrm{PGL}(n, \mathbb{C})$ e $\mathrm{PSL}(n, \mathbb{C})$ são isomorfos. Isso pode ser visto da seguinte forma. Se $A \in \mathrm{GL}(n, \mathbb{C})$, então A é equivalente à matriz $A' := \frac{1}{\alpha}A$, onde α é uma das raízes n -ésimas complexas de $\det(A) \in \mathbb{C}$. A matriz A' , porém, tem determinante 1 e, assim, um elemento de $\mathrm{SL}(n, \mathbb{C})$.

No caso real, porém, $\mathrm{PGL}(n, \mathbb{R})$ e $\mathrm{PSL}(n, \mathbb{R})$ somente são isomorfos caso n for ímpar.

E. 21.21 *Exercício.* Justifique! *

No caso n par, $\mathrm{PSL}(n, \mathbb{R})$ é isomorfo ao grupo projetivo associado a um subgrupo de $\mathrm{GL}(n, \mathbb{R})$, a saber, o grupo $\mathrm{GL}(n, \mathbb{R})^+$, formado pelas matrizes de determinante positivo (vide (21.13), página 1024).

Para ver isso, observemos que $\mathbf{Z}(\mathrm{GL}(n, \mathbb{R})^+) = \{\lambda\mathbb{1}_n, \lambda \in \mathbb{R} \setminus \{0\}\}$, caso n for par, e $\mathbf{Z}(\mathrm{GL}(n, \mathbb{R})^+) = \{\lambda\mathbb{1}_n, \lambda \in (0, \infty)\}$, caso n for ímpar. Podemos argumentar como no caso complexo que se $A \in \mathrm{GL}(n, \mathbb{R})^+$, então A é equivalente a $A' = \frac{1}{\alpha}A$, onde $\alpha = \det(A)^{1/n} \in (0, \infty)$, que tem determinante 1. Como $\mathrm{SL}(n, \mathbb{R})$ é um subgrupo de $\mathrm{GL}(n, \mathbb{R})^+$, concluímos que $\mathrm{P}(\mathrm{GL}(n, \mathbb{R})^+)$ e $\mathrm{PSL}(n, \mathbb{R})$ são isomorfos, e isso para todo $n \in \mathbb{N}$.

Disso concluímos também que, para n ímpar,

$$\mathrm{P}(\mathrm{GL}(n, \mathbb{R})^+) \simeq \mathrm{PSL}(n, \mathbb{R}) \simeq \mathrm{PGL}(n, \mathbb{R}).$$

Analogamente, temos também que $\mathrm{P}(\mathrm{GL}(n, \mathbb{C})^+)$ e $\mathrm{PSL}(n, \mathbb{C})$ são isomorfos, e isso para todo $n \in \mathbb{N}$.

Resumindo, temos:

1. $\mathrm{PGL}(n, \mathbb{C}) \simeq \mathrm{PSL}(n, \mathbb{C})$ para todo $n \in \mathbb{N}$.
2. $\mathrm{PGL}(n, \mathbb{R}) \simeq \mathrm{PSL}(n, \mathbb{R}) \simeq \mathrm{P}(\mathrm{GL}(n, \mathbb{R})^+)$ para todo n ímpar.
3. $\mathrm{PSL}(n, \mathbb{R}) \simeq \mathrm{P}(\mathrm{GL}(n, \mathbb{R})^+)$ para todo n par.

• O Grupo Modular

No caso n par uma construção similar pode ser feita com o grupo $\mathrm{SL}(n, \mathbb{Z})$, de matrizes $n \times n$ com entradas inteiras e determinante 1. Já observamos anteriormente que trata-se de um grupo (vide página 1023). Para n par, definimos

$$\mathrm{PSL}(n, \mathbb{Z}) := \mathrm{SL}(n, \mathbb{Z})/\{-\mathbb{1}_n, \mathbb{1}_n\}.$$

Trata-se do grupo de matrizes de ordem par com entradas inteiras e determinante 1 no qual identifica-se os elementos A e $-A$ de $\mathrm{SL}(n, \mathbb{Z})$.

No caso n ímpar, o conjunto das matrizes da forma $k\mathbb{1}_n$, $k \in \mathbb{Z}$ e com determinante 1 limita-se à matrix $\mathbb{1}_n$ e o quociente nada fornece de novo: $\mathrm{PSL}(n, \mathbb{Z}) \simeq \mathrm{SL}(n, \mathbb{Z})$.

O caso $n = 2$ é especialmente relevante: o grupo

$$\mathrm{PSL}(2, \mathbb{Z}) := \mathrm{SL}(2, \mathbb{Z})/\{-\mathbb{1}_2, \mathbb{1}_2\}$$

é denominado *grupo modular* e esse grupo desempenha um papel muito importante na Análise Complexa e na teoria dos chamados espaços modulares.

21.3.2 O Grupo de Borel e o Grupo de Heisenberg

Uma matriz A , complexa, $n \times n$, é dita ser uma *matriz triangular superior* se seus elementos de matriz A_{ij} satisfizerem $A_{ij} = 0$ se $i > j$. Tais matrizes têm a forma

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1(n-1)} & A_{1n} \\ 0 & A_{22} & \cdots & A_{2(n-1)} & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{(n-1)(n-1)} & A_{(n-1)n} \\ 0 & 0 & \cdots & 0 & A_{nn} \end{pmatrix},$$

onde os elementos abaixo da diagonal principal são nulos. Aqueles que ficam acima da diagonal principal podem ser nulos ou não.

De acordo com a Proposição 10.29, página 507, o conjunto das matrizes complexas $n \times n$ triangulares superiores inversíveis forma um grupo, denominado por alguns autores *Grupo de Borel*¹¹ de ordem n e denotado por $GB_n(\mathbb{C})$.

E. 21.22 Exercício-exemplo. Para duas matrizes triangulares superiores inversíveis 2×2

$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} d & e \\ 0 & f \end{pmatrix}$$

verifique que

$$AB = \begin{pmatrix} ad & ae + bf \\ 0 & cf \end{pmatrix},$$

que é novamente uma matriz triangular superior, e verifique que

$$A^{-1} = \begin{pmatrix} \frac{1}{a} & -\frac{b}{ac} \\ 0 & \frac{1}{c} \end{pmatrix}.$$

*

Um caso particular do grupo de Borel é o grupo de Heisenberg, que agora discutiremos.

21.3.2.1 O Grupo de Heisenberg

• O grupo de Heisenberg $GH_3(\mathbb{C})$

O chamado *grupo de Heisenberg*¹², denotado por $GH_3(\mathbb{C})$ (os grupos $GH_n(\mathbb{C})$ com $n \geq 3$ são definidos adiante), é

¹¹Armand Borel (1923–2003). A noção de grupo de Borel é mais geral. As matrizes $n \times n$ triangulares superiores inversíveis compõem o grupo de Borel associado ao grupo $GL(\mathbb{C}, n)$.

¹²Werner Karl Heisenberg (1901–1976).

definido como o grupo formado por todas as matrizes 3×3 da forma

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

onde $a, b, c \in \mathbb{C}$, com o produto usual de matrizes (se $a, b, c \in \mathbb{R}$ temos o grupo $\text{GH}_3(\mathbb{R})$). A matriz identidade é um elemento de $\text{GH}_3(\mathbb{C})$ pois $H(0, 0, 0) = \mathbb{1}$ e tem-se

$$H(a, b, c)H(a', b', c') = H(a + a', b + b', c + c' + ab'). \tag{21.15}$$

Essa relação, em particular, diz que o produto de duas matrizes de $\text{GH}_3(\mathbb{C})$ é novamente uma matriz de $\text{GH}_3(\mathbb{C})$. Tem-se também que

$$H(a, b, c)^{-1} = H(-a, -b, ab - c) = \begin{pmatrix} 1 & -a & ab - c \\ 0 & 1 & -b \\ 0 & 0 & 1 \end{pmatrix}, \tag{21.16}$$

que mostra que toda matriz de $\text{GH}_3(\mathbb{C})$ tem inversa e que essa inversa é também uma matriz de $\text{GH}_3(\mathbb{C})$. Assim, $\text{GH}_3(\mathbb{C})$ é um grupo matricial. De (21.15) constata-se facilmente que $\text{GH}_3(\mathbb{C})$ não é um grupo Abelian.

E. 21.23 *Exercício.* Verifique essas afirmações. ✱

E. 21.24 *Exercício.* Mostre que o centro do grupo de Heisenberg é formado pelas matrizes do tipo $H(0, 0, c)$ com $c \in \mathbb{C}$. O conceito de centro de um grupo foi introduzido à página 138. ✱

Como é fácil de ver, o grupo de Heisenberg é um grupo de Lie (grupos de Lie serão tratados no Capítulo 22, página 1162) que, como variedade analítica, é difeomorfo a \mathbb{C}^3 . O exercício seguinte discute três de seus subgrupos uniparamétricos.

E. 21.25 *Exercício.* Verifique que as matrizes $H_1(t) := H(t, 0, 0)$, $H_2(t) := H(0, t, 0)$, $H_3(t) := H(0, 0, t)$ satisfazem $H_j(t)H_j(t') = H_j(t + t')$ e $H_j(0) = \mathbb{1}$, $j = 1, 2, 3$. Assim, para cada j , as matrizes $H_j(t)$ representam subgrupos uniparamétricos de $\text{GH}_3(\mathbb{C})$. Os geradores infinitesimais desses subgrupos são $\mathfrak{h}_j := \frac{d}{dt}H_j(t)|_{t=0}$. Verifique que

$$\mathfrak{h}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathfrak{h}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathfrak{h}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

e mostre explicitamente que para todo t vale

$$H_1(t) = e^{t\mathfrak{h}_1}, \quad H_2(t) = e^{t\mathfrak{h}_2} \quad \text{e} \quad H_3(t) = e^{t\mathfrak{h}_3}.$$

✱

Mostraremos agora que esses geradores infinitesimais formam uma álgebra de Lie, a chamada álgebra de Heisenberg $\mathfrak{gh}_3(\mathbb{C})$. Adiante explicaremos por que o nome de Heisenberg é associado ao grupo $\text{GH}_3(\mathbb{C})$ e à álgebra $\mathfrak{gh}_3(\mathbb{C})$.

• **A álgebra de Heisenberg $\mathfrak{gh}_3(\mathbb{C})$**

Considere matrizes $h(a, b, c)$ da forma $a\mathfrak{h}_1 + b\mathfrak{h}_2 + c\mathfrak{h}_3$ (\mathfrak{h}_k são os geradores infinitesimais definidos no Exercício E.

21.25), ou seja,

$$h(a, b, c) = \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}, \tag{21.17}$$

onde $a, b, c \in \mathbb{C}$. Calculando-se o comutador de duas de tais matrizes tem-se

$$[h(a, b, c), h(a', b', c')] = h(0, 0, ab' - a'b), \tag{21.18}$$

(verifique!) que é novamente da forma (21.17). Assim, o conjunto de matrizes da forma (21.17) forma uma álgebra de Lie com o produto definido pelo comutador de matrizes. Essa álgebra de Lie, denotada por $\mathfrak{gh}_3(\mathbb{C})$, é denominada *álgebra de Heisenberg*.

A razão dessa denominação é a seguinte. Podemos encontrar em $\mathfrak{gh}_3(\mathbb{C})$ uma base especial formada por três matrizes que, por razões “psicológicas”, denotaremos por p, q e \hbar :

$$p = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \hbar = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{21.19}$$

É um exercício fácil (e fortemente recomendado) verificar que essas matrizes satisfazem as seguintes regras de comutação:

$$[p, \hbar] = 0, \quad [q, \hbar] = 0, \quad [p, q] = -i\hbar. \tag{21.20}$$

Para aqueles familiarizados com a Mecânica Quântica as relações acima justificam a denominação dessa álgebra em honra a Heisenberg: as relações de comutação acima são precisamente iguais às relações canônicas de comutação satisfeitas pelos operadores associados ao momento (p) e posição (q) de uma partícula se movendo em uma dimensão. No caso da Mecânica Quântica, p é o operador $-i\hbar \frac{\partial}{\partial x}$, $q = x$ e \hbar representa um número (a constante de Planck¹³), que obviamente comuta com os operadores p e q .

Nota. O estudante deve, porém, observar que as matrizes p, q e \hbar , acima, não são autoadjuntas, ao contrário dos operadores correspondentes da Mecânica Quântica. Essa observação é relevante, pois é possível provar que as relações canônicas de comutação não podem ser satisfeitas por operadores autoadjuntos agindo em espaços de Hilbert de dimensão finita ou por operadores autoadjuntos limitados agindo em espaços de Hilbert de dimensão infinita. De fato, no espaço de Hilbert $L^2(\mathbb{R}, dx)$ os operadores $p = -i\hbar \frac{\partial}{\partial x}$ e $q = x$ são autoadjuntos (em um domínio conveniente), mas não são limitados.

A argumentação para tal é a seguinte. Vamos supor que p, q sejam operadores autoadjuntos e limitados agindo em um espaço de Hilbert e que satisfaçam (21.20), com $\hbar > 0$. Então, tem-se $\hbar = \|[p, q]\| \leq 2\|p\| \|q\|$. É fácil provar por indução (faça-o!) que vale $[p, q^n] = -in\hbar q^{n-1}$ para todo $n \in \mathbb{N}$. Assim, tem-se $n\hbar \|q^{n-1}\| \leq \|[p, q^n]\| \leq 2\|p\| \|q^n\|$. Agora, como q é autoadjunto, vale $\|q^n\| = \|q\|^n$ para todo $n \in \mathbb{N}$ (há diversas provas desse fato. Para uma prova geral no contexto de álgebras C^* , vide Corolário 39.10, página 2053). Assim, temos a desigualdade $n\hbar \|q\|^{n-1} \leq 2\|p\| \|q\|^n$. Como, naturalmente, supomos que $\|q\| \neq 0$, podemos cancelar um fator $\|q\|^{n-1}$ de ambos os lados e obter $n\hbar \leq 2\|p\| \|q\|$. Ora, essa desigualdade não pode ser satisfeita para todo $n \in \mathbb{N}$, a menos que $\hbar = 0$ (um caso trivial, pois implicaria que as relações (21.20) são comutativas) ou que p ou q (ou ambos) seja(m) não limitado(s), provando o que desejávamos. A referência [41] atribui esse resultado a Wielandt¹⁴. ♣

O que faz $\mathfrak{gh}_3(\mathbb{C})$ especial como álgebra de Lie é a propriedade expressa no seguinte exercício:

E. 21.26 *Exercício importante.* Verifique que para quaisquer três elementos l_1, l_2 e l_3 da álgebra de Heisenberg $\mathfrak{gh}_3(\mathbb{C})$ tem-se

$$[l_1, [l_2, l_3]] = 0. \tag{21.21}$$

Sugestão: use as relações de comutação de p, q e \hbar , dadas acima ou use diretamente (21.18). A relação (21.21) mostra que $\mathfrak{gh}_3(\mathbb{C})$ é o que se chama uma álgebra de Lie nilpotente (de grau 2). ♣

¹³Max Karl Ernst Ludwig Planck (1858–1947).

¹⁴Helmut Wielandt (1910–2001).

Para entender a relação da álgebra de Heisenberg $\mathfrak{gh}_3(\mathbb{C})$ com o grupo de Heisenberg $\text{GH}_3(\mathbb{C})$, façamos o seguinte. Notemos em primeiro lugar que as matrizes $h(a, b, c)$ são matrizes nilpotentes de grau 3, ou seja,

$$h(a, b, c)^3 = 0.$$

(Mostre isso!). É fácil com isso verificar que se calcularmos a exponencial de $h(a, b, c)$ teremos

$$\exp(h(a, b, c)) = \mathbb{1} + h(a, b, c) + \frac{1}{2}h(a, b, c)^2 = \begin{pmatrix} 1 & a & c + \frac{ab}{2} \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = H\left(a, b, c + \frac{ab}{2}\right) \quad (21.22)$$

e disso conclui-se que

$$H(a, b, c) = \exp\left(h\left(a, b, c - \frac{ab}{2}\right)\right). \quad (21.23)$$

E. 21.27 *Exercício.* Escreva $h(a, b, c - \frac{ab}{2})$ como combinação linear de p, q e \hbar . ✦

Pelo que vimos, todos os elementos do grupo de Heisenberg $\text{GH}_3(\mathbb{C})$ são obtidos pela exponenciação de elementos da álgebra de Lie $\mathfrak{gh}_3(\mathbb{C})$, ou seja, a exponenciação é uma aplicação sobrejetora de $\mathfrak{gh}_3(\mathbb{C})$ em seu grupo de Lie $\text{GH}_3(\mathbb{C})$. Em verdade, é fácil constatar que essa aplicação é também injetora (faça isso!). A aplicação exponencial é, portanto, uma bijeção de $\mathfrak{gh}_3(\mathbb{C})$ em $\text{GH}_3(\mathbb{C})$.

E. 21.28 *Exercício importante.* Usando a fórmula de Baker-Campbell-Hausdorff (equações (11.4), página 568, ou (11.68), página 598) e as relações (21.18) e (21.21), mostre que

$$\exp(h(a, b, c)) \exp(h(a', b', c')) = \exp\left(h\left(a + a', b + b', c + c' + \frac{ab' - a'b}{2}\right)\right). \quad (21.24)$$

Usando (21.22) e (21.23), reobtenha de (21.24) a regra de produto (21.15). ✦

Comentário. Esse exercício ilustra uma aplicação da fórmula de Baker-Campbell-Hausdorff. Note-se que, devido ao fato de $\mathfrak{gh}_3(\mathbb{C})$ ser uma álgebra de Lie nilpotente (vide (21.21)), a série de Baker-Campbell-Hausdorff é composta apenas por um número finito de termos e, portanto, converge sempre. ♣

• **O grupo de Heisenberg $\text{GH}_n(\mathbb{C})$, $n \geq 3$**

Vamos agora generalizar o grupo $\text{GH}_3(\mathbb{C})$. Para $n \geq 3$, os chamados *grupos de Heisenberg* $\text{GH}_n(\mathbb{C})$ são definidos como sendo os grupos formados por todas as matrizes $n \times n$ da forma

$$H(a, b, c) = \begin{pmatrix} 1 & a^T & c \\ \mathbb{0}_m & \mathbb{1}_m & b \\ 0 & \mathbb{0}_m^T & 1 \end{pmatrix}$$

com o produto usual de matrizes, sendo $m = n - 2$, onde $a, b \in \mathbb{C}^{n-2}$ e $c \in \mathbb{C}$. Acima, a e b representam matrizes-coluna com $m = n - 2$ linhas, enquanto que a^T e b^T , as transpostas de a e b , respectivamente, representam matrizes-linha com $m = n - 2$ colunas:

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_{n-2} \end{pmatrix}, \quad a^T = (a_1 \quad \dots \quad a_{n-2}), \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_{n-2} \end{pmatrix}, \quad b^T = (b_1 \quad \dots \quad b_{n-2}),$$

sendo $\mathbb{0}_m = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ a matriz coluna identicamente nula com $m = n - 2$ linhas e sendo $\mathbb{1}_m$ a matriz identidade $m \times m$.

Por exemplo, no caso $n = 4$, para $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{C}^2$, a matriz $H(a, b, c)$ é $H(a, b, c) = \begin{pmatrix} 1 & a_1 & a_2 & c \\ 0 & 1 & 0 & b_1 \\ 0 & 0 & 1 & b_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Para simplificar a notação, iremos doravante escrever $H(a, b, c)$ na forma

$$H(a, b, c) = \begin{pmatrix} 1 & a^T & c \\ 0 & \mathbb{1} & b \\ 0 & 0 & 1 \end{pmatrix}.$$

A matriz identidade é um elemento de $\text{GH}_n(\mathbb{C})$, pois $H(0, 0, 0) = \mathbb{1}$ e tem-se

$$H(a, b, c)H(a', b', c') = H(a + a', b + b', c + c' + a^T b'), \tag{21.25}$$

sendo que definimos a forma bilinear $a^T b' := \langle a, b' \rangle_{\mathbb{R}} = a_1 b'_1 + \dots + a_{n-2} b'_{n-2}$. Essa relação, em particular, diz-nos que o produto de duas matrizes de $\text{GH}_n(\mathbb{C})$ é novamente uma matriz de $\text{GH}_n(\mathbb{C})$. Vale também que

$$H(a, b, c)^{-1} = H(-a, -b, a^T b - c) = \begin{pmatrix} 1 & -a & a^T b - c \\ 0 & \mathbb{1}_m & -b \\ 0 & 0 & 1 \end{pmatrix}, \tag{21.26}$$

que mostra que toda matriz de $\text{GH}_n(\mathbb{C})$ tem inversa e que essa inversa é também um elemento de $\text{GH}_n(\mathbb{C})$. Assim, $\text{GH}_n(\mathbb{C})$ é um grupo matricial.

• **A álgebra de Heisenberg $\text{gh}_n(\mathbb{C})$, $n \geq 3$**

Para $n \geq 3$, considere matrizes de $\text{Mat}(n, \mathbb{C})$ da forma

$$h(a, b, c) = \begin{pmatrix} 0 & a^T & c \\ \mathbb{0}_m & \mathbb{0}_{mm} & b \\ 0 & \mathbb{0}_m^T & 0 \end{pmatrix}, \tag{21.27}$$

com $m = n - 2$, onde $\mathbb{0}_{mm}$ é a matriz $m \times m$ identicamente nula e onde $a, b \in \mathbb{C}^{n-2}$ e $c \in \mathbb{C}$, como acima. Por exemplo,

no caso $n = 4$, para $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{C}^2$, a matriz $h(a, b, c)$ é $h(a, b, c) = \begin{pmatrix} 0 & a_1 & a_2 & c \\ 0 & 0 & 0 & b_1 \\ 0 & 0 & 0 & b_2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Calculando-se o comutador de duas de tais matrizes tem-se

$$[h(a, b, c), h(a', b', c')] = h(0, 0, a^T b' - a'^T b), \tag{21.28}$$

(verifique!) que é novamente da forma (21.27). Assim, o conjunto de matrizes da forma (21.27) forma uma álgebra de Lie com o produto definido pelo comutador de matrizes. Essa álgebra de Lie, denotada por $\text{gh}_n(\mathbb{C})$, é igualmente denominada *álgebra de Heisenberg*.

E. 21.29 Exercício importante. Verifique que para quaisquer três elementos h_1, h_2 e h_3 da álgebra de Heisenberg $\text{gh}_n(\mathbb{C})$ tem-se

$$[h_1, [h_2, h_3]] = 0. \tag{21.29}$$

A relação (21.29) mostra que $\text{gh}_n(\mathbb{C})$ é o que se chama uma álgebra de Lie nilpotente (de grau 2). ★

Podemos encontrar em $\mathfrak{gh}_n(\mathbb{C})$ uma base especial formada pelas matrizes \hbar e $p_k, q_k, k = 1, \dots, n - 2$ definidas por

$$p_k = \begin{pmatrix} 0 & e_k^T & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad q_k = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & e_k \\ 0 & 0 & 0 \end{pmatrix}, \quad \hbar = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (21.30)$$

sendo $e_k, k = 1, \dots, n - 2$ as matrizes-coluna definidas por

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_{n-2} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

ou seja, todas as linhas de e_j são nulas, exceto a j -ésima, que vale 1. No caso $n = 4$, por exemplo, tem-se

$$p_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$q_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad q_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \hbar = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Em analogia com o caso do grupo $\text{GH}_3(\mathbb{C})$, é fácil constatar que as matrizes p_k, q_k e $-i\hbar$ são geradores infinitesimais de subgrupos uniparamétricos de $\text{GH}_n(\mathbb{C})$.

E. 21.30 *Exercício.* Verifique a afirmação do último parágrafo. Determine os subgrupos uniparamétricos de $\text{GH}_n(\mathbb{C})$ mencionados.
*

Como $e_k^T e_l = \delta_{k,l}$ para todos k e l , é um exercício fácil (e fortemente recomendado!) verificar que essas matrizes satisfazem as seguintes regras de comutação:

$$[p_k, q_l] = -i\hbar \delta_{k,l}, \quad [p_k, \hbar] = [q_k, \hbar] = [p_k, p_l] = [q_k, q_l] = 0, \quad (21.31)$$

para todos $k, l = 1, \dots, n - 2$. Como o estudante familiarizado com a Mecânica Quântica percebe, essas são as relações canônicas de comutação de um sistema com $n - 2$ graus de liberdade.

Para entender a relação da álgebra de Heisenberg $\mathfrak{gh}_n(\mathbb{C})$ com o grupo de Heisenberg $\text{GH}_n(\mathbb{C})$, notemos em primeiro lugar que, assim como no caso $n = 3$, as matrizes $h(a, b, c)$ são matrizes nilpotentes de grau 3, ou seja,

$$h(a, b, c)^3 = 0.$$

(Mostre isso!). É fácil com isso verificar que

$$\exp(h(a, b, c)) = \mathbb{1} + h(a, b, c) + \frac{1}{2}h(a, b, c)^2 = \begin{pmatrix} 1 & a^T & c + \frac{a^T b}{2} \\ 0 & \mathbb{1} & b \\ 0 & 0 & 1 \end{pmatrix} = H\left(a, b, c + \frac{a^T b}{2}\right), \quad (21.32)$$

ou seja,

$$H(a, b, c) = \exp\left(h\left(a, b, c - \frac{a^T b}{2}\right)\right). \quad (21.33)$$

Pelo que vimos, todos os elementos do grupo de Heisenberg $\text{GH}_n(\mathbb{C})$ são obtidos pela exponenciação de elementos da álgebra de Lie $\mathfrak{gh}_n(\mathbb{C})$, ou seja, a exponenciação é uma aplicação sobrejetora de $\mathfrak{gh}_n(\mathbb{C})$ em seu grupo de Lie $\text{GH}_n(\mathbb{C})$. Em verdade, é fácil constatar que essa aplicação é também injetora (faça isso!). A aplicação exponencial é, portanto, uma bijeção de $\mathfrak{gh}_n(\mathbb{C})$ em $\text{GH}_n(\mathbb{C})$.

E. 21.31 *Exercício importante.* Usando a fórmula de Baker-Campbell-Hausdorff (equações (11.4), página 568, ou (11.68), página 598) e as relações (21.28) e (21.29), mostre que

$$\exp(h(a, b, c)) \exp(h(a', b', c')) = \exp\left(h\left(a + a', b + b', c + c' + \frac{a^T b' - a'^T b}{2}\right)\right). \quad (21.34)$$

Usando (21.32) e (21.33), reobtenha de (21.34) a regra de produto (21.25). ✦

• As relações de Weyl em $\text{GH}_n(\mathbb{C})$

É evidente pelas definições (21.30) que podemos escrever

$$h(a, b, c) = \sum_{k=1}^{n-2} (a_k p_k + b_k q_k) - ic\hbar. \quad (21.35)$$

Por outro lado, as relações de comutação (21.31) afirmam que \hbar comuta com cada p_k e com cada q_k . Logo, tem-se

$$\exp(h(a, b, c)) = \exp(-ic\hbar) \exp\left(\sum_{k=1}^{n-2} (a_k p_k + b_k q_k)\right).$$

Por (21.35), tem-se $\sum_{k=1}^{n-2} (a_k p_k + b_k q_k) = h(a, b, 0)$. Assim, vale a identidade

$$\exp(h(a, b, c)) = \exp(-ic\hbar) \exp(h(a, b, 0)).$$

Usando isso no lado direito da relação (21.34), a mesma fica

$$\exp(h(a, b, c)) \exp(h(a', b', c')) = \exp\left(-i\left(c + c' + \frac{a^T b' - a'^T b}{2}\right)\hbar\right) \exp\left(h(a + a', b + b', 0)\right).$$

Efetuada-se nessa expressão as trocas $a \leftrightarrow a', b \leftrightarrow b', c \leftrightarrow c'$, obtemos também

$$\exp(h(a', b', c')) \exp(h(a, b, c)) = \exp\left(-i\left(c + c' + \frac{a'^T b - a^T b'}{2}\right)\hbar\right) \exp\left(h(a + a', b + b', 0)\right).$$

Comparando-se as duas últimas expressões, conclui-se facilmente que

$$\exp(h(a, b, c)) \exp(h(a', b', c')) = \exp(-i(a^T b' - a'^T b)\hbar) \exp(h(a', b', c')) \exp(h(a, b, c)). \quad (21.36)$$

Essa expressão é denominada *relação de Weyl*¹⁵, ou *relação de comutação de Weyl*, para $\text{GH}_n(\mathbb{C})$.

Definindo-se

$$U(a) := \exp(h(ia, 0, 0)) = \exp\left(i \sum_{k=1}^{n-2} a_k p_k\right) \quad e$$

$$V(b) := \exp(h(0, ib, 0)) = \exp\left(i \sum_{k=1}^{n-2} b_k q_k\right),$$

é fácil checar por (21.36) (faça-o!) que valem

$$U(a)V(b) = \exp\left(i(a^T b)\hbar\right) V(b)U(a), \tag{21.37}$$

$$U(a)U(a') = U(a')U(a) = U(a + a') \quad e \tag{21.38}$$

$$V(b)V(b') = V(b')V(b) = V(b + b'). \tag{21.39}$$

A relação (21.37) é também denominada *relação de Weyl*. Não por acaso, as relações (21.37)–(21.39) são essencialmente idênticas às relações (37.50)–(37.52), página 1861, válidas em um espaço de funções (o espaço de Schwartz). De fato, se a e b acima forem vetores reais, a expressão $a^T b$, que ocorre em (21.37), é o produto escalar usual entre eles.

E. 21.32 *Exercício.* Você sabe explicar o porquê dessa coincidência? ✦

As relações de Weyl na forma (21.37)–(21.39) ou na forma (37.50)–(37.52) são a motivação da definição das chamadas *Álgebras CCR* (vide, e.g., [61]), cujo estudo é relevante tanto na Mecânica Quântica quanto na Teoria Quântica de Campos.

21.3.3 Grupos Associados a Formas Bilineares e Sesquilineares

Na corrente seção vamos estudar grupos compostos por operadores lineares que preservam formas bilineares ou sesquilineares. Tais grupos são por vezes denominados *Grupos Clássicos*. O interesse específico por operadores lineares é consequência de dois resultados, o Lema 21.3 e o Teorema 21.7, abaixo, cujo interesse vai além de nossas pretensões iniciais, como teremos oportunidade de comentar alhures. O Teorema 21.7, por exemplo, é relevante na discussão sobre a invariância do intervalo na Teoria da Relatividade restrita e no estudo do movimento de corpos rígidos na Mecânica Clássica.

Lema 21.3 *Seja V um espaço vetorial real dotado de uma forma bilinear simétrica e não degenerada $\omega : V \times V \rightarrow \mathbb{R}$. Seja q a forma quadrática associada a ω , definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $R : V \rightarrow V$ uma função sobrejetora. Então, R satisfaz as condições*

$$R(0) = 0 \quad e \quad q(R(u) - R(v)) = q(u - v) \quad \text{para todos } u, v \in V \tag{21.40}$$

se e somente se satisfizer

$$\omega(R(u), R(v)) = \omega(u, v), \tag{21.41}$$

também para todos $u, v \in V$.

Se ω for um produto escalar a condição de $R : V \rightarrow V$ ser sobrejetora pode ser omitida. □

Prova do Lema 21.3. A prova é dividida em duas partes:

I. Vamos supor a validade de (21.40). Tomando-se $v = 0$ na segunda igualdade de (21.40) e usando a primeira ($R(0) = 0$), tem-se $q(R(u)) = q(u)$ para todo u . Agora, Por um lado

$$q(R(u)) + q(R(v)) - q(R(u) - R(v)) = q(u) + q(v) - q(u - v) = \omega(u, u) + \omega(v, v) - \omega(u - v, u - v) = \omega(u, v) + \omega(v, u).$$

¹⁵Hermann Klaus Hugo Weyl (1885–1955).

Por outro lado,

$$\begin{aligned} q(R(u)) + q(R(v)) - q(R(u) - R(v)) &= \omega(R(u), R(u)) + \omega(R(v), R(v)) - \omega(R(u) - R(v), R(u) - R(v)) \\ &= \omega(R(u), R(v)) + \omega(R(v), R(u)). \end{aligned}$$

Isso estabeleceu que

$$\omega(R(u), R(v)) + \omega(R(v), R(u)) = \omega(u, v) + \omega(v, u).$$

Se ω é simétrica, isso implica

$$\omega(R(u), R(v)) = \omega(u, v).$$

II. Vamos supor a validade de (21.41). Tomando-se nela $v = 0$, temos que $\omega(R(u), R(0)) = 0$ para todo $u \in V$, o que implica $R(0) = 0$ pois ω é não degenerada e R é sobrejetora (se ω for um produto escalar tomamos $u = v = 0$ e (21.41) diz-nos que $\omega(R(0), R(0)) = 0$, o que igualmente implica $R(0) = 0$). Além disso,

$$\begin{aligned} q(R(u) - R(v)) &= \omega(R(u) - R(v), R(u) - R(v)) = \omega(R(u), R(u)) + \omega(R(v), R(v)) - 2\omega(R(u), R(v)) \\ &\stackrel{(21.41)}{=} \omega(u, u) + \omega(v, v) - 2\omega(u, v) = \omega(u - v, u - v) = q(u - v), \end{aligned} \quad (21.42)$$

estabelecendo a validade de (21.40), completando a prova do Lema 21.3. ■

O Teorema que segue é muito importante no estudo de funções que mantêm formas quadráticas invariantes.

Teorema 21.7 *Seja V um espaço vetorial real dotado de uma forma bilinear simétrica e não degenerada $\omega : V \times V \rightarrow \mathbb{R}$. Seja q a forma quadrática associada a ω , definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \rightarrow V$ uma função sobrejetora definida em V e tal que $q(A(u) - A(v)) = q(u - v)$ para todos $u, v \in V$. Então, existem: um vetor constante $c \in V$ e um operador linear sobrejetor $R : V \rightarrow V$ satisfazendo $\omega(Ru, Rv) = \omega(u, v)$ para todos $u, v \in V$, tais que*

$$A(u) = Ru + c$$

para todo $u \in V$. O operador linear R e a constante c são univocamente determinados.

Se ω for um produto escalar a condição de $A : V \rightarrow V$ ser sobrejetora pode ser omitida. □

Prova do Teorema 21.7. Seja V um espaço vetorial real dotado de uma forma bilinear simétrica e não degenerada $\omega : V \times V \rightarrow \mathbb{R}$. Seja q a forma quadrática associada a ω , definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \rightarrow V$ uma função sobrejetora definida em V tal que $q(A(u) - A(v)) = q(u - v)$ para todos $u, v \in V$.

Defina-se $c := A(0) \in V$ e defina-se uma nova função $R : V \rightarrow V$ dada por $R(v) := A(v) - c = A(v) - A(0)$ para todo $v \in V$. É claro que R é sobrejetora, por A o ser. É claro também que $R(0) = 0$ e que $q(R(u) - R(v)) = q(u - v)$ para todos $u, v \in V$.

Vamos agora mostrar que se $R : V \rightarrow V$ satisfaz (21.40) (e, portanto, (21.41), pelo Lema 21.3), então R é linear. Desejamos provar que para todos $\alpha_1, \alpha_2 \in \mathbb{R}$ e todos $u_1, u_2 \in V$ vale $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$. Temos que, para $w \in V$, arbitrário,

$$\begin{aligned} &\omega\left(R(w), R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2)\right) \\ &= \omega\left(R(w), R(\alpha_1 u_1 + \alpha_2 u_2)\right) - \alpha_1 \omega\left(R(w), R(u_1)\right) - \alpha_2 \omega\left(R(w), R(u_2)\right) \\ &\stackrel{(21.41)}{=} \omega\left(w, \alpha_1 u_1 + \alpha_2 u_2\right) - \alpha_1 \omega\left(w, u_1\right) - \alpha_2 \omega\left(w, u_2\right) \\ &= \omega\left(w, \underbrace{(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2}_{=0}\right) = 0. \end{aligned} \quad (21.43)$$

Como R é sobrejetora e w é arbitrário, a não degenerescência de ω implica que $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$ para todos $\alpha_1, \alpha_2 \in \mathbb{R}$ e todos $u_1, u_2 \in V$, significando que $R : V \rightarrow V$ é linear.

Se ω for um produto escalar, prova-se que $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$ mostrando-se que

$$\omega\left(R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2), R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2)\right) = 0,$$

o que é feito expandindo-se o lado esquerdo em nove termos do tipo $\omega(R(x), R(y)) = \omega(x, y)$, como feito em (21.43), e reagrupando-os para obter $\omega\left((\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2, (\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2\right) = 0$.

Resta apenas demonstrar a unicidade de R e c . Vamos supor que para todo $u \in V$ tenhamos $Ru + c = R'u + c'$. Tomando-se $u = 0$ concluiríamos que $c = c'$. A relação restante $Ru = R'u$, válida para todo $u \in V$, significa que $R = R'$. Isso completa a prova do Teorema 21.7. ■

Uma generalização do Lema 21.3 e do Teorema 21.7 para o caso complexo pode ser encontrada na Seção 21.A, página 1149.

• Grupos Associados a Formas Bilineares e Sesquilineares

Seja \mathcal{E} um espaço vetorial sobre os reais ou sobre os complexos. Vamos denotar por $\text{GL}(\mathcal{E})$ o conjunto de todos os operadores lineares bijetores (e portanto inversíveis) de \mathcal{E} em \mathcal{E} . É bem claro que $\text{GL}(\mathcal{E})$ forma um grupo, tendo como produto o produto de operadores.

Seja $\omega : \mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}$ (ou \mathbb{C}) uma função, em princípio arbitrária, definida em $\mathcal{E} \times \mathcal{E}$ e assumindo valores reais ou complexos. Em palavras mais simples, seja ω uma função de duas variáveis que a cada par de vetores $x, y \in \mathcal{E}$ associa um número real ou complexo $\omega(x, y)$.

Denotaremos por $\Omega(\mathcal{E}, \omega)$ o subconjunto de $\text{GL}(\mathcal{E})$ formado por todos os operadores lineares $O : \mathcal{E} \rightarrow \mathcal{E}$, inversíveis, tais que $\omega(Ox, Oy) = \omega(x, y)$ para todos $x, y \in \mathcal{E}$:

$$\Omega(\mathcal{E}, \omega) := \left\{ O \in \text{GL}(\mathcal{E}) \mid \omega(Ox, Oy) = \omega(x, y) \text{ para todos } x, y \in \mathcal{E} \right\}. \quad (21.44)$$

Dizemos que $\Omega(\mathcal{E}, \omega)$ é composto pelos operadores lineares inversíveis agindo em \mathcal{E} que preservam a função ω . Vamos mostrar que $\Omega(\mathcal{E}, \omega)$ é um subgrupo de $\text{GL}(\mathcal{E})$, muita vezes denominado o *grupo de invariância associado a ω* .

Primeiramente, é claro que $\mathbb{1} \in \Omega(\mathcal{E}, \omega)$. Em segundo lugar, sejam O_1 e O_2 dois operadores de $\Omega(\mathcal{E}, \omega)$. Teremos pelas hipóteses que

$$\omega(O_1 O_2 x, O_1 O_2 y) = \omega(O_2 x, O_2 y) = \omega(x, y)$$

para todos $x, y \in \mathcal{E}$ e, portanto, $O_1 O_2 \in \Omega(\mathcal{E}, \omega)$. Resta mostrar que se $O \in \Omega(\mathcal{E}, \omega)$ então $O^{-1} \in \Omega(\mathcal{E}, \omega)$. De fato,

$$\omega(O^{-1}x, O^{-1}y) = \omega(OO^{-1}x, OO^{-1}y) = \omega(x, y)$$

para todos $x, y \in \mathcal{E}$, que é o que queríamos provar.

• Grupos de invariância de formas bilineares e sesquilineares em um espaço vetorial

Há um grande interesse nos grupos $\Omega(\mathcal{E}, \omega)$ nos importantes casos particulares (nos quais nos especializaremos logo em seguida) em que ω é uma forma bilinear ou sesquilinear definida em \mathcal{E} (esse último caso se \mathcal{E} for um espaço vetorial sobre os complexos). As noções de forma bilinear ou sesquilinear foram introduzidas no Capítulo 3, página 219. Também de interesse são as situações em que \mathcal{E} é um espaço vetorial (real ou complexo) de dimensão finita. Muitos dos grupos clássicos de relevância na Física e na Geometria são grupos de invariância de formas em espaços vetoriais de dimensão finita, como veremos mais adiante.

• O caso de formas sesquilineares Hermitianas

Uma simplificação relevante na definição de $\Omega(\mathcal{E}, \omega)$ ocorre no caso de ω ser uma forma sesquilinear: podemos identificar o grupo $\Omega(\mathcal{E}, \omega)$ com o conjunto

$$\Omega(\mathcal{E}, \omega) := \left\{ O \in \text{GL}(\mathcal{E}) \mid \omega(Ox, Ox) = \omega(x, x) \text{ para todo } x \in \mathcal{E} \right\}. \quad (21.45)$$

Compare-se com (21.44). Primeiramente, é evidente que se $O \in \text{GL}(\mathcal{E})$ e satisfaz $\omega(Ox, Oy) = \omega(x, y)$ para todo $x, y \in \mathcal{E}$, então também satisfaz $\omega(Ox, Ox) = \omega(x, x)$ para todo $x \in \mathcal{E}$. A recíproca, porém, é também verdadeira, para formas sesquilineares, devido à chamada *identidade de polarização*, relação (3.10), página 225, que claramente diz-nos que se $\omega(Ox, Ox) = \omega(x, x)$ para todo $x \in \mathcal{E}$, então $\omega(Ox, Oy) = \omega(x, y)$ para todos $x, y \in \mathcal{E}$, justificando (21.45).

E. 21.33 *Exercício.* Justifique essas afirmações. ✱

• **O caso de formas bilineares simétricas**

Algo semelhante se dá no caso de formas bilineares simétricas. Se ω for uma forma bilinear simétrica (ou seja, se valer $\omega(x, y) = \omega(y, x)$ para todos $x, y \in \mathcal{E}$), então podemos também identificar o grupo $\Omega(\mathcal{E}, \omega)$ com o conjunto

$$\Omega(\mathcal{E}, \omega) := \left\{ O \in \text{GL}(\mathcal{E}) \mid \omega(Ox, Ox) = \omega(x, x) \text{ para todo } x \in \mathcal{E} \right\}. \quad (21.46)$$

Compare-se com (21.44). A justificativa é análoga ao caso de formas sesquilineares, acima, sendo, porém, que agora a identidade de polarização assume agora a forma da relação (3.3), página 220, a qual claramente diz-nos que se $\omega(Ox, Ox) = \omega(x, x)$ para todo $x \in \mathcal{E}$, então $\omega(Ox, Oy) = \omega(x, y)$ para todos $x, y \in \mathcal{E}$, justificando (21.46).

E. 21.34 *Exercício.* Justifique essas afirmações. ✱

Vamos agora considerar casos particulares em que \mathcal{E} é o espaço \mathbb{R}^n ou \mathbb{C}^n . Esses são os casos mais relevantes de grupos de invariância associados a formas.

• **Grupos de invariância de formas bilineares em dimensão finita**

Seja $\mathcal{E} = \mathbb{R}^n$ e seja ω_A uma forma bilinear em \mathbb{R}^n , que pelas considerações da Seção 3.4 é da forma $\omega_A(x, y) = \langle x, Ay \rangle_{\mathbb{R}}$ para alguma matriz real A . Neste caso, $\Omega(\mathbb{R}^n, \omega_A)$ é o conjunto de todas as matrizes M inversíveis reais $n \times n$ tais que

$$\langle Mx, AMy \rangle_{\mathbb{R}} = \langle x, Ay \rangle_{\mathbb{R}}$$

para todos $x, y \in \mathbb{R}^n$. Essa relação diz-nos que $\langle x, M^T AMy \rangle_{\mathbb{R}} = \langle x, Ay \rangle_{\mathbb{R}}$ para todos $x, y \in \mathbb{R}^n$, o que implica

$$M^T AM = A$$

(por quê?). Assim,

$$\Omega(\mathbb{R}^n, \omega_A) := \left\{ M \in \text{GL}(n, \mathbb{R}), \det(M) \neq 0 \text{ e } M^T AM = A \right\}. \quad (21.47)$$

Caso $\det A \neq 0$, (ou seja, se ω_A for não degenerada), a condição $M^T AM = A$ implica $(\det M)^2 = 1$, e, portanto, implica por si só que M tem inversa. Nesse caso, portanto, é suficiente escrevermos

$$\Omega(\mathbb{R}^n, \omega_A) := \left\{ M \in \text{GL}(n, \mathbb{R}), M^T AM = A \right\}.$$

Se a matriz A for inversível (ou seja, se ω_A for não degenerada), então podemos escrever também

$$\Omega(\mathbb{R}^n, \omega_A) = \left\{ M \in \text{GL}(n, \mathbb{R}), M^{-1} = A^{-1} M^T A \right\}.$$

Se A possui uma inversa, a relação $M^T AM = A$ implica que $\det M = \pm 1$ (justifique!). Nesse caso, o grupo $\Omega(\mathbb{R}^n, \omega_A)$ possui um subgrupo de particular interesse, composto por suas matrizes de determinante igual a +1:

$$\text{SO}(\mathbb{R}^n, \omega_A) := \left\{ M \in \text{GL}(n, \mathbb{R}), M^{-1} = A^{-1} M^T A \text{ e } \det(M) = +1 \right\}.$$

• **Grupos de invariância de formas sesquilineares em dimensão finita**

Seja $\mathcal{E} = \mathbb{C}^n$ e seja ω_A uma forma sesquilinear em \mathbb{C}^n , que pelas considerações da Seção 3.4 é da forma $\omega_A(x, y) = \langle x, Ay \rangle_{\mathbb{C}}$ para alguma matriz complexa A . Neste caso $\Omega(\mathbb{C}^n, \omega_A)$ é o conjunto de todas as matrizes M inversíveis complexas $n \times n$ tais que

$$\langle Mx, AMy \rangle_{\mathbb{C}} = \langle x, Ay \rangle_{\mathbb{C}}$$

para todos $x, y \in \mathbb{C}^n$. Essa relação nos diz que $\langle x, M^*AMy \rangle_{\mathbb{C}} = \langle x, Ay \rangle_{\mathbb{C}}$ para todos $x, y \in \mathbb{C}^n$, o que implica

$$M^*AM = A.$$

Acima, $M^* = \overline{M^T}$. Assim,

$$\Omega(\mathbb{C}^n, \omega_A) = \left\{ M \in \text{GL}(n, \mathbb{C}), \det(M) \neq 0 \text{ e } M^*AM = A \right\}. \tag{21.48}$$

Se A possui uma inversa (ou seja, se ω_A for não degenerada), a relação $M^*AM = A$ implica que $|\det M| = 1$ (justifique!), mostrando que M é inversível. Nesse caso, portanto, é suficiente escrevermos

$$\Omega(\mathbb{C}^n, \omega_A) := \left\{ M \in \text{GL}(n, \mathbb{C}), M^*AM = A \right\}. \tag{21.49}$$

Se a matriz A for inversível (ou seja, se ω_A for não degenerada), então podemos escrever também

$$\Omega(\mathbb{C}^n, \omega_A) = \left\{ M \in \text{GL}(n, \mathbb{C}), M^{-1} = A^{-1}M^*A \right\}.$$

Caso A seja inversível o grupo $\Omega(\mathbb{C}^n, \omega_A)$ possui um subgrupo de particular interesse, composto por suas matrizes de determinante igual a 1:

$$\text{S}\Omega(\mathbb{C}^n, \omega_A) := \left\{ M \in \text{GL}(n, \mathbb{C}), M^{-1} = A^{-1}M^*A \text{ e } \det(M) = 1 \right\}.$$

• **Grupos de invariância de formas sesquilineares e bilineares em dimensão finita são grupos de Lie**

Façamos neste ponto um comentário talvez relevante a um estudante mais avançado: os grupos de invariância de formas bilineares ou sesquilineares não degeneradas em dimensão finita são grupos de Lie. Nas seções que seguiremos verificaremos de forma explícita essa asserção em diversos exemplos, mas é relevante saber da sua validade geral.

Proposição 21.3 *Os grupos $\Omega(\mathbb{C}^n, \omega_A)$ (ou $\Omega(\mathbb{R}^n, \omega_A)$), com $\det A \neq 0$, definidos acima, e que mantêm invariantes as formas sesquilineares (ou bilineares) definidas por A , são subgrupos fechados topologicamente de $\text{GL}(n, \mathbb{C})$ e, portanto (pelo Teorema 22.2, página 1177), são grupos de Lie. O mesmo se dá para os grupos $\text{S}\Omega(\mathbb{C}^n, \omega_A)$ (e $\text{S}\Omega(\mathbb{R}^n, \omega_A)$). □*

Postergamos a demonstração para a Seção 22.3.2, página 1167 (Proposição 22.6, página 1170).

* * * * *

Vamos no que segue estudar diversos casos particulares dentre os grupos descritos acima, que mantêm invariantes formas bilineares ou sesquilineares, especialmente aqueles mais relevantes em Física, na Geometria e em outras áreas, como os grupos ortogonais, os unitários e os simpléticos.

21.3.3.1 Os Grupos Ortogonais

Apresentamos agora a definição dos grupos ortogonais e ortogonais especiais. Mais detalhes sobre esses grupos serão apresentados nas seções subsequentes. Recordamos que, devido à Proposição 21.3, página 1038, todos esses grupos são topologicamente fechados e são grupos de Lie.

• **Os grupos $O(n)$ e $SO(n)$**

Um caso de particular interesse é aquele onde $\mathcal{E} = \mathbb{R}^n$ e $A = \mathbb{1}$, ou seja, $\omega_A(x, y) = \langle x, y \rangle_{\mathbb{R}}$. Neste caso o grupo $\Omega(\mathbb{R}^n, \omega_A)$ é denotado por $O(n)$ e tem-se

$$O(n) := \left\{ M \in \text{GL}(n, \mathbb{R}), M^{-1} = M^T \right\}.$$

$O(n)$ é o grupo das matrizes ditas ortogonais $n \times n$.

Se M é uma matriz ortogonal, tem-se que $MM^T = \mathbb{1}$. Daí, $1 = \det(\mathbb{1}) = \det(MM^T) = \det(M)\det(M^T) = (\det(M))^2$. Concluimos que se uma matriz M é ortogonal, vale $\det(M) = \pm 1$.

$O(n, \mathbb{C})$ não pode ser confundido com o grupo $U(n)$ das matrizes unitárias. É fácil ver também que se $M \in O(n, \mathbb{C})$, então $\det(M) = \pm 1$. Daí, define-se

$$SO(n, \mathbb{C}) := \left\{ M \in GL(n, \mathbb{C}) \mid M^T = M^{-1} \text{ e } \det(M) = 1 \right\}.$$

Como é fácil de se ver, $SO(n, \mathbb{C})$ é um subgrupo de $O(n, \mathbb{C})$. Esses grupos são denominados *grupos ortogonais especiais complexos*.

E. 21.35 *Exercício.* Mostre que

$$SO(2, \mathbb{C}) = \left\{ \begin{pmatrix} \cos z & \operatorname{sen} z \\ -\operatorname{sen} z & \cos z \end{pmatrix}, z \in \mathbb{C} \right\}.$$

Isso mostra, em particular, que $SO(2, \mathbb{C})$ não é um grupo compacto, em contraste (como veremos) com o grupo real $SO(2)$. Observe também que as matrizes $\begin{pmatrix} \cos z & \operatorname{sen} z \\ -\operatorname{sen} z & \cos z \end{pmatrix}$ com $z \in \mathbb{C}$ em geral não são unitárias. ✱

21.3.3.2 Os Grupos Unitários

Apresentamos agora a definição dos grupos unitários e unitários especiais. Mais detalhes sobre esses grupos serão apresentados nas seções subsequentes. Recordamos que, devido à Proposição 21.3, página 1038, todos esses grupos são topologicamente fechados e são grupos de Lie.

• Os grupos $U(n)$ e $SU(n)$

Mais um caso importante é aquele onde $\mathcal{E} = \mathbb{C}^n$ e ω_A é a forma sesquilinear associada a $A = \mathbb{1}$, ou seja, $\omega_A(x, y) = \langle x, y \rangle_{\mathbb{C}}$. Neste caso o grupo $\Omega(\mathbb{C}^n, \omega_A)$ é denotado por $U(n)$ e tem-se

$$U(n) := \{ U \in GL(n, \mathbb{C}), U^{-1} = U^* \}.$$

$U(n)$ é o grupo das matrizes ditas *unitárias* $n \times n$.

Se U é uma matriz unitária, tem-se que $UU^* = \mathbb{1}$. Daí,

$$1 = \det(\mathbb{1}) = \det(UU^*) = \det(U) \det(U^*) = \det(U) \det(\overline{U^T}) = \det(U) \overline{\det(U^T)} = \det(U) \overline{\det(U)} = |\det(U)|^2.$$

Concluimos que se $U \in U(n)$, vale $|\det(U)| = 1$.

$U(n)$ possui um subgrupo, denotado por $SU(n)$, que é composto pelas matrizes unitárias com determinante igual a 1:

$$SU(n) := \left\{ U \in GL(n, \mathbb{C}), U^{-1} = U^* \text{ e } \det(U) = 1 \right\}.$$

Esses grupos são denominados *grupos unitários especiais*.

Os grupos $SO(n)$ e $SU(n)$ serão estudados com alguma generalidade na Seção 21.5, página 1087. Outras seções anteriores tratarão de grupos de interesse particular, como os grupos $SO(3)$ e $SU(2)$, os quais desempenham um papel muito importante na Mecânica Quântica e na Física das Partículas Elementares. Na Seção 21.5 veremos, por exemplo, que os elementos do grupo $SO(n)$ são caracterizados por $n(n-1)/2$ parâmetros reais e os elementos de $SU(n)$ por $n^2 - 1$ parâmetros reais. Disso segue que os grupos $SO(3)$ e $SU(2)$ são caracterizados pelo mesmo número de parâmetros reais, a saber, 3. Mais adiante entenderemos o porquê dessa coincidência.

• Os grupos $U(p, n)$ e $SU(p, n)$

Mais um caso é aquele onde $\mathcal{E} = \mathbb{C}^{p+n}$, $p, n \in \mathbb{N}_0$, e $\omega(x, y) = \langle x, \eta(p, n)y \rangle_{\mathbb{C}}$, onde $\eta(p, n)$ foi definida em (21.50), sendo $m = p + n$. Neste caso o grupo $\Omega(\mathbb{C}^{p+n}, \omega)$ é denotado por $U(p, n)$ e tem-se

$$U(p, n) := \left\{ M \in GL(p+n, \mathbb{C}), M^{-1} = \eta(p, n)M^*\eta(p, n) \right\}.$$

Se $M \in U(p, n)$, tem-se que $M\eta(p, n)M^*\eta(p, n) = \mathbb{1}$. Daí,

$$1 = \det(\mathbb{1}) = \det(M\eta(p, n)M^*\eta(p, n)) = \det(M)\det(M^*)\left(\det(\eta(p, n))\right)^2 = \det(M)\det(\overline{M^T}) = \det(M)\overline{\det(M^T)} = \det(M)\overline{\det(M)} = |\det(M)|^2.$$

Concluimos que se $M \in U(p, n)$, vale $|\det(M)| = 1$.

$U(p, n)$ possui um subgrupo, denominado $SU(p, n)$, que é composto pelas matrizes de $U(p, n)$ com determinante igual a 1:

$$SU(p, n) := \left\{ M \in GL(p+n, \mathbb{C}), M^{-1} = \eta(p, n)M^*\eta(p, n) \text{ e } \det(M) = 1 \right\}.$$

21.3.3.3 Os Grupos Simpléticos

Faremos agora uma primeira apresentação dos chamados *grupos simpléticos*¹⁶. Seu tratamento será continuado com resultados mais técnicos na Seção 21.9, página 1133, à qual o leitor poderá passar sem perdas ao término da presente.

Em termos gerais, grupos simpléticos são grupos que mantêm invariantes *formas simpléticas*, ou seja, formas bilineares alternantes (*i.e.*, antissimétricas) e não degeneradas (noção introduzida à página 221). Assim, uma forma simplética em um espaço vetorial \mathcal{E} sobre um corpo \mathbb{K} é uma forma bilinear para a qual $\omega(u, v) = -\omega(v, u)$ para todos os vetores $u, v \in \mathcal{E}$ e tal que se $\omega(u, v) = 0$ para todo $v \in \mathcal{E}$, então $u = 0$.

Grupos simpléticos desempenham um papel importante em tratamentos matemáticos da Mecânica Clássica (como em [17], [2] ou [129]) e também na Teoria da Informação Quântica. A teoria dos grupos simpléticos está na base da área da Matemática denominada *Geometria Simplética*, a qual recebeu grandes impulsos nas últimas décadas.

Para usos do grupo simplético na Mecânica Clássica, vide a discussão sobre transformações canônicas na Seção 44.5.4.1, página 2446. Para usos do grupo simplético na Mecânica Quântica, vide a discussão sobre as Relações de Incerteza (na versão de Schrödinger) na Seção 46.4, página 2486, e, especialmente, na Seção 46.4.3.1, página 2496.

Tendo essas motivações em mente, estaremos no que segue interessados no casos em que $\mathcal{E} = \mathbb{R}^{2n}$ ou \mathbb{C}^{2n} . Como discutido à página 221 toda forma simplética ω nesse caso é da forma

$$\omega(u, v) = \langle u, Av \rangle_{\mathbb{R}}$$

onde A é uma matriz $(2n) \times (2n)$ antissimétrica ($A^T = -A$) e inversível $\det(A) \neq 0$. A discussão à página 221 esclarece o porquê de a dimensão de \mathcal{E} ter de ser par.

De especial interesse nesse contexto são as matrizes $(2n) \times (2n)$ dadas por

$$J_{2n} := \begin{pmatrix} \mathbb{0}_n & \mathbb{1}_n \\ -\mathbb{1}_n & \mathbb{0}_n \end{pmatrix},$$

onde $\mathbb{1}_n$ e $\mathbb{0}_n$ são a matriz identidade e a matriz identicamente nula $n \times n$, respectivamente. Elas são antissimétricas, $J_{2n}^T = -J_{2n}$, e inversíveis com $J_{2n}^{-1} = -J_{2n}$, sendo ainda $\det J_{2n} = 1$, como facilmente se vê. Dessa forma, as matrizes J_{2n} definem formas simpléticas em \mathbb{C}^{2n} e \mathbb{R}^{2n} . Com tais matrizes podemos definir protótipos de grupos simpléticos.

Definimos o grupo simplético $Sp(2n, \mathbb{C})$ com $n \in \mathbb{N}$, como o grupo composto pelas matrizes inversíveis de $Mat(\mathbb{C}, 2n)$ que mantêm invariante a forma simplética $\omega(u, v) := \langle u, J_{2n}v \rangle_{\mathbb{R}}$ definida em \mathbb{C}^{2n} . Assim, pelas nossas considerações anteriores,

$$Sp(2n, \mathbb{C}) = \left\{ A \in GL(\mathbb{C}, 2n) \mid A^T J_{2n} A = J_{2n} \right\}.$$

Analogamente, definimos

$$Sp(2n, \mathbb{R}) := \left\{ A \in GL(\mathbb{R}, 2n) \mid A^T J_{2n} A = J_{2n} \right\}.$$

¹⁶Do grego *symplektikós*: que serve para ligar, trançado, enlaçado. O termo foi cunhado por Hermann Klaus Hugo Weyl (1885–1955) em [405].

O grupo $\text{Sp}(2n, \mathbb{R})$ (ou $\text{Sp}(2n, \mathbb{C})$) é denominado *grupo simplético real (complexo)* e seus elementos são denominados *matrizes simpléticas* (reais ou complexas). Ambos os grupos são também denominados *grupos simpléticos não compactos*, para contrastá-los com certos subgrupos compactos que serão introduzidos logo adiante.

Façamos agora alguns pequenos comentários sobre as definições acima.

1. De acordo com a Proposição 22.6, página 1170, os grupos simpléticos $\text{Sp}(2n, \mathbb{R})$ e $\text{Sp}(2n, \mathbb{C})$ são grupos de Lie.
2. É relevante notarmos que se $A \in \text{Mat}(\mathbb{R}, 2n)$, então também $A \in \text{Mat}(\mathbb{C}, 2n)$, pois $\text{Mat}(\mathbb{R}, 2n)$ é um subespaço real de $\text{Mat}(\mathbb{C}, 2n)$. Assim, pelas definições, podemos considerar $\text{Sp}(2n, \mathbb{R})$ como um subgrupo de $\text{Sp}(2n, \mathbb{C})$. Vamos frequentemente fazê-lo.
3. O leitor deve ser advertido quanto ao fato de que o grupo $\text{Sp}(2n, \mathbb{C})$ é por vezes denotado como $\text{Sp}(n, \mathbb{C})$, como $\text{Sp}(2n)$ ou até mesmo como $\text{Sp}(n)$, em certos textos. Verdade seja dita, o estudo de grupos simpléticos padece muito com a falta de uniformidade de definições e de notações. Cuidado se faz necessário, portanto, ao se comparar resultados de textos diferentes.
4. Tanto no caso real ou complexo, a condição $A^T J_{2n} A = J_{2n}$ equivale a $J_{2n} A = (A^{-1})^T J_{2n}$. Tomando-se a inversa de ambos os lados, temos $A^{-1} J_{2n} = J_{2n} A^T$ e, portanto,

$$A J_{2n} A^T = J_{2n}, \tag{21.51}$$

ou seja. $(A^T)^T J_{2n} A^T = J_{2n}$. Isso mostrou que se $A \in \text{Sp}(2n, \mathbb{C})$, então $A^T \in \text{Sp}(2n, \mathbb{C})$ e vice-versa. Essas afirmações claramente também valem com \mathbb{C} substituído por \mathbb{R} .

5. Observe-se que se $A \in \text{Sp}(2n, \mathbb{C})$, então $\bar{A} \in \text{Sp}(2n, \mathbb{C})$, pois J_{2n} é uma matriz real.
6. Os dois últimos itens informam que se $A \in \text{Sp}(2n, \mathbb{C})$ então $A^* \in \text{Sp}(2n, \mathbb{C})$.
7. A condição $A^T J_{2n} A = J_{2n}$ implica $A^T J_{2n} A J_{2n} = -\mathbb{1}_{2n}$ e, portanto, implica

$$(A^T)^{-1} = -J_{2n} A J_{2n}. \tag{21.52}$$

8. Vale notar também que $J_{2n} \in \text{Sp}(2n, \mathbb{R})$, já que $J_{2n}^T J_{2n} J_{2n} = J_{2n}$, pois $J_{2n}^T = J_{2n}^{-1}$.
9. É importante também notar que se $A \in \text{Sp}(2n, \mathbb{C})$, então a condição $A^T J_{2n} A = J_{2n}$ implica $(\det A)^2 = 1$ e, portanto, $\det A = \pm 1$. A Proposição 21.33, página 1133, mostra, porém, que apenas o caso $\det A = +1$ é admissível em grupos simpléticos!
10. O fato que $\det(A) = 1$ para todo $A \in \text{Sp}(2n, \mathbb{C})$ significa que $\text{Sp}(2n, \mathbb{C})$ é um subgrupo de $\text{SL}(2n, \mathbb{C})$. Veremos mais adiante que no caso $n = 1$ vale até $\text{Sp}(2, \mathbb{C}) = \text{SL}(2, \mathbb{C})$. No caso $n \geq 2$, $\text{Sp}(2n, \mathbb{C})$ é um subgrupo próprio de $\text{SL}(2n, \mathbb{C})$. Para ver isso, considere-se as matrizes diagonais (caso $n = 2$) da forma $A = \text{diag}(\alpha, \alpha^{-1}, \alpha, \alpha^{-1})$, com $\alpha \neq 0$. A matriz A tem determinante 1, mas não é simplética para $\alpha \neq 1$, como é fácil verificar. Exemplos como esse podem ser produzidos para todo $n \geq 2$.

E. 21.36 Exercício. Para $n = 2$, considere as matrizes diagonais da forma $\text{diag}(\alpha, \beta, \alpha^{-1}, \beta^{-1})$, com $\alpha, \beta \neq 0$. Verifique que elas são simpléticas. *

• O significado geométrico dos grupos simpléticos $\text{Sp}(2n, \mathbb{R})$

Antes de prosseguirmos apresentando propriedades algébricas dos grupos simpléticos é importante apresentarmos algumas ideias sobre o significado geométrico da ação do grupo simplético $\text{Sp}(2n, \mathbb{R})$ em vetores de \mathbb{R}^{2n} .

Sabemos que os grupos ortogonais podem ser entendidos como grupos de transformações que preservam os produtos escalares entre pares de vetores (e, portanto, as normas de vetores). Esse é, em essência, o significado geométrico desses grupos. Os grupos simpléticos possuem uma interpretação geométrica diferente: eles são grupos de transformações que, num sentido a ser melhor precisado, preservam áreas varridas por pares de vetores.

Vejamos isso primeiro no caso mais simples para depois tratar do caso geral. Considere-se o espaço vetorial \mathbb{R}^2 e considere-se dois vetores¹⁷ $v = \begin{pmatrix} q \\ p \end{pmatrix}$ e $v' = \begin{pmatrix} q' \\ p' \end{pmatrix} \in \mathbb{R}^2$. A área $A(v, v')$ varrida por ambos, descrita graficamente na Figura 21.1, página 1043, é sabidamente dada por

$$A(v, v') = pq' - qp' = \begin{pmatrix} q' & p' \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \langle v', J_2 v \rangle_{\mathbb{R}}.$$

Vemos com isso que essa área é preservada pelas transformações lineares do grupo simplético $\text{Sp}(2, \mathbb{R})$, ou seja, pelas transformações $v \mapsto Av, v' \mapsto Av'$, para uma mesma matriz $A \in \text{Sp}(2, \mathbb{R})$.

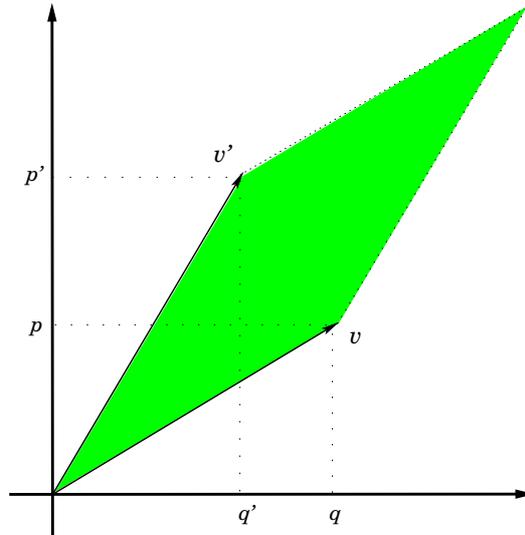


Figura 21.1: A área $A(v, v') = pq' - qp'$ varrida pelos vetores $v, v' \in \mathbb{R}^2$ com coordenadas (q, p) e (q', p') , respectivamente.

Consideremos agora o caso geral do espaço vetorial \mathbb{R}^{2n} , $n \in \mathbb{N}$. Um vetor genérico $v \in \mathbb{R}^{2n}$ é da forma

$$v = \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix}, \quad q_i, p_i \in \mathbb{R}, \forall i \in \{1, \dots, n\}.$$

Um tal vetor pode ser escrito de modo único como

$$v = v_1 + \dots + v_n,$$

onde v_j é o vetor cuja j -ésima e $(n + j)$ -ésima componentes são q_j e p_j , respectivamente:

$$v_1 = \begin{pmatrix} q_1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ p_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ q_2 \\ 0 \\ \vdots \\ 0 \\ p_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad v_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ q_n \\ 0 \\ 0 \\ \vdots \\ p_n \end{pmatrix}.$$

¹⁷O uso das letras q e p para denotar as componentes provém da Mecânica Clássica.

Vemos com isso que \mathbb{R}^{2n} pode ser escrito como soma direta de n subespaços bidimensionais: $\mathbb{R}^{2n} = F_1^{2n} \oplus \dots \oplus F_n^{2n}$, onde F_k^{2n} é o espaço gerado pelos vetores v_k , acima, ou seja, por vetores que possuem todas as componentes nulas, exceto possivelmente a k -ésima e a $(n+k)$ -ésima.

Cada subespaço bidimensional F_k^{2n} é dito ser um *plano conjugado*, ou *espaço de componentes canonicamente conjugadas*, novamente uma nomenclatura originária da Mecânica Clássica.

Um ponto simples mas muito importante é a seguinte observação: se $w_k \in F_k^{2n}$ e $v_l \in F_l^{2n}$ com $k \neq l$, então

$$\langle w_k, J_{2n}v_l \rangle_{\mathbb{R}} = 0.$$

De fato, sejam $q_l, p_l \in \mathbb{R}$ a l -ésima e a $(n+l)$ -ésima componentes, respectivamente, de v_l , todas as demais sendo nulas. Então, devido à forma da matriz J_{2n} , as únicas componentes eventualmente não nulas de $J_{2n}v_l$ são a l -ésima, que vale p_l , e a $(n+l)$ -ésima, que vale $-q_l$. Como, porém, $k \neq l$, essas mesmas componentes são nulas em w_k , provando a afirmação.

Colhemos disso a seguinte informação útil: se $v, w \in \mathbb{R}^{2n}$ possuem as decomposições $v = v_1 + \dots + v_n$ e $w = w_1 + \dots + w_n$, com $v_k, w_k \in F_k^{2n}$, para todo $k \in \{1, \dots, n\}$, então

$$\langle w, J_{2n}v \rangle_{\mathbb{R}} = \sum_{k=1}^n \langle w_k, J_{2n}v_k \rangle_{\mathbb{R}}.$$

Em palavras, a forma simplética definida por J_{2n} é diretamente decomposta pela soma direta $\mathbb{R}^{2n} = F_1^{2n} \oplus \dots \oplus F_n^{2n}$.

Chegamos agora ao fulcro da questão. Cada termo da soma acima é

$$\langle w_k, J_{2n}v_k \rangle_{\mathbb{R}} = (v_k)_{n+k}(w_k)_k - (v_k)_k(w_k)_{n+k}$$

que é a área varrida no plano F_k^{2n} pela projecção nele dos vetores v e w . Por definição, cada termo $\langle w_k, J_{2n}v_k \rangle_{\mathbb{R}}$ é invariante por transformações lineares $v_k \mapsto Av_k, w_k \mapsto Aw_k$ com o mesmo $A \in \text{Sp}(\mathbb{R}, 2n)$.

Concluimos que a área varrida pelas projeções dos vetores v e w em cada plano $F_k^{2n}, k = 1, \dots, n$, é preservada pelas transformações lineares $v \mapsto Av, w \mapsto Aw$ com o mesmo $A \in \text{Sp}(\mathbb{R}, 2n)$. Essa propriedade de preservação das áreas das projeções de pares de vetores sobre os planos conjugados $F_k^{2n}, k = 1, \dots, n$, é a propriedade geométrica característica do grupo simplético $\text{Sp}(\mathbb{R}, 2n)$, generalizando assim nossa interpretação anterior sobre o significado geométrico do grupo $\text{Sp}(\mathbb{R}, 2)$.



Todos os comentários acima, ainda que aparentemente simples, são fundamentais para diversos desenvolvimentos mais profundos da Geometria Simplética e da Mecânica Clássica, como o célebre *Teorema de Gromow*¹⁸ (também conhecido como “*teorema do camelo simplético*”), tema intimamente ligado ao princípio da incerteza da Física Quântica, como pode ser encontrado na literatura pertinente¹⁹.

• **A forma em blocos de matrizes simpléticas**

Cada matriz $A \in \text{Sp}(2n, \mathbb{C})$ pode ser escrita em forma de blocos como

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{sendo que} \quad A^T = \begin{pmatrix} a^T & c^T \\ b^T & d^T \end{pmatrix}, \tag{21.53}$$

onde a, b, c e d são matrizes $n \times n$. É um exercício fácil constatar que a condição de A ser simplética, a saber, $A^T J_{2n} A = J_{2n}$, equivale às seguintes relações:

$$a^T c = c^T a, \tag{21.54}$$

$$b^T d = d^T b, \tag{21.55}$$

$$a^T d - c^T b = \mathbb{1}_n. \tag{21.56}$$

¹⁸Mikhail Leonidovich Gromov (1943–).

¹⁹Para uma exposição geral, vide Maurice A. de Gosson, “The Symplectic Camel and the Uncertainty Principle: The Tip of an Iceberg?” *Foundations of Physics* **39**, Issue 2, pp 194–214 (2009) e outras referências lá citadas. Vide também I. Stewart, “The symplectic camel”. *Nature* **329**, 17–18 (1987). Esse último artigo revela a origem da expressão “teorema do camelo simplético” e sua anedótica relação com uma famosa parábola bíblica.

Adicionalmente, obtém-se também as relação $d^T a - b^T c = \mathbb{1}_n$, mas essa última relação é decorrência de (21.56), por transposição.

E. 21.37 *Exercício importante.* Verifique a validade das relações (21.54)–(21.56)! ✦

Como A^T é simplética se A o for, valem relações similares decorrentes da condição $AJ_{2n}A^T = J_{2n}$. Elas podem ser obtidas por cômputo direto, ou a partir de (21.54)–(21.56), procedendo-se as trocas $a \rightarrow a^T$, $b \rightarrow c^T$, $c \rightarrow b^T$ e $d \rightarrow d^T$. Com isso, obtemos as seguintes relações adicionais (também equivalentes a A ser simplética):

$$ab^T = ba^T, \tag{21.57}$$

$$cd^T = dc^T, \tag{21.58}$$

$$ad^T - bc^T = \mathbb{1}_n. \tag{21.59}$$

E. 21.38 *Exercício importante.* Verifique! ✦

As relações (21.54)–(21.56) e (21.57)–(21.59) são importantes em certos desenvolvimentos da teoria das matrizes simpléticas e encontraremos alguns adiante. Um resultado que pode ser obtido delas e de (21.52) é uma expressão para a matriz inversa A^{-1} :

$$A^{-1} = \begin{pmatrix} d^T & -b^T \\ -c^T & a^T \end{pmatrix}. \tag{21.60}$$

De fato,

$$\begin{pmatrix} d^T & -b^T \\ -c^T & a^T \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d^T a - b^T c & d^T b - b^T d \\ -c^T a + a^T c & -c^T b + a^T d \end{pmatrix} \stackrel{(21.54)-(21.56)}{=} \begin{pmatrix} \mathbb{1}_n & 0_n \\ 0_n & \mathbb{1}_n \end{pmatrix} = \mathbb{1}_{2n}.$$

É relevante frisar que as condições (21.54)–(21.56) ou (21.57)–(21.59) são necessárias e suficientes para que $A \in \text{Sp}(2n, \mathbb{C})$.

Um caso particular que merece comentário é aquele em que $n = 1$. Nessa situação a , b , c e d são matrizes complexas 1×1 , ou seja, são apenas números complexos. A duas primeiras condições em (21.54)–(21.56) ou (21.57)–(21.59) são triviais, mas as últimas afirmam que $ad - cb = 1$. Ora, isso coincide com a afirmação que $\det A = \det \begin{pmatrix} a & c \\ c & d \end{pmatrix} = 1$. Vemos com isso que $A \in \text{Sp}(2, \mathbb{C})$ se e somente se $A \in \text{SL}(2, \mathbb{C})$. Assim, estabelecemos que

$$\text{Sp}(2, \mathbb{C}) = \text{SL}(2, \mathbb{C}). \tag{21.61}$$

Analogamente,

$$\text{Sp}(2, \mathbb{R}) = \text{SL}(2, \mathbb{R}). \tag{21.62}$$

• **Os grupos simpléticos compactos**

Além dos grupos simpléticos $\text{Sp}(2n, \mathbb{C})$ e $\text{Sp}(2n, \mathbb{R})$ introduzidos acima, há também interesse em alguns de seus subgrupos. O mais notório é

$$\text{Sp}(2n) := \text{Sp}(2n, \mathbb{C}) \cap \text{U}(2n),$$

o subgrupo das matrizes unitárias de $\text{Sp}(2n, \mathbb{C})$. Como comentaremos, trata-se de um grupo compacto (por ser um subgrupo fechado do grupo compacto $\text{U}(2n)$). O grupo $\text{Sp}(2n)$ é denominado *grupo simplético unitário*, ou *grupo simplético compacto*²⁰.

É relevante observar que $\text{Sp}(2n, \mathbb{C}) \cap \text{U}(2n) = \text{Sp}(2n, \mathbb{C}) \cap \text{SU}(2n)$ pois, como já afirmamos, o determinante de matrizes simpléticas é sempre igual a 1.

²⁰Fazemos novamente notar que, lamentavelmente, diversas outras denominações e notações são oferecidas ao grupo $\text{Sp}(2n)$ na literatura.

Note-se que, devido a (21.61), tem-se

$$\mathrm{Sp}(2) = \mathrm{SU}(2), \tag{21.63}$$

pois $\mathrm{SL}(2, \mathbb{C}) \cap \mathrm{U}(2) = \mathrm{SU}(2)$.

Analogamente, podemos considerar o grupo das matrizes simpléticas reais ortogonais:

$$\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n).$$

Novamente, é relevante observar que $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) = \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{SO}(2n)$ pois, como já afirmamos, o determinante de matrizes simpléticas é sempre igual a 1.

Similarmente a (21.63),

$$\mathrm{Sp}(2, \mathbb{R}) \cap \mathrm{O}(2) = \mathrm{SO}(2), \tag{21.64}$$

pois $\mathrm{SL}(2, \mathbb{R}) \cap \mathrm{O}(2) = \mathrm{SO}(2)$.

O grupo $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ é relevante em aplicações. Porém, de um ponto de vista estrutural, esse grupo não manifesta ser uma novidade, pois é isomorfo a $\mathrm{U}(n)$, o grupo das matrizes unitárias agindo em \mathbb{C}^n , isto é, $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) \simeq \mathrm{U}(n)$. Por exemplo, vê-se de (21.64) que $\mathrm{Sp}(2, \mathbb{R}) \cap \mathrm{O}(2) \simeq \mathrm{U}(1)$, pois $\mathrm{SO}(2)$ e $\mathrm{U}(1)$ são grupos isomorfos.

O isomorfismo entre $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ e $\mathrm{U}(n)$ é parte do conteúdo do Teorema 21.8, a seguir, que será revisitado na Seção 21.9.1, página 1134, (Proposição 21.34, página 1137).

Antes de apresentarmos esse teorema, façamos uma observação. Cada elemento A de $\mathrm{Sp}(2n, \mathbb{R})$ satisfaz $A^T J_{2n} A = J_{2n}$. Se adicionalmente A for ortogonal, ou seja, se também valer $A^T = A^{-1}$, teremos $J_{2n} A = A J_{2n}$. Assim, podemos escrever

$$\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) = \{A \in \mathrm{Mat}(\mathbb{R}, 2n) \mid A J_{2n} = J_{2n} A \text{ e } A^T = A^{-1}\}. \tag{21.65}$$

Teorema 21.8 *Todo elemento de $A \in \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ assume a forma em blocos*

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

com $a, b \in \mathrm{Mat}(n, \mathbb{R})$ satisfazendo

$$ab^T = ba^T, \tag{21.66}$$

$$a^T a + b^T b = \mathbb{1}_n, \tag{21.67}$$

$$a^T b = b^T a \quad e \tag{21.68}$$

$$aa^T + bb^T = \mathbb{1}_n. \tag{21.69}$$

Além disso, $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ e $\mathrm{U}(n)$ são grupos isomorfos, com o isomorfismo $\mathcal{U} : \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) \rightarrow \mathrm{U}(n)$ dado por

$$\mathcal{U} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} := a + bi,$$

com a e $b \in \mathrm{Mat}(n, \mathbb{R})$ satisfazendo (21.66)–(21.69). □

A afirmação do Teorema 21.8 sobre o isomorfismo entre $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ e $\mathrm{U}(n)$ é um resultado profundo da teoria dos grupos simpléticos. Na Seção 21.9.1, página 1134, apresentaremos uma outra demonstração dessa afirmação que torna transparente o fato de o isomorfismo entre os grupos $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ e $\mathrm{U}(n)$ decorrer de uma identificação entre o espaço vetorial real \mathbb{R}^{2n} com o espaço vetorial complexo \mathbb{C}^n que respeita a estrutura complexa.

Demonstração do Teorema 21.8. Na sua representação em blocos, os elementos de $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ são matrizes A , reais $(2n) \times (2n)$, da forma (21.53) com a, b, c e d satisfazendo (21.54)–(21.56) ou (21.57)–(21.59). Por serem também ortogonais, a condição $A^T A = \mathbb{1}_{2n}$ implica ainda as relações

$$a^T a + c^T c = \mathbb{1}_n, \tag{21.70}$$

$$b^T b + d^T d = \mathbb{1}_n, \tag{21.71}$$

$$a^T b = -c^T d. \tag{21.72}$$

Verifique! As condições (21.54)–(21.56), (21.57)–(21.59) e (21.70)–(21.72) juntas implicam

$$b \stackrel{(21.70)}{=} b(a^T a + c^T c) = (ba^T)a + (bc^T)c \stackrel{(21.57)}{=} ab^T a + (bc^T)c \stackrel{(21.59)}{=} ab^T a + (ad^T - \mathbb{1}_n)c \stackrel{(21.72)}{=} ab^T a - ab^T a - c = -c.$$

Além disso,

$$d \stackrel{(21.70)}{=} d(a^T a + c^T c) = (da^T)a + dc^T c \stackrel{(21.59)}{=} (\mathbb{1}_n + cb^T)a + dc^T c \stackrel{(21.58)}{=} a + cb^T a + cd^T c = a + c(b^T a + d^T c) \stackrel{(21.72)}{=} a.$$

Com isso, concluímos que todo elemento de $A \in \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ assume a forma em blocos

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

com

$$ab^T = ba^T \quad \text{e} \tag{21.73}$$

$$a^T a + b^T b = \mathbb{1}_n, \tag{21.74}$$

devido a (21.57) e a (21.70) (dado que $b = -c$), respectivamente. Como $A^T = \begin{pmatrix} a^T & -b^T \\ b^T & a^T \end{pmatrix}$ é também elemento de $A \in \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ obtemos de (21.73)–(21.74) através das substituições $a \mapsto a^T, b \mapsto -b^T$ que valem adicionalmente as relações

$$a^T b = b^T a \quad \text{e} \tag{21.75}$$

$$aa^T + bb^T = \mathbb{1}_n. \tag{21.76}$$

Sob as condições (21.73)–(21.76) é fácil constatar que a matrizes da forma $a+ib$ são unitárias. De fato, $(a+ib)^*(a+ib) = (a^T - ib^T)(a+ib) = (a^T a + b^T b) + i(a^T b - b^T a)$. Agora, por (21.74) temos que $a^T a + b^T b = \mathbb{1}_n$ e por (21.75) temos $a^T b - b^T a = 0_n$, o que estabelece que $(a+ib)^*(a+ib) = \mathbb{1}_n$ e prova que $a+ib \in \text{U}(n)$.

Defina-se, dessa forma, $\mathcal{U} : \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n) \rightarrow \text{U}(n)$ por

$$\mathcal{U} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} := a + bi.$$

Afirmamos que \mathcal{U} é um isomorfismo entre esses grupos. Para vermos que \mathcal{U} é um homomorfismo de grupos, tomemos $A, A' \in \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ na forma de blocos $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ e $A' = \begin{pmatrix} a' & b' \\ -b' & a' \end{pmatrix}$, como acima. Teremos que

$$AA' = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} a' & b' \\ -b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & ab' + ba' \\ -(ab' + ba') & aa' - bb' \end{pmatrix}$$

e, portanto,

$$\mathcal{U}(AA') = (aa' - bb') + i(ab' + ba') = (a + ib)(a' + ib') = \mathcal{U}(A)\mathcal{U}(A').$$

Isso provou que \mathcal{U} é um homomorfismo de grupos. Vamos agora provar que \mathcal{U} é injetora e sobrejetora. A injetividade é evidente, pois se $\mathcal{U}(A) = \mathcal{U}(A')$, então $a + ib = a' + ib'$. Como as quatro matrizes a , a' , b e b' são reais, isso implica $a = a'$ e $b = b'$ e, portanto, $A = A'$. Para ver que \mathcal{U} é sobrejetora, notemos que se $U \in \mathbf{U}(n)$, então podemos escrever $U = U_R + iU_I$, onde $U_R, U_I \in \text{Mat}(n, \mathbb{R})$ são as matrizes reais cujos elementos de matriz são as parte real e imaginária, respectivamente, dos elementos de matriz de U : $(U_R)_{ij} = \text{Re}(U_{ij})$ e $(U_I)_{ij} = \text{Im}(U_{ij})$. Assim,

$$U = \mathcal{U} \begin{pmatrix} U_R & U_I \\ -U_I & U_R \end{pmatrix}. \quad (21.77)$$

Como U é unitária, valem $U^*U = \mathbb{1}_n$ e $UU^* = \mathbb{1}_n$, ou seja $(U_R^T - iU_I^T)(U_R + iU_I) = \mathbb{1}_n$ e $(U_R + iU_I)(U_R^T - iU_I^T) = \mathbb{1}_n$. Portanto,

$$U_R^T U_R + U_I^T U_I = \mathbb{1}_n, \quad (21.78)$$

$$U_R^T U_I = U_I^T U_R, \quad (21.79)$$

$$U_R U_R^T + U_I U_I^T = \mathbb{1}_n, \quad (21.80)$$

$$U_R U_I^T = U_I U_R^T. \quad (21.81)$$

Uma simples inspeção revela que essas quatro condições coincidem com as condições (21.73)–(21.76) e que, portanto, $\begin{pmatrix} U_R & U_I \\ -U_I & U_R \end{pmatrix}$ é um elemento de $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$. Assim, (21.77) revela que todo elemento de $\mathbf{U}(n)$ é a imagem por \mathcal{U} de um elemento de $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$, estabelecendo que \mathcal{U} é sobrejetora. Isso completou a prova que \mathcal{U} é um isomorfismo de grupos. ■

A relação (21.77) informa-nos que o isomorfismo inverso $\mathcal{U}^{-1} : \mathbf{U}(n) \rightarrow \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ é dado por

$$\mathcal{U}^{-1}(U) = \begin{pmatrix} U_R & U_I \\ -U_I & U_R \end{pmatrix}, \quad U \in \mathbf{U}(n), \quad (21.82)$$

onde $U = U_R + iU_I$, onde $U_R, U_I \in \text{Mat}(n, \mathbb{R})$ são as matrizes reais cujos elementos de matriz são as parte real e imaginária, respectivamente, dos elementos de matriz de U , $(U_R)_{ij} = \text{Re}(U_{ij})$ e $(U_I)_{ij} = \text{Im}(U_{ij})$, e satisfazem (21.78)–(21.81).

• Homomorfismos de $\text{Sp}(2n, \mathbb{R})$ em $\text{Sp}(2n, \mathbb{C})$

Para $U \in \text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$, fixo, defina-se

$$\mathbf{Ad}_U(A) = U^*AU, \quad A \in \text{Mat}(\mathbb{R}, 2n).$$

Tem-se $\mathbf{Ad}_U(AA') = U^*AA'U = (U^*AU)(U^*A'U) = \mathbf{Ad}_U(A)\mathbf{Ad}_U(A')$ para todos $A, A' \in \text{Mat}(\mathbb{R}, 2n)$.

Afirmamos que cada \mathbf{Ad}_U é um homomorfismo de $\text{Sp}(2n, \mathbb{R})$ em $\text{Sp}(2n, \mathbb{C})$, assim como é um homomorfismo de $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ em $\text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$.

A primeira afirmação decorre trivialmente do fato que $\text{Sp}(2n, \mathbb{R})$ e $\text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$ são subgrupos de $\text{Sp}(2n, \mathbb{C})$, o que implica que $U^*AU \in \text{Sp}(2n, \mathbb{C})$ se $A \in \text{Sp}(2n, \mathbb{R})$ e $U \in \text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$.

A segunda afirmação decorre trivialmente do fato que $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ ser um subgrupo de $\text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$, o que implica que $U^*AU \in \text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$ se $A \in \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ e $U \in \text{Sp}(2n, \mathbb{C}) \cap \mathbf{U}(2n)$.

E. 21.39 *Exercício.* Mostre que a matriz $K \in \text{Mat}(\mathbb{C}, 2n)$ dada por

$$K := \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1}_n & i\mathbb{1}_n \\ i\mathbb{1}_n & \mathbb{1}_n \end{pmatrix}$$

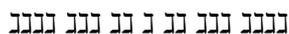
é um elemento de $\text{Sp}(2n, \mathbb{C}) \cap \text{U}(2n)$ e mostre que para $U \in \text{U}(n)$ tem-se

$$\text{Ad}_K(\mathcal{U}^{-1}(U)) = \begin{pmatrix} U^* & \mathbb{0}_n \\ \mathbb{0}_n & U \end{pmatrix},$$

onde \mathcal{U}^{-1} é o isomorfismo dado em (21.82). Verifique explicitamente que o lado direito é elemento de $\text{Sp}(2n, \mathbb{C}) \cap \text{U}(2n)$. Vê-se com isso que $\text{Ad}_K \circ \mathcal{U}^{-1} : \text{U}(n) \mapsto \text{Sp}(2n, \mathbb{C}) \cap \text{U}(2n)$ é um homomorfismo. *

• **Representações irredutíveis dos grupos simpléticos**

Um estudo das representações irredutíveis dos grupos simpléticos pode ser encontrado em [156].



O estudo de grupos simpléticos será continuado na Seção 21.9, página 1133, à qual o leitor poderá passar sem perdas.

21.4 Os Grupos $\text{SO}(2)$, $\text{SO}(3)$, $\text{SU}(2)$ e $\text{SL}(2, \mathbb{C})$

Em função de sua particular importância na Física, em especial na Física Quântica, vamos discutir aqui com algum detalhe os grupos $\text{SO}(3)$, $\text{SU}(2)$ e $\text{SL}(2, \mathbb{C})$, os quais, ademais, como veremos, são intimamente relacionados. Por razões pedagógicas, ilustraremos o estudo dos grupos $\text{SO}(3)$ e $\text{SU}(2)$ tratando antes do grupo $\text{SO}(2)$.

21.4.1 Os Grupos $\text{SO}(2)$, $\text{O}(2)$, $\text{SO}(1, 1)$ e $\text{O}(1, 1)$

Vamos nesta seção conhecer pessoalmente alguns grupos de matrizes 2×2 de grande interesse.

• **Os grupos $\text{SO}(2)$ e $\text{O}(2)$**

Conforme já definimos, o grupo $\text{SO}(2)$ é o grupo das matrizes ortogonais 2×2 reais com determinante igual a 1: $\text{SO}(2) = \{R \in \text{Mat}(\mathbb{R}, 2) \mid R^T = R^{-1} \text{ e } \det(R) = 1\}$. Vamos começar estudando a forma geral de tais matrizes.

Como toda matriz 2×2 real, uma matriz genérica $R \in \text{SO}(2)$ é da forma $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, onde $a, b, c, d \in \mathbb{R}$. Vamos estudar a condição $R^{-1} = R^T$. Podemos calcular R^{-1} usando a regra de Laplace, expressão (10.20), página 453: R^{-1} é dada pela transposta da matriz dos cofatores de R dividida pelo determinante de R , que é 1, neste caso. Ou seja, $R^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Assim, $R^{-1} = R^T$ significa nesse caso

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

ou seja, $c = -b$ e $d = a$. Logo, $R = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. A condição $\det(R) = 1$ implica, portanto, $a^2 + b^2 = 1$. Podemos, então, escrever a e b na forma $a = \cos(\theta)$, $b = -\text{sen}(\theta)$, com $\theta \in (-\pi, \pi]$. Resumindo:

$$\text{SO}(2) = \left\{ \begin{pmatrix} \cos \theta & -\text{sen} \theta \\ \text{sen} \theta & \cos \theta \end{pmatrix}, \text{ onde } \theta \in (-\pi, \pi] \right\}.$$

Seja

$$R(\theta) := \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{pmatrix}.$$

Como $R(\theta) = R(\theta + 2\pi)$ vemos que $\text{SO}(2)$ é homeomorfo ao círculo unitário \mathbb{S}^1 , que é uma variedade diferenciável. Como o produto e a inversa são contínuos em $\text{SO}(2)$, isso diz que $\text{SO}(2)$ é um grupo de Lie. É fácil constatar que $R(0) = \mathbb{1}$ e que vale a regra de produto $R(\theta)R(\theta') = R(\theta + \theta')$ (faça!). $\text{SO}(2)$ é, portanto, um grupo uniparamétrico homomorfo ao grupo $(\mathbb{R}, +)$ e isomorfo ao grupo $(\mathbb{R}, + \bmod 2\pi)$.

O gerador infinitesimal J de $\text{SO}(2)$ é definido por

$$J := \left. \frac{d}{d\theta} R(\theta) \right|_{\theta=0} = \left. \frac{d}{d\theta} \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{pmatrix} \right|_{\theta=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

É igualmente elementar constatar que $J^2 = -\mathbb{1}$. Daí

$$\begin{aligned} \exp(\theta J) &= \sum_{m=0}^{\infty} \frac{\theta^m}{m!} J^m \\ &= \sum_{k=0}^{\infty} \frac{\theta^{2k}}{(2k)!} J^{2k} + \sum_{k=0}^{\infty} \frac{\theta^{2k+1}}{(2k+1)!} J^{2k+1} \\ &= \left(\sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} \right) \mathbb{1} + \left(\sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} \right) J \\ &= \cos(\theta) \mathbb{1} + \operatorname{sen}(\theta) J \\ &= R(\theta). \end{aligned}$$

Assim,

$$\text{SO}(2) = \left\{ \exp(\theta J), \text{ onde } \theta \in (-\pi, \pi] \right\}. \tag{21.83}$$

Com isso, (21.83) está nos dizendo que todo elemento de $\text{SO}(2)$ pode ser escrito como exponencial do seu gerador infinitesimal. Veremos que algo semelhante também se dá nos grupos $\text{SO}(3)$ e $\text{SU}(2)$.

O grupo $\text{O}(2)$ é o grupo das matrizes ortogonais 2×2 reais: $\text{O}(2) = \{R \in \text{Mat}(\mathbb{R}, 2) \mid R^T = R^{-1}\}$. Se $R \in \text{O}(2)$ então $\det(R) = \pm 1$. O caso $\det(R) = 1$ corresponde a $\text{SO}(2)$, que tratamos acima. Vamos considerar o caso $\det(R) = -1$.

Como toda matriz 2×2 real, uma matriz genérica $R \in \text{O}(2)$ com $\det(R) = -1$ é da forma $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, onde $a, b, c, d \in \mathbb{R}$. Neste caso, como $\det(R) = -1$, teremos $R^{-1} = \begin{pmatrix} -d & b \\ c & -a \end{pmatrix}$. Assim, a condição $R^{-1} = R^T$ significa nesse caso

$$\begin{pmatrix} -d & b \\ c & -a \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

ou seja, $c = b$ e $d = -a$. Logo, $R = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$. A condição $\det(R) = -1$ implica novamente $a^2 + b^2 = 1$. Podemos então escrever a e b na forma $a = \cos \theta$, $b = -\operatorname{sen} \theta$, com $\theta \in (-\pi, \pi]$. Assim, R é da forma

$$R = \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ -\operatorname{sen} \theta & -\cos \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{pmatrix}.$$

Resumindo:

$$O(2) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^P \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{pmatrix}, \text{ onde } P \in \{0, 1\} \text{ e } \theta \in (-\pi, \pi] \right\}.$$

• **O grupo U(1)**

E. 21.40 Exercício. Mostre que o grupo $U(1) := \{z \in \mathbb{C}, |z| = 1\}$ é isomorfo ao grupo $SO(2)$. ✱

• **O grupo O(1, 1) (o grupo de Lorentz em 1+1 dimensões)**

Aqui estudaremos com algum detalhe o grupo $O(1, 1)$, também denominado Grupo de Lorentz em 1+1 dimensões. A leitura deste tópico pode servir de introdução à leitura da Seção 21.7 que tratará do Grupo de Lorentz em 3+1 dimensões.

Seja M matriz inversível real 2×2 na forma $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, onde $a, b, c, d \in \mathbb{R}$. Tem-se que, $M^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, onde $\det(M) = ad - bc$. Se $\eta := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, então $\eta M^T \eta = \begin{pmatrix} a & -c \\ -b & d \end{pmatrix}$, como facilmente se vê.

Se $M \in SO(1, 1)$, então $M^{-1} = \eta M^T \eta$ e $\det(M) = 1$. Isso significa que $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a & -c \\ -b & d \end{pmatrix}$. Assim, devemos ter $a = d$ e $b = c$. A condição $\det(M) = 1$ significa $a^2 - b^2 = 1$. Logo,

$$SO(1, 1) = \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, a, b \in \mathbb{R}, \text{ com } a^2 - b^2 = 1 \right\}.$$

Como se vê, $SO(1, 1)$ é homeomorfo ao conjunto $H_+ \cup H_-$ formado por duas hipérbolas

$$H_{\pm} := \{(x, y) \in \mathbb{R}^2 \mid x = \pm \sqrt{1 + y^2}, y \in \mathbb{R}\}.$$

$SO(1, 1)$ tem, portanto, duas componentes conexas, que denotaremos por \mathcal{L}_+^{\uparrow} e $\mathcal{L}_+^{\downarrow}$:

$$\mathcal{L}_+^{\uparrow} := \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} \sqrt{1+b^2} & b \\ b & \sqrt{1+b^2} \end{pmatrix}, b \in \mathbb{R} \right\},$$

$$\mathcal{L}_+^{\downarrow} := \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} -\sqrt{1+b^2} & b \\ b & -\sqrt{1+b^2} \end{pmatrix}, b \in \mathbb{R} \right\}.$$

Note-se que apenas \mathcal{L}_+^{\uparrow} é conexa à identidade e, portanto, apenas a componente \mathcal{L}_+^{\uparrow} é um subgrupo de $SO(1, 1)$.

Parametrizando $b \in \mathbb{R}$ na forma $b = -\operatorname{senh}(z)$, com $z \in \mathbb{R}$, constatamos que

$$\mathcal{L}_+^{\uparrow} = \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} \cosh(z) & -\operatorname{senh}(z) \\ -\operatorname{senh}(z) & \cosh(z) \end{pmatrix}, z \in \mathbb{R} \right\}, \tag{21.84}$$

$$\mathcal{L}_+^{\downarrow} = \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} -\cosh(z) & -\operatorname{senh}(z) \\ -\operatorname{senh}(z) & -\cosh(z) \end{pmatrix}, z \in \mathbb{R} \right\}. \tag{21.85}$$

Os elementos de $O(1, 1)$ que não são de $SO(1, 1)$ têm determinante -1 . Assim, são matrizes que satisfazem $\begin{pmatrix} -d & b \\ c & -a \end{pmatrix} = \begin{pmatrix} a & b \\ -c & -a \end{pmatrix}$ sendo, portanto, da forma $\begin{pmatrix} a & b \\ -b & -a \end{pmatrix}$ com $a^2 - b^2 = 1$. O conjunto de tais matrizes é igualmente homeomorfo ao conjunto $H_+ \cup H_-$ e consta também de duas componentes conexas, a saber, os conjuntos

$$\mathcal{L}_-^{\uparrow} := \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} \sqrt{1+b^2} & b \\ -b & -\sqrt{1+b^2} \end{pmatrix}, b \in \mathbb{R} \right\},$$

$$\mathcal{L}_-^{\downarrow} := \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} -\sqrt{1+b^2} & b \\ -b & \sqrt{1+b^2} \end{pmatrix}, b \in \mathbb{R} \right\}.$$

É claro que nem \mathcal{L}_-^{\uparrow} nem $\mathcal{L}_-^{\downarrow}$ são subgrupos de $O(1, 1)$. Parametrizando $b \in \mathbb{R}$ novamente na forma $b = -\operatorname{senh}(z)$, com $z \in \mathbb{R}$, constatamos que

$$\mathcal{L}_-^{\uparrow} = \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} \cosh(z) & -\operatorname{senh}(z) \\ \operatorname{senh}(z) & -\cosh(z) \end{pmatrix}, z \in \mathbb{R} \right\}, \tag{21.86}$$

$$\mathcal{L}_-^{\downarrow} = \left\{ M \in \operatorname{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} -\cosh(z) & -\operatorname{senh}(z) \\ \operatorname{senh}(z) & \cosh(z) \end{pmatrix}, z \in \mathbb{R} \right\}. \tag{21.87}$$

O grupo $O(1, 1)$ é, portanto, a união de quatro componentes conexas:

$$O(1, 1) = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow \cup \mathcal{L}_-^\uparrow \cup \mathcal{L}_-^\downarrow,$$

sendo cada componente disjunta das demais. Dentre elas apenas \mathcal{L}_+^\uparrow é um grupo.

Definindo as matrizes $T := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathcal{L}_-^\downarrow$ e $P := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathcal{L}_-^\uparrow$, podemos escrever

$$\mathcal{L}_+^\downarrow = \left\{ M \in \text{Mat}(\mathbb{R}, 2) \mid M = P \begin{pmatrix} \cosh(z) & -\sinh(z) \\ -\sinh(z) & \cosh(z) \end{pmatrix} T, z \in \mathbb{R} \right\},$$

$$\mathcal{L}_-^\uparrow = \left\{ M \in \text{Mat}(\mathbb{R}, 2) \mid M = P \begin{pmatrix} \cosh(z) & -\sinh(z) \\ -\sinh(z) & \cosh(z) \end{pmatrix}, z \in \mathbb{R} \right\},$$

$$\mathcal{L}_-^\downarrow = \left\{ M \in \text{Mat}(\mathbb{R}, 2) \mid M = \begin{pmatrix} \cosh(z) & -\sinh(z) \\ -\sinh(z) & \cosh(z) \end{pmatrix} T, z \in \mathbb{R} \right\},$$

o que exhibe a relação entre as matrizes dessas três componentes conexas e as matrizes de \mathcal{L}_+^\uparrow .

E. 21.41 Exercício importante. Mostre que

$$\mathcal{L}_+^\uparrow = \left\{ M \in \text{Mat}(\mathbb{R}, 2) \mid M = \exp(z\mathcal{M}_1), z \in \mathbb{R} \right\},$$

onde $\mathcal{M}_1 := \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$.

✱

*

O grupo $O(1, 1)$ é por vezes denominado Grupo de Lorentz em 1+1 dimensões. \mathcal{L}_+^\uparrow é denominado Grupo de Lorentz próprio ortócrono em 1+1 dimensões. O Grupo de Lorentz em 3+1 dimensões será estudado em detalhe na Seção 21.7, página 1099.

Para fazermos contacto com a teoria da relatividade restrita, façamos uma outra parametrização de \mathcal{L}_+^\uparrow , definindo $v = c \tanh(z)$. Com isso $-c < v < c$, $\cosh(z) = \gamma(v)$ e $\sinh(z) = \frac{v}{c}\gamma(v)$, onde $\gamma(v) := (1 - (v/c)^2)^{-1/2}$. Assim,

$$B_1(v) := \begin{pmatrix} \gamma(v) & -\frac{v}{c}\gamma(v) \\ -\frac{v}{c}\gamma(v) & \gamma(v) \end{pmatrix} = \begin{pmatrix} \cosh(z) & -\sinh(z) \\ -\sinh(z) & \cosh(z) \end{pmatrix}, \tag{21.88}$$

teremos

$$\mathcal{L}_+^\uparrow = \left\{ B_1(v), -c < v < c \right\}.$$

$B_1(v) \in \mathcal{L}_+^\uparrow$ age em um vetor $\begin{pmatrix} ct \\ x \end{pmatrix}$ como $B_1(v) \begin{pmatrix} ct \\ x \end{pmatrix} = \begin{pmatrix} ct' \\ x' \end{pmatrix}$, onde

$$t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}},$$

que são as bem conhecidas *transformações de Lorentz* da Teoria da Relatividade Restrita.

E. 21.42 Exercício. Qual a interpretação física das matrizes P e T introduzidas acima?

✱

E. 21.43 Exercício. Constate que para todos $v, v' \in (-c, c)$ vale

$$B_1(v)B_1(v') = B_1\left(\frac{v + v'}{1 + \frac{vv'}{c^2}}\right).$$

Trata-se da bem conhecida regra relativística de composição de velocidades.

✱

21.4.2 O Grupo SO(3)

A presente seção é dedicada ao grupo SO(3), o subgrupo de O(3) composto por matrizes de determinante 1. Ao restante do grupo O(3) dedicamos a Seção 21.4.3, página 1072.

Conforme já definimos, SO(3) é o grupo formado por todas as matrizes 3×3 reais R tais que $R^T = R^{-1}$ e tais que $\det(R) = 1$. Vamos começar seu estudo mostrando que toda a matriz $R \neq \mathbb{1}$ de SO(3) representa uma rotação por algum ângulo em torno de algum eixo que passa pela origem. A essa interpretação seremos conduzidos pelas Proposições 21.4 e 21.5, logo adiante, cujas demonstrações iniciam-se com o seguinte resultado, válido para os grupos SO(n) com n ímpar:

Lema 21.4 *Seja $R \in \text{SO}(n)$ com n ímpar. Então, 1 é um autovalor de R e, portanto, existe ao menos um vetor não nulo $v \in \mathbb{R}^n$ tal que $Rv = v$.* □

Prova. Como $R \in \text{SO}(n)$, tem-se $RR^T = \mathbb{1}$ e $\det(R) = 1$. Logo,

$$\det(\mathbb{1} - R) = \det(R(R^T - \mathbb{1})) = \underbrace{\det(R)}_{=1} \det((R - \mathbb{1})^T) = \det(R - \mathbb{1}) = \det(-(\mathbb{1} - R)) = -\det(\mathbb{1} - R),$$

o que implica $\det(\mathbb{1} - R) = 0$. Acima, usamos o fato que se M for uma matriz $n \times n$ com n ímpar, então $\det(-M) = -\det(M)$ (essa afirmação é falsa se n for par).

O fato que $\det(\mathbb{1} - R) = 0$ implica que 1 é um autovalor de R e que $R - \mathbb{1}$ não possui inversa. Pelo Corolário 10.1, página 452, existe um vetor não nulo $v \in \mathbb{R}^n$ que é anulado $R - \mathbb{1}$, ou seja, que é tal que $Rv = v$. ■

No caso específico de SO(3) tem-se a seguinte afirmação:

Proposição 21.4 *Para cada matriz $R \in \text{SO}(3)$, $R \neq \mathbb{1}$, o subespaço $V \equiv V_R$ de \mathbb{R}^3 formado por vetores que são deixados invariantes por R (ou seja, que satisfazem $Rv = v$ para todo $v \in V$) é unidimensional.* □

Note que o subespaço V pode não ser o mesmo para matrizes R distintas. Note também que excluimos $R = \mathbb{1}$ por razões óbvias: todo vetor de \mathbb{R}^3 é invariante por $\mathbb{1}$ e não apenas um subespaço unidimensional.

Prova da Proposição 21.4. Seja $R \neq \mathbb{1}$ uma matriz qualquer de SO(3), fixa daqui por diante. Seja V o subespaço de \mathbb{R}^3 formado por todos os vetores v que são autovetores de R com autovalor 1: $V = \{v \in \mathbb{R}^3 \mid Rv = v\}$. Sabemos pelo Lema 21.4, página 1053, que $V \neq \{0\}$ e sua dimensão pode, portanto, ser 1, 2 ou 3.

Notemos de passagem que se $v \in V$ então vale também que $R^T v = v$. De fato, se aplicarmos R^T à esquerda na igualdade $v = Rv$ e lembrarmos que $R^T R = \mathbb{1}$, segue que $R^T v = v$. Notemos também que V^\perp , o subespaço formado por todos os vetores ortogonais a todos os vetores de V , é também deixado invariante por R , ou seja, se $u \in V^\perp$ então $Ru \in V^\perp$. De fato, se $v \in V$ e $u \in V^\perp$

$$\langle Ru, v \rangle_{\mathbb{R}} = \langle u, R^T v \rangle_{\mathbb{R}} = \langle u, v \rangle_{\mathbb{R}} = 0.$$

Como isso vale para todo $v \in V$, concluímos que $Ru \in V^\perp$, como queríamos.

Como dissemos, a dimensão de V pode ser igual a 1, 2 ou 3. Vamos mostrar que os dois últimos casos não são possíveis.

Se a dimensão de V fosse 3, V seria idêntico ao espaço \mathbb{R}^3 . Nesse caso então $Rv = v$ para todo vetor $v \in \mathbb{R}^3$, ou seja, $R = \mathbb{1}$, situação que excluimos.

Vamos supor, então, que a dimensão de V é 2. Nesse caso a dimensão de seu complemento ortogonal V^\perp é 1. Agora, como V^\perp é unidimensional e é invariante pela ação de R , teremos para $u \in V^\perp$ que $Ru = \lambda u$, para algum $\lambda \in \mathbb{R}$. Mas isso diz que

$$\langle u, u \rangle_{\mathbb{R}} = \langle Ru, Ru \rangle_{\mathbb{R}} = \langle \lambda u, \lambda u \rangle_{\mathbb{R}} = \lambda^2 \langle u, u \rangle_{\mathbb{R}}$$

e, portanto, $\lambda = \pm 1$. O caso $\lambda = +1$ já está excluído (pois aí $u \in V$). Logo, $\lambda = -1$ e $Ru = -u$. Assim, R possui um autovalor $\lambda = +1$ duplamente degenerado e um autovalor $\lambda = -1$, simples. Logo, $\det(R) = -1$, uma contradição com a hipótese que $R \in \text{SO}(3)$. Logo, a dimensão de V dever ser igual a 1, e isso completa a prova. ■

Seja $R \neq \mathbb{1}$ um elemento de $SO(3)$ e seja V_R o subespaço unidimensional formado pelos vetores deixados invariantes por R e cuja existência foi estabelecida na proposição que acabamos de provar. Como também vimos, R também deixa invariante o subespaço bidimensional V_R^\perp , que é ortogonal a V_R .

Isso significa que se escolhermos em \mathbb{R}^3 uma base ortonormal v, u_1, u_2 com $v \in V_R$ e $u_i \in V_R^\perp$, a matriz R terá a forma

$$R := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \boxed{} \\ 0 & \phantom{\boxed{}} & \phantom{\phantom{\boxed{}}} \end{pmatrix}, \tag{21.89}$$

onde r é uma matriz real 2×2 . Que propriedades tem r ? Como veremos, $r \in SO(2)$. De fato, pela definição de R , teremos para qualquer vetor u , que $\langle u, u \rangle_{\mathbb{R}} = \langle Ru, Ru \rangle_{\mathbb{R}}$, mas se escolhermos $u \in V_R^\perp$, teremos $Ru = ru$ em V_R^\perp e a relação acima significa $\langle u, u \rangle_{\mathbb{R}} = \langle ru, ru \rangle_{\mathbb{R}}$. Logo, $r \in O(2)$. Fora isso, (21.89) mostra que $1 = \det(R) = \det(r)$, provando que $r \in SO(2)$. Como sabemos a forma geral de uma matriz de $SO(2)$ é

$$r = \begin{pmatrix} \cos \varphi & -\text{sen} \varphi \\ \text{sen} \varphi & \cos \varphi \end{pmatrix},$$

com $\varphi \in (-\pi, \pi]$. Isso está também dizendo que R representa uma rotação de φ em torno do eixo (que passa pela origem) representado por V_R .

Concluimos então o seguinte:

Proposição 21.5 *Para cada $R \in SO(3)$ existe uma base ortonormal de \mathbb{R}^3 onde R é da forma*

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\text{sen} \varphi \\ 0 & \text{sen} \varphi & \cos \varphi \end{pmatrix} \tag{21.90}$$

com $\varphi \in (-\pi, \pi]$. □

A discussão precedente nos informa que cada elemento R de $SO(3)$ representa uma rotação de um certo ângulo $\theta \in (-\pi, \pi]$ em torno de um eixo (que passa pela origem) definido por um vetor $\vec{\eta} \in \mathbb{R}^3$ (que podemos sempre escolher como tendo norma 1, ou seja, $\|\vec{\eta}\| = 1$)²¹. Essa afirmação é por vezes denominada *Teorema da Rotação de Euler*, que o demonstrou pela primeira vez (com outros argumentos) em 1776.

• **Os geradores infinitesimais de $SO(3)$**

Pela discussão precedente, se considerarmos os elementos de $SO(3)$ que correspondem a rotações por um ângulo φ *no sentido anti-horário* em torno dos eixos canônicos 1, 2 e 3 do espaço tridimensional \mathbb{R}^3 (eixos esses que suporemos

²¹No caso $R = \mathbb{1}$ tomamos $\theta = 0$ e $\vec{\eta}$ arbitrário.

orientados positivamente, como usual), teremos que as respectivas matrizes de rotação são dadas por

$$R_1(\varphi) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\operatorname{sen} \varphi \\ 0 & \operatorname{sen} \varphi & \cos \varphi \end{pmatrix}, \quad R_2(\varphi) := \begin{pmatrix} \cos \varphi & 0 & \operatorname{sen} \varphi \\ 0 & 1 & 0 \\ -\operatorname{sen} \varphi & 0 & \cos \varphi \end{pmatrix} \quad \text{e} \quad R_3(\varphi) := \begin{pmatrix} \cos \varphi & -\operatorname{sen} \varphi & 0 \\ \operatorname{sen} \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{21.91}$$

com $\varphi \in (-\pi, \pi]$. O estudante deve convencer-se do fato que a diferença de sinal nos senos entre $R_2(\phi)$ e as outras duas matrizes em (21.91) se deve ao fato de que as três matrizes representam rotações *no sentido anti-horário*.

É um exercício elementar (faça-o!) verificar que cada matriz $R_i(\theta)$ representa um subgrupo uniparamétrico de $SO(3)$: $R_i(0) = 1$ e $R_i(\theta)R_i(\theta') = R_i(\theta + \theta')$. Os geradores infinitesimais desses subgrupos são dados por

$$J_1 := \left. \frac{d}{d\varphi} R_1(\varphi) \right|_{\varphi=0} = \left. \frac{d}{d\varphi} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\operatorname{sen} \varphi \\ 0 & \operatorname{sen} \varphi & \cos \varphi \end{pmatrix} \right|_{\varphi=0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{21.92}$$

$$J_2 := \left. \frac{d}{d\varphi} R_2(\varphi) \right|_{\varphi=0} = \left. \frac{d}{d\varphi} \begin{pmatrix} \cos \varphi & 0 & \operatorname{sen} \varphi \\ 0 & 1 & 0 \\ -\operatorname{sen} \varphi & 0 & \cos \varphi \end{pmatrix} \right|_{\varphi=0} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tag{21.93}$$

$$J_3 := \left. \frac{d}{d\varphi} R_3(\varphi) \right|_{\varphi=0} = \left. \frac{d}{d\varphi} \begin{pmatrix} \cos \varphi & -\operatorname{sen} \varphi & 0 \\ \operatorname{sen} \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \right|_{\varphi=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{21.94}$$

Para ilustrar a discussão que faremos na Seção 21.5.2, página 1091, notemos aqui que as três matrizes J_k compõem uma base no espaço das matrizes 3×3 antissimétricas.

É relevante constatar, contemplando (21.92)–(21.94), que os elementos de matriz das matrizes J_a , $a = 1, 2, 3$, são dados por

$$(J_a)_{bc} = -\varepsilon_{abc}, \tag{21.95}$$

onde ε_{abc} , com $a, b, c = 1, 2, 3$, é o chamado *símbolo* (ou *tensor*) de Levi-Civita²², definido da seguinte forma:

$$\varepsilon_{abc} := \begin{cases} 1, & \text{se } abc \text{ for uma permutação par de } 123, \\ -1, & \text{se } abc \text{ for uma permutação ímpar de } 123, \\ 0, & \text{se quaisquer dois índices forem iguais.} \end{cases} \tag{21.96}$$

O exercício a seguir revela um fato importante sobre as matrizes J_k :

E. 21.44 Exercício importantíssimo. Verifique que as matrizes J_1 , J_2 e J_3 satisfazem as relações de comutação

$$[J_a, J_b] = \sum_{c=1}^3 \varepsilon_{abc} J_c. \tag{21.97}$$

²²Tullio Levi-Civita (1873–1941).

Sugestão 1: use a fórmula explícita das matrizes J_k dada em (21.92)–(21.94), calcule seus diversos comutadores e constate a validade de (21.97). **Sugestão 2:** use (21.95) e a identidade de Jacobi para os símbolos de Levi-Civita, relação (4.10), página 259. *

E. 21.45 *Exercício.* Sejam $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ e $\vec{\beta} = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3$. Usando (21.97), mostre que

$$[\vec{\alpha} \cdot \vec{J}, \vec{\beta} \cdot \vec{J}] = (\vec{\alpha} \times \vec{\beta}) \cdot \vec{J}, \tag{21.98}$$

sendo que “ \times ” denota o produto vetorial em \mathbb{R}^3 e $\vec{\alpha} \cdot \vec{J}$ é uma abreviação sugestiva para $\alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3$. *

Os Exercícios E. 21.44 e E. 21.45 dizem-nos que as matrizes J_1, J_2 e J_3 geram uma álgebra de Lie, denominada álgebra de Lie $\mathfrak{so}(3)$ (com letras minúsculas), para lembrar sua associação com o grupo $\text{SO}(3)$.

E. 21.46 *Exercício.* Verifique que as matrizes J_1, J_2 e J_3 satisfazem

$$(J_1)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} =: E_1, \quad (J_2)^2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} =: E_2, \quad \text{e} \quad (J_3)^2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: E_3, \tag{21.99}$$

e que vale

$$(J_1)^2 + (J_2)^2 + (J_3)^2 = -2\mathbb{1}.$$

*

E. 21.47 *Exercício.* Verifique que com as matrizes E_1, E_2 e E_3 acima podemos escrever

$$R_a(\varphi) = \mathbb{1} + (1 - \cos(\varphi))E_a + \text{sen}(\varphi)J_a \tag{21.100}$$

para $a = 1, 2$ e 3 . *

Com o uso de (21.99) podemos facilmente provar o seguinte fato: para $a = 1, 2$ ou 3 tem-se

$$R_a(\varphi) = \exp(\varphi J_a).$$

Vamos mostrar isso. Por (21.99) é evidente que $J_a^3 = E_a J_a = -J_a$ (verifique!). Logo,

$$J_a^{2k} = (-1)^{k+1} E_a, \quad \forall k \in \mathbb{N} \quad \text{e} \quad J_a^{2k+1} = (-1)^k J_a, \quad \forall k \in \mathbb{N}_0. \tag{21.101}$$

Assim, temos para $a = 1, 2$ ou 3 ,

$$\begin{aligned} \exp(\varphi J_a) &= \mathbb{1} + \sum_{m=1}^{\infty} \frac{\varphi^m}{m!} J_a^m \\ &= \mathbb{1} + \sum_{k=1}^{\infty} \frac{\varphi^{2k}}{(2k)!} J_a^{2k} + \sum_{k=0}^{\infty} \frac{\varphi^{2k+1}}{(2k+1)!} J_a^{2k+1} \\ &\stackrel{(21.101)}{=} \mathbb{1} + \left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1} \varphi^{2k}}{(2k)!} \right) E_a + \left(\sum_{k=0}^{\infty} \frac{(-1)^k \varphi^{2k+1}}{(2k+1)!} \right) J_a \\ &= \mathbb{1} + (1 - \cos(\varphi))E_a + \text{sen}(\varphi)J_a \\ &\stackrel{(21.100)}{=} R_a(\varphi), \end{aligned}$$

que é o que queríamos mostrar.

* * * *

E. 21.48 *Exercício.* Sejam as matrizes de rotação definidas em (21.91). Mostre que vale a relação

$$R_3(-\pi/2)R_1(-\pi/2)R_2(-\pi/2) = R_1(-\pi/2)$$

e interprete-a geometricamente. ✱

• **Os geradores infinitesimais de $SO(3)$ e a aplicação exponencial**

Vamos agora mostrar que todo elemento de $SO(3)$ pode ser escrito como exponencial de uma combinação linear das matrizes J_a . Antes, porém, precisamos estabelecer uma certa convenção.

Como comentamos anteriormente, cada elemento R de $SO(3)$ representa uma rotação de um certo ângulo $\theta \in (-\pi, \pi]$ em torno de um eixo definido por um vetor $\vec{\eta} \in \mathbb{R}^3$ (que podemos sempre escolher como tendo norma 1, ou seja, $\|\vec{\eta}\| = 1$). Agora, uma rotação de um ângulo θ em sentido anti-horário em relação a $\vec{\eta}$ corresponde a uma rotação de um ângulo θ em sentido horário em torno de η ou a uma rotação de ângulo $-\theta$ em sentido anti-horário em relação a $\vec{\eta}$. Para evitar essas ambiguidades, vamos doravante adotar a convenção de considerar apenas rotações em sentido anti-horário e limitar o ângulo θ ao intervalo $[0, \pi]$.

Designaremos por $R(\theta, \vec{\eta})$ a matriz de rotação de um ângulo $\theta \in [0, \pi]$ em sentido anti-horário em torno de $\vec{\eta} \in \mathbb{R}^3$ com $\|\vec{\eta}\| = 1$. Convencionamos também que a rotação identidade $\mathbb{1}$ é entendida como uma rotação de um ângulo nulo em torno de um eixo arbitrário: $\mathbb{1} = R(0, \vec{\eta})$. Um comentário elementar que será relevante no que segue é que se duas matrizes de $SO(3)$ realizam uma rotação de um mesmo ângulo em torno de um mesmo eixo, então essas matrizes são iguais.

E. 21.49 *Exercício.* Justifique essa última afirmação! ✱

Temos, então, a seguinte proposição:

Proposição 21.6 *Seja $R(\theta, \vec{\eta}) \in SO(3)$ uma rotação em sentido anti-horário de um ângulo $\theta \in [0, \pi]$ em torno de um eixo definido por um vetor $\vec{\eta} \in \mathbb{R}^3$, $\vec{\eta} = (\eta_1, \eta_2, \eta_3)$, com $\|\vec{\eta}\| = 1$. Então, podemos escrever*

$$R(\theta, \vec{\eta}) = \exp\left(\theta \vec{\eta} \cdot \vec{J}\right), \tag{21.102}$$

onde $\vec{\eta} \cdot \vec{J} := \eta_1 J_1 + \eta_2 J_2 + \eta_3 J_3$. Calculando-se a exponencial acima obtém-se

$$R(\theta, \vec{\eta}) = \mathbb{1} + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{J})^2 + \text{sen}(\theta) (\vec{\eta} \cdot \vec{J}), \tag{21.103}$$

ou seja, escrevendo-se explicitamente,

$$R(\theta, \vec{\eta}) = \begin{pmatrix} (1 - \cos(\theta))\eta_1^2 + \cos(\theta) & (1 - \cos(\theta))\eta_1\eta_2 - \text{sen}(\theta)\eta_3 & (1 - \cos(\theta))\eta_1\eta_3 + \text{sen}(\theta)\eta_2 \\ (1 - \cos(\theta))\eta_1\eta_2 + \text{sen}(\theta)\eta_3 & (1 - \cos(\theta))\eta_2^2 + \cos(\theta) & (1 - \cos(\theta))\eta_3\eta_2 - \text{sen}(\theta)\eta_1 \\ (1 - \cos(\theta))\eta_1\eta_3 - \text{sen}(\theta)\eta_2 & (1 - \cos(\theta))\eta_3\eta_2 + \text{sen}(\theta)\eta_1 & (1 - \cos(\theta))\eta_3^2 + \cos(\theta) \end{pmatrix}. \tag{21.104}$$

Por fim, para $\vec{\alpha} \in \mathbb{R}^3$ arbitrário, tem-se que

$$R(\theta, \vec{\eta})\vec{\alpha} = \vec{\alpha} + (1 - \cos(\theta)) (\vec{\eta} \times (\vec{\eta} \times \vec{\alpha})) + \text{sen}(\theta) \vec{\eta} \times \vec{\alpha} \tag{21.105}$$

$$= \cos(\theta)\vec{\alpha} + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} + \text{sen}(\theta) \vec{\eta} \times \vec{\alpha}. \tag{21.106}$$

As expressões (21.103) e (21.105)–(21.106) são denominadas fórmulas de Rodrigues²³ para o grupo $SO(3)$. ■

²³Benjamin Olinde Rodrigues (1794–1851). Rodrigues foi banqueiro, socialista, e matemático amador, nascido na França, mas de origem judeo-portuguesa. Seu nome é mais conhecido por uma identidade sobre polinômios de Legendre, a equação (16.40), página 806, (vide também Seção 16.1.2, página 803 para generalizações).

Antes da demonstração, façamos alguns

Comentários. A Proposição 21.6 revela-nos que $R(\theta, \vec{\eta})$ depende do ângulo θ e do vetor $\vec{\eta}$ através da combinação $\theta\vec{\eta}$. O vetor $\theta\vec{\eta}$ é, por isso, denominado *vetor de rotação* associado ao elemento $R(\theta, \vec{\eta}) \in \text{SO}(3)$.

A fórmula de Rodrigues (21.103) também pode ser obtida com o uso do Cálculo Funcional (Teorema Espectral). Vide Exercício E. 21.123, página 1147.

Para uma prova mais geométrica de (21.106), siga os passos indicados no Exercício E. 21.124, página 1147. As expressões (21.105)–(21.106) são interessantes por exibirem explicitamente a ação da rotação $R(\theta, \vec{\eta})$ sobre cada vetor $\vec{\alpha} \in \mathbb{R}^3$. ♣

Prova da Proposição 21.6. A equação (21.102) é evidente no caso $\theta = 0$. Vamos supor $R(\theta, \vec{\eta}) \neq \mathbb{1}$ e, portanto, $\theta \neq 0$. Pela Proposição 21.4, página 1053, existe um subespaço unidimensional V_R que é deixado invariante por $R(\theta, \vec{\eta})$. É óbvio que $R(\theta, \vec{\eta})\vec{\eta} = \vec{\eta}$.

O que faremos para demonstrar nossa proposição é mostrar que $\exp(\theta\vec{\eta} \cdot \vec{J})$ é uma matriz ortogonal, de determinante 1, e que mantém $\vec{\eta}$ invariante e roda os vetores perpendiculares a $\vec{\eta}$ de um ângulo θ (no sentido anti-horário) em torno do eixo definido por $\vec{\eta}$. Com isso, podemos identificar $R(\theta, \vec{\eta}) = \exp(\theta\vec{\eta} \cdot \vec{J})$, como queremos.

Vamos abaixo calcular de modo mais explícito o que é a matriz $\exp(\theta\vec{\eta} \cdot \vec{J})$ mas, antes disso, vamos demonstrar que $\exp(\theta\vec{\eta} \cdot \vec{J}) \in \text{SO}(3)$. Para isso, começamos com a observação que

$$\vec{\eta} \cdot \vec{J} := \eta_1 J_1 + \eta_2 J_2 + \eta_3 J_3 = \begin{pmatrix} 0 & -\eta_3 & \eta_2 \\ \eta_3 & 0 & -\eta_1 \\ -\eta_2 & \eta_1 & 0 \end{pmatrix} \quad (21.107)$$

é uma matriz antissimétrica, ou seja, $(\vec{\eta} \cdot \vec{J})^T = -\vec{\eta} \cdot \vec{J}$. Note-se que seus elementos de matriz são dados por

$$(\vec{\eta} \cdot \vec{J})_{ij} = -\sum_{k=1}^3 \varepsilon_{ijk} \eta_k. \quad (21.108)$$

Assim,

$$\left[\exp(\theta\vec{\eta} \cdot \vec{J}) \right]^T = \sum_{m=0}^{\infty} \frac{\theta^m}{m!} \left([\vec{\eta} \cdot \vec{J}]^T \right)^m = \sum_{m=0}^{\infty} \frac{(-\theta)^m}{m!} (\vec{\eta} \cdot \vec{J})^m = \exp(-\theta\vec{\eta} \cdot \vec{J}) = \left[\exp(\theta\vec{\eta} \cdot \vec{J}) \right]^{-1}.$$

Isso demonstrou que $\exp(\theta\vec{\eta} \cdot \vec{J})$ é uma matriz ortogonal, ou seja, sua transposta é igual a sua inversa. Resta-nos mostrar que $\det(\exp(\theta\vec{\eta} \cdot \vec{J})) = 1$. Como $\exp(\theta\vec{\eta} \cdot \vec{J})$ é ortogonal, seu determinante é ± 1 . Assim, como $\det(\exp(\theta\vec{\eta} \cdot \vec{J}))$ depende continuamente de θ (para isso, vide, por exemplo, a expressão (21.111), abaixo), temos que $\det(\exp(\theta\vec{\eta} \cdot \vec{J}))$ é constante para todo $\theta \in (-\pi, \pi]$. Calculando em $\theta = 0$, teremos

$$\det(\exp(\theta\vec{\eta} \cdot \vec{J})) = \det(\exp(0\vec{\eta} \cdot \vec{J})) = \det(\mathbb{1}) = 1.$$

Um segundo argumento que pode ser empregado para se provar o mesmo fato é evocar a Proposição 10.14, página 466, segundo a qual vale

$$\det(\exp(\theta\vec{\eta} \cdot \vec{J})) = e^{\theta \text{Tr}(\vec{\eta} \cdot \vec{J})}.$$

Como $\vec{\eta} \cdot \vec{J}$ é uma matriz antissimétrica, segue que $\text{Tr}(\vec{\eta} \cdot \vec{J}) = 0$ (evidente por (21.107)) e, portanto, $\det(\exp(\theta\vec{\eta} \cdot \vec{J})) = 1$. Estabelecemos, assim, $\exp(\theta\vec{\eta} \cdot \vec{J}) \in \text{SO}(3)$ para todo θ e todo $\vec{\eta}$.

Vamos agora expressar de modo mais explícito a matriz $\exp(\theta\vec{\eta} \cdot \vec{J})$. Para isso, será importante mostrar que

$$(\vec{\eta} \cdot \vec{J})^3 = -(\vec{\eta} \cdot \vec{J}). \quad (21.109)$$

A maneira pedestre de mostrar isso é por verificação explícita. De fato, por (21.107),

$$\left(\vec{\eta} \cdot \vec{J}\right)^2 = \begin{pmatrix} \eta_1^2 - 1 & \eta_1 \eta_2 & \eta_1 \eta_3 \\ \eta_1 \eta_2 & \eta_2^2 - 1 & \eta_3 \eta_2 \\ \eta_1 \eta_3 & \eta_3 \eta_2 & \eta_3^2 - 1 \end{pmatrix}. \quad (21.110)$$

Multiplicando-se novamente por $\vec{\eta} \cdot \vec{J}$, obtém-se (21.109). Temos, então, o seguinte: para todo $k \in \mathbb{N}$, vale

$$\left(\vec{\eta} \cdot \vec{J}\right)^{2k} = (-1)^{k+1} \left(\vec{\eta} \cdot \vec{J}\right)^2 \quad \text{e} \quad \left(\vec{\eta} \cdot \vec{J}\right)^{2k+1} = (-1)^k \left(\vec{\eta} \cdot \vec{J}\right).$$

Logo,

$$\begin{aligned} \exp\left(\theta \vec{\eta} \cdot \vec{J}\right) &= \mathbb{1} + \sum_{m=1}^{\infty} \frac{\theta^m}{m!} \left(\vec{\eta} \cdot \vec{J}\right)^m \\ &= \mathbb{1} + \sum_{k=1}^{\infty} \frac{\theta^{2k}}{(2k)!} \left(\vec{\eta} \cdot \vec{J}\right)^{2k} + \sum_{k=0}^{\infty} \frac{\theta^{2k+1}}{(2k+1)!} \left(\vec{\eta} \cdot \vec{J}\right)^{2k+1} \\ &= \mathbb{1} + \left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1} \theta^{2k}}{(2k)!}\right) \left(\vec{\eta} \cdot \vec{J}\right)^2 + \left(\sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!}\right) \vec{\eta} \cdot \vec{J} \\ &= \mathbb{1} + (1 - \cos(\theta)) \left(\vec{\eta} \cdot \vec{J}\right)^2 + \sin(\theta) \left(\vec{\eta} \cdot \vec{J}\right). \end{aligned}$$

Resumindo,

$$\exp\left(\theta \vec{\eta} \cdot \vec{J}\right) = \mathbb{1} + (1 - \cos(\theta)) \left(\vec{\eta} \cdot \vec{J}\right)^2 + \sin(\theta) \left(\vec{\eta} \cdot \vec{J}\right). \quad (21.111)$$

Seja agora $\vec{\alpha} \in \mathbb{R}^3$, arbitrário. Tem-se que

$$\left(\vec{\eta} \cdot \vec{J}\right) \vec{\alpha} = \vec{\eta} \times \vec{\alpha}. \quad (21.112)$$

E. 21.50 *Exercício.* Verifique-a usando a forma explícita (21.107) de $\vec{\eta} \cdot \vec{J}$ ou usando (21.108). *

Segue imediatamente de (21.112) que

$$\left(\vec{\eta} \cdot \vec{J}\right)^2 \vec{\alpha} = \vec{\eta} \times (\vec{\eta} \times \vec{\alpha}) \stackrel{(4.14)}{=} (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} - \vec{\alpha}, \quad (21.113)$$

onde, na segunda igualdade, usou-se a bem conhecida identidade (4.14), página 260 (que afirma que $\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$ para quaisquer vetores $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$), bem como o fato que $\|\vec{\eta}\| = 1$. Retornando com isso a (21.111), temos imediatamente que

$$\exp\left(\theta \vec{\eta} \cdot \vec{J}\right) \vec{\alpha} = \vec{\alpha} + (1 - \cos(\theta)) \left(\vec{\eta} \times (\vec{\eta} \times \vec{\alpha})\right) + \sin(\theta) \vec{\eta} \times \vec{\alpha} \quad (21.114)$$

$$= \cos(\theta) \vec{\alpha} + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} + \sin(\theta) \vec{\eta} \times \vec{\alpha}. \quad (21.115)$$

Vamos agora provar que $\exp\left(\theta \vec{\eta} \cdot \vec{J}\right)$ mantém o vetor $\vec{\eta}$ invariante e roda os vetores ortogonais a $\vec{\eta}$ por um ângulo θ no sentido horário, o que nos permite identificar $R(\theta, \vec{\eta}) = \exp\left(\theta \vec{\eta} \cdot \vec{J}\right)$.

De (21.114) segue imediatamente que

$$\exp\left(\theta\vec{\eta} \cdot \vec{J}\right)\vec{\eta} = \vec{\eta},$$

ou seja, tal como $R(\theta, \vec{\eta})$, a matriz $\exp\left(\theta\vec{\eta} \cdot \vec{J}\right)$ mantém $\vec{\eta}$ invariante para qualquer θ . Tomemos agora um vetor unitário $\vec{\zeta}$ que seja ortogonal a $\vec{\eta}$. Os vetores $\vec{\zeta}$ e $\vec{\eta} \times \vec{\zeta}$ são ambos unitários e formam uma base ortonormal no subespaço ortogonal a $\vec{\eta}$. Evocando (21.115), e usando que $\vec{\eta} \cdot \vec{\zeta} = 0$, que $\vec{\eta} \cdot (\vec{\eta} \times \vec{\zeta}) = 0$ e que $\vec{\eta} \times (\vec{\eta} \times \vec{\zeta}) = -\vec{\zeta}$, temos

$$\exp\left(\theta\vec{\eta} \cdot \vec{J}\right)\vec{\zeta} = \cos(\theta)\vec{\zeta} + \sin(\theta)(\vec{\eta} \times \vec{\zeta}),$$

$$\exp\left(\theta\vec{\eta} \cdot \vec{J}\right)(\vec{\eta} \times \vec{\zeta}) = -\sin(\theta)\vec{\zeta} + \cos(\theta)(\vec{\eta} \times \vec{\zeta}),$$

como facilmente se constata (faça-o!). Essas relações mostram-nos claramente que $\exp\left(\theta\vec{\eta} \cdot \vec{J}\right)$ age no subespaço gerado por $\vec{\zeta}$ e $\vec{\eta} \times \vec{\zeta}$ como uma rotação de um ângulo θ no sentido anti-horário. Assim, podemos identificar $R(\theta, \vec{\eta}) = \exp\left(\theta\vec{\eta} \cdot \vec{J}\right)$, o que completa a demonstração da Proposição 21.6. ■

Resumindo nossas conclusões,

$$\text{SO}(3) = \left\{ \exp\left(\theta\vec{\eta} \cdot \vec{J}\right), \theta \in [0, \pi], \vec{\eta} \in \mathbb{R}^3 \text{ com } \|\vec{\eta}\| = 1 \right\}.$$

O leitor deve atentar para o fato que alguns elementos listados no conjunto à direita podem ser idênticos, vide comentário adiante.

A Proposição 21.6, página 1057, está dizendo-nos que todo elemento de $\text{SO}(3)$ pode ser escrito como a exponencial de um elemento de sua álgebra de Lie. Isso constata um teorema geral (vide, por exemplo, [347]) que diz que se um grupo de Lie é compacto²⁴ e conexo e sua álgebra de Lie é semissimples, a aplicação exponencial da sua álgebra de Lie é sobrejetora no grupo. De fato, $\text{SO}(3)$ é compacto, conexo e $\mathfrak{so}(3)$ é semissimples.

Finalizamos com uma breve observação, um tanto evidente, mas que será evocada mais adiante:

Proposição 21.7 *Seja $\vec{\eta} \in \mathbb{R}^3$, $\vec{\eta} = (\eta_1, \eta_2, \eta_3)$. Então, a matriz $\exp\left(\theta\vec{\eta} \cdot \vec{J}\right)$ é igual à matriz identidade se e somente se $\theta = 2\pi n$ para $n \in \mathbb{Z}$.* □

Prova. De (21.103) vemos que $\exp\left(\theta\vec{\eta} \cdot \vec{J}\right) = \mathbb{1}$ se e somente se $(1 - \cos(\theta))\left(\vec{\eta} \cdot \vec{J}\right)^2 + \sin(\theta)\left(\vec{\eta} \cdot \vec{J}\right) = 0$. Agora, vemos por (21.107) e (21.110) que as matrizes $\vec{\eta} \cdot \vec{J}$ e $\left(\vec{\eta} \cdot \vec{J}\right)^2$ são linearmente independentes (observe que todos os elementos da diagonal de $\vec{\eta} \cdot \vec{J}$ são nulos, ao contrário de $\left(\vec{\eta} \cdot \vec{J}\right)^2$, cujos elementos da diagonal nunca podem ser todos nulos). Logo, devemos ter $\cos(\theta) = 1$ e $\sin(\theta) = 0$, o que se dá se e somente se $\theta = 2\pi n$ para $n \in \mathbb{Z}$. ■

• O espaço dos parâmetros do grupo $\text{SO}(3)$

Para continuar esta exposição sobre o grupo $\text{SO}(3)$, vamos descrever sua estrutura enquanto variedade. Como vimos, os elementos de $\text{SO}(3)$ são parametrizados por pontos $\theta\vec{\eta}$ de \mathbb{R}^3 , sendo que $\theta \in [0, \pi]$ e $\|\vec{\eta}\| = 1$. O conjunto de todos os pontos desse tipo compreende a esfera fechada de raio π centrada na origem. Para cada $\vec{\eta}$ fixo, os dois pontos antípodas da superfície dessa esfera que estão na direção definida por $\vec{\eta}$ são $\pm\pi\vec{\eta}$. É claro, porém, que tais pontos correspondem à mesma rotação: uma rotação de π em torno de um $\vec{\eta}$ é o mesmo que uma rotação de π em torno de $-\vec{\eta}$. De fato, é trivial verificar por (21.111) que $\exp\left(\pi\vec{\eta} \cdot \vec{J}\right) = \exp\left(-\pi\vec{\eta} \cdot \vec{J}\right)$. Assim, $\text{SO}(3)$ corresponde nessa imagem ao espaço obtido tomando-se uma esfera fechada e identificando-se todos os pares de pontos antípodas de sua superfície.

Na linguagem da topologia e das variedades, o conjunto obtido por essa construção coincide com o chamado *espaço projetivo real*, denotado por \mathbb{RP}^3 , que é definido como o conjunto de todas as linhas retas em \mathbb{R}^4 que passam pela origem.

²⁴Para a definição da noção de compacidade e suas propriedades, vide Seção 32.3, página 1514.

Com mais generalidade, o conjunto $\mathbb{R}P^n$ é a variedade topológica (diferenciável) n -dimensional formada pelo conjunto de todas as linhas retas de \mathbb{R}^{n+1} que passam pela origem. O espaço projetivo real $\mathbb{R}P^n$ é definido e analisado na Seção 33.4.6, página 1637. À página 1639 esclarecemos por que $\mathbb{R}P^n$ pode ser também entendido como o conjunto obtido tomando-se uma bola fechada em \mathbb{R}^{n+1} e identificando-se os pontos antípodas de sua superfície.

$SO(3)$ é homeomorfo, enquanto variedade, ao espaço projetivo $\mathbb{R}P^3$. Como veremos na próxima seção, o grupo $SU(2)$, que é fortemente aparentado a $SO(3)$, tem outra estrutura: $SU(2)$ é homeomorfo a S^3 , a superfície da bola de raio 1 em \mathbb{R}^4 . Para uma introdução ao estudo de variedades, vide Capítulo 33, página 1582, e outras referências lá citadas.

21.4.2.1 Mais Propriedades das Matrizes de $SO(3)$

As fórmulas de Rodrigues (21.103) e (21.105)–(21.106) permitem a obtenção de outros resultados relevantes sobre as matrizes que compõem o grupo $SO(3)$. Nesta seção exibiremos alguns deles.

Proposição 21.8 *Sejam $\vec{\eta}_1$ e $\vec{\eta}_2$ dois vetores unitários de \mathbb{R}^3 . Então, existe $R_{12} \equiv R_{12}(\theta_{12}, \vec{\eta}_{12}) \in SO(3)$ tal que $R_{12}\vec{\eta}_1 = \vec{\eta}_2$, ou seja, existe uma matriz de $SO(3)$ que leva $\vec{\eta}_1$ em $\vec{\eta}_2$.*

1. No caso em que $\vec{\eta}_1 = \vec{\eta}_2$ podemos escolher (evidentemente) $R_{12} = \mathbb{1}$, ou seja, $\theta_{12} = 0$ e $\vec{\eta}_{12}$ arbitrário.
2. No caso em que $\vec{\eta}_1 = -\vec{\eta}_2$ podemos escolher $\theta_{12} = \pi$ e $\vec{\eta}_{12}$ qualquer vetor que seja ortogonal a $\vec{\eta}_1$ (e, portanto, a $\vec{\eta}_2$).
3. No caso em que $\vec{\eta}_1 \neq \vec{\eta}_2$ e $\vec{\eta}_1 \neq -\vec{\eta}_2$, podemos escolher

$$R_{12} = R(\theta_{12}, \vec{\eta}_{12}), \quad \text{com } \theta_{12} = \arccos(\vec{\eta}_1 \cdot \vec{\eta}_2) \in [0, \pi] \quad \text{e} \quad \vec{\eta}_{12} = \frac{\vec{\eta}_1 \times \vec{\eta}_2}{\|\vec{\eta}_1 \times \vec{\eta}_2\|}.$$

□

Prova. O caso em que $\vec{\eta}_1 = \vec{\eta}_2$ é trivial. No caso em que $\vec{\eta}_1 = -\vec{\eta}_2$ a escolha $\theta_{12} = \pi$ fornece por (21.106)

$$R(\theta_{12}, \vec{\eta}_{12})\vec{\eta}_1 = -\vec{\eta}_1 + 2(\vec{\eta}_{12} \cdot \vec{\eta}_1)\vec{\eta}_{12},$$

que vale $-\vec{\eta}_1 = \vec{\eta}_2$ se tivermos $\vec{\eta}_{12} \cdot \vec{\eta}_1 = 0$, ou seja, se $\vec{\eta}_{12}$ for ortogonal a $\vec{\eta}_1$ (e, portanto, a $\vec{\eta}_2$).

O único caso que realmente requer algum esforço de demonstração é o caso em que $\vec{\eta}_1 \neq \vec{\eta}_2$ e $\vec{\eta}_1 \neq -\vec{\eta}_2$. A escolha $\theta_{12} = \arccos(\vec{\eta}_1 \cdot \vec{\eta}_2) \in [0, \pi]$ significa que θ_{12} é o ângulo formado entre $\vec{\eta}_1$ e $\vec{\eta}_2$. Logo, $\cos\theta_{12} = \vec{\eta}_1 \cdot \vec{\eta}_2$ e $\sin\theta_{12} = \|\vec{\eta}_1 \times \vec{\eta}_2\|$. Portanto, por (21.106), temos

$$R(\theta_{12}, \vec{\eta}_{12})\vec{\eta}_1 \stackrel{(21.106)}{=} (\vec{\eta}_1 \cdot \vec{\eta}_2)\vec{\eta}_1 + (\vec{\eta}_1 \times \vec{\eta}_2) \times \vec{\eta}_1 \stackrel{(4.14)}{=} (\vec{\eta}_1 \cdot \vec{\eta}_2)\vec{\eta}_1 - (\vec{\eta}_1 \cdot \vec{\eta}_2)\vec{\eta}_1 + \vec{\eta}_2 = \vec{\eta}_2,$$

onde na primeira igualdade usamos também que $\vec{\eta}_{12} \cdot \vec{\eta}_1 = 0$ (por ambos os vetores serem mutuamente ortogonais) e na penúltima usamos também o fato que $\|\vec{\eta}_1\| = 1$. ■

O que essa proposição diz é que $SO(3)$ age transitivamente na esfera de raio 1 de \mathbb{R}^3 : qualquer vetor unitário pode ser levado a qualquer outro vetor unitário por meio da ação de algum elemento de $SO(3)$.

O resultado a seguir é óbvio sob vários aspectos, mas ainda assim apresentamos uma prova do mesmo.

Lema 21.5 *Se $R \in SO(3)$, vale*

$$R(\vec{\alpha} \times \vec{\beta}) = (R\vec{\alpha}) \times (R\vec{\beta}), \tag{21.116}$$

para todos $\vec{\alpha}, \vec{\beta} \in \mathbb{R}^3$.

□

Prova. Seja $R \equiv R(\theta, \vec{\eta})$ para algum vetor unitário $\vec{\eta}$ e algum $\theta \in [0, \pi]$. Usando a bem conhecida identidade (4.14), página 260, é fácil ver que

$$\begin{aligned} (R(\vec{\alpha} \times \vec{\beta})) \times ((R\vec{\alpha}) \times (R\vec{\beta})) &\stackrel{(4.14)}{=} \left((R(\vec{\alpha} \times \vec{\beta})) \cdot R\vec{\beta} \right) R\vec{\alpha} - \left((R(\vec{\alpha} \times \vec{\beta})) \cdot R\vec{\alpha} \right) R\vec{\beta} \\ &= \left((\vec{\alpha} \times \vec{\beta}) \cdot \vec{\beta} \right) R\vec{\alpha} - \left((\vec{\alpha} \times \vec{\beta}) \cdot \vec{\alpha} \right) R\vec{\beta} = 0. \end{aligned}$$

Na segunda igualdade usamos o fato que R preserva produtos escalares e na terceira igualdade usamos o fato que $\vec{\alpha} \times \vec{\beta}$ é ortogonal a $\vec{\alpha}$ e a $\vec{\beta}$. Isso mostrou que $R(\vec{\alpha} \times \vec{\beta}) = \lambda((R\vec{\alpha}) \times (R\vec{\beta}))$ para alguma constante $\lambda \in \mathbb{R}$.

Como R preserva a norma dos vetores, temos

$$\|R(\vec{\alpha} \times \vec{\beta})\| = \|\vec{\alpha} \times \vec{\beta}\| = \|\vec{\alpha}\| \|\vec{\beta}\| \sin\varphi,$$

onde φ é o ângulo entre $\vec{\alpha}$ e $\vec{\beta}$. Temos também

$$\|(R\vec{\alpha}) \times (R\vec{\beta})\| = \|R\vec{\alpha}\| \|R\vec{\beta}\| \sin\varphi' = \|\vec{\alpha}\| \|\vec{\beta}\| \sin\varphi',$$

onde φ' é o ângulo entre $R\vec{\alpha}$ e $R\vec{\beta}$. Como R também preserva o ângulo entre dois vetores (pois preserva suas normas e seu produto escalar), tem-se $\varphi' = \varphi$. Disso segue imediatamente que $\lambda = \pm 1$. A matriz $R \equiv R(\theta, \vec{\eta})$ é contínua em θ . Como λ pode assumir apenas dois valores, a continuidade diz-nos que λ não se altera em uma vizinhança aberta de $\theta = 0$ (em verdade em parte alguma, pela compacidade e conexidade do grupo). Para $\theta = 0$ temos $R = \mathbb{1}$, em cujo caso vale evidentemente $\lambda = 1$. Esse valor de λ , portanto, é o mesmo para todo θ e, assim, estabeleceu-se que $R(\vec{\alpha} \times \vec{\beta}) = (R\vec{\alpha}) \times (R\vec{\beta})$. ■

E. 21.51 *Exercício*. Há outras provas do Lema 21.5. Prove-o, por exemplo, usando diretamente (21.105)–(21.106). Vide também Exercício E. 21.53, página 1063. *

Uma consequência do Lema 21.5 é a seguinte proposição importante:

Proposição 21.9 *Se $R_0 \in \text{SO}(3)$ vale*

$$R_0 R(\theta, \vec{\eta}) R_0^{-1} = R(\theta, R_0 \vec{\eta}) \tag{21.117}$$

para todo vetor unitário $\vec{\eta} \in \mathbb{R}^3$ e todo $\theta \in [0, \pi]$. Assim, tem-se

$$\exp\left(\theta \left((R_0 \vec{\eta}) \cdot \vec{J} \right)\right) = R_0 \exp\left(\theta \vec{\eta} \cdot \vec{J}\right) R_0^{-1} = \exp\left(\theta \vec{\eta} \cdot \left(R_0 \vec{J} R_0^{-1} \right)\right). \tag{21.118}$$

□

Prova. Seja $\vec{\alpha} \in \mathbb{R}^3$. Temos por (21.106)

$$\begin{aligned} R(\theta, R_0 \vec{\eta}) R_0 \vec{\alpha} &\stackrel{(21.106)}{=} \cos(\theta) (R_0 \vec{\alpha}) + (1 - \cos(\theta)) \left((R_0 \vec{\eta}) \cdot (R_0 \vec{\alpha}) \right) R_0 \vec{\eta} + \sin(\theta) \left((R_0 \vec{\eta}) \times (R_0 \vec{\alpha}) \right) \\ &= \cos(\theta) R_0 \vec{\alpha} + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{\alpha}) R_0 \vec{\eta} + \sin(\theta) R_0 (\vec{\eta} \times \vec{\alpha}) \\ &= R_0 \left[\cos(\theta) (\vec{\alpha}) + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} + \sin(\theta) (\vec{\eta} \times \vec{\alpha}) \right] \\ &\stackrel{(21.106)}{=} R_0 R(\theta, \vec{\eta}) \vec{\alpha}. \end{aligned}$$

sendo que na segunda igualdade usamos que R_0 preserva produtos escalares e usamos (21.116). Como isso vale para todo $\vec{\alpha} \in \mathbb{R}^3$, estabelecemos que $R(\theta, R_0 \vec{\eta}) R_0 = R_0 R(\theta, \vec{\eta})$, completando a demonstração de (21.117). A relação (21.118)

segue disso e do fato provado alhures que $P \exp(A) P^{-1} = \exp(PAP^{-1})$ para quaisquer matrizes quadradas A e P de mesma ordem (P inversível). ■

As Proposições 21.8 e 21.9 têm o seguinte corolário evidente:

Corolário 21.1 *Sejam $\vec{\eta}_1$ e $\vec{\eta}_2$ dois vetores unitários de \mathbb{R}^3 . Então, existe $R_{12} \in \text{SO}(3)$ tal que*

$$R(\theta, \vec{\eta}_2) = R_{12} R(\theta, \vec{\eta}_1) R_{12}^{-1}.$$

A expressão para a matriz $R_{12} \in \text{SO}(3)$ é a mesma descrita na Proposição 21.8. □

Outra consequência útil da Proposição 21.9 é a identidade

$$(R_0 \vec{\eta}) \cdot \vec{J} = R_0(\vec{\eta} \cdot \vec{J}) R_0^{-1} = \vec{\eta} \cdot (R_0 \vec{J} R_0^{-1}), \tag{21.119}$$

que pode ser provada diferenciando-se os termos da igualdade (21.118) em relação a θ e tomando-se $\theta = 0$. A relação (21.119) é válida para qualquer $R_0 \in \text{SO}(3)$ e qualquer $\vec{\eta} \in \mathbb{R}^3$ (aqui não é necessário impor $\|\vec{\eta}\| = 1$, por razões óbvias).

E. 21.52 *Exercício.* A relação (21.119) diz-nos que $R(\theta, \vec{\eta})(\vec{\alpha} \cdot \vec{J})R(-\theta, \vec{\eta}) = (R(\theta, \vec{\eta})\vec{\alpha}) \cdot \vec{J}$. Usando a fórmula de Rodrigues (21.106), obtenha a relação

$$R(\theta, \vec{\eta})(\vec{\alpha} \cdot \vec{J})R(-\theta, \vec{\eta}) = \cos(\theta)(\vec{\alpha} \cdot \vec{J}) + (1 - \cos(\theta))(\vec{\eta} \cdot \vec{\alpha})(\vec{\eta} \cdot \vec{J}) + \sin(\theta)(\vec{\eta} \times \vec{\alpha}) \cdot \vec{J}. \tag{21.120}$$

Essa relação é particularmente útil no caso em que $\vec{\alpha}$ e $\vec{\eta}$ são vetores da base canônica $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ de \mathbb{R}^3 . Mostre que

$$R_k(\theta) J_l R_k(-\theta) = (\delta_{k,l} + (1 - \delta_{k,l}) \cos(\theta)) J_l + \sin(\theta) \sum_{m=1}^3 \varepsilon_{klm} J_m, \tag{21.121}$$

tomando $\vec{\eta} = \mathbf{e}_k$ e $\vec{\alpha} = \mathbf{e}_l$ na relação (21.120), com $k, l \in \{1, 2, 3\}$. ✦

E. 21.53 *Exercício.* Demonstre a relação (21.116) usando (21.119) e as relações de comutação (21.98). ✦

• **Fixando rotações com pares de vetores ortonormais**

Vamos denotar por \mathbf{e}_k , $k = 1, 2, 3$, os vetores da base canônica de \mathbb{R}^3 : $\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ e $\mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Naturalmente, $(\mathbf{e}_k)_l = \delta_{kl}$.

Se $R \in \text{SO}(3)$, suas colunas são as imagens da base canônica por R : se

$$R = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix}, \quad \text{então} \quad \begin{pmatrix} R_{1k} \\ R_{2k} \\ R_{3k} \end{pmatrix} = R \mathbf{e}_k.$$

Assim, a imagem de R nos três vetores da base canônica fixa R univocamente. Em verdade, é suficiente considerar a ação em apenas dois desses vetores pois tem-se, por exemplo, $R \mathbf{e}_3 = R(\mathbf{e}_1 \times \mathbf{e}_2) = (R \mathbf{e}_1) \times (R \mathbf{e}_2)$ (pelo Lema 21.5, página 1061). A recíproca dessa afirmação é também verdadeira:

Proposição 21.10 *Sejam \mathbf{e}'_1 e \mathbf{e}'_2 dois vetores ortonormais, ou seja, satisfazendo $\mathbf{e}'_a \cdot \mathbf{e}'_b = \delta_{ab}$. Então, existe uma matriz $R \in \text{SO}(3)$, única, tal que $R \mathbf{e}_1 = \mathbf{e}'_1$, $R \mathbf{e}_2 = \mathbf{e}'_2$ e $R \mathbf{e}_3 = \mathbf{e}'_1 \times \mathbf{e}'_2$. □*

Prova. Seja $\mathbf{e}'_3 := \mathbf{e}'_1 \times \mathbf{e}'_2$. Como $\mathcal{E} \equiv \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ e $\mathcal{E}' \equiv \{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ são bases em \mathbb{R}^3 podemos expandir qualquer vetor $\vec{x} \in \mathbb{R}^3$ em ambas: $\vec{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = x'_1\mathbf{e}'_1 + x'_2\mathbf{e}'_2 + x'_3\mathbf{e}'_3$, com $x_k = \vec{x} \cdot \mathbf{e}_k$ e $x'_k = \vec{x} \cdot \mathbf{e}'_k$, $k = 1, 2, 3$.

Afirmamos que a matriz R procurada tem elementos de matriz dados por

$$R_{ij} = \mathbf{e}_i \cdot \mathbf{e}'_j, \quad i, j = 1, 2, 3. \tag{21.122}$$

Antes de provarmos que a mesma pertence a $SO(3)$, calculemos sua ação sobre os vetores da base canônica. Teremos

$$(R\mathbf{e}_k)_i = \sum_{j=1}^3 R_{ij}(\mathbf{e}_k)_j = \sum_{j=1}^3 R_{ij} \delta_{jk} = R_{ik} = \mathbf{e}_i \cdot \mathbf{e}'_k.$$

Logo,

$$R\mathbf{e}_k = \sum_{i=1}^3 (R\mathbf{e}_k)_i \mathbf{e}_i = \sum_{i=1}^3 (\mathbf{e}_i \cdot \mathbf{e}'_k) \mathbf{e}_i = \mathbf{e}'_k, \quad k = 1, 2, 3.$$

Como $\mathbf{e}_3 = \mathbf{e}_1 \times \mathbf{e}_2$, o fato que $R\mathbf{e}_3 = \mathbf{e}'_3$ já seguiria diretamente do Lema 21.5, página 1061.

Da definição vê-se que a k -ésima coluna de R é o vetor \mathbf{e}'_k (na base \mathcal{E}). Com isso, a condição de ortogonalidade $RR^T = \mathbb{1}$ segue da ortonormalidade dos elementos de \mathcal{E}' . A condição $\det(R) = 1$ segue do fato de \mathcal{E}' ser positivamente orientado. ■

A Proposição 21.10 será relevante para nossa discussão sobre os ângulos de Euler, a ser feita mais abaixo.

Resumimos a discussão acima da seguinte forma:

Corolário 21.2 *O grupo $SO(3)$ pode ser bijetivamente associado ao conjunto de todos pares ordenados $(\mathbf{e}'_1, \mathbf{e}'_2) \in \mathbb{R}^3 \times \mathbb{R}^3$ com $\mathbf{e}'_i \cdot \mathbf{e}'_j = \delta_{ij}$, $i, j = 1, 2$.* □

Os vetores unitários \mathbf{e}'_1 podem ser associados bijetivamente a pontos da esfera unitária \mathbb{S}^2 . Para cada \mathbf{e}'_1 , os vetores \mathbf{e}'_2 que são ortogonais a \mathbf{e}'_1 podem ser bijetivamente associados ao círculo unitário \mathbb{S}^1 . Assim, concluímos que $SO(3) \simeq \mathbb{S}^2 \times \mathbb{S}^1$.

21.4.2.2 $SO(3)$ e os Ângulos de Euler

Nesta seção expomos mais uma forma de expressar as matrizes de rotação $R \in SO(3)$, desta vez em termos dos chamados *ângulos de Euler*. Essa nova forma é relevante na descrição do movimento de corpos rígidos na Mecânica Clássica, na Astronomia (Dinâmica Planetária) e também surge no contexto da Mecânica Quântica. Algumas dessas aplicações à Mecânica Clássica serão discutidas na Seção 44.4, página 2407. Usaremos diversos resultados anteriores, especialmente da Seção 21.4.2.1, página 1061, mas começamos com mais alguns resultados preparatórios.

• Mais alguns resultados úteis sobre rotações

Proposição 21.11 *Sejam $\vec{\alpha}$ e $\vec{\beta}$ dois vetores unitários de \mathbb{R}^3 e seja $R(\theta, \vec{\eta}) \in SO(3)$ a matriz de rotação por um ângulo θ em torno de um terceiro vetor unitário $\vec{\eta}$. Então, existirá um ângulo θ_0 tal que $R(\theta_0, \vec{\eta})\vec{\alpha}$ é ortogonal a $\vec{\beta}$, ou seja, $\langle R(\theta_0, \vec{\eta})\vec{\alpha}, \vec{\beta} \rangle_{\mathbb{R}} = 0$, se e somente se a condição*

$$(\vec{\alpha} \cdot \vec{\beta})^2 - 2(\vec{\alpha} \cdot \vec{\beta})(\vec{\eta} \cdot \vec{\alpha})(\vec{\eta} \cdot \vec{\beta}) + ((\vec{\alpha} \times \vec{\beta}) \cdot \vec{\eta})^2 \geq 0 \tag{21.123}$$

for satisfeita. □

Prova. Por (21.106) tem-se

$$\langle R(\theta, \vec{\eta})\vec{\alpha}, \vec{\beta} \rangle_{\mathbb{R}} = \cos(\theta)\vec{\alpha} \cdot \vec{\beta} + (1 - \cos(\theta))(\vec{\eta} \cdot \vec{\alpha})(\vec{\eta} \cdot \vec{\beta}) + \sin(\theta)(\vec{\alpha} \times \vec{\beta}) \cdot \vec{\eta}.$$

(Acima usamos que $(\vec{\eta} \times \vec{\alpha}) \cdot \vec{\beta} = (\vec{\alpha} \times \vec{\beta}) \cdot \vec{\eta}$, identidade (4.13), página 260). Desejamos, portanto, saber se existe um ângulo θ tal que

$$\cos(\theta) \left[\vec{\alpha} \cdot \vec{\beta} - (\vec{\eta} \cdot \vec{\alpha}) (\vec{\eta} \cdot \vec{\beta}) \right] + \sin(\theta) \left[(\vec{\alpha} \times \vec{\beta}) \cdot \vec{\eta} \right] = -(\vec{\eta} \cdot \vec{\alpha}) (\vec{\eta} \cdot \vec{\beta}). \quad (21.124)$$

Afirmamos que, para A e B constantes reais, a imagem da função

$$f(\theta) = A \cos(\theta) + B \sin(\theta), \quad \theta \in \mathbb{R},$$

é o intervalo fechado $\left[-\sqrt{A^2 + B^2}, \sqrt{A^2 + B^2} \right]$. Se $\sqrt{A^2 + B^2} = 0$ essa afirmação é óbvia (pois aí $A = B = 0$ e f é identicamente nula). Se $\sqrt{A^2 + B^2} \neq 0$, defina-se $\sin(\delta) := A/\sqrt{A^2 + B^2}$ e $\cos(\delta) := B/\sqrt{A^2 + B^2}$ e teremos

$$f(\theta) = \sqrt{A^2 + B^2} (\cos(\theta) \sin(\delta) + \sin(\theta) \cos(\delta)) = \sqrt{A^2 + B^2} \sin(\theta + \delta),$$

o que torna a afirmação evidente.

Adotando-se $A = \vec{\alpha} \cdot \vec{\beta} - (\vec{\eta} \cdot \vec{\alpha}) (\vec{\eta} \cdot \vec{\beta})$ e $B = (\vec{\alpha} \times \vec{\beta}) \cdot \vec{\eta}$, a equação (21.124) será satisfeita para algum ângulo θ_0 se e somente se

$$-(\vec{\eta} \cdot \vec{\alpha}) (\vec{\eta} \cdot \vec{\beta}) \in \left[-\sqrt{A^2 + B^2}, \sqrt{A^2 + B^2} \right],$$

ou seja, se e somente se $(\vec{\eta} \cdot \vec{\alpha})^2 (\vec{\eta} \cdot \vec{\beta})^2 \leq A^2 + B^2$. Expandindo-se o lado direito, vê-se facilmente que essa condição é idêntica à condição (21.123) ■

O corolário a seguir será usado quando da nossa discussão sobre a parametrização de $\text{SO}(3)$ em termos dos chamados *ângulos de Euler*.

Corolário 21.3 *Sejam $\vec{\alpha}$ e $\vec{\beta}$ dois vetores unitários de \mathbb{R}^3 e seja $R(\theta, \vec{\eta}) \in \text{SO}(3)$ a matriz de rotação por um ângulo θ em torno de um terceiro vetor unitário $\vec{\eta}$. Assuma que $\vec{\eta}$ é ortogonal a $\vec{\alpha}$. Então, existirá um ângulo θ_0 tal que $R(\theta_0, \vec{\eta})\vec{\alpha}$ é ortogonal a $\vec{\beta}$, ou seja, $\langle R(\theta_0, \vec{\eta})\vec{\alpha}, \vec{\beta} \rangle_{\mathbb{R}} = 0$. □*

Prova. Sob a hipótese que $\vec{\eta} \cdot \vec{\alpha} = 0$ a condição (21.123) fica $(\vec{\alpha} \cdot \vec{\beta})^2 + \left((\vec{\alpha} \times \vec{\beta}) \cdot \vec{\eta} \right)^2 \geq 0$, a qual, evidentemente, é sempre satisfeita. ■

O Corolário 21.3 é geometricamente intuitivo, mas é preciso destacar que a condição $\vec{\eta} \cdot \vec{\alpha} = 0$ é importante. Se os três vetores $\vec{\eta}$, $\vec{\alpha}$ e $\vec{\beta}$ forem próximos, ou seja, quase-paralelos, então $R(\theta, \vec{\eta})\vec{\alpha}$ pode nunca ser ortogonal a $\vec{\beta}$: nesse caso, o cone descrito por $R(\theta, \vec{\eta})\vec{\alpha}$ quando θ varia pode nunca interceptar o plano ortogonal a $\vec{\beta}$ que passa pela origem.

• Parametrização de rotações em termos de ângulos de Euler

Vamos agora apresentar o resultado principal da presente seção: a parametrização de rotações em \mathbb{R}^3 em termos dos chamados *ângulos de Euler*.

Seja $R \in \text{SO}(3)$, sejam \mathbf{e}_k , $k = 1, 2, 3$, os vetores da base canônica de \mathbb{R}^3 . No que segue, denotaremos por \mathbf{e}'_k o vetor unitário obtido após a ação de R sobre \mathbf{e}_k , ou seja, $\mathbf{e}'_k := R\mathbf{e}_k$, $k = 1, 2, 3$. Os vetores $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ compõem também, evidentemente, uma base ortonormal em \mathbb{R}^3 . Para o acompanhamento do que segue a Figura 21.2, página 1067, deve ser de utilidade.

Teorema 21.9 (Parametrização de Rotações em Termos dos Ângulos de Euler) *Seja $R \in \text{SO}(3)$ e sejam \mathbf{e}_k , $k = 1, 2, 3$, os vetores da base canônica de \mathbb{R}^3 e vamos assumir que $\mathbf{e}'_3 = R\mathbf{e}_3 \neq \pm\mathbf{e}_3$.*

Então, existem três ângulos $\theta \in (0, \pi)$, $\varphi \in [0, 2\pi)$ e $\psi \in [0, 2\pi)$, denominados ângulos de Euler, tais que R pode ser escrita na forma

$$R = R(\varphi, \mathbf{e}_3)R(\theta, \mathbf{e}_1)R(\psi, \mathbf{e}_3) = R_3(\varphi)R_1(\theta)R_3(\psi), \quad (21.125)$$

(na segunda igualdade apenas empregamos a notação de (21.91), página 1055). A expressão (21.125) afirma que R é um produto de três rotações sucessivas: uma rotação de um ângulo ψ em torno do eixo 3 seguida de uma rotação de um ângulo θ em torno do eixo 1 e, finalmente, seguida de uma rotação de um ângulo φ em torno do eixo 3.

Além disso, R também pode ser escrita na forma

$$R = R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})R(\varphi, \mathbf{e}_3), \quad (21.126)$$

onde $\mathbf{e}'_3 := R\mathbf{e}_3 = R(\theta, \vec{n})\mathbf{e}_3$ e onde

$$\vec{n} := R(\varphi, \mathbf{e}_3)\mathbf{e}_1 = R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1 \quad (21.127)$$

com $\mathbf{e}'_1 := R\mathbf{e}_1$. A parametrização (21.126) afirma que R pode ser escrita como o produto de três rotações sucessivas: uma rotação de um ângulo φ em torno do eixo canônico 3 seguida de uma rotação de um ângulo θ em torno de um certo vetor unitário \vec{n} situado no plano definido por \mathbf{e}_1 e \mathbf{e}_2 e seguida, finalmente, por uma rotação de um ângulo ψ em torno do eixo definido pelo vetor $\mathbf{e}'_3 = R\mathbf{e}_3$. Vide Figura 21.2, página 1067.

No caso em que $\mathbf{e}'_3 = R\mathbf{e}_3 = +\mathbf{e}_3$, podemos escrever R na forma

$$R = R(\varphi, \mathbf{e}_3), \quad (21.128)$$

com $\varphi \in [0, 2\pi)$, e no caso em que $\mathbf{e}'_3 = R\mathbf{e}_3 = -\mathbf{e}_3$, podemos escrever R na forma

$$R = R(\varphi, \mathbf{e}_3)R(\pi, \mathbf{e}_1), \quad (21.129)$$

com $\varphi \in [0, \pi)$. □

Antes de apresentarmos a demonstração desse teorema, façamos alguns comentários.

1. A forma explícita da matriz R em termos dos ângulos de Euler θ , φ e ψ fornecida pela relação (21.125) é exibida na expressão (21.137), página 1069.
2. Os três ângulos de Euler que surgem em (21.125) e (21.126) têm denominações especiais na Mecânica Clássica e na Dinâmica Planetária. O ângulo θ é denominado *ângulo de nutação*, o ângulo φ é denominado *ângulo de precessão* e o ângulo ψ é denominado *ângulo de rotação (intrínseca)*. Vide [231], [17] ou [141] para o emprego dessa nomenclatura na Mecânica Clássica.
3. O vetor unitário \vec{n} que surge em (21.126) é denominado *vetor nodal*. Ele está situado na intersecção do plano definido por \mathbf{e}_1 e \mathbf{e}_2 com o plano definido por $\mathbf{e}'_1 := R\mathbf{e}_1$ e $\mathbf{e}'_2 := R\mathbf{e}_2$. O vetor unitário \vec{n} pode ser obtido rodando o vetor \mathbf{e}_1 por um ângulo φ em torno do eixo canônico 3 (vide a definição (21.127) e vide Figura 21.2, página 1067.) ou rodando o vetor \mathbf{e}'_1 de um ângulo $-\psi$ em torno do eixo \mathbf{e}'_3 (vide a demonstração do Teorema 21.9, logo adiante, e vide Figura 21.2).
4. Na demonstração do Teorema 21.9, logo adiante, provaremos primeiramente a validade da parametrização (21.126) e dela obteremos a parametrização (21.125).
5. A parametrização (21.125) de uma rotação em termos de três rotações sucessivas em torno dos eixos 3, 1 e 3 pode ser estendida a quaisquer outros dois eixos canônicos distintos: se $a, b \in \{1, 2, 3\}$ e $a \neq b$, então existem ângulos θ_{ab} , φ_{ab} e ψ_{ab} tais

$$R = R(\varphi_{ab}, \mathbf{e}_a)R(\theta_{ab}, \mathbf{e}_b)R(\psi_{ab}, \mathbf{e}_a), \quad (21.130)$$

Os ângulos θ_{ab} , φ_{ab} e ψ_{ab} são também denominados ângulos de Euler, mas é importante não confundí-los. A relação (21.126) também se generaliza analogamente. Há ao todo seis convenções distintas possíveis para as escolhas dos dois eixos distintos. Adiante demonstraremos apenas a validade da parametrização (21.125) (eixos 3–1–3) e a extensão a outros pares de eixos distintos ficará óbvia.

6. A convenção que adotamos, com os ângulos de Euler θ , φ e ψ e rotações nos eixos sucessivos 3–1–3 é a mais comum em textos de Mecânica Clássica, sendo adotada, por exemplo em [231], em [17] e em [141] (assim como em alguns textos de Teoria de Grupos, como [284]). Na Mecânica Quântica e na Física Nuclear a convenção mais adotada parece ser a 3–2–3. Áreas específicas da Engenharia adotam outras convenções que não a convenção 3–1–3. A referência [141] discute diversas convenções em um apêndice.

Prova do Teorema 21.9. Seja $R \in SO(3)$, sejam \mathbf{e}_k , $k = 1, 2, 3$, os vetores da base canônica de \mathbb{R}^3 e sejam $\mathbf{e}'_k := R\mathbf{e}_k$, $k = 1, 2, 3$, os vetores da base obtida após a ação de R . Para o acompanhamento do que segue a Figura 21.2, página 1067, deve ser de utilidade.

Vamos primeiramente assumir que $\mathbf{e}'_3 = R\mathbf{e}_3 \neq \pm\mathbf{e}_3$ e deixar os casos $\mathbf{e}'_3 = R\mathbf{e}_3 = \pm\mathbf{e}_3$ para o final.

Como \mathbf{e}'_3 e \mathbf{e}'_1 são ortogonais, podemos evocar o Corolário 21.3, página 1065, (adotando-se $\vec{\eta} \equiv \mathbf{e}'_3$, $\vec{\alpha} \equiv \mathbf{e}'_1$ e $\vec{\beta} \equiv \mathbf{e}_3$) e afirmar que existe algum ângulo $-\psi$ tal que $R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1$ é ortogonal a \mathbf{e}_3 . Se assim for, o vetor $R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1$ está no plano gerado por \mathbf{e}_1 e \mathbf{e}_2 e, portanto, pode ser alcançado a partir de \mathbf{e}_1 por uma rotação de algum ângulo φ em torno de \mathbf{e}_3 . Vale, portanto,

$$R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1 = R(\varphi, \mathbf{e}_3)\mathbf{e}_1. \quad (21.131)$$

É fácil ao leitor convencer-se (vide Figura 21.2, página 1067) que podemos limitar φ e ψ aos domínios $\varphi \in [0, 2\pi)$ e $\psi \in [0, 2\pi)$ ou defini-los módulo 2π .

Vamos definir $\vec{n} \in \mathbb{R}^3$, denominado *vetor nodal*, como sendo o vetor unitário dado por

$$\vec{n} := R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1 = R(\varphi, \mathbf{e}_3)\mathbf{e}_1. \quad (21.132)$$

Afirmamos que \vec{n} é ortogonal a \mathbf{e}_3 e a \mathbf{e}'_3 . De fato,

$$\langle \vec{n}, \mathbf{e}_3 \rangle_{\mathbb{R}} = \langle R(\varphi, \mathbf{e}_3)\mathbf{e}_1, \mathbf{e}_3 \rangle_{\mathbb{R}} = \langle \mathbf{e}_1, R(-\varphi, \mathbf{e}_3)\mathbf{e}_3 \rangle_{\mathbb{R}} = \langle \mathbf{e}_1, \mathbf{e}_3 \rangle_{\mathbb{R}} = 0$$

e

$$\langle \vec{n}, \mathbf{e}'_3 \rangle_{\mathbb{R}} = \langle R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1, \mathbf{e}'_3 \rangle_{\mathbb{R}} = \langle \mathbf{e}'_1, R(\psi, \mathbf{e}'_3)\mathbf{e}'_3 \rangle_{\mathbb{R}} = \langle \mathbf{e}'_1, \mathbf{e}'_3 \rangle_{\mathbb{R}} = 0.$$

Acima, usamos os fatos evidentes que $R(-\varphi, \mathbf{e}_3)\mathbf{e}_3 = \mathbf{e}_3$, e $R(\psi, \mathbf{e}'_3)\mathbf{e}'_3 = \mathbf{e}'_3$, por serem rotações de vetores em torno de seus próprios eixos.

Agora, se \vec{n} é ortogonal a \mathbf{e}_3 e a \mathbf{e}'_3 , então é possível rodar o vetor \mathbf{e}_3 de um certo ângulo θ em torno de \vec{n} de modo a alcançar \mathbf{e}'_3 , ou seja, existe um ângulo θ tal que

$$\mathbf{e}'_3 = R(\theta, \vec{n})\mathbf{e}_3. \quad (21.133)$$

É fácil ao leitor convencer-se (vide Figura 21.2, página 1067) que podemos limitar θ ao domínio $\theta \in (0, \pi)$.

Seja agora $S \in SO(3)$ dada por

$$S := R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})R(\varphi, \mathbf{e}_3)$$

e calculemos sua ação nos vetores \mathbf{e}_1 e \mathbf{e}_3 . Temos

$$S\mathbf{e}_1 = R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})R(\varphi, \mathbf{e}_3)\mathbf{e}_1 \stackrel{(21.132)}{=} R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})\vec{n} = R(\psi, \mathbf{e}'_3)\vec{n} \stackrel{(21.132)}{=} R(\psi, \mathbf{e}'_3)R(-\psi, \mathbf{e}'_3)\mathbf{e}'_1 = \mathbf{e}'_1$$

e temos

$$S\mathbf{e}_3 = R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})R(\varphi, \mathbf{e}_3)\mathbf{e}_3 = R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})\mathbf{e}_3 \stackrel{(21.133)}{=} R(\psi, \mathbf{e}'_3)\mathbf{e}'_3 = \mathbf{e}'_3.$$

Isso estabeleceu que $S\mathbf{e}_1 = \mathbf{e}'_1 = R\mathbf{e}_1$ e $S\mathbf{e}_3 = \mathbf{e}'_3 = R\mathbf{e}_3$. Segue também que $S\mathbf{e}_2 = \mathbf{e}'_2$, pois $S\mathbf{e}_2 = S(\mathbf{e}_3 \times \mathbf{e}_1) \stackrel{(21.116)}{=} (S\mathbf{e}_3) \times (S\mathbf{e}_1) = \mathbf{e}'_3 \times \mathbf{e}'_1 = \mathbf{e}'_2$. Pela Proposição 21.10, página 1063, e pelo Corolário 21.2, página 1064, concluímos disso que $S = R$. Provamos, portanto, que

$$R = R(\psi, \mathbf{e}'_3)R(\theta, \vec{n})R(\varphi, \mathbf{e}_3). \quad (21.134)$$

O lado direito de (21.134) pode ainda ser escrito de uma forma mais limpa, envolvendo apenas rotações em torno de \mathbf{e}_1 e \mathbf{e}_3 . Por (21.133) $\mathbf{e}'_3 = R(\theta, \vec{n})\mathbf{e}_3$. Assim, usando-se (21.117), temos

$$R(\psi, \mathbf{e}'_3) = R(\psi, R(\theta, \vec{n})\mathbf{e}_3) \stackrel{(21.117)}{=} R(\theta, \vec{n})R(\psi, \mathbf{e}_3)R(-\theta, \vec{n}).$$

Substituindo isso no primeiro fator do lado direito de (21.134), temos

$$R = R(\theta, \vec{n})R(\varphi + \psi, \mathbf{e}_3). \quad (21.135)$$

Agora, por (21.132) temos $\vec{n} = R(\varphi, \mathbf{e}_3)\mathbf{e}_1$ e, novamente por (21.117), obtemos

$$R(\theta, \vec{n}) = R(\theta, R(\varphi, \mathbf{e}_3)\mathbf{e}_1) \stackrel{(21.117)}{=} R(\varphi, \mathbf{e}_3)R(\theta, \mathbf{e}_1)R(-\varphi, \mathbf{e}_3).$$

Substituindo isso no primeiro fator do lado direito de (21.135), temos

$$R = R(\varphi, \mathbf{e}_3)R(\theta, \mathbf{e}_1)R(\psi, \mathbf{e}_3). \tag{21.136}$$

As afirmações (21.128) e (21.129) para os casos em que $\mathbf{e}'_3 = R\mathbf{e}_3 = \mathbf{e}_3$ e $\mathbf{e}'_3 = R\mathbf{e}_3 = -\mathbf{e}_3$, respectivamente, são a esta altura evidentes. Isso completa a demonstração do Teorema 21.9. ■

• **Forma explícita de uma rotação em termos de ângulos de Euler**

Na notação das matrizes de rotação de (21.91) obtivemos em (21.136)

$$\begin{aligned} R &= R_3(\varphi)R_1(\theta)R_3(\psi) = \begin{pmatrix} \cos \varphi & -\text{sen} \varphi & 0 \\ \text{sen} \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\text{sen} \theta \\ 0 & \text{sen} \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \psi & -\text{sen} \psi & 0 \\ \text{sen} \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi \cos \psi - \cos \theta \text{sen} \varphi \text{sen} \psi & -\cos \varphi \text{sen} \psi - \cos \theta \text{sen} \varphi \cos \psi & \text{sen} \varphi \text{sen} \theta \\ \text{sen} \varphi \cos \psi + \cos \theta \cos \varphi \text{sen} \psi & \text{sen} \varphi \text{sen} \psi + \cos \theta \cos \varphi \cos \psi & -\cos \varphi \text{sen} \theta \\ \text{sen} \psi \text{sen} \theta & \cos \psi \text{sen} \theta & \cos \theta \end{pmatrix}. \tag{21.137} \end{aligned}$$

E. 21.54 *Exercício.* Verifique! ✦

A expressão (21.137) fornece uma fórmula explícita para os elementos de $SO(3)$ em termos dos ângulos de Euler.

21.4.2.3 A Parametrização de Cayley de $SO(3)$ (e de $SO(n)$)

• **A transformada de Cayley**

No contexto das transformadas de Cayley, da forma como as definiremos, uma matriz A é dita ser uma *matriz excepcional* se $\det(\mathbb{1} + A) = 0$, ou seja, se $\mathbb{1} + A$ não possuir inversa, ou seja, se não possuir -1 como autovalor.

Vamos denotar por $F(n, \mathbb{C}) \subset \text{Mat}(n, \mathbb{C})$ o conjunto de todas as matrizes complexas $n \times n$ não excepcionais, ou seja, de todas as matrizes $A \in \text{Mat}(n, \mathbb{C})$ que tenham a propriedade que $\mathbb{1} + A$ é inversível, ou seja, que não tenham -1 como autovalor. Definimos $F(n, \mathbb{R})$ analogamente.

A chamada *transformada de Cayley*²⁶, \mathcal{C} , é definida para matrizes de $F(n, \mathbb{C})$ (ou de $F(n, \mathbb{R})$) pela seguinte expressão:

$$\mathcal{C}(A) := (\mathbb{1} - A)(\mathbb{1} + A)^{-1}. \tag{21.138}$$

Observe-se que

$$\mathbb{1} + \mathcal{C}(A) = \left((\mathbb{1} + A) + (\mathbb{1} - A) \right) (\mathbb{1} + A)^{-1} = 2(\mathbb{1} + A)^{-1},$$

que, naturalmente, possui inversa, a saber $\frac{1}{2}(\mathbb{1} + A)$. Assim, \mathcal{C} transforma matrizes de $F(n, \mathbb{C})$ em matrizes de $F(n, \mathbb{C})$ (e analogamente para $F(n, \mathbb{R})$).

²⁶Arthur Cayley (1821–1895).

Afirmamos que a transformada de Cayley tem a si mesmo como inversa. De fato, para qualquer $A \in F(n, \mathbb{C})$ (ou $A \in F(n, \mathbb{R})$),

$$\begin{aligned} \mathcal{C}(\mathcal{C}(A)) &= (\mathbb{1} - (\mathbb{1} - A)(\mathbb{1} + A)^{-1})(\mathbb{1} + (\mathbb{1} - A)(\mathbb{1} + A)^{-1})^{-1} \\ &= ((\mathbb{1} + A) - (\mathbb{1} - A))(\mathbb{1} + A)^{-1}(\mathbb{1} + (\mathbb{1} - A)(\mathbb{1} + A)^{-1})^{-1} \\ &= 2A(\mathbb{1} + A)^{-1}(\mathbb{1} + (\mathbb{1} - A)(\mathbb{1} + A)^{-1})^{-1} \\ &= 2A \left[(\mathbb{1} + (\mathbb{1} - A)(\mathbb{1} + A)^{-1})(\mathbb{1} + A) \right]^{-1} \\ &= 2A \left[(\mathbb{1} + A) + (\mathbb{1} - A) \right]^{-1} = A. \end{aligned}$$

Concluimos disso que \mathcal{C} mapeia bijetivamente o conjunto $F(n, \mathbb{C})$ (ou $F(n, \mathbb{R})$) em si mesmo.

• **A transformada de Cayley e os grupos $SO(n)$**

Lembremos que $so(n)$, a álgebra de Lie (real) do grupo $SO(n)$, coincide com o conjunto das matrizes reais anti-simétricas $n \times n$.

O ponto importante para nós é que a transformada de Cayley \mathcal{C} mapeia bijetivamente matrizes de $so(n)$ em matrizes de $SO(n) \cap F(n, \mathbb{R})$, como passaremos a provar.

De fato, se $A \in so(n)$, ou seja, se $A^T = -A$, então iA é autoadjunta e, portanto, tem espectro real. Disso segue que o espectro de A é composto de números imaginários puros, incluindo eventualmente o zero. Assim, toda matriz real anti-simétrica é automaticamente elemento de $F(n, \mathbb{R})$ e sua transformada de Cayley está definida. Agora, para tal A tem-se

$$\begin{aligned} \mathcal{C}(A)^T \mathcal{C}(A) &= (\mathbb{1} + A^T)^{-1}(\mathbb{1} - A^T)(\mathbb{1} - A)(\mathbb{1} + A)^{-1} = (\mathbb{1} - A)^{-1}(\mathbb{1} + A)(\mathbb{1} - A)(\mathbb{1} + A)^{-1} \\ &= (\mathbb{1} + A)(\mathbb{1} - A)^{-1}(\mathbb{1} - A)(\mathbb{1} + A)^{-1} = \mathbb{1}, \end{aligned}$$

provando que $\mathcal{C}(A)$ é ortogonal. Para provar que seu determinante vale 1, observemos que

$$\det(\mathcal{C}(A)) = \frac{\det(\mathbb{1} - A)}{\det(\mathbb{1} + A)} = \frac{\det((\mathbb{1} - A)^T)}{\det(\mathbb{1} + A)} = \frac{\det(\mathbb{1} + A)}{\det(\mathbb{1} + A)} = 1,$$

onde usamos que o determinante de uma matriz é igual ao de sua transposta.

Reciprocamente, se $R \in SO(n) \cap F(n, \mathbb{R})$, então $\mathcal{C}(R)$ é anti-simétrica, o que é atestado pelo seguinte cômputo:

$$\begin{aligned} \mathcal{C}(R)^T &= (\mathbb{1} + R^T)^{-1}(\mathbb{1} - R^T) = (\mathbb{1} + R^{-1})^{-1}(\mathbb{1} - R^{-1}) = (\mathbb{1} + R)^{-1}R(\mathbb{1} - R^{-1}) \\ &= -(\mathbb{1} + R)^{-1}(\mathbb{1} - R) = -(\mathbb{1} - R)(\mathbb{1} + R)^{-1} = -\mathcal{C}(R). \end{aligned}$$

Vemos com isso que a transformada de Cayley \mathcal{C} mapeia bijetivamente $SO(n) \cap F(n, \mathbb{R})$ e o conjunto $so(m)$. Podemos, portanto, parametrizar $SO(n) \cap F(n, \mathbb{R})$ univocamente, por via da transformada de Cayley, por matrizes anti-simétricas reais $n \times n$. Note-se que $SO(n) \cap F(n, \mathbb{R})$ é um conjunto aberto. Essa parametrização é denominada *parametrização de Cayley de $SO(n)$* . Essa parametrização pode ser feita de forma bastante explícita nos casos $n = 2$ e $n = 3$, como veremos agora.

• A transformada de Cayley e SO(2)

Toda matriz $A \in \text{so}(2)$ é da forma $A = \begin{pmatrix} 0 & v \\ -v & 0 \end{pmatrix}$ com $v \in \mathbb{R}$. Para tais matrizes vale a relação algébrica $A^2 = -v^2 \mathbb{1}$, como facilmente se vê. Com isso é fácil verificar que

$$(\mathbb{1} + A)^{-1} = \frac{1}{1 + v^2}(\mathbb{1} - A).$$

E. 21.55 *Exercício.* Prove isso constatando que $\frac{1}{1+v^2}(\mathbb{1} - A)(\mathbb{1} + A) = \mathbb{1}$. Use $A^2 = -v^2 \mathbb{1}$. ✱

É agora fácil calcular $\mathcal{C}(A)$. O resultado é

$$\mathcal{C}(A) = (\mathbb{1} - A)(\mathbb{1} + A)^{-1} = \frac{1 - v^2}{1 + v^2} \mathbb{1} - \frac{2}{1 + v^2} A.$$

E. 21.56 *Exercício.* Mostre isso! ✱

Assim, temos de forma explícita

$$\mathcal{C}(A) = \frac{1}{1 + v^2} \begin{pmatrix} 1 - v^2 & -2v \\ 2v & 1 - v^2 \end{pmatrix}, \quad v \in \mathbb{R}. \tag{21.139}$$

Essa é a parametrização de Cayley de $\text{SO}(2) \cap \text{F}(\mathbb{R}, 2)$.

E. 21.57 *Exercício.* Mostre que adotando-se a a reparametrização $v = \tan(\theta/2)$, $-\pi < \theta < \pi$, a expressão (21.139) fica $\mathcal{C}(A) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$, que é a parametrização, mais familiar, de $\text{SO}(2)$ em termos do ângulo de rotação $\theta \in (-\pi, \pi)$. Observe-se que $\theta = \pi$ está excluído. Isso se dá, pois para esse valor de θ a matriz de $\text{SO}(2)$ correspondente seria $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, que é uma matriz excepcional, ou seja, não é elemento de $\text{SO}(2) \cap \text{F}(\mathbb{R}, 2)$. ✱

• A transformada de Cayley e SO(3)

Toda matriz $A \in \text{so}(3)$ é da forma $A = \vec{v} \cdot \vec{J}$, com $\vec{v} \in \mathbb{R}^3$ e onde $\vec{J} = (J_1, J_2, J_3)$ são os geradores infinitesimais de $\text{SO}(3)$ introduzidos em (21.92)–(21.94), página 1055.

Para $\vec{v} \neq 0$ podemos escrever $A = \vec{v} \cdot \vec{J} = \|\vec{v}\| \vec{\eta} \cdot \vec{J}$, onde $\vec{\eta}$ é o vetor unitário $\vec{\eta} := \vec{v}/\|\vec{v}\|$. Assim, por (21.109), página 1058, tem-se $A^3 = (\vec{v} \cdot \vec{J})^3 = -\|\vec{v}\|^3 \vec{\eta} \cdot \vec{J} = -\|\vec{v}\|^2 A$. Essa relação

$$A^3 = -\|\vec{v}\|^2 A \tag{21.140}$$

para matrizes de $\text{so}(3)$ permite algumas simplificações.

Afirmamos, por exemplo, que se A é a matriz antissimétrica $\vec{v} \cdot \vec{J}$, então

$$(\mathbb{1} + A)^{-1} = \mathbb{1} - \frac{1}{1 + \|\vec{v}\|^2} (A - A^2).$$

Há diversas maneiras de demonstrar isso, a mais direta sendo por verificação:

$$\begin{aligned} \left(\mathbb{1} - \frac{1}{1 + \|\vec{v}\|^2} (A - A^2) \right) (\mathbb{1} + A) &= \mathbb{1} - \frac{1}{1 + \|\vec{v}\|^2} (A - A^2) + A - \frac{1}{1 + \|\vec{v}\|^2} (A^2 - A^3) \\ &= \mathbb{1} - \frac{1}{1 + \|\vec{v}\|^2} (A - A^2) + A - \frac{1}{1 + \|\vec{v}\|^2} (A^2 + \|\vec{v}\|^2 A) = \mathbb{1}, \end{aligned}$$

como facilmente se vê, após cancelamentos simples. Acima usamos o fato já mencionado em (21.140) que $A^3 = -\|\vec{v}\|^2 A$.

Com isso, podemos escrever $\mathcal{C}(A)$ em uma forma mais explícita:

$$\begin{aligned} \mathcal{C}(A) &= (\mathbb{1} - A)(\mathbb{1} + A)^{-1} = (\mathbb{1} - A) \left(\mathbb{1} - \frac{1}{1 + \|\vec{v}\|^2} (A - A^2) \right) \\ &= \mathbb{1} - \frac{1}{1 + \|\vec{v}\|^2} (A - A^2) - A + \frac{1}{1 + \|\vec{v}\|^2} (A^2 - A^3) = \mathbb{1} - \frac{2}{1 + \|\vec{v}\|^2} (A - A^2), \end{aligned}$$

onde usamos novamente (21.140). Verifique!

Assim, escrevendo $v \equiv \|\vec{v}\|$, temos

$$\mathcal{C}(A) = \mathbb{1} - \frac{2v}{1 + v^2} (\vec{\eta} \cdot \vec{J}) + \frac{2v^2}{1 + v^2} (\vec{\eta} \cdot \vec{J})^2. \tag{21.141}$$

Essa é a parametrização de Cayley de $SO(3) \cap F(\mathbb{R}, 3)$ em termos de um número $v \geq 0$ e de um vetor unitário $\vec{\eta} \in \mathbb{R}^3$.

Se realizarmos uma reparametrização, escrevendo $v = \tan(\theta/2)$, com $\theta \in [0, \pi)$, teremos

$$\frac{2v}{1 + v^2} = \text{sen}\theta \quad \text{e} \quad \frac{2v^2}{1 + v^2} = 1 - \cos\theta.$$

Verifique! Com isso

$$\mathcal{C}(A) = \mathbb{1} - \text{sen}(\theta) (\vec{\eta} \cdot \vec{J}) + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{J})^2. \tag{21.142}$$

Após a troca $\vec{\eta} \rightarrow \vec{\eta}$ (uma troca que sempre pode ser feita dada a natureza arbitrária desse vetor unitário) obtemos exatamente a *fórmula de Rodrigues* (21.103), página 1057. A expressão (21.104), página 1057, é sua versão matricial explícita.

E. 21.58 *Exercício.* O que ocorre com a parametrização em $\theta = \pi$? *

21.4.3 O Grupo $O(3)$

O grupo $O(3)$, das matrizes ortogonais reais de ordem 3, contém o subgrupo $SO(3)$, de matrizes ortogonais reais de ordem 3 e de determinante 1, estudado na Seção 21.4.2, página 1053, acima, assim como contém o conjunto complementar $O(3) \setminus SO(3)$ das matrizes ortogonais de determinante -1 . Na maior parte desta seção vamos tentar caracterizar os elementos de $O(3) \setminus SO(3)$, e tentar descrever sua ação sobre vetores do espaço tridimensional \mathbb{R}^3 .

Conforme vimos na Seção 21.4.2, $SO(3)$ é o grupo das rotações puras em \mathbb{R}^3 e, como veremos na presente seção, os elementos de $O(3) \setminus SO(3)$ são composições de rotações puras e de reflexões puras em \mathbb{R}^3 .

• Reflexões sobre planos e sobre a origem

Dentre os elementos de $O(3) \setminus SO(3)$ encontram-se as seguintes matrizes:

$$P_1 := \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_3 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{21.143}$$

$$P_0 := P_1 P_2 P_3 := \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -\mathbb{1}. \tag{21.144}$$

É trivial verificar que se trata, de fato, de matrizes ortogonais de determinante -1 . A matriz P_1 implementa uma reflexão sobre o plano 2-3, a matriz P_2 implementa uma reflexão sobre o plano 3-1, a matriz P_3 implementa uma reflexão sobre o plano 1-2 e a matriz $P_0 \equiv -\mathbb{1}$ implementa uma reflexão sobre a origem.

A ideia pode ser generalizada. Seja $\vec{\zeta} \in \mathbb{R}^3$ um vetor unitário e seja $\{\vec{\zeta}\}^\perp$ o subespaço bidimensional ortogonal a $\vec{\zeta}$. Pelo Teorema da Decomposição Ortogonal, todo $\vec{v} \in \mathbb{R}^3$ pode ser escrito de forma única como $\vec{v} = v_1\vec{\zeta} + \vec{v}_2$, com $v_1 \in \mathbb{R}$ e $\vec{v}_2 \in \{\vec{\zeta}\}^\perp$. Defina-se o operador linear $P_{\vec{\zeta}}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ por

$$P_{\vec{\zeta}}\vec{v} = P_{\vec{\zeta}}(v_1\vec{\zeta} + \vec{v}_2) := -v_1\vec{\zeta} + \vec{v}_2. \tag{21.145}$$

Claro está que $\vec{\zeta}$ é um autovetor de $P_{\vec{\zeta}}$ com autovalor -1 e que o subespaço bidimensional $\{\vec{\zeta}\}^\perp$ é o subespaço dos autovetores com autovalor $+1$. Logo, $\det(P_{\vec{\zeta}}) = -1$. Além disso é um exercício elementar (faça-o!) provar que para quaisquer $\vec{v}, \vec{u} \in \mathbb{R}^3$ vale $\langle P_{\vec{\zeta}}\vec{v}, P_{\vec{\zeta}}\vec{u} \rangle_{\mathbb{R}} = \langle \vec{v}, \vec{u} \rangle_{\mathbb{R}}$, mostrando que $P_{\vec{\zeta}}$ é ortogonal. Logo, $P_{\vec{\zeta}}$ é um elemento de $O(3)$ com determinante -1 e que representa uma reflexão sobre o plano $\{\vec{\zeta}\}^\perp$.

Claro está que $P_k = P_{\mathbf{e}_k}$ para cada $k = 1, 2, 3$. Claro está também que P_0 não coincide com nenhum $P_{\vec{\zeta}}$, pois para P_0 o autovalor -1 é triplamente degenerado, enquanto que para as matrizes $P_{\vec{\zeta}}$ é simplesmente degenerado. Por fim, é evidente da definição que $(P_{\vec{\zeta}})^2 = \mathbb{1}$. Chamamos também a atenção para o fato que $P_{\vec{\zeta}} = P_{-\vec{\zeta}}$ (verifique!).

Elementos de $O(3) \setminus SO(3)$ como P_0 ou como as matrizes $P_{\vec{\zeta}}$ são denominados *reflexões puras* em \mathbb{R}^3 . Vamos agora encontrar a forma geral dos elementos de $O(3) \setminus SO(3)$ e entender como eles agem em \mathbb{R}^3 .

• Os elementos gerais de $O(3) \setminus SO(3)$

O primeiro fato de constatação elementar é que todo elemento de $O(3) \setminus SO(3)$ é obtido pelo produto de qualquer uma das matrizes $P_{\vec{\zeta}}$ ou P_0 acima com uma matriz de $SO(3)$. De fato, se $P \in O(3) \setminus SO(3)$, então $P_{\vec{\zeta}}P$ é também um elemento de $O(3)$, por ser igualmente ortogonal, e ser tal que $\det(P_{\vec{\zeta}}P) = \det(P_{\vec{\zeta}})\det(P) = (-1)^2 = 1$. Logo, $P_{\vec{\zeta}}P \in SO(3)$ e, dado que $(P_{\vec{\zeta}})^2 = \mathbb{1}$, segue que $P = P_{\vec{\zeta}}R$ para alguma $R \in SO(3)$. Note-se que a associação entre P e R é unívoca, pois se $P = P_{\vec{\zeta}}R = P_{\vec{\zeta}}R_2$, valerá $R = R_2$, pois $(P_{\vec{\zeta}})^2 = \mathbb{1}$. Note-se também que, *ipso facto*, $PP_{\vec{\zeta}}$ é também um elemento de $O(3)$ e, portanto, também podemos escrever $P = R'P_{\vec{\zeta}}$ para algum $R' \in SO(3)$, único (mas não necessariamente igual a R). Os mesmos raciocínios aplicam-se a P_0 .

Concluimos disso que os elementos de $O(3) \setminus SO(3)$ representam rotações de algum ângulo em torno de um eixo que passa pela origem, seguidas por reflexões sobre algum plano bidimensional ou sobre a origem (ou na ordem reversa). Provamos, assim, as seguintes afirmações:

Proposição 21.12 *Seja $P \in O(3) \setminus SO(3)$. Então, para qualquer $\vec{\zeta}$ unitário em \mathbb{R}^3 existem R e $R' \in SO(3)$ tais que*

$$P = P_{\vec{\zeta}}R = R'P_{\vec{\zeta}},$$

assim como existem R_0 e $R'_0 \in SO(3)$ tais que $P = P_0R_0 = R'_0P_0$.

Para cada vetor unitário $\vec{\zeta} \in \mathbb{R}^3$, as aplicações de $SO(3)$ em $O(3) \setminus SO(3)$ dadas por

$$SO(3) \ni R \mapsto P_{\vec{\zeta}}R \in O(3) \setminus SO(3) \quad e \quad SO(3) \ni R' \mapsto R'P_{\vec{\zeta}} \in O(3) \setminus SO(3)$$

são bijetoras, com inversas dadas por $O(3) \setminus SO(3) \ni P \mapsto P_{\vec{\zeta}}P \in SO(3)$ e $O(3) \setminus SO(3) \ni P \mapsto PP_{\vec{\zeta}} \in SO(3)$, respectivamente. Ambas as aplicações são contínuas e com inversa contínua e, portanto, são homeomorfismos entre $SO(3)$ e $O(3) \setminus SO(3)$. As mesmas afirmações valem para as aplicações $SO(3) \ni R_0 \mapsto P_0R_0 \in O(3) \setminus SO(3)$ e $SO(3) \ni R'_0 \mapsto R'_0P_0 \in O(3) \setminus SO(3)$. \square

As afirmações sobre continuidade, acima, são óbvias e não requerem demonstração. A Proposição 21.12 mostra que $SO(3)$ e $O(3) \setminus SO(3)$ são homeomorfos e, com a discussão da página 1060 sobre o espaço de parâmetros de $SO(3)$, podemos afirmar que também $O(3) \setminus SO(3)$ pode ser identificado com o espaço projetivo tridimensional \mathbb{RP}^3 (para a definição deste, vide Seção 33.4.6, página 1637).

Reunindo alguns fatos previamente apresentados temos a seguinte afirmação sobre grupo $O(3)$ como um todo:

Proposição 21.13 *O grupo $O(3)$ é composto por duas componentes conexas, $SO(3)$ e $O(3) \setminus SO(3)$, ambas homeomorfas ao espaço projetivo tridimensional real $\mathbb{R}P^3$.* \square

E. 21.59 *Exercício.* Sejam $\vec{\zeta}_1$ e $\vec{\zeta}_2$, dois vetores unitários de \mathbb{R}^3 . Determine explicitamente em função de $\vec{\zeta}_1$ e $\vec{\zeta}_2$ as matrizes $R(\vec{\zeta}_1, \vec{\zeta}_2)$ e $R'(\vec{\zeta}_1, \vec{\zeta}_2)$ de $SO(3)$ que satisfazem $P_{\vec{\zeta}_1} = P_{\vec{\zeta}_2} R(\vec{\zeta}_1, \vec{\zeta}_2)$ e $P_{\vec{\zeta}_1} = R'(\vec{\zeta}_1, \vec{\zeta}_2) P_{\vec{\zeta}_2}$, respectivamente.

Sugestões: No caso em que $\vec{\zeta}_1$ e $\vec{\zeta}_2$ são vetores linearmente dependentes não há o que provar pois, evidentemente, $R(\vec{\zeta}_1, \vec{\zeta}_2) = R'(\vec{\zeta}_1, \vec{\zeta}_2) = \mathbb{1}$. Considere, então, que $\vec{\zeta}_1$ e $\vec{\zeta}_2$ são vetores linearmente independentes. Expresse um vetor geral \vec{v} de \mathbb{R}^3 na base $\{\vec{\zeta}_1, \vec{\zeta}_2, \vec{\zeta}_1 \times \vec{\zeta}_2\}$ (que não é necessariamente uma base ortogonal!) e, usando (21.145), expresse $P_{\vec{\zeta}_1} P_{\vec{\zeta}_2}$ e $P_{\vec{\zeta}_2} P_{\vec{\zeta}_1}$. Convença-se que ambas $P_{\vec{\zeta}_1} P_{\vec{\zeta}_2}$ e $P_{\vec{\zeta}_2} P_{\vec{\zeta}_1}$ representam rotações em torno do eixo $\vec{\zeta}_1 \times \vec{\zeta}_2$ (pois a componente de \vec{v} na direção $\vec{\zeta}_1 \times \vec{\zeta}_2$ é mantida invariante por $P_{\vec{\zeta}_1}$ e por $P_{\vec{\zeta}_2}$) e, usando (21.105)–(21.106), determine os ângulos de rotação a que $R(\vec{\zeta}_1, \vec{\zeta}_2)$ e $R'(\vec{\zeta}_1, \vec{\zeta}_2)$ correspondem. \star

• **Mais sobre os elementos de $O(3) \setminus SO(3)$**

As Proposições 21.14 e 21.15, a seguir, fornecem uma visão talvez mais precisa de $O(3) \setminus SO(3)$ e da ação dos seus elementos.

Proposição 21.14 *Se $P \in O(3) \setminus SO(3)$ com $P \neq -\mathbb{1}$, então podemos encontrar um vetor unitário $\vec{\eta}$ e um ângulo $\theta \in (-\pi, \pi]$ tais que valem*

$$P = R(\theta, \vec{\eta}) P_{\vec{\eta}} = P_{\vec{\eta}} R(\theta, \vec{\eta}) = R(-\theta, -\vec{\eta}) P_{-\vec{\eta}} = P_{-\vec{\eta}} R(-\theta, -\vec{\eta}). \tag{21.146}$$

Para cada $P \in O(3) \setminus SO(3)$, o ângulo θ e o vetor $\vec{\eta}$ são únicos a menos das identificações indicadas em (21.146). Assim tais elementos $P \in O(3) \setminus SO(3)$ com $P \neq -\mathbb{1}$ representam uma reflexão sobre um plano bidimensional específico seguida de uma rotação sobre o eixo ortogonal a esse plano (ou na ordem reversa).

Para o elemento $-\mathbb{1}$ temos as representações $-\mathbb{1} = R(\pm\pi, \vec{\eta}) P_{\vec{\eta}} = P_{\vec{\eta}} R(\pm\pi, \vec{\eta})$, com $\vec{\eta} \in \mathbb{R}^3$ unitário qualquer. Assim, a reflexão sobre a origem também pode ser vista como uma reflexão sobre um plano bidimensional qualquer seguida de uma rotação de $\pm\pi$ sobre o eixo ortogonal a esse plano (ou na ordem reversa). \square

Prova da Proposição 21.14. Seja $P \in O(3) \setminus SO(3)$. Seja $p(x) = \det(x\mathbb{1} - P)$ seu polinômio característico. O polinômio $p(x)$ é de grau 3 e o termo de maior grau é $+x^3$. Logo, $\lim_{x \rightarrow -\infty} p(x) = -\infty$. Por outro lado, $p(0) = \det(-P) = (-1)^3 \det(P) = 1$. Como p é contínuo, positivo em $x = 0$ e vai a $-\infty$ quando $x \rightarrow -\infty$, segue que P deve ter um autovalor λ (uma raiz de p) que seja negativo. Seja $\vec{\eta}$ um autovetor desse autovalor negativo, i.e., $P\vec{\eta} = \lambda\vec{\eta}$ e escolhamos $\|\vec{\eta}\| = 1$. Como P é ortogonal, ela preserva o comprimento de $\vec{\eta}$ e, portanto, $\lambda = -1$ e temos $P\vec{\eta} = -\vec{\eta}$.

Para esse $\vec{\eta}$, seja a matriz $P_{\vec{\eta}}$ como definida acima, que reflete vetores sobre o plano $\{\vec{\eta}\}^\perp$. O produto $PP_{\vec{\eta}}$ é um elemento R de $SO(3)$, como já observamos. Logo, $P = RP_{\vec{\eta}}$. Agora, sabemos que $P\vec{\eta} = -\vec{\eta}$ e que $P_{\vec{\eta}}\vec{\eta} = -\vec{\eta}$. Logo, $R\vec{\eta} = PP_{\vec{\eta}}\vec{\eta} = \vec{\eta}$. Portanto, $R = R(\theta, \vec{\eta})$ para algum $\theta \in (-\pi, \pi]$.

No caso em que $P = -\mathbb{1}$ temos $R(\theta, \vec{\eta}) = -P_{\vec{\eta}}$. Segundo a definição (21.145), se \vec{v} é um vetor arbitrário de \mathbb{R}^3 escrito na forma $\vec{v} = v_1\vec{\eta} + \vec{v}_2$, com $v_1 \in \mathbb{R}$ e $\vec{v}_2 \in \{\vec{\eta}\}^\perp$, então $R(\theta, \vec{\eta})\vec{v} = -P_{\vec{\eta}}(v_1\vec{\eta} + \vec{v}_2) = v_1\vec{\eta} - \vec{v}_2$. Isso mostra que $R(\theta, \vec{\eta})$ é uma rotação de $\pm\pi$ em torno de $\vec{\eta}$, e concluímos que $-\mathbb{1} = R(\pm\pi, \vec{\eta}) P_{\vec{\eta}}$. Ainda sobre o caso $P = -\mathbb{1}$ é importante notar que a escolha de $\vec{\eta}$ entre os vetores unitários de \mathbb{R}^3 é totalmente arbitrária, pois são todos autovetores de $-\mathbb{1}$ com autovalor -1 , obviamente.

Retornando ao caso geral, a igualdade $R(\theta, \vec{\eta}) P_{\vec{\eta}} = P_{\vec{\eta}} R(\theta, \vec{\eta})$ que ocorre no lado direito de (21.148) pode ser facilmente constatada geometricamente, mas uma prova formal pode ser obtida com uso de (21.106), segundo a qual, valem para todo $\vec{\alpha} \in \mathbb{R}^3$

$$\begin{aligned} R(\theta, \vec{\eta}) P_{\vec{\eta}} \vec{\alpha} &= \cos(\theta) P_{\vec{\eta}} \vec{\alpha} + (1 - \cos(\theta)) (\vec{\eta} \cdot (P_{\vec{\eta}} \vec{\alpha})) \vec{\eta} + \sin(\theta) \vec{\eta} \times (P_{\vec{\eta}} \vec{\alpha}) \\ &= \cos(\theta) P_{\vec{\eta}} \vec{\alpha} - (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} + \sin(\theta) \vec{\eta} \times \vec{\alpha} \end{aligned}$$

e

$$\begin{aligned} P_{\vec{\eta}}R(\theta, \vec{\eta})\vec{\alpha} &= \cos(\theta)P_{\vec{\eta}}\vec{\alpha} + (1 - \cos(\theta))(\vec{\eta} \cdot \vec{\alpha})P_{\vec{\eta}}\vec{\eta} + \sin(\theta)P_{\vec{\eta}}(\vec{\eta} \times \vec{\alpha}) \\ &= \cos(\theta)P_{\vec{\eta}}\vec{\alpha} - (1 - \cos(\theta))(\vec{\eta} \cdot \vec{\alpha})\vec{\eta} + \sin(\theta)\vec{\eta} \times \vec{\alpha}. \end{aligned}$$

Acima, além de alguns fatos óbvios, usamos que $P_{\vec{\eta}}(\vec{\eta} \times \vec{\alpha}) = \vec{\eta} \times \vec{\alpha}$, pois $\vec{\eta} \times \vec{\alpha}$ é um vetor de $\{\vec{\eta}\}^\perp$. Comparando, vemos que $R(\theta, \vec{\eta})P_{\vec{\eta}} = P_{\vec{\eta}}R(\theta, \vec{\eta})$, como queríamos mostrar.

Concluimos disso, em particular, que temos

$$-\mathbb{1} = R(\pm\pi, \vec{\eta})P_{\vec{\eta}} = P_{\vec{\eta}}R(\pm\pi, \vec{\eta})$$

com os dois sinais independentes e para qualquer $\vec{\eta} \in \mathbb{R}^3$ unitário.

No caso geral há ainda algumas identificações a se fazer. Notemos que $-\vec{\eta}$ também é autovetor de P com autovalor -1 . Logo, temos também

$$P = R(\theta', -\vec{\eta})P_{-\vec{\eta}} = P_{-\vec{\eta}}R(\theta', -\vec{\eta})$$

para algum $\theta' \in (\pi, \pi]$. Como $P_{\vec{\eta}} = P_{-\vec{\eta}}$ e $R(\theta', -\vec{\eta}) = R(-\theta', \vec{\eta})$, conclui-se facilmente que $R(-\theta', \vec{\eta}) = R(\theta, \vec{\eta})$. Logo, $R(\theta + \theta', \vec{\eta}) = \mathbb{1}$, o que implica que $\theta' + \theta \pmod{2\pi}$. Para $\theta, \theta' \in (-\pi, \pi]$ isso implica $\theta' = -\theta$.

Com isso estabelecemos que todo $P \in O(3) \setminus SO(3)$ se deixa escrever na forma

$$P = R(\theta, \vec{\eta})P_{\vec{\eta}} = P_{\vec{\eta}}R(\theta, \vec{\eta}) = R(-\theta, -\vec{\eta})P_{-\vec{\eta}} = P_{-\vec{\eta}}R(-\theta, -\vec{\eta}), \tag{21.147}$$

para ao menos um par $(\theta, \vec{\eta}) \in [-\pi, \pi] \times \mathbb{S}^2$. Aqui, \mathbb{S}^2 é a esfera unitária de \mathbb{R}^3 : o conjunto de todos os vetores $\vec{\eta} \in \mathbb{R}^3$ com $\|\vec{\eta}\| = 1$.

Vamos agora estabelecer que se $P \in O(3) \setminus SO(3)$ e $P \neq P_0 = -\mathbb{1}$, então o vetor $\vec{\eta}$ acima é único a menos de um sinal.

Suponhamos que $P \in O(3) \setminus SO(3)$ seja tal que haja dois vetores unitários linearmente independentes $\vec{\eta}_1$ e $\vec{\eta}_2 \in \mathbb{R}^3$ tais que $P = R(\theta_1, \vec{\eta}_1)P_{\vec{\eta}_1}$ e $R(\theta_2, \vec{\eta}_2)P_{\vec{\eta}_2}$, com $\theta_1, \theta_2 \in [-\pi, \pi]$. Mas isso informa que $\vec{\eta}_1$ e $\vec{\eta}_2$ são ambos autovetores de P com autovalor -1 . Como P é uma matriz ortogonal, P mantém invariantes os subespaços $\{\vec{\eta}_1\}^\perp$ e $\{\vec{\eta}_2\}^\perp$, e portanto, mantém invariante o subespaço unidimensional $\{\vec{\eta}_1\}^\perp \cap \{\vec{\eta}_2\}^\perp$, que é o subespaço gerado por $\vec{\eta}_1 \times \vec{\eta}_2$. Logo, $P(\vec{\eta}_1 \times \vec{\eta}_2) = \lambda\vec{\eta}_1 \times \vec{\eta}_2$. Como P preserva os comprimentos de vetores (por ser uma matriz ortogonal), devemos ter $\lambda = \pm 1$. Assim, $\vec{\eta}_1, \vec{\eta}_2$ e $\vec{\eta}_1 \times \vec{\eta}_2$ são três autovetores linearmente independentes de P com autovalores $-1, -1$ e $\lambda = \pm 1$. Logo, como $\det(P) = -1$, devemos ter $\lambda = -1$, e como $\vec{\eta}_1, \vec{\eta}_2$ e $\vec{\eta}_1 \times \vec{\eta}_2$ formam uma base em \mathbb{R}^3 isso implicaria que $P = -\mathbb{1}$. Portanto, como $P \neq -\mathbb{1}$, concluimos que $\vec{\eta}_1$ e $\vec{\eta}_2$ não podem ser linearmente independentes e, portanto, $\vec{\eta}_2 = -\vec{\eta}_1$. Em (21.147) vemos que a troca $\vec{\eta} \rightarrow -\vec{\eta}$ tem de ser acompanhada pela troca de θ por $-\theta$ no intervalo $[-\pi, \pi]$. Isso completa a demonstração. ■

A Proposição 21.14, pode ser reformulada em termos um tanto mais sucintos com uso de algumas noções topológicas, como faremos na Proposição 21.15, logo abaixo. Para enunciá-la precisamos de uma definição. Alguma familiaridade com a construção de espaços quociente, como discutido na Seção 33.1.1, página 1588, será útil ao leitor para o que segue.

Considere-se o conjunto $[-\pi, \pi] \times \mathbb{S}^2$ e vamos introduzir no mesmo a seguinte relação de equivalência:

1. Todo elemento $(\theta, \vec{\eta}) \in [-\pi, \pi] \times \mathbb{S}^2$ é declarado equivalente a si mesmo.
2. Dois pares $(\theta_1, \vec{\eta}_1)$ e $(\theta_2, \vec{\eta}_2)$ são equivalentes para quaisquer $\vec{\eta}_1$ e $\vec{\eta}_2 \in \mathbb{S}^2$ se $|\theta_1| = |\theta_2| = \pi$.
3. Cada elemento $(\theta, \vec{\eta}) \in [-\pi, \pi] \times \mathbb{S}^2$ é declarado equivalente ao elemento $(-\theta, -\vec{\eta}) \in [-\pi, \pi] \times \mathbb{S}^2$.

Denotaremos por $[(\theta, \vec{\eta})]$ a classe de equivalência de $[-\pi, \pi] \times \mathbb{S}^2$ pela relação de equivalência acima. Temos,

$$[(\theta, \vec{\eta})] = \begin{cases} \{(\theta, \vec{\eta}), (-\theta, -\vec{\eta})\}, & \text{caso } \theta \neq \pm\pi, \\ \{(-\pi, \vec{\eta}'), \vec{\eta}' \in \mathbb{S}^2\} \cup \{(\pi, \vec{\eta}'), \vec{\eta}' \in \mathbb{S}^2\}, & \text{caso } \theta = \pm\pi. \end{cases}$$

Seja $\mathbb{L}^3 := ([-\pi, \pi] \times \mathbb{S}^2) / \sim$ a coleção de todas as classes de equivalência acima. Com alguma ginástica mental é possível perceber que \mathbb{L}^3 pode ser compreendido como o espaço obtido tomando-se a bola fechada de raio π e identificando-se os pontos antípodas de sua superfície. Essa foi precisamente a descrição que fizemos de $\text{SO}(3)$ o qual, como já comentamos, identifica-se também com o espaço projetivo real tridimensional $\mathbb{R}\mathbb{P}^3$.

Proposição 21.15 *A aplicação $P : \mathbb{L}^3 \rightarrow \text{O}(3) \setminus (\text{SO}(3) \cup \{-\mathbb{1}\})$ dada por*

$$P([\theta, \vec{\eta}]) := R(\theta, \vec{\eta})P_{\vec{\eta}} = P_{\vec{\eta}}R(\theta, \vec{\eta}) \tag{21.148}$$

é uma aplicação bijetora entre \mathbb{L}^3 e $\text{O}(3) \setminus \text{SO}(3)$. □

A prova da Proposição 21.15 é uma mera releitura das afirmações da Proposição 21.14. Comentemos também que, como \mathbb{L}^3 pode ser identificado com $\mathbb{R}\mathbb{P}^3$ e com $\text{SO}(3)$, reobtemos com a Proposição 21.15 a afirmação já feita na Proposição 21.12, e comentários que se lhe seguem, sobre a identificação de $\text{O}(3) \setminus \text{SO}(3)$ com esses espaços.

21.4.4 O Grupo $\text{SU}(2)$

Esta seção é dedicada ao grupo $\text{SU}(2)$, de grande relevância na Mecânica Quântica, na Física Nuclear, na Mecânica Quântica Relativística, na Teoria Quântica de Campos e na Física das Partículas Elementares.

• **As matrizes de Pauli**

De grande importância no estudo do grupo $\text{SU}(2)$ são as chamadas *matrizes de Pauli*²⁷, definidas como

$$\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{e} \quad \sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{21.149}$$

As matrizes de Pauli satisfazem as seguintes relações algébricas: para todos $a, b = 1, 2, 3$ valem

$$[\sigma_a, \sigma_b] := \sigma_a \sigma_b - \sigma_b \sigma_a = 2i \sum_{c=1}^3 \varepsilon_{abc} \sigma_c, \tag{21.150}$$

$$\{\sigma_a, \sigma_b\} := \sigma_a \sigma_b + \sigma_b \sigma_a = 2\delta_{ab} \mathbb{1}, \tag{21.151}$$

$$\sigma_a \sigma_b = \delta_{ab} \mathbb{1} + i \sum_{c=1}^3 \varepsilon_{abc} \sigma_c. \tag{21.152}$$

E. 21.60 *Exercício importantíssimo (todo estudante deve fazê-lo pelo menos uma vez na vida).* Verifique as relações algébricas acima. Note que (21.152) segue diretamente de (21.150)–(21.151) e vice-versa. ★

Note também que as matrizes de Pauli são autoadjuntas (pois $\sigma_a^* = \sigma_a$, $a = 1, 2, 3$) e unitárias (pois $\sigma_a^* = (\sigma_a)^{-1}$, $a = 1, 2, 3$). Note ainda que as quatro matrizes $\mathbb{1}, \sigma_1, \sigma_2, \sigma_3$ formam uma base em $\text{Mat}(\mathbb{C}, 2)$: toda matriz complexa 2×2 pode ser escrita como uma combinação linear complexa das mesmas.

E. 21.61 *Exercício.* Mostre que as matrizes $\mathbb{1}, \sigma_1, \sigma_2, \sigma_3$ são ortonormais em relação ao seguinte produto escalar definido em $\text{Mat}(\mathbb{C}, 2)$: $\langle A, B \rangle := \frac{1}{2} \text{Tr}(A^* B)$. ★

As matrizes de Pauli desempenham um papel importante na Mecânica Quântica, estando associadas aos operadores de spin para partículas de spin $1/2$, tais como o elétron, o próton, o neutron, os quarks e outras.

²⁷Wolfgang Ernst Pauli (1900–1958).

• **A forma geral das matrizes de SU(2)**

Conforme já definimos, o grupo SU(2) é o grupo das matrizes unitárias complexas 2×2 com determinante igual a 1: $SU(2) = \{U \in \text{Mat}(\mathbb{C}, 2) \mid U^* = U^{-1} \text{ e } \det(U) = 1\}$. Vamos começar estudando a forma geral de tais matrizes, procurando uma parametrização conveniente para as mesmas que permitirá estudar as propriedades de SU(2) como um grupo de Lie.

Como toda matriz 2×2 complexa, uma matriz genérica $U \in SU(2)$ é da forma $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, onde $a, b, c, d \in \mathbb{C}$. Vamos estudar a condição $U^{-1} = U^*$. Podemos calcular U^{-1} usando a regra de Laplace, expressão (10.20), página 453: U^{-1} é dada pela transposta da matriz dos cofatores de U dividida pelo determinante de U , que é 1, neste caso. Ou seja, $U^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Assim, $U^{-1} = U^*$ significa nesse caso

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix},$$

ou seja, $c = -\bar{b}$ e $d = \bar{a}$. Logo, $U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$. A condição $\det(U) = 1$ implica, portanto, $|a|^2 + |b|^2 = 1$.

Resumindo:

$$SU(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \text{ onde } a, b \in \mathbb{C} \text{ com } |a|^2 + |b|^2 = 1 \right\}.$$

Escrevendo os números complexos a e b como soma de suas partes real e imaginária: $a = a_1 + ia_2$ e $b = b_1 + ib_2$, com $a_1, a_2, b_1, b_2 \in \mathbb{R}$, poderemos escrever U como uma combinação linear de matrizes de Pauli (e da unidade):

$$U = \begin{pmatrix} a_1 + ia_2 & b_1 + ib_2 \\ -b_1 + ib_2 & a_1 - ia_2 \end{pmatrix} = a_1 \mathbb{1} + i(b_2 \sigma_1 + b_1 \sigma_2 + a_2 \sigma_3). \tag{21.153}$$

Essa expressão será usada adiante.

Vamos agora nos voltar para a condição $|a|^2 + |b|^2 = 1$. A mesma significa $a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1$. Temos então,

$$SU(2) = \left\{ \begin{pmatrix} a_1 + ia_2 & b_1 + ib_2 \\ -b_1 + ib_2 & a_1 - ia_2 \end{pmatrix}, \text{ onde } (a_1, a_2, b_1, b_2) \in \mathbb{R}^4 \text{ com } a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1 \right\}. \tag{21.154}$$

Lembremos que para todo inteiro $n \geq 1$, o conjunto de pontos

$$\mathbb{S}^n := \{(x_1, \dots, x_{n+1}) \in \mathbb{R}^{n+1} \text{ com } x_1^2 + \dots + x_{n+1}^2 = 1\} \subset \mathbb{R}^{n+1}$$

designa a superfície da esfera unitária de \mathbb{R}^{n+1} . Assim, vemos que SU(2) é homeomorfo a \mathbb{S}^3 , a superfície da esfera unitária do espaço quadridimensional \mathbb{R}^4 . Isso ilustra o fato que SU(2) é uma variedade diferenciável. Como o produto e a inversa são contínuos em SU(2), o mesmo é um grupo de Lie.

Comentário. Conforme discutimos à página 213, as relações (21.153) e (21.154) permitem ver que o grupo SU(2) é isomorfo ao grupo \mathbb{H}_1 , o grupo dos quatérnios unitários (vide página 213). Por essa razão, (21.153) é por vezes denominada *representação quaterniônica das matrizes do grupo SU(2)*. ♣

Vamos tentar agora parametrizar de outra forma o vetor $(a_1, a_2, b_1, b_2) \in \mathbb{S}^3$ que aparece do lado direito de (21.154). Claramente, a condição $a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1$ diz que a_1, a_2, b_1 e b_2 são números reais contidos no intervalo $[-1, 1]$. Podemos assim definir um ângulo $\theta \in [-\pi, \pi]$ de forma que

$$a_1 = \cos \theta.$$

Fora isso, para $\cos(\theta) \neq \pm 1$, podemos definir

$$\eta_1 := \frac{b_2}{\text{sen}\theta}, \quad \eta_2 := \frac{b_1}{\text{sen}\theta}, \quad \eta_3 := \frac{a_2}{\text{sen}\theta}.$$

A condição $a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1$ implica então (verifique!) que $\eta_1^2 + \eta_2^2 + \eta_3^2 = 1$. Assim, o vetor $\vec{\eta} := (\eta_1, \eta_2, \eta_3)$ de \mathbb{R}^3 é um vetor de comprimento 1. Com esses novos parâmetros θ e $\vec{\eta}$ podemos reescrever (21.153) como

$$U = \cos(\theta)\mathbb{1} + i \operatorname{sen}(\theta)\vec{\eta} \cdot \vec{\sigma},$$

onde

$$\vec{\eta} \cdot \vec{\sigma} := \eta_1\sigma_1 + \eta_2\sigma_2 + \eta_3\sigma_3 = \begin{pmatrix} \eta_3 & \eta_1 - i\eta_2 \\ \eta_1 + i\eta_2 & -\eta_3 \end{pmatrix}.$$

Assim,

$$\text{SU}(2) = \left\{ \cos(\theta)\mathbb{1} + i \operatorname{sen}(\theta)\vec{\eta} \cdot \vec{\sigma}, \text{ onde } \theta \in [-\pi, \pi] \text{ e } \vec{\eta} \in \mathbb{R}^3 \text{ com } \|\vec{\eta}\| = 1 \right\}.$$

A importância de se expressar $U \in \text{SU}(2)$ dessa forma, em termos de θ e $\vec{\eta}$, provém da seguinte identidade:

$$\cos(\theta)\mathbb{1} + i \operatorname{sen}(\theta)\vec{\eta} \cdot \vec{\sigma} = \exp(i\theta\vec{\eta} \cdot \vec{\sigma}).$$

Vamos provar isso expandindo o lado direito e verificando que é igual ao lado esquerdo. De fato, pela definição da exponencial de matrizes,

$$\exp(i\theta\vec{\eta} \cdot \vec{\sigma}) = \sum_{m=0}^{\infty} \frac{(i\theta)^m}{m!} (\vec{\eta} \cdot \vec{\sigma})^m = \sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} (\vec{\eta} \cdot \vec{\sigma})^{2k} + \sum_{k=0}^{\infty} \frac{(i\theta)^{2k+1}}{(2k+1)!} (\vec{\eta} \cdot \vec{\sigma})^{2k+1},$$

onde, na última linha, apenas fizemos separar a soma em m da primeira linha nos casos m par e m ímpar. É um exercício muito fácil (faça!) verificar que

$$(\vec{\eta} \cdot \vec{\sigma})^2 = \begin{pmatrix} \eta_3 & \eta_1 - i\eta_2 \\ \eta_1 + i\eta_2 & -\eta_3 \end{pmatrix}^2 = \mathbb{1}.$$

Portanto, $(\vec{\eta} \cdot \vec{\sigma})^{2k} = \mathbb{1}$ e $(\vec{\eta} \cdot \vec{\sigma})^{2k+1} = \vec{\eta} \cdot \vec{\sigma}$. Logo,

$$\begin{aligned} \exp(i\theta\vec{\eta} \cdot \vec{\sigma}) &= \left(\sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} \right) \mathbb{1} + \left(\sum_{k=0}^{\infty} \frac{(i\theta)^{2k+1}}{(2k+1)!} \right) \vec{\eta} \cdot \vec{\sigma} \\ &= \cos(\theta)\mathbb{1} + i \operatorname{sen}(\theta)\vec{\eta} \cdot \vec{\sigma}, \end{aligned} \tag{21.155}$$

que é o que queríamos mostrar.

Resumindo nossas conclusões,

$$\text{SU}(2) = \left\{ \exp(i\theta\vec{\eta} \cdot \vec{\sigma}) \text{ onde } \theta \in [-\pi, \pi] \text{ e } \vec{\eta} \in \mathbb{R}^3 \text{ com } \|\vec{\eta}\| = 1 \right\}. \tag{21.156}$$

Se tomarmos $\vec{\eta}_1 = (1, 0, 0)$, $\vec{\eta}_2 = (0, 1, 0)$ ou $\vec{\eta}_3 = (0, 0, 1)$, obtemos três subgrupos uniparamétricos distintos de $\text{SU}(2)$:

$$U_1(\theta) := \exp(i\theta\sigma_1) = \begin{pmatrix} \cos \theta & i \operatorname{sen} \theta \\ i \operatorname{sen} \theta & \cos \theta \end{pmatrix}, \tag{21.157}$$

$$U_2(\theta) := \exp(i\theta\sigma_2) = \begin{pmatrix} \cos \theta & \operatorname{sen} \theta \\ -\operatorname{sen} \theta & \cos \theta \end{pmatrix}, \tag{21.158}$$

$$U_3(\theta) := \exp(i\theta\sigma_3) = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \tag{21.159}$$

respectivamente. Isso nos permite identificar as matrizes $i\sigma_1$, $i\sigma_2$ e $i\sigma_3$ como os geradores infinitesimais desses subgrupos uniparamétricos. As relações (21.150) são as relações de comutação satisfeitas por essas matrizes, como elementos de uma álgebra de Lie, que é denominada álgebra de Lie $\mathfrak{su}(2)$.

Com isso, (21.156) está nos dizendo que todo elemento de $SU(2)$ pode ser escrito como exponencial de um elemento de sua álgebra de Lie. Isso constata um teorema geral (vide, por exemplo, [347]) que diz que se um grupo de Lie é compacto e sua álgebra de Lie é semissimples, a aplicação exponencial da sua álgebra de Lie é sobrejetora no grupo. De fato, tal como $SO(3)$, $SU(2)$ é compacto e $\mathfrak{su}(2)$ é semissimples.

E. 21.62 *Exercício.* Mostre que

$$U(2) = \left\{ \exp(i\alpha\mathbb{1} + i\theta\vec{\eta} \cdot \vec{\sigma}) \text{ onde } \alpha, \theta \in [-\pi, \pi] \text{ e } \vec{\eta} \in \mathbb{R}^3 \text{ com } \|\vec{\eta}\| = 1 \right\}.$$

★

• Parametrização de elementos de $SU(2)$ em termos de ângulos de Euler

Sabemos que o grupo $SU(2)$ é formado por matrizes da forma $U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$ com $a, b \in \mathbb{C}$ satisfazendo $|a|^2 + |b|^2 = 1$. Vamos escrever $a = a_1 + ia_2$ e $b = b_1 + ib_2$, com a_k e b_k reais. A condição $|a|^2 + |b|^2 = 1$ equivale a podermos escrever $|a| = \cos(\theta)$ e $|b| = \sin(\theta)$ para algum $\theta \in [0, \pi/2]$. Assim, podemos escrever $a = \cos(\theta)e^{i\alpha}$ e $b = i\sin(\theta)e^{i\beta}$, com α e $\beta \in (-\pi, \pi]$ sendo as fases de a e $-ib$, respectivamente. Definamos $\varphi = (\alpha + \beta)/2 \pmod{2\pi}$ e $\psi = (\alpha - \beta)/2 \pmod{2\pi}$. Então, temos $\varphi, \psi \in (-\pi, \pi]$ e podemos escrever $a = \cos(\theta)e^{i(\varphi+\psi)}$ e $b = i\sin(\theta)e^{i(\varphi-\psi)}$. Com isso, podemos escrever $U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$ na forma

$$U \equiv U(\varphi, \theta, \psi) := \begin{pmatrix} \cos(\theta)e^{i(\varphi+\psi)} & i\sin(\theta)e^{i(\varphi-\psi)} \\ i\sin(\theta)e^{-i(\varphi-\psi)} & \cos(\theta)e^{-i(\varphi+\psi)} \end{pmatrix}. \tag{21.160}$$

Temos, por (21.160),

$$U(\varphi, 0, 0) = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix} \stackrel{(21.159)}{=} \exp(i\varphi\sigma_3),$$

$$U(0, \theta, 0) = \begin{pmatrix} \cos(\theta) & i\sin(\theta) \\ i\sin(\theta) & \cos(\theta) \end{pmatrix} \stackrel{(21.157)}{=} \exp(i\theta\sigma_1),$$

$$U(0, 0, \psi) = \begin{pmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{pmatrix} \stackrel{(21.159)}{=} \exp(i\psi\sigma_3).$$

É claro disso que $\{U(\varphi, 0, 0), \varphi \in (-\pi, \pi]\}$, $\{U(0, \theta, 0), \theta \in [0, \pi]\}$ e $\{U(0, 0, \psi), \psi \in (-\pi, \pi]\}$ são três subgrupos uniparamétricos de $SU(2)$. Importante, para nós, porém, é notar que podemos escrever

$$U(\varphi, \theta, \psi) = \exp(i\varphi\sigma_3) \exp(i\theta\sigma_1) \exp(i\psi\sigma_3). \tag{21.161}$$

E. 21.63 *Exercício.* Verifique!

★

A equação (21.161) é a versão para o grupo $SU(2)$ da parametrização em termos de *ângulos de Euler* que discutimos com detalhe, no caso do grupo $SO(3)$, na Seção 21.4.2.2, página 1064. A forma explícita da parametrização (21.161) é dada em (21.160).

Outras parametrizações também são possíveis. Se tivéssemos escrito $a = \cos(\theta)e^{i(\varphi+\psi)}$ e $b = \sin(\theta)e^{i(\varphi-\psi)}$ obteríamos a parametrização

$$U(\varphi, \theta, \psi) = \exp(i\varphi\sigma_3) \exp(i\theta\sigma_2) \exp(i\psi\sigma_3) . \tag{21.162}$$

E. 21.64 *Exercício.* Verifique! ✱

Essas várias parametrizações foram discutidas no caso do grupo $SO(3)$ na Seção 21.4.2.2, página 1064. O fato de tanto $SO(3)$ quanto $SU(2)$ possuírem parametrizações em termos de ângulos de Euler deve-se à relação próxima entre esses dois grupos, a ser precisada na Seção 21.4.5, página 1081.

• **Parametrização de elementos de $SU(2)$ em termos de ângulos de Tait-Bryan**

No exercício que segue vamos provar que todo elemento de $SU(2)$ pode também ser escrito na forma de um produto do tipo $\exp(i\phi_1\sigma_1) \exp(i\phi_2\sigma_2) \exp(i\phi_3\sigma_3)$. Essa parametrização é denominada *parametrização de Tait-Bryan de $SU(2)$* ²⁸. A forma explícita dessa parametrização é dada em (21.164), abaixo.

E. 21.65 *Exercício dirigido.* Sabemos que o grupo $SU(2)$ é formado por matrizes da forma $U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$ com $a, b \in \mathbb{C}$ satisfazendo $|a|^2 + |b|^2 = 1$. Mostre que escolhendo

$$a = (\cos(\phi_1) \cos(\phi_2) - i \sin(\phi_1) \sin(\phi_2))e^{i\phi_3} \quad e \quad b = (\cos(\phi_1) \sin(\phi_2) + i \sin(\phi_1) \cos(\phi_2))e^{-i\phi_3} \tag{21.163}$$

com $\phi_j \in (-\pi, \pi]$, a condição $|a|^2 + |b|^2 = 1$ é satisfeita.

Com isso, podemos parametrizar os elementos de $SU(2)$ como

$$U(\phi_1, \phi_2, \phi_3) = \begin{pmatrix} (\cos(\phi_1) \cos(\phi_2) - i \sin(\phi_1) \sin(\phi_2))e^{i\phi_3} & (\cos(\phi_1) \sin(\phi_2) + i \sin(\phi_1) \cos(\phi_2))e^{-i\phi_3} \\ (-\cos(\phi_1) \sin(\phi_2) + i \sin(\phi_1) \cos(\phi_2))e^{i\phi_3} & (\cos(\phi_1) \cos(\phi_2) + i \sin(\phi_1) \sin(\phi_2))e^{-i\phi_3} \end{pmatrix}, \tag{21.164}$$

com $\phi_j \in (-\pi, \pi]$.

Em seguida, mostre que $\{U(\phi_1, 0, 0), \phi_1 \in (-\pi, \pi]\}$, $\{U(0, \phi_2, 0), \phi_2 \in (-\pi, \pi]\}$ e $\{U(0, 0, \phi_3), \phi_3 \in (-\pi, \pi]\}$ são três subgrupos uniparamétricos de $SU(2)$.

Mostre que os geradores infinitesimais desses subgrupos uniparamétricos são $i\sigma_1, i\sigma_2$ e $i\sigma_3$, respectivamente.

Por fim, mostre que

$$U(\phi_1, \phi_2, \phi_3) = \exp(i\phi_1\sigma_1) \exp(i\phi_2\sigma_2) \exp(i\phi_3\sigma_3) \tag{21.165}$$

usando (21.157)–(21.159) e calculando explicitamente o produto do lado direito.

A justificação da parametrização (21.163) se dá como segue. Escrevamos $a = a_1 + ia_2$ e $b = b_1 + ib_2$, com a_k e b_k reais. Definamos $a', b' \in \mathbb{C}$ por $a' = ae^{-i\phi_3}$ e $b' = be^{i\phi_3}$, com $\phi_3 \in \mathbb{R}$ a ser fixado adiante. Com isso, escrevemos

$$a = (a'_1 + ia'_2)e^{i\phi_3} \quad e \quad a = (b'_1 + ib'_2)e^{-i\phi_3} . \tag{21.166}$$

É claro que $|a|^2 + |b|^2 = 1$ se e somente se $|a'|^2 + |b'|^2 = 1$. Escrevamos $a' = a'_1 + ia'_2$ e $b' = b'_1 + ib'_2$ com a'_k e b'_k reais para todo k . A condição $|a|^2 + |b|^2 = 1$ equivale à condição $(a'_1)^2 + (a'_2)^2 + (b'_1)^2 + (b'_2)^2 = 1$. Definamos $c, d \in \mathbb{C}$ por $c = a'_1 + ib'_1$ e $d = b'_2 - ia'_2$.

Vamos agora fixar ϕ_3 através da imposição que c e d tenham a mesma fase: $b'_1/a'_1 = -a'_2/b'_2$, ou seja, que $b'_1b'_2 = -a'_1a'_2$. Escrevendo a'_1, a'_2, b'_1 e b'_2 em termos de a_1, a_2, b_1 e b_2 e ϕ_3 , mostre que a condição $b'_1b'_2 = -a'_1a'_2$ equivale a

$$(a_1a_2 + b_1b_2) \cos(2\phi_3) + (a_1^2 + b_1^2 - a_2^2 - b_2^2) \sin(2\phi_3) = 0 .$$

Constata-se que essa condição sempre pode ser satisfeita para algum $\phi_3 \in \mathbb{R}$.

Se a condição $(a'_1)^2 + (a'_2)^2 + (b'_1)^2 + (b'_2)^2 = 1$ for satisfeita, podemos escrever $(a'_1)^2 + (b'_1)^2 = (\cos(\phi_1))^2$ e $(a'_2)^2 + (b'_2)^2 = (\sin(\phi_1))^2$ para algum ângulo ϕ_1 . Se a condição $(a'_1)^2 + (b'_1)^2 = (\cos(\phi_1))^2$ for satisfeita, podemos escrever

$$a'_1 = \cos(\phi_1) \cos(\phi_2) \quad e \quad b'_1 = \cos(\phi_1) \sin(\phi_2)$$

²⁸Peter Guthrie Tait (1831–1901). George Hartley Bryan (1864–1928).

para algum ângulo ϕ_2 . Se a condição $(a'_2)^2 + (b'_2)^2 = (\text{sen}(\phi_1))^2$ for satisfeita, então podemos escrever

$$a'_2 = -\text{sen}(\phi_1)\text{sen}(\phi_4) \quad \text{e} \quad b'_2 = \text{sen}(\phi_1)\cos(\phi_4)$$

para algum ângulo ϕ_4 . Agora, a escolha de ϕ_3 impõe a condição $b'_1 b'_2 = -a'_1 a'_2$, ou seja, $\cos(\phi_1)\text{sen}(\phi_2)\text{sen}(\phi_1)\cos(\phi_4) = \cos(\phi_1)\cos(\phi_2)\text{sen}(\phi_1)$. Isto é, $\text{sen}(\phi_2)\cos(\phi_4) = \cos(\phi_2)\text{sen}(\phi_4)$. Isso implica que devemos ter $\phi_4 = \phi_2$ e, portanto,

$$a'_1 = \cos(\phi_1)\cos(\phi_2), \quad a'_2 = -\text{sen}(\phi_1)\text{sen}(\phi_2), \quad b'_1 = \cos(\phi_1)\text{sen}(\phi_2), \quad b'_2 = \text{sen}(\phi_1)\cos(\phi_2).$$

Retornando com isso a (21.166), temos a parametrização (21.163). *

21.4.5 A Relação Entre SO(3) e SU(2)

O leitor que acompanhou com atenção as exposições precedentes sobre os grupos SO(3) e SU(2) certamente apercebeu-se da existência de uma série de semelhanças entre ambos. Vamos agora precisá-las.

Em primeiro lugar, note-se que os geradores infinitesimais de SO(3) são matrizes 3×3 satisfazendo as relações algébricas $[J_a, J_b] = \sum_{c=1}^3 \varepsilon_{abc} J_c$, enquanto que geradores infinitesimais de SU(2) são matrizes 2×2 satisfazendo as relações algébricas $[\sigma_a, \sigma_b] = 2i \sum_{c=1}^3 \varepsilon_{abc} \sigma_c$. Se, porém, definirmos $j_a := -i\sigma_a/2$, obtemos $[j_a, j_b] = \sum_{c=1}^3 \varepsilon_{abc} j_c$, as mesmas relações de comutação dos geradores infinitesimais de SO(3).

Seja

$$\text{so}(3) := \left\{ L \in \text{Mat}(\mathbb{R}, 3) : L = \alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3, \alpha_k \in \mathbb{R}, k = 1, 2, 3 \right\}$$

a álgebra de Lie (real) associada aos geradores infinitesimais de SO(3) e seja

$$\text{su}(2) := \left\{ l \in \text{Mat}(\mathbb{C}, 2) : l = \alpha_1 j_1 + \alpha_2 j_2 + \alpha_3 j_3, \alpha_k \in \mathbb{R}, k = 1, 2, 3 \right\}$$

a álgebra de Lie (real) associada aos geradores infinitesimais de SU(2).

E. 21.66 *Exercício.* Mostre que $\text{so}(3)$ coincide com a álgebra de Lie de todas as matrizes reais 3×3 antissimétricas. (Vide exercício à página 108). *

E. 21.67 *Exercício.* Mostre que $\text{su}(2)$ coincide com a álgebra de Lie de todas as matrizes complexas 2×2 antiautoadjuntas e de traço nulo. (Vide exercício à página 108). *

É muito fácil constatar que a aplicação linear $\varphi : \text{su}(2) \rightarrow \text{so}(3)$ dada por

$$\varphi(\alpha_1 j_1 + \alpha_2 j_2 + \alpha_3 j_3) = \alpha_1 J_1 + \alpha_2 J_2 + \alpha_3 J_3$$

é um *isomorfismo de álgebras de Lie*, ou seja, é bijetora e satisfaz

$$\varphi([l_a, l_b]) = [\varphi(l_a), \varphi(l_b)] \tag{21.167}$$

para todos $l_a, l_b \in \text{su}(2)$.

E. 21.68 *Exercício importante.* Prove as afirmativas acima. *

Assim, as álgebras de Lie $\text{so}(3)$ e $\text{su}(2)$ são isomorfas. Discutiremos agora que implicações isso traz sobre a relação entre os grupos SO(3) e SU(2).

O isomorfismo φ definido acima sugere considerar-se a seguinte aplicação $\phi : \text{SU}(2) \rightarrow \text{SO}(3)$ dada por

$$\phi(\exp(l)) := \exp(\varphi(l)), \quad \forall l \in \text{su}(2),$$

ou seja,

$$\phi\left(\exp(\theta \vec{\eta} \cdot \vec{j})\right) := \exp\left(\theta \vec{\eta} \cdot \vec{J}\right), \tag{21.168}$$

para todos $\theta \in (-2\pi, 2\pi]$, e $\vec{\eta} \in \mathbb{R}^3$ com $\|\vec{\eta}\| = 1$. Antes de estudarmos as propriedades da aplicação (21.168) é preciso mostrar que a mesma está bem definida enquanto aplicação de $SU(2)$ em $SO(3)$, no seguinte sentido: se $\exp(\theta\vec{\eta} \cdot \vec{j})$ e $\exp(\theta'\vec{\eta}' \cdot \vec{j})$ representarem um mesmo elemento de $SU(2)$ então a imagem de ambos por ϕ é a mesma, ou seja, se

$$\exp(\theta\vec{\eta} \cdot \vec{j}) = \exp(\theta'\vec{\eta}' \cdot \vec{j}), \quad \text{então} \quad \exp(\theta\vec{\eta} \cdot \vec{J}) = \exp(\theta'\vec{\eta}' \cdot \vec{J}).$$

Sabemos por (21.155) que $\exp(\theta\vec{\eta} \cdot \vec{j}) = \exp(-i\frac{\theta}{2}\vec{\eta} \cdot \vec{\sigma}) = \cos(\frac{\theta}{2})\mathbb{1} - i\sin(\frac{\theta}{2})\vec{\eta} \cdot \vec{\sigma}$. Assim, $\exp(\theta\vec{\eta} \cdot \vec{j}) = \exp(\theta'\vec{\eta}' \cdot \vec{j})$ ocorre se e somente se

$$\cos\left(\frac{\theta}{2}\right) = \cos\left(\frac{\theta'}{2}\right) \quad \text{e} \quad \sin\left(\frac{\theta}{2}\right)\vec{\eta} = \sin\left(\frac{\theta'}{2}\right)\vec{\eta}', \quad (21.169)$$

pois as matrizes $\mathbb{1}$, σ_1 , σ_2 e σ_3 compõem uma base em $\text{Mat}(\mathbb{C}, 2)$. A primeira relação em (21.169) garante que $\sin(\frac{\theta}{2}) = \pm \sin(\frac{\theta'}{2})$ e, portanto, temos pela segunda relação que $\vec{\eta} = \pm\vec{\eta}'$.

Agora, pela fórmula de Rodrigues (21.103), temos

$$\begin{aligned} \exp(\theta\vec{\eta} \cdot \vec{J}) &\stackrel{(21.103)}{=} \mathbb{1} + (1 - \cos(\theta))(\vec{\eta} \cdot \vec{J})^2 + \sin(\theta)(\vec{\eta} \cdot \vec{J}) \\ &= \mathbb{1} + 2\left[\sin\left(\frac{\theta}{2}\right)(\vec{\eta} \cdot \vec{J})\right]^2 + 2\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right)(\vec{\eta} \cdot \vec{J}) \\ &\stackrel{(21.169)}{=} \mathbb{1} + 2\left[\sin\left(\frac{\theta'}{2}\right)(\vec{\eta}' \cdot \vec{J})\right]^2 + 2\cos\left(\frac{\theta'}{2}\right)\sin\left(\frac{\theta'}{2}\right)(\vec{\eta}' \cdot \vec{J}) \\ &= \mathbb{1} + (1 - \cos(\theta'))(\vec{\eta}' \cdot \vec{J})^2 + \sin(\theta')(\vec{\eta}' \cdot \vec{J}) \\ &\stackrel{(21.103)}{=} \exp(\theta'\vec{\eta}' \cdot \vec{J}). \end{aligned}$$

Na segunda (e na quarta) igualdade, acima, usamos identidades trigonométricas bem conhecidas e na terceira igualdade usamos as relações (21.169).

Assim, provamos que se $\exp(\theta\vec{\eta} \cdot \vec{j}) = \exp(\theta'\vec{\eta}' \cdot \vec{j})$, então $\exp(\theta\vec{\eta} \cdot \vec{J}) = \exp(\theta'\vec{\eta}' \cdot \vec{J})$ e vemos com isso que ϕ dada em (21.168) está realmente bem definida como aplicação de $SU(2)$ em $SO(3)$. Vamos agora estudar suas propriedades objetivando provar que ϕ é um homomorfismo.

Em primeiro lugar, é fácil ver que ϕ é sobrejetora (por quê?), mas não é injetora, pois para $U_1 := \exp(-i\frac{0}{2}\vec{\eta} \cdot \vec{\sigma}) = \mathbb{1}$ e $U_2 := \exp(-i\frac{2\pi}{2}\vec{\eta} \cdot \vec{\sigma}) = -\mathbb{1}$ tem-se $\phi(U_1) = \phi(U_2) = \mathbb{1}$. Verifique!²⁹ A questão agora é: como se comporta ϕ em relação ao produto dos elementos do grupo? A resposta encontra-se na afirmativa da proposição seguinte.

Proposição 21.16 *A aplicação $\phi : SU(2) \rightarrow SO(3)$ definida em (21.168) é um homomorfismo do grupo $SU(2)$ no grupo $SO(3)$, ou seja, $\phi(\mathbb{1}) = \mathbb{1}$ e para todos $U_a, U_b \in SU(2)$ vale $\phi(U_a)\phi(U_b) = \phi(U_a U_b)$. \square*

Como ϕ é sobrejetora, a Proposição 21.16 estabelece que ϕ é um *epimorfismo* de $SU(2)$ em $SO(3)$. Vide definição à página 123. Para um homomorfismo alternativo de $SU(2)$ em $SO(3)$, vide Exercício E. 21.126, página 1148.

Prova da Proposição 21.16. Que $\phi(\mathbb{1}) = \mathbb{1}$ é trivial. Provemos que $\phi(U_a)\phi(U_b) = \phi(U_a U_b)$ para todos $U_a, U_b \in SU(2)$. Sejam U_a e U_b da forma

$$U_a = \exp\left(\sum_{k=1}^3 \alpha_k j_k\right), \quad U_b = \exp\left(\sum_{k=1}^3 \beta_k j_k\right),$$

com $\alpha_k, \beta_k \in \mathbb{R}$, $k = 1, 2, 3$, e limitemos provisoriamente os valores dos α_k 's e β_k 's a uma vizinhança \mathcal{O} suficientemente pequena de zero de modo que as matrizes $a = \sum_{k=1}^3 \alpha_k j_k$ e $b = \sum_{k=1}^3 \beta_k j_k$ tenham ambas normas menores que

²⁹Por essa razão, não podemos definir uma aplicação $\phi' : SO(3) \rightarrow SU(2)$ por $\phi'(\exp(\theta\vec{\eta} \cdot \vec{J})) := \exp(\theta\vec{\eta} \cdot \vec{j})$, revertendo (21.168). Tal aplicação não é bem definida.

$\frac{1}{2} \ln \left(2 - \frac{\sqrt{2}}{2} \right)$. Essa restrição provisória às normas de a e b é útil pois coloca-nos no domínio de validade da fórmula de Baker-Campbell-Hausdorff (eq. (11.68) à página 598. Vide também (11.69)). Isso justifica, então, escrevermos

$$U_a U_b = e^a e^b = \exp(a * b),$$

onde $a * b$ está definida em (11.68). Como a série que define $a * b$ é convergente e envolve comutadores múltiplos de elementos da álgebra de Lie $\mathfrak{su}(2)$, é evidente que $a * b$ é também um elemento de $\mathfrak{su}(2)$ e, mais que isso, tem-se

$$a * b = \sum_{k=1}^3 \gamma_k j_k = \sum_{k=1}^3 \gamma_k(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) j_k, \tag{21.170}$$

onde cada γ_k é uma função analítica das variáveis $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ em um aberto suficientemente pequeno próximo a zero. A analiticidade se deve ao fato de que a série que define $a * b$ é absolutamente convergente e envolve, em cada termo, polinômios nas variáveis α e β .

E. 21.69 Exercício. Lance um olhar meditativo sobre a fórmula de Baker-Campbell-Hausdorff (11.68) e convença-se da veracidade das afirmações feitas no último parágrafo sobre a analiticidade das funções γ_k . De modo mais iluminante, mostre usando (11.69) e as relações de comutação (21.98), que os primeiros termos de $\vec{\gamma} = (\gamma_1, \gamma_2, \gamma_3)$ são

$$\vec{\gamma} = \vec{\alpha} + \vec{\beta} + \frac{1}{2}(\vec{\alpha} \times \vec{\beta}) + \frac{1}{12}(\vec{\alpha} \times (\vec{\alpha} \times \vec{\beta}) + \vec{\beta} \times (\vec{\beta} \times \vec{\alpha})) + \dots,$$

onde $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ e $\vec{\beta} = (\beta_1, \beta_2, \beta_3)$. *

Retomando, sejam agora

$$\phi(U_a) = \exp\left(\sum_{k=1}^3 \alpha_k J_k\right), \quad \phi(U_b) = \exp\left(\sum_{k=1}^3 \beta_k J_k\right),$$

e $A = \varphi(a)$, $B = \varphi(b)$, ou seja, $A = \sum_{k=1}^3 \alpha_k J_k$ e $B = \sum_{k=1}^3 \beta_k J_k$. Novamente, tem-se que

$$\phi(U_a)\phi(U_b) = e^A e^B = \exp(A * B).$$

Como φ é um isomorfismo entre $\mathfrak{su}(2)$ e $\mathfrak{so}(3)$, tem-se $\varphi(a * b) = \varphi(a) * \varphi(b)$, já que o produto “*” envolve apenas combinações lineares de comutadores (múltiplos) de elementos de $\mathfrak{su}(2)$ (vide (21.167)). Logo, segue que

$$\begin{aligned} A * B &= \varphi(a) * \varphi(b) = \varphi(a * b) = \varphi\left(\sum_{k=1}^3 \gamma_k(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) j_k\right) \\ &= \sum_{k=1}^3 \gamma_k(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) \varphi(j_k) = \sum_{k=1}^3 \gamma_k(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) J_k. \end{aligned}$$

Isso concluiu que, pelo menos quando $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ são suficientemente próximos de zero, vale

$$\phi(U_a)\phi(U_b) = \exp(\varphi(a * b)) = \phi(\exp(a * b)) = \phi(U_a U_b).$$

O que nos falta agora é um argumento que justifique que essa igualdade vale não apenas para $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$ suficientemente próximos de zero, mas para quaisquer valores desses parâmetros. Esse argumento é a analiticidade.

Cada elemento de matriz de $\exp\left(\sum_{k=1}^3 \alpha_k J_k\right)$ é uma função analítica (inteira) de α_1, α_2 e α_3 (pois a série que define a exponencial converge absolutamente em toda parte). O mesmo vale para os elementos de matriz de $\exp\left(\sum_{k=1}^3 \beta_k J_k\right)$. Assim, cada elemento de matriz do produto $\exp\left(\sum_{k=1}^3 \alpha_k J_k\right) \exp\left(\sum_{k=1}^3 \beta_k J_k\right)$ é uma função analítica (inteira) de $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$ e β_3 . Igualmente, cada elemento de matriz de $\exp\left(\sum_{k=1}^3 \gamma_k J_k\right)$ é uma função analítica das variáveis $\alpha_1,$

$\alpha_2, \alpha_3, \beta_1, \beta_2$ e β_3 quando estas estão próximas a zero (pois a composição de funções analíticas é também uma função analítica). Portanto, provamos acima que as funções analíticas $\exp\left(\sum_{k=1}^3 \alpha_k J_k\right) \exp\left(\sum_{k=1}^3 \beta_k J_k\right)$ e $\exp\left(\sum_{k=1}^3 \gamma_k J_k\right)$ coincidem em um aberto suficientemente pequeno. Por um teorema geral da teoria de funções de variáveis complexas, isso implica que essas funções são iguais em toda parte. Assim, vale para todos $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2$ e β_3 reais ou complexos que $\phi(U_a)\phi(U_b) = \phi(U_a U_b)$, completando a prova. ■

Note que a aplicação ϕ não pode ser um isomorfismo de grupos pois, como vimos, não é bijetora. Vale, no entanto, o seguinte resultado, importante, por exemplo, na Mecânica Quântica:

Proposição 21.17 *Os grupos $SU(2)/\{\mathbb{1}, -\mathbb{1}\}$ e $SO(3)$ são isomorfos.* □

Prova. Coloquemos a seguinte questão: qual é o núcleo³⁰ $\text{Ker}(\phi)$ do homomorfismo $\phi : SU(2) \rightarrow SO(3)$ definido em (21.168), ou seja, quais elementos de $SU(2)$ são mapeados na identidade de $SO(3)$ por ϕ ? De acordo com (21.168) e com a Proposição 21.7, página 1060, esses são exatamente os elementos da forma $\exp(\theta \vec{\eta} \cdot \vec{j}) = \exp(-i \frac{\theta}{2} \vec{\eta} \cdot \vec{\sigma})$ com $\theta = 2\pi n$, $n \in \mathbb{Z}$. Agora, por (21.155) temos que $\exp(-i \frac{2\pi n}{2} \vec{\eta} \cdot \vec{\sigma}) = \cos(n\pi) \mathbb{1} - i \text{sen}(n\pi) \vec{\eta} \cdot \vec{\sigma} = (-1)^n \mathbb{1}$. Logo, concluímos que $\text{Ker}(\phi) = \{\mathbb{1}, -\mathbb{1}\}$.

Como ϕ é sobrejetora tem-se que sua imagem é $\text{Ran}(\phi) = SO(3)$. Assim, evocando o Teorema 2.6, página 136, para o homomorfismo $\phi : SU(2) \rightarrow SO(3)$ dado em (21.168), concluímos imediatamente que $SU(2)/\{\mathbb{1}, -\mathbb{1}\}$ e $SO(3)$ são grupos isomorfos. ■

*

As considerações acima sobre a relação entre os grupos $SO(3)$ e $SU(2)$ são de importância na Física Quântica, particularmente no que concerne à parametrização do grupo de rotações $SO(3)$ para partículas de spin 1/2. Ainda mais profunda é a relação entre o grupo $SL(2, \mathbb{C})$ e o grupo de Lorentz, relação essa que discutiremos na Seção 45.1, página 2464.

• **Mais alguns resultados sobre o grupo $SU(2)$**

A proposição que segue estende ao grupo $SU(2)$ resultados estabelecidos para o grupo $SO(3)$ na Proposição 21.9, página 1062. Ela pode ser demonstrada usando-se o isomorfismo acima, mas apresentamos uma demonstração que faz uso instrutivo da série de Lie, expressão (11.60)–(11.61), página 594.

Proposição 21.18 *Valem as relações*

$$\exp\left(-i \frac{\theta}{2} \vec{\eta} \cdot \vec{\sigma}\right) (\vec{\xi} \cdot \vec{\sigma}) \exp\left(i \frac{\theta}{2} \vec{\eta} \cdot \vec{\sigma}\right) = \left(R(\theta, \vec{\eta}) \vec{\xi}\right) \cdot \vec{\sigma} \tag{21.171}$$

e

$$\exp\left(-i \frac{\theta}{2} \vec{\eta} \cdot \vec{\sigma}\right) \exp\left(-i \frac{\theta'}{2} \vec{\xi} \cdot \vec{\sigma}\right) \exp\left(i \frac{\theta}{2} \vec{\eta} \cdot \vec{\sigma}\right) = \exp\left(-i \frac{\theta'}{2} \left(R(\theta, \vec{\eta}) \vec{\xi}\right) \cdot \vec{\sigma}\right), \tag{21.172}$$

para $\theta, \theta' \in \mathbb{R}$ e para dois vetores unitários $\vec{\eta} = (\eta_1, \eta_2, \eta_3) \in \mathbb{R}^3$ e $\vec{\xi} = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$. □

Prova. A relação (21.172) segue diretamente de (21.171), pois $A \exp\left(-i \frac{\theta'}{2} \vec{\xi} \cdot \vec{\sigma}\right) A^{-1} = \exp\left(-i \frac{\theta'}{2} A(\vec{\xi} \cdot \vec{\sigma}) A^{-1}\right)$ para qualquer matriz invertível A . Resta-nos demonstrar (21.171).

Definamos $j_a := -i\sigma_a/2$. Valem, portanto, relações de comutação $[j_a, j_b] = \varepsilon_{abc} j_c$ (aqui e no restante desta demonstração usaremos a convenção de implicitamente somarmos (de 1 a 3) sobre índices repetidos). O lado esquerdo de (21.171) fica

$$2i \xi_j e^{\theta \vec{\eta} \cdot \vec{j}} j_j e^{-\theta \vec{\eta} \cdot \vec{j}}.$$

³⁰A noção de núcleo de um homomorfismo foi definida em (2.50), página 132.

Pela *série de Lie*, expressão (11.60)–(11.61), página 594, temos

$$e^{\theta \vec{\eta} \cdot \vec{j}} j_j e^{-\theta \vec{\eta} \cdot \vec{j}} = j_j + \sum_{a=1}^{\infty} \frac{\theta^a}{a!} \left[\underbrace{[\vec{\eta} \cdot \vec{j}, [\vec{\eta} \cdot \vec{j}, \dots, [\vec{\eta} \cdot \vec{j}, j_j] \dots]]}_{a \text{ vezes}} \right].$$

Para o primeiro comutador, temos

$$[\vec{\eta} \cdot \vec{j}, j_j] = \eta_i [j_i, j_j] = \eta_i \varepsilon_{ijk} j_k.$$

Para o segundo comutador, teremos

$$[\vec{\eta} \cdot \vec{j}, [\vec{\eta} \cdot \vec{j}, j_j]] = \eta_l \eta_i \varepsilon_{ijk} [j_l, j_k] = \eta_l \eta_i \varepsilon_{ijk} \varepsilon_{lkm} j_m \stackrel{(4.7)}{=} \eta_l \eta_i (\delta_{im} \delta_{jl} - \delta_{il} \delta_{jm}) j_m = \eta_j (\vec{\eta} \cdot \vec{j}) - j_j.$$

Assim, o terceiro comutador será igual a $[\vec{\eta} \cdot \vec{j}, \eta_j (\vec{\eta} \cdot \vec{j}) - j_j]$. Agora, o primeiro termo é

$$[\vec{\eta} \cdot \vec{j}, \eta_j (\vec{\eta} \cdot \vec{j})] = \eta_l \eta_j \eta_k [j_l, j_k] = \eta_l \eta_j \eta_k \varepsilon_{lkn} j_n = 0,$$

devido à antissimetria de ε_{lkn} . Logo, o terceiro comutador será igual a $-\left[\vec{\eta} \cdot \vec{j}, j_j\right] = -\eta_i \varepsilon_{ijk} j_k$ por ser igual ao primeiro comutador, mas com o sinal trocado.

Segue disso uma expressão geral para os comutadores múltiplos:

$$\left[\underbrace{[\vec{\eta} \cdot \vec{j}, [\vec{\eta} \cdot \vec{j}, \dots, [\vec{\eta} \cdot \vec{j}, j_j] \dots]]}_{a \text{ vezes}} \right] = \begin{cases} (-1)^{a/2} (\eta_j (\vec{\eta} \cdot \vec{j}) - j_j), & a \text{ par,} \\ (-1)^{(a-1)/2} \eta_i \varepsilon_{ijk} j_k, & a \text{ ímpar,} \end{cases}$$

a qual pode ser facilmente provada por indução (faça-o!).

Assim,

$$\begin{aligned} \exp(\theta \vec{\eta} \cdot \vec{j}) j_j \exp(-\theta \vec{\eta} \cdot \vec{j}) &= j_j + \left(\sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} \right) \eta_i \varepsilon_{ijk} j_k + \left(\sum_{k=1}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} \right) (\eta_j (\vec{\eta} \cdot \vec{j}) - j_j) \\ &= j_j + \text{sen}(\theta) \eta_i \varepsilon_{ijk} j_k + (1 - \text{cos}(\theta)) (\eta_j (\vec{\eta} \cdot \vec{j}) - j_j). \end{aligned}$$

Logo,

$$\begin{aligned} \exp(\theta \vec{\eta} \cdot \vec{j}) (\vec{\alpha} \cdot \vec{j}) \exp(-\theta \vec{\eta} \cdot \vec{j}) &= \vec{\alpha} \cdot \vec{j} + \text{sen}(\theta) (\vec{\eta} \times \vec{\alpha}) \cdot \vec{j} + (1 - \text{cos}(\theta)) ((\vec{\eta} \cdot \vec{\alpha}) (\vec{\eta} \cdot \vec{j}) - (\vec{\alpha} \cdot \vec{j})) \\ &= \left(\vec{\alpha} + \text{sen}(\theta) (\vec{\eta} \times \vec{\alpha}) + (1 - \text{cos}(\theta)) ((\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} - \vec{\alpha}) \right) \cdot \vec{j} \\ &= \left(\text{cos}(\theta) \vec{\alpha} + \text{sen}(\theta) (\vec{\eta} \times \vec{\alpha}) + (1 - \text{cos}(\theta)) (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} \right) \cdot \vec{j} \\ &\stackrel{(21.106)}{=} \left(R(\theta, \vec{\eta}) \vec{\alpha} \right) \cdot \vec{j}. \end{aligned}$$

Isso completa a demonstração de (21.171). ■

21.4.6 O Grupo $SL(2, \mathbb{C})$

Vamos aqui tratar de um grupo fortemente aparentado ao grupo $SU(2)$ e ao grupo de Lorentz, cujo estudo é importante na teoria dos spinors, particularmente no estudo de representações do grupo de Lorentz para partículas de spin $1/2$.

Trata-se do grupo $SL(2, \mathbb{C})$, o grupo das matrizes complexas 2×2 de determinante 1. Mais sobre o grupo $SL(2, \mathbb{C})$, em especial, sua relação com o grupo de Lorentz, é apresentado na Seção 2.B, página 214 e na Seção 45.1, página 2464.

O grupo $SL(2, \mathbb{C})$ é definido como o grupo formado pelas matrizes complexas 2×2 de determinante igual a 1. Como as matrizes $\mathbb{1}, \sigma_1, \sigma_2, \sigma_3$ formam uma base em $\text{Mat}(\mathbb{C}, 2)$, podemos escrever toda matriz $A \in SL(2, \mathbb{C})$ na forma

$$A = a_0 \mathbb{1} + a_1 \sigma_1 + a_2 \sigma_2 + a_3 \sigma_3 = \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix},$$

com $a_0, a_1, a_2, a_3 \in \mathbb{C}$. A condição $\det(A) = 1$ implica $a_0^2 - a_1^2 - a_2^2 - a_3^2 = 1$. Assim,

$$SL(2, \mathbb{C}) = \left\{ \begin{pmatrix} a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3 \end{pmatrix} \text{ com } a_0, a_1, a_2, a_3 \in \mathbb{C} \text{ e } a_0^2 - a_1^2 - a_2^2 - a_3^2 = 1 \right\}. \quad (21.173)$$

Como a_0 é um número complexo arbitrário, podemos escrever

$$a_0 = \cosh z,$$

para algum $z \in \mathbb{C}$, pois a função $\cosh : \mathbb{C} \rightarrow \mathbb{C}$ é sobrejetora. Fora isso, para $z \neq in\pi, n \in \mathbb{Z}$, podemos definir três números complexos η_1, η_2, η_3 por

$$\eta_1 := \frac{a_1}{\sinh z}, \quad \eta_2 := \frac{a_2}{\sinh z}, \quad \eta_3 := \frac{a_3}{\sinh z}.$$

A condição $a_0^2 - a_1^2 - a_2^2 - a_3^2 = 1$ implica (verifique!) que os números complexos η_1, η_2, η_3 satisfazem $\eta_1^2 + \eta_2^2 + \eta_3^2 = 1$. Com isso vemos que

$$SL(2, \mathbb{C}) = \left\{ \cosh(z)\mathbb{1} + \sinh(z)(\vec{\eta} \cdot \vec{\sigma}), \text{ onde } z \in \mathbb{C} \text{ e } \vec{\eta} \in \mathbb{C}^3 \text{ com } \eta_1^2 + \eta_2^2 + \eta_3^2 = 1 \right\}. \quad (21.174)$$

Note-se que a parametrização usada em (21.174) para os elementos de $SL(2, \mathbb{C})$ não é unívoca, pois $\cosh(z) = \cosh(z')$ e $\sinh(z) = \sinh(z')$ caso $z - z' = 2\pi mi$, com $m \in \mathbb{Z}$.

Como vimos anteriormente quando tratamos de $SU(2)$, a condição $\eta_1^2 + \eta_2^2 + \eta_3^2 = 1$ implica que $(\vec{\eta} \cdot \vec{\sigma})^2 = \mathbb{1}$ (mesmo para vetores $\vec{\eta}$ complexos!). Portanto,

$$\begin{aligned} \exp(z \vec{\eta} \cdot \vec{\sigma}) &= \sum_{m=0}^{\infty} \frac{z^m}{m!} (\vec{\eta} \cdot \vec{\sigma})^m \\ &= \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} (\vec{\eta} \cdot \vec{\sigma})^{2k} + \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} (\vec{\eta} \cdot \vec{\sigma})^{2k+1} \\ &= \left(\sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!} \right) \mathbb{1} + \left(\sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} \right) (\vec{\eta} \cdot \vec{\sigma}) \\ &= \cosh(z)\mathbb{1} + \sinh(z)(\vec{\eta} \cdot \vec{\sigma}). \end{aligned}$$

Assim, todo elemento $A \in SL(2, \mathbb{C})$ é da forma $\exp(z \vec{\eta} \cdot \vec{\sigma})$. Em resumo,

$$SL(2, \mathbb{C}) = \left\{ \exp(z \vec{\eta} \cdot \vec{\sigma}), \text{ onde } z \in \mathbb{C} \text{ e } \vec{\eta} \in \mathbb{C}^3 \text{ com } \eta_1^2 + \eta_2^2 + \eta_3^2 = 1 \right\}. \quad (21.175)$$

Como já observamos logo após (21.174), a parametrização usada em (21.175) para os elementos de $SL(2, \mathbb{C})$ não é unívoca.

Como já vimos, o subgrupo $SU(2)$ de $SL(2, \mathbb{C})$ corresponde a $z = i\theta, \theta \in \mathbb{R}$, e $\vec{\eta} \in \mathbb{R}^3$. Como vemos, $SU(2)$ e $SL(2, \mathbb{C})$ têm ambas álgebras de Lie geradas pelas matrizes de Pauli, mas em $SU(2)$ essa álgebra de Lie é uma álgebra sobre o corpo dos *reais*, enquanto que em $SL(2, \mathbb{C})$ é sobre o corpo dos *complexos*.

Mais material sobre o grupo $SL(2, \mathbb{C})$, em especial, sobre sua relação com o grupo de Lorentz, será apresentado na Seção 45.1, página 2464.

21.5 Generalidades Sobre os Grupos $SU(n)$ e $SO(n)$

Nesta seção discutiremos algumas qualidades gerais dos grupos $SU(n)$ e $SO(n)$. Para esta seção recomenda-se a leitura prévia de partes do Capítulo 22, página 1162. Começaremos com os grupos $SU(n)$ pois seu tratamento é ligeiramente mais simples que o dos grupos $SO(n)$. O caso fisicamente importante do grupo $SU(3)$ será brevemente discutido na Seção 21.5.1.1, página 1089.

21.5.1 Os Grupos $SU(n)$

Após termos adquirido algum conhecimento sobre o grupo $SU(2)$, vamos estudar alguns aspectos gerais dos grupos $SU(n)$, $n \geq 2$. Vimos acima de modo explícito que os elementos de $SU(2)$ podem ser escritos como exponenciais de elementos de sua álgebra de Lie. Veremos que esse fato é também válido para $SU(n)$.

Lembremos a definição: para $n \geq 1$,

$$SU(n) := \left\{ U \in GL(\mathbb{C}, n) \mid U^* = U^{-1} \text{ e } \det(U) = 1 \right\}.$$

Começemos com a seguinte observação.

Proposição 21.19 $SU(n)$ é um subgrupo compacto de $GL(\mathbb{C}, n)$. □

Prova. Provemos primeiramente que $SU(n)$ é um subconjunto (topologicamente) fechado de $GL(\mathbb{C}, n)$. Em verdade, essa afirmação já se encontra demonstrada na Proposição 22.6, página 1170, mas apresentamos agora uma outra prova mais específica.

Seja $U_n, n \in \mathbb{N}$, uma sequência de matrizes de $SU(n)$ que converge em norma a uma matriz $U \in \text{Mat}(n, \mathbb{C})$, ou seja, $\lim_{n \rightarrow \infty} \|U_n - U\|_{\mathbb{C}} = 0$, onde $\|\cdot\|_{\mathbb{C}}$ é a norma operatorial de matrizes. Desejamos provar que $U \in SU(n)$.

Em primeiro lugar, notemos que podemos escrever

$$U^*U = (U - U_n + U_n)^*(U - U_n + U_n) = (U - U_n)^*(U - U_n) + U_n^*(U - U_n) + (U - U_n)^*U_n + U_n^*U_n.$$

Como os U_n são unitários, $U_n^*U_n = \mathbb{1}$ e conclui-se que $U^*U - \mathbb{1} = (U - U_n)^*(U - U_n) + U_n^*(U - U_n) + (U - U_n)^*U_n$. Assim

$$\begin{aligned} \|U^*U - \mathbb{1}\|_{\mathbb{C}} &= \left\| (U - U_n)^*(U - U_n) + U_n^*(U - U_n) + (U - U_n)^*U_n \right\|_{\mathbb{C}} \\ &\leq \left\| (U - U_n)^*(U - U_n) \right\|_{\mathbb{C}} + \left\| U_n^*(U - U_n) \right\|_{\mathbb{C}} + \left\| (U - U_n)^*U_n \right\|_{\mathbb{C}} \\ &\leq \left\| (U - U_n)^* \right\|_{\mathbb{C}} \|U - U_n\|_{\mathbb{C}} + \|U_n^*\|_{\mathbb{C}} \|U - U_n\|_{\mathbb{C}} + \left\| (U - U_n)^* \right\|_{\mathbb{C}} \|U_n\|_{\mathbb{C}} \\ &\leq \|U - U_n\|_{\mathbb{C}}^2 + 2\|U - U_n\|_{\mathbb{C}}. \end{aligned} \tag{21.176}$$

(Ao estudante deve ser claro que acima usamos os fatos que, para quaisquer matrizes A, B , complexas $n \times n$, valem $\|A + B\|_{\mathbb{C}} \leq \|A\|_{\mathbb{C}} + \|B\|_{\mathbb{C}}$, $\|AB\|_{\mathbb{C}} \leq \|A\|_{\mathbb{C}}\|B\|_{\mathbb{C}}$, $\|A\|_{\mathbb{C}} = \|A^*\|_{\mathbb{C}}$ e que $\|A\|_{\mathbb{C}} = 1$ se A é unitária. Se não for claro, justifique esses fatos como exercício ou leia o Capítulo 39).

Agora, como o extremo direito da sequência de desigualdades (21.176) pode ser feito arbitrariamente pequeno para $n \rightarrow \infty$, concluímos que o extremo esquerdo é nulo, ou seja, $U^*U = \mathbb{1}$. Analogamente, prova-se que $UU^* = \mathbb{1}$. Isso estabelece que U é unitário.

Que o determinante de U vale 1, segue do fato que $U_n \in SU(n)$ converge a U na norma operatorial e da continuidade do determinante estabelecida na Proposição 11.18, página 607. Isso estabelece que $U \in SU(n)$ e isso prova que $SU(n)$ é um subconjunto topologicamente fechado de $GL(\mathbb{C}, n)$, como queríamos.

Para provarmos que $SU(n)$ é compacto, resta apenas provar que $SU(n)$ é um conjunto limitado³¹. A condição $U^*U = \mathbb{1}$

³¹Para a definição da noção de compacidade e suas propriedades, vide Seção 32.3, página 1514.

implica $\text{Tr}(U^*U) = n$. Assim, vale

$$\sum_{a,b=1}^n |U_{ab}|^2 = n,$$

para todo $U \in \text{SU}(n)$. Isso mostra que $\text{SU}(n)$ é limitado e, portanto, compacto. ■

Seja agora $\{U(t) \in \text{SU}(n), t \in \mathbb{R}\}$, um subgrupo uniparamétrico de $\text{SU}(n)$ (ou seja, $U(0) = \mathbb{1}$ e $U(t)U(t') = U(t+t')$, sendo $t \mapsto U(t)$ contínua). Pela Proposição 22.7, página 1171, $U(t) = \exp(tA)$ para alguma matriz A . Agora, sejam u, v dois vetores arbitrários de \mathbb{C}^n . Temos que, para todo t vale $\langle u, v \rangle_{\mathbb{C}} = \langle U(t)u, U(t)v \rangle_{\mathbb{C}}$. Diferenciando essa igualdade em relação a t , escrevendo-se $U(t) = \exp(tA)$ e calculando a derivada em $t = 0$, tem-se $0 = \langle Au, v \rangle_{\mathbb{C}} + \langle u, Av \rangle_{\mathbb{C}}$, ou seja, $\langle u, (A + A^*)v \rangle_{\mathbb{C}} = 0$. Como isso vale para todo u, v em \mathbb{C}^n , segue que $A^* = -A$. Fora isso³², como $1 = \det(\exp(tA)) = \exp(t\text{Tr}(A))$, segue que A tem traço nulo.

Assim, vimos que os geradores infinitesimais dos subgrupos uniparamétricos de $\text{SU}(n)$ são antiautoadjuntos e têm traço nulo. Podemos nos perguntar se a recíproca é válida, ou seja, se todas as matrizes antiautoadjuntas e de traço nulo são geradoras infinitesimais de subgrupos uniparamétricos de $\text{SU}(n)$. Para responder isso, precisamos da seguinte proposição:

Proposição 21.20 *Se $A \in \text{Mat}(n, \mathbb{C})$ é antiautoadjunta (ou seja, $A^* = -A$) satisfazendo também $\text{Tr}(A) = 0$, então a matriz $\exp(A)$ é um elemento de $\text{SU}(n)$.* □

Prova. Precisamos provar que $\exp(A)$ é unitária e que seu determinante é igual a 1. Pela definição da exponencial de matrizes em termos de uma série de potências (a série de Taylor da função exponencial), sabe-se que $\exp(M)^* = \exp(M^*)$ para qualquer matriz $n \times n$ complexa M . Assim, $\exp(A)^* = \exp(A^*) = \exp(-A) = \exp(A)^{-1}$, provando que $\exp(A)$ é unitária. Além disso, para nossa matriz A , tem-se (pela Proposição 10.14, página 466) $\det(\exp(A)) = \exp(\text{Tr}(A)) = \exp(0) = 1$, o que prova que $\exp(A) \in \text{SU}(n)$, como queríamos. ■

Essa proposição diz-nos que, se $A \in \text{Mat}(n, \mathbb{C})$ é antiautoadjunta e tem traço nulo, então $U(t) = \exp(tA)$, $t \in \mathbb{R}$ é um subgrupo uniparamétrico de $\text{SU}(n)$. Em resumo, concluímos que o conjunto de todas as matrizes $n \times n$ complexas antiautoadjuntas e de traço nulo é idêntico ao conjunto de todos os geradores infinitesimais de subgrupos uniparamétricos de $\text{SU}(n)$.

Como $\text{SU}(n)$ é um subgrupo fechado de $\text{GL}(\mathbb{C}, n)$, segue do Teorema 22.1 que o conjunto de seus geradores infinitesimais é uma álgebra de Lie. Essa álgebra de Lie é dita ser a álgebra de Lie de $\text{SU}(n)$, e é denotada por $\text{su}(n)$ (assim, com letras minúsculas). Como vimos, $\text{su}(n)$ coincide com o conjunto de todas as matrizes $n \times n$ complexas antiautoadjuntas de traço nulo.

De passagem, notemos que o conjunto de todas as matrizes $n \times n$ complexas antiautoadjuntas de traço nulo forma uma álgebra de Lie real, fato já visto independentemente nos exercícios da página 109.

Provemos agora uma outra proposição, a qual essencialmente diz-nos que todo elemento de $\text{SU}(n)$ pode ser obtido como exponencial de um elemento de $\text{su}(n)$. No caso de $\text{SU}(2)$ isso foi provado explicitamente, quando mostramos que todo elemento de $\text{SU}(2)$ é da forma $\exp(i\theta\vec{\eta} \cdot \vec{\sigma})$.

Proposição 21.21 *Todo elemento U de $\text{SU}(n)$ pode ser escrito na forma $U = e^A$, onde $A \in \text{Mat}(n, \mathbb{C})$ é antiautoadjunta (ou seja, $A^* = -A$) e de traço nulo (ou seja, $\text{Tr}(A) = 0$). Segue disso que $\text{SU}(n)$ é um grupo a $n^2 - 1$ parâmetros reais.* □

Prova. Seja $U \in \text{SU}(n)$. Como toda matriz unitária, U é normal, pois vale $UU^* = U^*U (= \mathbb{1})$. Uma das consequências do Teorema Espectral para matrizes diz-nos que toda matriz normal pode ser diagonalizada por uma matriz unitária (vide Teorema 10.17 e as páginas que o antecedem).

Assim, existe V , matriz unitária, tal que $U = VDV^*$, onde $D = \text{diag}(u_1, \dots, u_n)$, e onde os u_k são números complexos (os autovalores de U). Da condição $UU^* = \mathbb{1}$ segue imediatamente que $DD^* = \mathbb{1}$, o que implica que cada u_k é um número complexo de módulo 1: $|u_k|^2 = 1$. Assim, podemos escrever $u_k = e^{i\lambda_k}$, onde $\lambda_k \in \mathbb{R}$, sendo que cada λ_k é determinado a menos de um termo $2\pi m$, com m inteiro.

³²Aqui usamos a Proposição 10.14, página 466, ou a Proposição 11.7, página 578.

Note-se que, como U tem determinante 1, segue que $1 = \det(U) = \det(VDV^*) = \det(D) = \exp(i \sum_{k=1}^n \lambda_k)$. Assim, $\sum_{k=1}^n \lambda_k = 2\pi m_0$, com m_0 inteiro. Podemos redefinir, digamos, λ_n , subtraindo-lhe $2\pi m_0$. Com essa nova escolha teremos

$$\sum_{k=1}^n \lambda_k = 0. \tag{21.177}$$

Definamos agora a matriz $L = \text{diag}(i\lambda_1, \dots, i\lambda_n)$. Note-se que, como os λ_k são reais, vale $L^* = -L$. É claro que $D = e^L$ e também que $U = \exp(A)$, onde $A = VLV^*$. É agora elementar constatar que $A^* = -A$. Fora isso, por (21.177) segue que $\text{Tr}(A) = \text{Tr}(VLV^*) = \text{Tr}(L) = i \sum_{k=1}^n \lambda_k = 0$.

O número de parâmetros reais que determina os elementos de $SU(n)$ é, naturalmente, a dimensão (real) do espaço das matrizes antiautoadjuntas de traço nulo. Os elementos independentes de tais matrizes podem ser contados da seguinte forma. Para uma matriz antiautoadjunta A temos $A_{ij} = -\overline{A_{ji}}$ para $i \neq j$ e $A_{jj} = -\overline{A_{jj}}$. Assim, são independentes os elementos complexos acima da diagonal principal e os elementos imaginários puros da diagonal principal. Acima da diagonal principal há $1 + \dots + n - 1 = n(n - 1)/2$ elementos complexos e, portanto, os mesmos são determinados por $n(n - 1)$ parâmetros reais. Na diagonal há mais n parâmetros reais independentes, mas a condição de traço nulo reduz esse número para $n - 1$. Assim, o número total de parâmetros reais independentes é $n(n - 1) + n - 1 = n^2 - 1$. Isso completa a prova. ■

A Proposição 21.21 tem o seguinte corolário simples:

Corolário 21.4 *O grupo $SU(n)$ é conexo por caminhos e, portanto, é um espaço conexo.* □

Prova. Pelo que vimos, se $U \in SU(n)$, então U é da forma $U = e^A$, para alguma $A \in \mathfrak{su}(n)$. Logo, U pertence ao subgrupo uniparamétrico de $SU(n)$ gerado por A : $\{\exp(tA), t \in \mathbb{R}\}$. Esse subgrupo conecta continuamente U à identidade $\mathbb{1}$ (que corresponde a $t = 0$). ■

A Proposição 21.21 diz-nos que a exponenciação é uma aplicação sobrejetora de $\mathfrak{su}(n)$ em $SU(n)$. Isso é um caso particular de um teorema mais geral que diz que isso é válido para qualquer grupo de Lie compacto, conexo e cuja álgebra de Lie seja de dimensão finita.

E. 21.70 Exercício. Pelo que vimos $\mathfrak{su}(2)$ coincide com a álgebra de Lie real de todas as matrizes complexas 2×2 , antiautoadjuntas e de traço zero. Mostre que as matrizes $i\sigma_1, i\sigma_2$ e $i\sigma_3$ formam uma base nesse espaço de matrizes. Conclua que todo elemento de $SU(2)$ é da forma $\exp(i\alpha_1\sigma_1 + i\alpha_2\sigma_2 + i\alpha_3\sigma_3)$ com $\alpha_k \in \mathbb{R}$, fato esse que já estabelecemos por outros meios (vide (21.156), página 1078). ★

21.5.1.1 Um Pouco Sobre o Grupo $SU(3)$

O grupo $SU(3)$ é de grande importância na Física das Partículas Elementares, estando associado à uma simetria aproximada, dita de “sabor”, e a uma simetria exata, dita de “cor”. Não nos deteremos nesses aspectos aqui, e remetemos o estudante aos bons livros sobre Física das Partículas Elementares e Teoria Quântica de Campos (por exemplo, [408]-[409]). Vide também [128].

O grupo $SU(3)$ é um grupo a $3^2 - 1 = 8$ parâmetros reais. Pelo que vimos, $\mathfrak{su}(3)$ coincide com o espaço das matrizes complexas 3×3 , antiautoadjuntas e de traço zero. Para o estudo do grupo $SU(3)$ no contexto da Física das Partículas Elementares é conveniente introduzir-se uma base explícita nesse espaço. Como toda matriz antiautoadjunta pode ser escrita como $i\lambda$, onde λ é autoadjunta, basta-nos procurar uma base no espaço das matrizes autoadjuntas de traço zero.

Comunmente adota-se as chamadas *Matrizes de Gell-Mann*³³ $\lambda_i, i = 1, \dots, 8$, que são as seguintes matrizes:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Note que todas as matrizes λ_i são autoadjuntas e de traço zero, formando uma base no espaço das matrizes complexas autoadjuntas e de traço nulo (mostre isso!). As mesmas são normalizadas de modo que $\text{Tr}(\lambda_a \lambda_b) = 2\delta_{ab}$.

E. 21.71 *Exercício.* Prove as afirmativas do último parágrafo. *

A álgebra de Lie de $\mathfrak{su}(3)$ pode ser expressa para as matrizes de Gell-Mann da seguinte forma:

$$[\lambda_a, \lambda_b] = 2i \sum_{c=1}^8 f_{abc} \lambda_c,$$

onde f_{abc} , as chamadas *constantes de estrutura*³⁴ de $\mathfrak{su}(3)$, são totalmente antissimétricas, ou seja,

$$f_{abc} = f_{bca} = f_{cab} = -f_{bac} = -f_{acb} = -f_{cba},$$

sendo

$$f_{123} = 1,$$

$$f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = \frac{1}{2},$$

$$f_{458} = f_{678} = \frac{\sqrt{3}}{2},$$

e as demais constantes independentes são nulas.

E. 21.72 *Exercício.* Verifique isso. *Sugestão:* tire uma tarde livre. *

Pelo que aprendemos da nossa discussão geral sobre grupos $SU(n)$, todo elemento U de $SU(3)$ pode ser escrito na forma

$$U = \exp \left(i \sum_{k=1}^8 \alpha_k \lambda_k \right),$$

onde os α_k 's são números reais.

³³Murray Gell-Mann (1929–2019).

³⁴A noção geral de “constante de estrutura” de uma álgebra de dimensão finita foi introduzida à página 105.

21.5.2 Os Grupos $SO(n)$

Primeiramente lembremos a definição: para $n \geq 2$,

$$SO(n) := \left\{ R \in GL(\mathbb{R}, n) \mid R^T = R^{-1} \text{ e } \det(R) = 1 \right\}.$$

Sob vários aspectos os grupos $SO(n)$ podem ser tratados de modo semelhante aos grupos $SU(n)$, exceto por um ponto importante: por agirem em um espaço vetorial real (\mathbb{R}^n), não podemos aplicar o teorema espectral às matrizes ortogonais, tal como fizemos na prova da Proposição 21.21. Por isso, um desvio mais longo deverá ser seguido, ainda que as conclusões sejam as mesmas, em essência.

Analogamente ao que fizemos no caso $SU(n)$, comecemos com a seguinte observação.

Proposição 21.22 $SO(n)$ é um subgrupo compacto de $GL(\mathbb{R}, n)$. □

Prova. A prova é uma mera imitação da demonstração correspondente no caso $SU(n)$ e poupamo-nos de reproduzi-la. ■

Seja agora $\{R(t) \in SO(n), t \in \mathbb{R}\}$, um subgrupo uniparamétrico contínuo de $SO(n)$ (ou seja, $R(0) = \mathbb{1}$ e $R(t)R(t') = R(t+t')$). Pela Proposição 22.7, página 1171, $R(t) = \exp(tA)$ para alguma matriz A . Agora, sejam u, v dois vetores arbitrários de \mathbb{R}^n . Temos que, para todo t vale $\langle u, v \rangle_{\mathbb{R}} = \langle R(t)u, R(t)v \rangle_{\mathbb{R}}$. Diferenciando essa igualdade em relação a t , escrevendo-se $R(t) = \exp(tA)$ e calculando a derivada em $t = 0$, tem-se $0 = \langle Au, v \rangle_{\mathbb{R}} + \langle u, Av \rangle_{\mathbb{R}}$, ou seja, $\langle u, (A + A^T)v \rangle_{\mathbb{R}} = 0$. Como isso vale para todo u, v em \mathbb{R}^n , segue que $A^T = -A$. Assim, A é uma matriz antissimétrica, o que implica que seus elementos diagonais são nulos. Assim, é automático que $\text{Tr}(A) = 0$.

Assim, vimos que os geradores infinitesimais dos subgrupos uniparamétricos de $SO(n)$ são antissimétricos. Podemos nos perguntar se a recíproca é válida, ou seja, se todas as matrizes antissimétricas são geradores infinitesimais de subgrupos uniparamétricos de $SU(n)$. Para responder isso, precisamos da seguinte proposição:

Proposição 21.23 Se $A \in \text{Mat}(n, \mathbb{R})$ é antissimétrica (ou seja, $A^T = -A$), então a matriz $\exp(A)$ é um elemento de $SO(n)$. □

Prova. Precisamos provar que $\exp(A)$ é ortogonal e que seu determinante é igual a 1. Pela definição da exponencial de matrizes em termos de uma série de potências (a série de Taylor da função exponencial), sabe-se que $\exp(M)^T = \exp(M^T)$ para qualquer matriz $n \times n$ real ou complexa M . Assim, $\exp(A)^T = \exp(A^T) = \exp(-A) = \exp(A)^{-1}$, provando que $\exp(A)$ é ortogonal.

Como observamos, $\text{Tr}(A) = 0$. Logo, para nossa matriz A , tem-se $\det(\exp(A)) = \exp(\text{Tr}(A)) = \exp(0) = 1$, o que prova que $\exp(A) \in SO(n)$, como queríamos. ■

Essa proposição diz-nos que, se $A \in \text{Mat}(n, \mathbb{R})$ é antissimétrica, então $R(t) = \exp(tA)$, $t \in \mathbb{R}$ é um subgrupo uniparamétrico de $SO(n)$. Em resumo, concluímos que o conjunto de todas as matrizes $n \times n$ reais antissimétricas é idêntico ao conjunto de todos os geradores infinitesimais de subgrupos uniparamétricos de $SO(n)$.

Como $SO(n)$ é um subgrupo fechado de $GL(\mathbb{R}, n)$, segue do Teorema 22.1 que o conjunto de seus geradores infinitesimais é uma álgebra de Lie. Essa álgebra de Lie é dita ser a álgebra de Lie de $SO(n)$, e é denotada por $\mathfrak{so}(n)$. Como vimos, $\mathfrak{so}(n)$ coincide com o conjunto de todas as matrizes $n \times n$ reais antissimétricas.

De passagem, notemos que o fato de o conjunto de todas as matrizes $n \times n$ reais antissimétricas formar uma álgebra de Lie real já fora visto independentemente nos exercícios da página 109.

Provemos agora uma outra proposição, a qual essencialmente diz-nos que todo elemento de $SO(n)$ pode ser obtido como exponencial de um elemento de $\mathfrak{so}(n)$. Nos casos de $SO(2)$ e $SO(3)$ isso foi provado explicitamente nas páginas acima.

Proposição 21.24 Todo elemento R de $SO(n)$ pode ser escrito na forma $R = e^A$, onde $A \in \text{Mat}(n, \mathbb{R})$ é antissimétrica (ou seja, $A^T = -A$). Segue disso que $SO(n)$ é um grupo a $n(n-1)/2$ parâmetros reais. □

Prova. Como dissemos não podemos aqui seguir exatamente os passos da prova da Proposição 21.21, pois o teorema espectral não se aplica de modo direto a matrizes reais.

Seja $R \in SO(n)$, com elementos de matriz reais R_{ij} . Normalmente R age no espaço real \mathbb{R}^n , mas podemos fazê-la agir em \mathbb{C}^n da maneira usual: para um vetor $u \in \mathbb{C}^n$ com componentes $u_i \in \mathbb{C}$, tem-se $(Ru)_i = \sum_{j=1}^n R_{ij}u_j$. Como tal, R é uma matriz unitária de determinante 1, ou seja, um elemento de $SU(n)$, pois $(R^*)_{ij} = \overline{(R)_{ji}} = (R)_{ji} = (R^T)_{ij} = (R^{-1})_{ij}$. Aqui usamos que os R_{ij} são reais e o fato óbvio (por quê?) que a inversa de R em \mathbb{C}^n é a mesma que em \mathbb{R}^n .

Dado que R é unitária, seus autovalores são números eventualmente complexos mas de módulo 1. Notemos, porém, que os autovalores são raízes do polinômio característico $p(x) = \det(x\mathbb{1} - R)$, $x \in \mathbb{C}$. Como os R_{ij} são reais, esse polinômio tem coeficientes reais. É um fato elementar e bem conhecido que se x é raiz de um polinômio com coeficientes reais, então seu complexo conjugado \bar{x} também o é.

Se n é par, os autovalores são, portanto, pares de números complexos de módulo 1 complexo-conjugados: $e^{i\theta}$ e $e^{-i\theta}$. Como o determinante de R é o produto de seus autovalores, isso automaticamente garante que $\det(R) = 1$ desde que -1 , se for autovalor, o seja com multiplicidade algébrica par.

Se n é ímpar, os autovalores são pares de números complexos de módulo 1 complexo-conjugados: $e^{\pm i\theta}$, mas um deles pode ser real, podendo, portanto, ser ± 1 . Como o determinante de R é o produto de seus autovalores, a condição $\det(R) = 1$ implica que um dos autovalores deve ser $+1$ e que -1 , se for autovalor, o é com multiplicidade algébrica par.

Em resumo:

1. Se n é par, o conjunto de autovalores de R é do tipo $\{e^{\pm i\theta_k}, k = 1, \dots, n/2, \text{ sendo } \theta_k \in \mathbb{R}\}$.
2. Se n é ímpar, o conjunto de autovalores de R é do tipo $\{1\} \cup \{e^{\pm i\theta_k}, k = 1, \dots, (n-1)/2, \text{ sendo } \theta_k \in \mathbb{R}\}$.

Em ambos os casos -1 pode ser autovalor e, se o for, o é com multiplicidade algébrica par.

Seja o autovalor $e^{i\theta_k}$. Há dois casos a considerar.

Caso I. $e^{i\theta_k} \neq \pm 1$, de modo que $e^{i\theta_k}$ é não real e, portanto, distinto de $e^{-i\theta_k}$.

Seja $\mathbf{v}^k \in \mathbb{C}^n$ um autovetor de R com autovalor $e^{i\theta_k}$: $R\mathbf{v}^k = e^{i\theta_k}\mathbf{v}^k$, normalizado de modo que $\|\mathbf{v}^k\|_{\mathbb{C}}^2 = \langle \mathbf{v}^k, \mathbf{v}^k \rangle_{\mathbb{C}} = 1$. Segue que $R\overline{\mathbf{v}^k} = e^{-i\theta_k}\overline{\mathbf{v}^k}$, ou seja, $\overline{\mathbf{v}^k}$ é um autovetor de R com autovalor $e^{-i\theta_k}$. Como R é unitária, segue que autovetores que correspondem a autovalores distintos são ortogonais (em \mathbb{C}^n). Logo,

$$\langle \overline{\mathbf{v}^k}, \mathbf{v}^k \rangle_{\mathbb{C}} = 0 \quad \text{e, portanto,} \quad \langle \mathbf{v}^k, \mathbf{v}^k \rangle_{\mathbb{R}} = \langle \overline{\mathbf{v}^k}, \mathbf{v}^k \rangle_{\mathbb{C}} = 0. \quad (21.178)$$

Escrevamos \mathbf{v}^k separando componente a componente suas partes real e imaginária: $\mathbf{v}^k = \mathbf{a}^k + i\mathbf{b}^k$, com $\mathbf{a}^k, \mathbf{b}^k \in \mathbb{R}^n$. As relações $R\mathbf{v}^k = e^{i\theta_k}\mathbf{v}^k$ e $R\overline{\mathbf{v}^k} = e^{-i\theta_k}\overline{\mathbf{v}^k}$ tornam-se

$$R\mathbf{a}^k = (\cos \theta_k)\mathbf{a}^k - (\sin \theta_k)\mathbf{b}^k,$$

$$R\mathbf{b}^k = (\sin \theta_k)\mathbf{a}^k + (\cos \theta_k)\mathbf{b}^k.$$

Note-se que, como $\sin \theta_k \neq 0$, essas duas relações implicam que não se pode ter $\mathbf{a}^k = 0$, pois isso implicaria $\mathbf{b}^k = 0$ e vice-versa. Porém, \mathbf{a}^k e \mathbf{b}^k são vetores ortogonais em \mathbb{R}^n . De fato,

$$\begin{aligned} \langle \mathbf{a}^k, \mathbf{b}^k \rangle_{\mathbb{R}} &= \frac{1}{4} \langle (\mathbf{v}^k + \overline{\mathbf{v}^k}), (\mathbf{v}^k - \overline{\mathbf{v}^k}) \rangle_{\mathbb{R}} \\ &= \frac{1}{4} \left(\langle \mathbf{v}^k, \mathbf{v}^k \rangle_{\mathbb{R}} - \langle \mathbf{v}^k, \overline{\mathbf{v}^k} \rangle_{\mathbb{R}} + \langle \overline{\mathbf{v}^k}, \mathbf{v}^k \rangle_{\mathbb{R}} - \langle \overline{\mathbf{v}^k}, \overline{\mathbf{v}^k} \rangle_{\mathbb{R}} \right) \\ &= \frac{1}{4} \left(\langle \overline{\mathbf{v}^k}, \mathbf{v}^k \rangle_{\mathbb{C}} - \langle \overline{\mathbf{v}^k}, \overline{\mathbf{v}^k} \rangle_{\mathbb{C}} + \langle \mathbf{v}^k, \mathbf{v}^k \rangle_{\mathbb{C}} - \langle \mathbf{v}^k, \overline{\mathbf{v}^k} \rangle_{\mathbb{C}} \right) \\ &\stackrel{(21.178)}{=} \frac{1}{4} (0 - 1 + 1 - 0) = 0. \end{aligned}$$

Assim, concluímos que no subespaço real gerado pelos vetores ortogonais não nulos \mathbf{a}^k e \mathbf{b}^k , a matriz R age (na base gerada por \mathbf{a}^k e \mathbf{b}^k , como a matriz $\begin{pmatrix} \cos \theta_k & \text{sen} \theta_k \\ -\text{sen} \theta_k & \cos \theta_k \end{pmatrix}$, elemento de $\text{SO}(2)$.

É importante notar também que os vetores \mathbf{a}^k e \mathbf{b}^k são também ortogonais entre si para k 's diferentes. Isso é mostrado na proposição seguinte.

Proposição 21.25 *Se $\mathbf{v}^j = \mathbf{a}^j + i\mathbf{b}^j$ e $\mathbf{v}^k = \mathbf{a}^k + i\mathbf{b}^k$ são vetores de \mathbb{C}^n com $\mathbf{a}^j, \mathbf{a}^k, \mathbf{b}^j, \mathbf{b}^k \in \mathbb{R}^n$ e se valerem $\langle \mathbf{v}^j, \mathbf{v}^k \rangle_{\mathbb{C}} = 0$ e $\langle \overline{\mathbf{v}^j}, \mathbf{v}^k \rangle_{\mathbb{C}} = 0$, então tem-se*

$$\langle \mathbf{a}^j, \mathbf{a}^k \rangle_{\mathbb{R}} = \langle \mathbf{a}^j, \mathbf{b}^k \rangle_{\mathbb{R}} = \langle \mathbf{b}^j, \mathbf{a}^k \rangle_{\mathbb{R}} = \langle \mathbf{b}^j, \mathbf{b}^k \rangle_{\mathbb{R}} = 0.$$

□

Prova. De $\langle \mathbf{v}^j, \mathbf{v}^k \rangle_{\mathbb{C}} = 0$ segue facilmente que

$$\langle \mathbf{a}^j, \mathbf{a}^k \rangle_{\mathbb{R}} + \langle \mathbf{b}^j, \mathbf{b}^k \rangle_{\mathbb{R}} = 0 \quad \text{e} \quad \langle \mathbf{b}^j, \mathbf{a}^k \rangle_{\mathbb{R}} - \langle \mathbf{a}^j, \mathbf{b}^k \rangle_{\mathbb{R}} = 0.$$

Como $\overline{\mathbf{v}^j} = \mathbf{a}^j - i\mathbf{b}^j$, tem-se de $\langle \overline{\mathbf{v}^j}, \mathbf{v}^k \rangle_{\mathbb{C}} = 0$ que

$$\langle \mathbf{a}^j, \mathbf{a}^k \rangle_{\mathbb{R}} - \langle \mathbf{b}^j, \mathbf{b}^k \rangle_{\mathbb{R}} = 0 \quad \text{e} \quad \langle \mathbf{b}^j, \mathbf{a}^k \rangle_{\mathbb{R}} + \langle \mathbf{a}^j, \mathbf{b}^k \rangle_{\mathbb{R}} = 0.$$

Disso, o resultado desejado segue imediatamente. ■

O número de parâmetros reais que determina os elementos de $\text{SO}(n)$ é, naturalmente, a dimensão (real) do espaço das matrizes antissimétricas reais. Os elementos independentes de tais matrizes podem ser contados da seguinte forma. Para uma matriz antissimétrica A temos $A_{ij} = -A_{ji}$ para $i \neq j$ e $A_{jj} = 0$. Assim, são independentes os elementos reais acima da diagonal principal, que são em número $1 + \dots + n - 1 = n(n - 1)/2$, este sendo, portanto, o número de parâmetros reais independentes que determina os elementos de $\text{SO}(n)$.

O fato demonstrado nessa proposição acima mostra que os subespaços gerados por pares $\mathbf{a}^j, \mathbf{b}^j$ são ortogonais em \mathbb{R}^n . Na base formada por esses vetores, R tem a forma de blocos diagonais $\begin{pmatrix} \cos \theta_j & \text{sen} \theta_j \\ -\text{sen} \theta_j & \cos \theta_j \end{pmatrix}$. Resta-nos ainda discutir o que se passa com os autovalores reais.

Caso II. $e^{i\theta_k} = \pm 1$.

Como comentamos, o autovalor -1 tem multiplicidade algébrica par em \mathbb{C}^n . Como R é unitária em \mathbb{C}^n , R é simples (vide definição à página 463), concluímos que a multiplicidade geométrica desse autovalor em \mathbb{C}^n é igualmente par. Os autovalores reais de R correspondem a autovetores reais (por quê?). Assim, há um subespaço real de dimensão par onde R age como $-\mathbb{1}$. Como a dimensão é par, podemos escrever R nesse subespaço como uma série de blocos diagonais como $\begin{pmatrix} \cos \theta_j & \text{sen} \theta_j \\ -\text{sen} \theta_j & \cos \theta_j \end{pmatrix}$, mas para $\theta_j = \pi$.

Para o autovalor $+1$ a conclusão é a mesma, exceto que se n for ímpar a multiplicidade geométrica é ímpar. Assim, R age nesse subespaço como uma série de blocos diagonais como $\begin{pmatrix} \cos \theta_j & \text{sen} \theta_j \\ -\text{sen} \theta_j & \cos \theta_j \end{pmatrix}$, mas para $\theta_j = 0$ e um bloco 1×1 com elemento de matriz 1 .

A conclusão é a seguinte: para $R \in \text{SO}(n)$ existe uma matriz ortogonal³⁵ V tal que $R = VB V^{-1}$, onde B é a seguinte

³⁵ A matriz é ortogonal pois faz a mudança de base para a base dos vetores $\mathbf{a}^j, \mathbf{b}^j$ e dos autovetores de autovalor ± 1 , os quais são todos ortogonais entre si, como provamos acima. Um fato crucial, como se vê.

matriz: quando n é par, ou seja, $n = 2m$, para algum $m > 0$ inteiro, B é a matriz bloco-diagonal dada por

$$B = \begin{pmatrix} \boxed{\begin{matrix} \cos \theta_1 & \text{sen} \theta_1 \\ -\text{sen} \theta_1 & \cos \theta_1 \end{matrix}} & & 0 & \cdots & 0 \\ & & & & & \\ & 0 & \boxed{\begin{matrix} \cos \theta_2 & \text{sen} \theta_2 \\ -\text{sen} \theta_2 & \cos \theta_2 \end{matrix}} & & 0 \\ & \vdots & & \ddots & \\ & 0 & 0 & & \boxed{\begin{matrix} \cos \theta_m & \text{sen} \theta_m \\ -\text{sen} \theta_m & \cos \theta_m \end{matrix}} \end{pmatrix}, \quad (21.179)$$

que é formada por $m = n/2$ blocos 2×2 , como indicado acima, sendo os demais elementos de matriz nulos. Quando n é ímpar, ou seja, $n = 2m + 1$, para algum $m > 0$ inteiro, B é a matriz bloco-diagonal dada por

$$B = \begin{pmatrix} \boxed{\begin{matrix} \cos \theta_1 & \text{sen} \theta_1 \\ -\text{sen} \theta_1 & \cos \theta_1 \end{matrix}} & & 0 & \cdots & 0 & 0 \\ & & & & & \\ & 0 & \boxed{\begin{matrix} \cos \theta_2 & \text{sen} \theta_2 \\ -\text{sen} \theta_2 & \cos \theta_2 \end{matrix}} & & 0 & 0 \\ & \vdots & & \ddots & & \vdots \\ & 0 & 0 & & \boxed{\begin{matrix} \cos \theta_m & \text{sen} \theta_m \\ -\text{sen} \theta_m & \cos \theta_m \end{matrix}} & 0 \\ & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}, \quad (21.180)$$

que é formada por $m = (n - 1)/2$ blocos 2×2 , como indicado acima, sendo o elemento B_{nn} igual a 1, e os demais elementos de são matriz nulos.

Definamos agora (tanto para o caso em que n é par ou ímpar)

$$J_k := \frac{\partial}{\partial \theta_k} R \Big|_{\theta_1 = \dots = \theta_m = 0} .$$

É claro que cada J_k é a matriz antissimétrica composta pelo bloco $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ colocado na k -ésima posição, os demais elementos de matriz sendo iguais a zero. Deve ser também claro que $J_k J_l = J_l J_k$ para todos $k, l = 1, \dots, m$ e que

$$B = \exp(\theta_1 J_1 + \dots + \theta_m J_m) .$$

E. 21.73 *Exercício.* Complete os detalhes. ✦

Do comentado acima, temos, então, que $R = V B V^{-1} = \exp(A)$, onde

$$A := V(\theta_1 J_1 + \dots + \theta_m J_m) V^{-1} .$$

Agora, como V é ortogonal e as J_k são antissimétricas, é elementar verificar que $A^T = -A$. Isso completa a prova da Proposição 21.24. ■

A Proposição 21.24 diz-nos que a exponenciação é uma aplicação sobrejetora de $\mathfrak{so}(n)$ em $\mathrm{SO}(n)$. Isso é um caso particular de um teorema mais geral que diz que isso é válido para qualquer grupo de Lie compacto, conexo e cuja álgebra de Lie seja de dimensão finita.

A Proposição 21.24 tem os dois seguintes corolários simples:

Corolário 21.5 *Para n ímpar existe para cada $R \in \mathrm{SO}(n)$ um vetor $\vec{\eta} \in \mathbb{R}^n$ tal que $R\vec{\eta} = \vec{\eta}$.* □

O vetor $\vec{\eta}$ é o autovetor com autovalor 1. Se n é par pode não haver um tal vetor invariante. Esse corolário, junto com a Proposição 21.24, generaliza a Proposição 21.6, página 1057, que era restrita ao caso $\mathrm{SO}(3)$.

Corolário 21.6 *O grupo $\mathrm{SO}(n)$ é conexo por caminhos e, portanto, é conexo.* □

Prova. Pelo que vimos, se $R \in \mathrm{SO}(n)$, R é da forma $R = e^A$, para alguma $A \in \mathfrak{so}(n)$. Logo, R pertence ao subgrupo uniparamétrico de $\mathrm{SO}(n)$ gerado por A : $\{\exp(tA), t \in \mathbb{R}\}$. Esse subgrupo conecta continuamente U à identidade $\mathbb{1}$ (que corresponde a $t = 0$). ■

21.6 O Grupo Afim e o Grupo Euclidiano

Seja V um espaço vetorial (que, lembremos, é um grupo Abelian em relação à operação de adição de vetores). Vamos denotar por $\mathrm{GL}(V)$ o conjunto dos operadores lineares bijetores (e, portanto, inversíveis) de V em V . Também sabemos que $\mathrm{GL}(V)$ é um grupo.

Existe uma ação à esquerda natural de $\mathrm{GL}(V)$ em V , a saber $\alpha : \mathrm{GL}(V) \times V \rightarrow V$ dada por $\alpha(M, v) := Mv$ onde $M \in \mathrm{GL}(V)$ e $v \in V$. (Mostre que isso define uma ação à esquerda).

Dessa forma podemos definir o produto semidireto de $\mathrm{GL}(V)$ e V , denotado por $\mathrm{GL}(V) \circledast V$, ou simplesmente por $\mathrm{GL}(V) \circledast V$, definindo em $\mathrm{GL}(V) \times V$ o produto

$$(M, u) \cdot (M', u') := (MM', Mu' + u) ,$$

onde $M, M' \in \mathrm{GL}(V)$ e $u, u' \in V$. (A noção de produto semidireto de dois grupos foi definida na Seção 2.2.4.2, página 145).

$\mathrm{GL}(V) \circledast V$ é denominado o *grupo afim* do espaço vetorial V .

Se G for um subgrupo de $\mathrm{GL}(V)$, o produto semidireto $G \circledast V$ é definido analogamente $(M, u) \cdot (M', u') := (MM', Mu' + u)$, onde $M, M' \in G$ e $u, u' \in V$. É evidente que $G \circledast V$ é um subgrupo de $\mathrm{GL}(V) \circledast V$.

E. 21.74 *Exercício.* Mostre que o conjunto de translações puras formado pelos pares $(\mathbb{1}, v)$, $v \in V$ é um subgrupo normal de $GL(V) \otimes V$. Sugestão: basta mostrar que trata-se de um subgrupo Abeliano. *

E. 21.75 *Exercício.* Se G é um subgrupo normal de $GL(V)$, mostre que $G \otimes V$ é um subgrupo normal de $GL(V) \otimes V$. *

E. 21.76 *Exercício.* Se G é um subgrupo de $GL(V)$, mostre que $V \ni u \mapsto Ru + v$, para $(R, v) \in G \otimes V$, define uma ação à esquerda de $G \otimes V$ em V . *

Consideraremos dois exemplos importantes, o grupo Euclidiano³⁶ e o grupo de Poincaré³⁷ o qual será tratado na Seção 21.8.

• **O grupo Euclidiano**

O chamado *grupo Euclidiano em dimensão n* é definido como sendo o grupo $\mathbf{E}_n := O(n) \otimes \mathbb{R}^n$. O grupo $\mathbf{SE}_n := SO(n) \otimes \mathbb{R}^n$ é denominado *grupo Euclidiano especial em dimensão n* . O produto em \mathbf{E}_n é definido por $(R_1, a_1) \cdot (R_2, a_2) = (R_1 R_2, a_1 + R_1 a_2)$, com $R_1, R_2 \in SO(n)$ e $a_1, a_2 \in \mathbb{R}^n$.

O grupo \mathbf{E}_n tem uma ação natural em \mathbb{R}^n dada por $\mathbb{R}^n \ni y \mapsto Ry + x$, para cada elemento $(R, x) \in \mathbf{E}_n$. Assim, \mathbf{E}_n implementa em \mathbb{R}^n translações, rotações e reflexões, as chamadas *transformações Euclidianas* de \mathbb{R}^n . Essa é, em verdade, a própria motivação da definição de \mathbf{E}_n .

E. 21.77 *Exercício.* Mostre que $\mathbb{R}^n \ni y \mapsto Ry + x =: \alpha_{(R, x)}(y)$, para $(R, x) \in \mathbf{E}_n$, define uma ação à esquerda de \mathbf{E}_n em \mathbb{R}^n . Constate que $\alpha_{(\mathbb{1}_n, x)}(0) = x$ e conclua que essa ação é transitiva. *

Há um subgrupo de $GL(n + 1, \mathbb{R})$ que é isomorfo a \mathbf{E}_n . Sejam as matrizes reais $(n + 1) \times (n + 1)$

$$E(R, x) := \begin{pmatrix} \boxed{R} & \boxed{x} \\ \boxed{0} & \boxed{1} \end{pmatrix}, \quad \text{com } R \in O(n) \text{ e } x \in \mathbb{R}^n. \quad (21.181)$$

Então, tem-se

$$E(R, x) E(R', x') = E(RR', Rx' + x). \quad (21.182)$$

E. 21.78 *Exercício importante.* Mostre isso. *

Assim, o conjunto de matrizes $\{E(R, x) \in GL(n + 1, \mathbb{R}), \text{ com } R \in O(n) \text{ e } x \in \mathbb{R}^n\}$ forma um subgrupo de $GL(n + 1, \mathbb{R})$ que é isomorfo a \mathbf{E}_n . Também denotaremos esse grupo por \mathbf{E}_n .

E. 21.79 *Exercício.* Prove essa última afirmativa. *

A expressão (21.181) define, devido a (21.182), uma representação de \mathbf{E}_n em \mathbb{R}^{n+1} .

E. 21.80 *Exercício.* Consideremos o subconjunto (não é um subespaço) $M_n \subset \mathbb{R}^{n+1}$ dado por

$$M_n := \left\{ \begin{pmatrix} y_1 \\ \vdots \\ y_n \\ 1 \end{pmatrix}, \text{ com } y_1, \dots, y_n \in \mathbb{R} \right\}.$$

³⁶Euclides de Alexandria (ci. 325 A.C. – ci. 265 A.C.).

³⁷Jules Henri Poincaré (1854–1912).

Denotemos $\begin{pmatrix} y_1 \\ \vdots \\ y_n \\ 1 \end{pmatrix} \equiv \begin{pmatrix} y \\ 1 \end{pmatrix}$. Mostre que

$$E(R, x)\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} Ry+x \\ 1 \end{pmatrix}$$

e conclua que $E(R, x)\begin{pmatrix} y \\ 1 \end{pmatrix}$ é uma ação de \mathbf{E}_n em M_n . Como $E(\mathbb{1}_n, x)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}$, conclua que essa ação é transitiva. *

• **Os geradores infinitesimais do grupo Euclidiano especial \mathbf{SE}_3**

De particular interesse é o caso $n = 3$. É possível identificar os seguintes subgrupos uniparamétricos de \mathbf{SE}_3 , aqueles gerados pelas matrizes $E(R_j, 0)$, $j = 1, 2, 3$, onde R_j são as matrizes introduzidas em (21.91) e que geram subgrupos uniparamétricos de $\mathbf{SO}(3)$ e aqueles gerados pelas matrizes $E(\mathbb{1}, x_k)$, $k = 1, 2, 3$, onde $x_1 = (x, 0, 0)$, $x_2 = (0, x, 0)$ e $x_3 = (0, 0, x)$ com $x \in \mathbb{R}$. Esses últimos subgrupos geram translações nas direções $k = 1, 2, 3$.

E. 21.81 Exercício importante. Mostre que esses seis subgrupos são subgrupos uniparamétricos. *

Como facilmente se verifica, os geradores infinitesimais desses subgrupos são as seguintes matrizes:

$$j_1 := \begin{pmatrix} \boxed{J_1} & 0 \\ & 0 \\ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad j_2 := \begin{pmatrix} \boxed{J_2} & 0 \\ & 0 \\ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad j_3 := \begin{pmatrix} \boxed{J_3} & 0 \\ & 0 \\ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

e

$$p_1 := \begin{pmatrix} \boxed{0} & 1 \\ & 0 \\ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad p_2 := \begin{pmatrix} \boxed{0} & 0 \\ & 1 \\ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad p_3 := \begin{pmatrix} \boxed{0} & 0 \\ & 0 \\ & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

sendo que J_1, J_2 e J_3 são os geradores infinitesimais de $\mathbf{SO}(3)$, definidos em (21.92)-(21.94), página 1055. Usando a forma das matrizes J_k dada em (21.92)-(21.94), é fácil constatar as seguintes relações de comutação entre os geradores infinitesimais acima:

$$[j_a, j_b] = \sum_{c=1}^3 \varepsilon_{abc} j_c, \quad [p_a, p_b] = 0, \quad [j_a, p_b] = \sum_{c=1}^3 \varepsilon_{abc} p_c. \tag{21.183}$$

E. 21.82 Exercício. Verifique! *

As relações (21.183) representam as relações de comutação da álgebra de Lie \mathfrak{e}_3 do grupo \mathbf{SE}_3 . Note que p_1, p_2 e p_3 formam uma subálgebra Abeliãna de \mathfrak{e}_3 e que essa subálgebra é um ideal de \mathfrak{e}_3 . Esse fato reflete a propriedade que o subgrupo de translações é um subgrupo normal de \mathbf{SE}_3 .

• **Os geradores infinitesimais do grupo Euclidiano especial \mathbf{SE}_2**

De maneira análoga podemos tratar o caso (mais simples) do grupo \mathbf{SE}_2 . Os elementos de $\mathbf{SO}(2) \otimes \mathbb{R}^2$ podem ser

parametrizados na forma

$$\begin{pmatrix} \cos \theta & -\operatorname{sen} \theta & x_1 \\ \operatorname{sen} \theta & \cos \theta & x_2 \\ 0 & 0 & 1 \end{pmatrix}, \quad \theta \in (-\pi, \pi], \quad x_1, x_2 \in \mathbb{R}.$$

Seus geradores infinitesimais serão

$$j_1 := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_1 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_2 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Como é fácil de verificar, as relações de comutação entre esses geradores infinitesimais são

$$[j_1, p_1] = p_2, \quad [j_1, p_2] = -p_1, \quad [p_1, p_2] = 0.$$

Um elemento genérico dessa álgebra de Lie é da forma

$$I(\mathcal{J}, t) := \begin{pmatrix} \boxed{\mathcal{J}} & \boxed{t} \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{onde } \mathcal{J} = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} \quad \text{e} \quad t = \begin{pmatrix} t_1 \\ t_2 \end{pmatrix},$$

com $-\pi < \theta \leq \pi$ e $t_1, t_2 \in \mathbb{R}$. É um exercício fácil (faça-o!) constatar que para todo $k \in \mathbb{N}$, tem-se

$$I(\mathcal{J}, t)^k = I(\mathcal{J}^k, \mathcal{J}^{k-1}t).$$

Consequentemente, vale que

$$\exp(I(\mathcal{J}, t)) = \mathbb{1} + \sum_{k=1}^{\infty} \frac{1}{k!} I(\mathcal{J}, t)^k = \mathbb{1} + \sum_{k=1}^{\infty} \frac{1}{k!} I(\mathcal{J}^k, \mathcal{J}^{k-1}t) = \begin{pmatrix} \boxed{R} & \boxed{t'} \\ 0 & 0 & 1 \end{pmatrix},$$

onde

$$R := e^{\mathcal{J}} = \begin{pmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{pmatrix} \quad \text{e} \quad t' = f(\mathcal{J})t,$$

sendo f a função analítica inteira definida pela série de Taylor

$$f(w) := 1 + \sum_{k=2}^{\infty} \frac{1}{k!} w^{k-1}, \quad w \in \mathbb{C}. \tag{21.184}$$

É fácil constatar que

$$f(w) = \begin{cases} \frac{e^w - 1}{w}, & w \neq 0, \\ 1, & w = 0. \end{cases}$$

A matriz $f(\mathcal{J})$ pode ser calculada facilmente usando-se o fato que, para $M = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, valem $M^{2k} = (-1)^k \mathbb{1}$ e $M^{2k+1} = (-1)^k M$, $k \in \mathbb{N}_0$, de onde se extrai

$$\begin{aligned} f(\mathcal{J}) &:= \mathbb{1} + \sum_{k=2}^{\infty} \frac{1}{k!} \mathcal{J}^{k-1} = \mathbb{1} + \sum_{m=1}^{\infty} \frac{1}{(2m)!} \mathcal{J}^{2m-1} + \sum_{m=1}^{\infty} \frac{1}{(2m+1)!} \mathcal{J}^{2m} \\ &= \sum_{m=1}^{\infty} \frac{(-1)^m \theta^{2m-1}}{(2m)!} M + \sum_{m=0}^{\infty} \frac{(-1)^m \theta^{2m}}{(2m+1)!} \mathbb{1} = \frac{\cos \theta - 1}{\theta} M + \frac{\text{sen} \theta}{\theta} \mathbb{1} \\ &= \begin{pmatrix} \frac{\text{sen} \theta}{\theta} & -\frac{\cos \theta - 1}{\theta} \\ \frac{\cos \theta - 1}{\theta} & \frac{\text{sen} \theta}{\theta} \end{pmatrix}. \end{aligned} \quad (21.185)$$

Notemos que $\det f(\mathcal{J}) = 2 \left(\frac{1 - \cos \theta}{\theta^2} \right) \neq 0$ para $-\pi < \theta \leq \pi$. Assim, $f(\mathcal{J})$ é inversível e se escolhermos $t = f(\mathcal{J})^{-1}x$, para qualquer $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$, teremos

$$\exp(I(\mathcal{J}, f(\mathcal{J})^{-1}x)) = \begin{pmatrix} \boxed{R} & \boxed{x} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\text{sen} \theta & x_1 \\ \text{sen} \theta & \cos \theta & x_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Isso prova que todo elemento do grupo $\mathbf{SE}_2 = \text{SO}(2) \otimes \mathbb{R}^2$ pode ser escrito como exponencial de um elemento da sua própria álgebra de Lie. Essa afirmação é igualmente válida para todo os grupos $\mathbf{SE}_n = \text{SO}(n) \otimes \mathbb{R}^n$. A demonstração segue passos análogos aos acima pois, como observamos na Seção 21.5.2, página 1091, os elementos de $\text{SO}(n)$ podem ser escritos em uma base conveniente na forma de blocos de matrizes de $\text{SO}(2)$. Isso implicará que também no caso geral a matriz $f(\mathcal{J})$ é inversível. Deixamos os detalhes da demonstração como exercício ao leitor.

21.7 O Grupo de Lorentz em 3 + 1-Dimensões

Para a leitura desta seção uma certa familiaridade com rudimentos da teoria da relatividade restrita é recomendável, mas não totalmente indispensável.

21.7.1 O Espaço-Tempo, a Noção de Intervalo e a Estrutura Causal

É um fato elementar da natureza ser possível descrever qualquer evento idealmente pontual e de duração instantânea por uma coleção de quatro números que especificam sua posição espacial e seu instante de tempo, medidos em algum sistema de referência. A coleção de todos os eventos pontuais de duração instantânea é denominada *espaço-tempo*, noção

introduzida por Minkowski³⁸. Assim, é natural (pelo menos na ausência de campos gravitacionais, que podem alterar a topologia global do espaço-tempo) identificar o mesmo com o espaço matemático \mathbb{R}^4 . Assim descrito, cada evento pode ser especificado em um sistema de referência que adote coordenadas espaciais Cartesianas, por uma quádrupla ordenada (x_0, x_1, x_2, x_3) , onde convencionamos que os três últimos números são coordenadas espaciais do evento e o primeiro sua coordenada temporal. O leitor deve ser advertido que muitos autores convencionam escrever as coordenadas espaço-temporais de um evento na forma (x_1, x_2, x_3, x_4) , onde x_4 é a coordenada temporal. Isso alteraria a forma das matrizes que serão manuseadas abaixo, mas não a essência dos resultados que apresentaremos.

Na Mecânica Clássica, a primeira lei de Newton³⁹ afirma existirem certos sistemas de referência dotados da seguinte propriedade: se um corpo encontra-se isolado do restante do universo, ou seja, se sobre ele não atuam forças externas, então em relação a esse sistema de referência esse corpo se move com velocidade constante. Tais sistemas de referência são denominados *sistemas de referência inerciais*, pois neles vale o princípio de inércia. É muito fácil concluir que se um sistema de referência se move com velocidade constante em relação a um sistema de referência inercial, então ele é também um sistema de referência inercial.

Sistemas de referência inerciais desempenham um papel central, pois neles as Leis da Física assumem um carácter universal. É um postulado fundamental da Física que suas leis básicas são as mesmas em todos os sistemas de referência inerciais. Na mesma linha, é um postulado fundamental da Física que também suas constantes fundamentais, tais como a velocidade da luz c , a constante de Planck⁴⁰ \hbar , a constante de gravitação universal G e outras tenham também o mesmo valor em todos os sistemas de referência inerciais. Mais que isso, os sistemas de referência inerciais concordam quanto às relações de causa e efeito entre todos os eventos ocorridos no espaço-tempo. Essa série de princípios aqui mal delineados é por vezes denominada *princípio da relatividade*. O princípio da relatividade tem sua origem nos trabalhos de Galilei⁴¹ sobre Dinâmica, mas foi com a Teoria da Relatividade de Einstein⁴² que suas reais consequências foram exploradas em sua máxima extensão.

Ao realizarmos transformações entre sistemas de coordenadas inerciais, as coordenadas dos eventos transformam-se linearmente⁴³. Esse postulado é familiar se nos lembramos da ação do grupo de translações, da ação do grupo de rotações no espaço tridimensional ou das transformações de Galilei da Mecânica Clássica (não relativista). Assim, cada transformação entre sistemas de coordenadas inerciais deve ser representada na forma $Lx + t$, onde L é uma matriz real 4×4 e x e t são vetores de \mathbb{R}^4 . Aqui, x e t são representados na forma de um vetor coluna, como $x \equiv \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$.

O vetor t representa uma translação (tanto no espaço quanto no tempo) entre os sistemas de coordenadas. Cada matriz $L \in \text{Mat}(\mathbb{R}, 4)$ deve depender das velocidades relativas entre os sistemas inerciais cuja transformação descreve, da direção dessas velocidades e dos ângulos relativos entre os eixos Cartesianos espaciais dos dois sistemas. L deve também conter informação sobre se os eixos Cartesianos espaciais dos dois sistemas têm a mesma orientação (positiva ou negativa) e sobre se os relógios dos dois sistemas correm na mesma direção.

Dados dois eventos quaisquer x, y no espaço-tempo (que doravante identificaremos com \mathbb{R}^4) e cujas coordenadas sejam $x = (x_0, x_1, x_2, x_3)$ e $y = (y_0, y_1, y_2, y_3)$ em um determinado sistema de referência inercial, define-se o *intervalo* entre ambos como sendo a quantidade⁴⁴

$$I(x, y) := I(x - y) := c^2(x_0 - y_0)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2,$$

onde c é a velocidade da luz no sistema de referência inercial em questão.

³⁸Hermann Minkowski (1864–1909). A expressão “espaço-tempo” provém do Alemão “Raumzeit”.

³⁹Sir Isaac Newton (1643–1727).

⁴⁰Max Karl Ernst Ludwig Planck (1858–1947).

⁴¹Galileo Galilei (1564–1642).

⁴²Albert Einstein (1879–1955).

⁴³Transformações que preservam a estrutura causal (vide discussão adiante) no espaço-tempo de Minkowski são necessariamente lineares. Essa afirmação foi demonstrada por diversos autores de forma independente. Citamos ao estudante interessado as seguintes referências:

1. A. D. Alexandrov, “On Lorentz Transformations”, Uspehi Mat. Nauk. **5**, 187 (1950).
2. A. D. Alexandrov, “Mappings of Spaces with Families of Cones and Space-Time Transformations”, Annali di Mat. Pura Appl. **103**, 229–257 (1975).
3. E. C. Zeeman, “Causality Implies the Lorentz Group”, J. Math. Phys. **5**, 490–493 (1964).
4. H.-J. Borchers and G. C. Hegerfelt, “The Structure of Space-Time Transformations”, Commun. Math. Phys. **28**, 259–266 (1972).

No comentário à página 1107 provaremos com base no Teorema 21.7, página 1035, que transformações que mantêm invariantes os intervalos entre todos os pares de eventos são necessariamente lineares.

⁴⁴Novamente supomos a ausência de campos gravitacionais, em cuja presença a definição de intervalo tem de ser modificada.

A noção de intervalo entre eventos é de grande importância. Para começar a explicar isso consideremos a situação na qual dois eventos distintos de coordenadas x e y representam a produção e a absorção de um mesmo raio luminoso, respectivamente. Se em um determinado sistema de referência inercial as coordenadas desses eventos são $x = (x_0, x_1, x_2, x_3)$ e $y = (y_0, y_1, y_2, y_3)$, então a velocidade de propagação da luz entre x e y satisfaz

$$c^2 = \frac{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}{(y_0 - x_0)^2}$$

e, portanto, $I(y, x) = I(y - x) = 0$. Um dos postulados fundamentais da Teoria da Relatividade Restrita é a afirmação que a velocidade de propagação da luz no vácuo é a mesma para qualquer sistema de referência inercial. Portanto, se em um outro sistema de referência inercial as coordenadas de x e y fossem $x' = (x'_0, x'_1, x'_2, x'_3)$ e $y' = (y'_0, y'_1, y'_2, y'_3)$ teríamos igualmente

$$c^2 = \frac{(y'_1 - x'_1)^2 + (y'_2 - x'_2)^2 + (y'_3 - x'_3)^2}{(y'_0 - x'_0)^2}$$

e, portanto, tem-se igualmente $I(y', x') = I(y' - x') = 0$ com o mesmo valor c para a velocidade de propagação da luz.

Compreendemos, então, que o postulado da constância da velocidade da luz pode ser traduzido matematicamente da seguinte forma: *se o intervalo entre dois eventos é nulo em um sistema de referência inercial, então é também nulo em todos os demais sistemas de referência inerciais*. Mais adiante provaremos que, sob certas hipóteses físicas adicionais, esse fato implica uma condição ainda mais geral de invariância: o intervalo entre dois eventos quaisquer é o mesmo em qualquer sistema de referência inercial, mesmo quando não é nulo.

Nota. Independente de ser um postulado teórico, a constância da velocidade da luz é um fato experimental que tem sofrido sucessivas confirmações ao longo de várias décadas. Para uma lista seguramente parcial de referências recentes (das últimas quatro décadas) contendo testes experimentais da constância da velocidade da luz e testes da velocidade da luz como velocidade limite, vide:

1. T. S. Jaseja, A. Javan, J. Murray and C. H. Townes. “Test of Special Relativity or of the Isotropy of Space by Use of Infrared Masers”. Phys. Rev. **A133**, A1221–A1125 (1964).
2. T. Alväger, F. J. M. Farley, J. Kjellman and I. Wallin. “Test of the Second Postulate of Special Relativity in the GeV Region”. Phys. Lett. **12**, 260–263 (1964).
3. D. I. Blokhintsev. “Basis for Special Relativity Theory Provided by Experiments in High Energy Physics”. Sov. Phys. Uspekhi, **9**, 405 (1966).
4. Z. G. T. Guiragossian, G. B. Rothbart, M. R. Yearian, R. A. Gearhart and J. J. Murray. “Relative Velocity Measurements of Electrons and Gamma Rays at 15 GeV”. Phys. Rev. Lett. **34**, 335–338 (1975).
5. K. Brecher. “Is the Speed of Light Independent of the Velocity of the Source?”. Phys. Rev. Lett. **39**, 1051–1054, 1236(E) (1977).
6. D. Newman, G. W. Ford, A. Rich and E. Sweetman. “Precision Experimental Verification of Special Relativity”. Phys. Rev. Lett. **40**, 1355–1358 (1978).
7. K. M. Baird, D. S. Smith and B. G. Whitford. “Confirmation of the Currently Accepted Value 299 792 458 Metres per Second for the Speed of Light”. Opt. Comm. **31**, 367–368 (1979).
8. G. L. Greene, M. Scott Dewey, E. G. Kessler, Jr. and E. Fischbach. “Test of Special Relativity by a Determination of the Lorentz Limiting Velocity: Does $E = mc^2$?”. Phys. Rev. D **44**, R2216–R2219 (1991).
9. Bradley E. Schaefer. “Severe Limits on Variations of the Speed of Light with Frequency”. Phys. Rev. Lett. **82**, 4964 (1999).

Para um texto recente, vide [424]⁴⁵. ♣

Notemos que o intervalo depende da diferença $x - y$. Assim, translações entre sistemas de referência automaticamente mantêm invariantes os intervalos entre eventos. Por essa razão vamos por ora nos interessar apenas por transformações entre sistemas de referência que sejam do tipo Lx , com $L \in \text{Mat}(\mathbb{R}, 4)$.

Para prosseguirmos precisamos introduzir uma importante classificação de intervalos.

• Intervalos de tipo luz, de tipo tempo e de tipo espaço

Em um sistema de referência, dois eventos distintos x e y são ditos ser⁴⁶

1. do *tipo luz* se $I(x, y) = 0$,
2. do *tipo tempo* se $I(x, y) > 0$,

⁴⁵Agradecemos à Profa. Renata Zukanovich Funchal pelas referências acima.

⁴⁶As expressões em Inglês são “light-like”, “time-like” e “space-like”, respectivamente. Essa nomenclatura provém do Alemão: “lichtartig”, “zeitartig” e “raumartig”.

3. do tipo espaço se $I(x, y) < 0$.

Se dois eventos distintos $x = (x_0, x_1, x_2, x_3)$ e $y = (y_0, y_1, y_2, y_3)$ são do tipo luz, então

$$\frac{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}{(y_0 - x_0)^2} = c^2 .$$

Se dois eventos distintos $x = (x_0, x_1, x_2, x_3)$ e $y = (y_0, y_1, y_2, y_3)$ são do tipo tempo, então

$$\frac{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}{(y_0 - x_0)^2} < c^2 .$$

Se dois eventos distintos $x = (x_0, x_1, x_2, x_3)$ e $y = (y_0, y_1, y_2, y_3)$ são do tipo espaço, então

$$\frac{(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2}{(y_0 - x_0)^2} > c^2 .$$

Com isso, entendemos que

1. Se dois eventos são separados por um intervalo do tipo luz pode haver um sinal físico conectando ambos e que se propagaria com a velocidade da luz.
2. Se dois eventos são separados por um intervalo do tipo tempo pode haver um sinal físico conectando ambos e que se propagaria com velocidade menor que a da luz.
3. Se dois eventos são separados por um intervalo do tipo espaço não pode haver um sinal físico conectando ambos, pois o mesmo se propagaria com velocidade maior que a da luz.

A importância dessas considerações é a seguinte. É uma crença da Física atual que as partículas elementares que compõem toda a matéria conhecida do universo não podem mover-se com velocidade maior que a da luz. Consequentemente, se dois eventos são separados por um intervalo do tipo espaço não pode haver nenhum processo físico que, iniciando-se em um evento, influencie o outro. Diz-se, então, que esses eventos são *causalmente desconectados*, ou seja, não pode haver nenhuma *relação causal* (isto é, de causa e efeito) entre ambos. Por outro lado, se dois eventos são separados por um intervalo do tipo tempo então pode haver alguma influência causal entre ambos, por exemplo, através de uma partícula ou corpo material que, movendo-se no espaço-tempo com velocidades inferiores à da luz, parta de um evento e influencie o outro. No caso de intervalos do tipo luz a situação é a mesma mas, então, a eventual influência de um no outro deve propagar-se com a velocidade da luz.

E. 21.83 *Exercício*. Passe vários dias meditando sobre os parágrafos acima. *

• **A estrutura causal. Transformações que preservam a estrutura causal**

Como se percebe, se aceitarmos a ideia que processos físicos não podem propagar-se com velocidades superiores à da luz, a noção de intervalo estabelece as possíveis relações de causalidade entre todos os eventos do espaço-tempo, ao dizer quais eventos podem eventualmente influenciar-se (aqueles que são do tipo tempo ou do tipo luz um em relação ao outro) e quais não podem de forma alguma influenciar-se (aqueles que são do tipo espaço um em relação ao outro).

É uma crença da Física atual que essas relações de causalidade devem ser as mesmas para todos os sistemas de referência inerciais, pois os mesmos descrevem as mesmas Leis Físicas e devem perceber as mesmas relações de causa e efeito entre os eventos que compõem o universo.

E. 21.84 *Exercício*. Mais alguns dias de meditação. *

Com isso, podemos introduzir a seguinte definição: dizemos que uma transformação linear L , que representa uma transformação entre dois sistemas de referência, *preserva a estrutura causal do espaço-tempo* se a mesma satisfizer todas as três condições seguintes:

1. $I(Lx, Ly) = 0$ sempre que $I(x, y) = 0$,

2. $I(Lx, Ly) > 0$ sempre que $I(x, y) > 0$,
3. $I(Lx, Ly) < 0$ sempre que $I(x, y) < 0$.

Em palavras, L preserva o tipo de intervalo que separa todos os eventos do espaço-tempo, levando todos os intervalos do tipo luz em intervalos do tipo luz, levando todos os intervalos do tipo tempo em intervalos do tipo tempo e levando todos os intervalos do tipo espaço em intervalos do tipo espaço.

Notemos que a condição que impõe que $I(Lx, Ly) = 0$ sempre que $I(x, y) = 0$ é a condição da invariância da velocidade da luz (já mencionada acima), mas as demais representam algo diferente: a invariância das relações de causalidade por mudança de sistemas de referência inerciais.

Um pouco mais abaixo exploraremos as consequências matemáticas que essas imposições têm sobre as transformações L e concluiremos que, sob as hipóteses acima (e sob uma hipótese adicional de ausência de dilatações), vale uma consequência mais forte, a saber, que $I(Lx, Ly) = I(x, y)$ para todos os eventos x e y . Assim, transformações que preservam a estrutura causal e não envolvem dilatações preservam o valor do intervalo entre dois eventos quaisquer do espaço-tempo.

Por fim, apenas a título de ilustração, exemplifiquemos como seria uma transformação que preserva os intervalos de tipo luz mas não os demais, preservando, portanto, a velocidade da luz mas violando a estrutura causal. Consideremos um espaço-tempo bidimensional, onde cada evento é descrito por uma coordenada espacial x_1 e uma temporal t . Seja a matriz $L = \begin{pmatrix} 0 & c^{-1} \\ c & 0 \end{pmatrix}$. O intervalo entre os eventos $x = \begin{pmatrix} t \\ x_1 \end{pmatrix}$ e $0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ seria $I(x, 0) = c^2 t^2 - x_1^2$. Porém, pela transformação L teríamos $\begin{pmatrix} t' \\ x'_1 \end{pmatrix} = L \begin{pmatrix} t \\ x_1 \end{pmatrix} = \begin{pmatrix} c^{-1} x_1 \\ ct \end{pmatrix}$. Assim,

$$I(Lx, L0) = c^2 (t')^2 - (x'_1)^2 = x_1^2 - c^2 t^2 = -I(x, 0).$$

Logo, como os intervalos $I(Lx, L0)$ e $I(x, 0)$ diferem por um sinal, teríamos para quaisquer eventos x e y

1. $I(Lx, Ly) = 0$ sempre que $I(x, y) = 0$,
2. $I(Lx, Ly) > 0$ sempre que $I(x, y) < 0$,
3. $I(Lx, Ly) < 0$ sempre que $I(x, y) > 0$.

Portanto, intervalos tipo luz seriam levados em intervalos tipo luz, mas intervalos tipo espaço seriam levados em intervalos tipo tempo e vice-versa. Como se vê por esse exemplo, em transformações que violam a estrutura causal deve haver algo como uma permutação entre coordenadas espaciais e temporais.

E. 21.85 *Exercício.* São tais transformações fisicamente aceitáveis? *

• Dilatações

Vamos agora discutir uma classe de transformações que preservam a estrutura causal: as dilatações.

Para $\lambda \in \mathbb{R}$, $\lambda \neq 0$, a matriz $D(\lambda) := \lambda \mathbb{1}$ simplesmente transforma cada $x \in \mathbb{R}^4$ em λx , ou seja, $D(\lambda)$ representa uma *dilatação* ou mudança de escala das coordenadas espaço-temporais de eventos. É evidente que $I(D(\lambda)x, D(\lambda)y) = \lambda^2 I(x, y)$, de modo que dilatações são transformações lineares que preservam a estrutura causal.

São as dilatações aceitáveis enquanto mudanças de sistemas de referência inerciais? Essa é uma questão muito interessante e sutil e demanda uma certa discussão.

Claramente, mudanças de escala podem ocorrer naturalmente no caso de tratarmos de dois sistemas de referência que adotam sistemas métricos diferentes, como no caso em que um sistema mede distâncias em metros e um outro em jardas (mas de modo que as medidas de tempo em um e outro sejam tais que ambos atribuem o mesmo valor numérico para c). Essas situações são triviais e poderiam ser contornadas se ambos os sistemas de referência concordassem no uso de uma mesma escala de distâncias. Mas para que isso seja possível é preciso que haja objetos físicos, em repouso em ambos os sistemas de referência, que possuam as mesmas dimensões. Poderíamos, por exemplo, adotar como unidade de distância o “tamanho médio” do átomo de Hidrogênio⁴⁷, ou o comprimento de onda de uma linha de emissão de um certo átomo ou molécula, fixos em cada sistema de referência.

⁴⁷A noção de “tamanho médio” de um átomo pode ser definida na Mecânica Quântica, mas não entraremos em detalhes aqui.

Mas o que garante que o tamanho médio de um átomo de hidrogênio parado na Terra é o mesmo que o de um átomo de hidrogênio parado em uma galáxia distante que se move em relação a nós com uma certa velocidade? A princípio, nada garante, mas a crença que sistemas de referência inerciais descrevem a mesma física envolve também a crença que certas escalas básicas de distância e de tempo, como o tamanho médio de um átomo em repouso, são as mesmas em todos os sistemas de referência inerciais. Por exemplo, o tamanho médio do átomo de hidrogênio em repouso depende de propriedades físicas que regem a interação entre o próton e o elétron que o constituem (a lei de Coulomb⁴⁸), das leis da Mecânica que regem seus movimentos (as leis da Mecânica Quântica), assim como dos valores das cargas elétricas e das massas de repouso dessas partículas. Essas grandezas e leis devem ser as mesmas em quaisquer sistemas de referência inerciais.

Intimamente associada a isso está a questão dos valores das massas de repouso das partículas elementares. Isso se deve ao fato seguinte. A Física Quântica ensina-nos que se m_0 é a massa de repouso de uma partícula elementar, digamos um elétron, então a quantidade $\hbar/(m_0c)$ tem dimensão de comprimento (verifique!). Esse é o chamado *comprimento de onda Compton*⁴⁹ da partícula de massa de repouso m_0 . Assim, para qualquer partícula de massa de repouso m_0 há uma escala de distância a ela associada.

É parte da crença associada ao princípio da relatividade que as massas em repouso das partículas elementares, como elétrons, quarks etc., são as mesmas quer na Terra quer em uma galáxia distante que se move em relação a nós com velocidade constante. Até onde se sabe, essa hipótese tem corroboração experimental, pois sua violação levaria a consequências observacionais em relação ao comportamento da matéria que nunca foram verificadas quer em observações astronômicas quer em experimentos com aceleradores de partículas feitos na Terra. Como \hbar e c são constantes físicas, devem também ser as mesmas em quaisquer sistemas de referência inerciais e, portanto, o comprimento de onda Compton de, digamos, um elétron em repouso deve ser o mesmo em qualquer sistema de referência inercial e com ele poderíamos estabelecer uma escala de distâncias universal.

Em um universo em que não houvessem escalas de distância ou de massa naturais, como por exemplo no caso de universos em que todas as partículas elementares têm massa nula e não formam estados ligados (como átomos) que possuam alguma escala de distância típica, não haveria maneira de sistemas de referência inerciais concordarem com escalas espaciais e temporais e, aí, a inclusão de dilatações seria inevitável nas transformações entre sistemas de referência. Esse não é o caso do universo em que vivemos, pois nele sabidamente habitam partículas massivas.

Assim, apesar de as dilatações satisfazerem a condição de não violarem a estrutura causal do espaço-tempo, as mesmas não devem ser consideradas como transformações legítimas de coordenadas espaço-temporais entre sistemas de referência inerciais no nosso universo, pois partimos da crença que esses sistemas podem sempre concordar quanto a certas escalas básicas de certos objetos físicos em repouso, tais como as massas de repouso de certas partículas elementares e seus comprimentos de onda Compton.

E. 21.86 *Exercício.* Mais meditação. ✱

- **A convenção que $c = 1$**

Daqui por diante adotaremos a convenção simplificadora que $c = 1$. Isso pode ser obtido pela escolha de um sistema de unidades métricas conveniente. Essa convenção, muito empregada atualmente em textos de Física Teórica⁵⁰, tem a vantagem de “limpar” as expressões matemáticas de fatores que dependam de c . Admitidamente, há uma certa “preguiça” na adoção dessa convenção, mas a mesma traz vantagens. De qualquer forma, os fatores c omitidos podem ser facilmente recuperados por considerações de análise dimensional.

- **Convenção sobre a notação matricial**

Nesta seção sobre o grupo de Lorentz convencionamos que índices matriciais e vetoriais variam de 0 a 3. Assim,

⁴⁸Charles Augustin de Coulomb (1736–1806).

⁴⁹Arthur Holly Compton (1892–1962). Compton recebeu o prêmio Nobel de Física de 1927 “for his discovery of the effect named after him”.

⁵⁰Em textos teóricos de Mecânica Quântica e Teoria Quântica de Campos, adota-se também $\hbar = 1$.

representaremos uma matriz $M \in \text{Mat}(\mathbb{R}, 4)$ na forma

$$M = \begin{pmatrix} M_{00} & M_{01} & M_{02} & M_{03} \\ M_{10} & M_{11} & M_{12} & M_{13} \\ M_{20} & M_{21} & M_{22} & M_{23} \\ M_{30} & M_{31} & M_{32} & M_{33} \end{pmatrix}. \tag{21.186}$$

• **O tensor métrico de Minkowski**

É muito conveniente escrever o intervalo entre dois eventos x e y com uso da seguinte notação matricial:

$$I(x - y) = (x_0 - y_0)^2 - (x_1 - y_1)^2 - (x_2 - y_2)^2 - (x_3 - y_3)^2 = \langle (x - y), \eta(x - y) \rangle_{\mathbb{R}},$$

onde

$$\eta := \eta(1, 3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{-\mathbb{1}} & & \\ 0 & & & \\ 0 & & & \end{pmatrix}. \tag{21.187}$$

E. 21.87 *Exercício.* Verifique! ✦

A matriz η é frequentemente denominada *tensor métrico de Minkowski* ou, especialmente entre Físicos, *métrica de Minkowski*.

21.7.2 A Invariância do Intervalo

Aprendemos acima que o postulado da invariância da velocidade de propagação da luz, visto como uma transformação linear entre sistemas de referência inerciais, implica que se x e y são dois eventos tais que

$$I(x, y) = \langle (x - y), \eta(x - y) \rangle_{\mathbb{R}} = 0, \tag{21.188}$$

então, tem-se também

$$I(Lx, Ly) = \langle L(x - y), \eta L(x - y) \rangle_{\mathbb{R}} = 0 \tag{21.189}$$

para qualquer transformação linear $L \in \text{Mat}(\mathbb{R}, 4)$ que represente uma mudança entre sistemas de referência inerciais.

Nesta seção iremos provar uma afirmação, o Teorema 21.10, adiante, que generaliza ainda mais o descrito no último parágrafo, a saber, provaremos que se $L \in \text{Mat}(\mathbb{R}, 4)$ representa uma mudança entre sistemas de referência inerciais que *preserva a estrutura causal* e *não envolve dilatações* (definições adiante), então $I(x, y) = I(Lx, Ly)$ para quaisquer eventos x e y , mesmo aqueles para os quais $I(x, y) \neq 0$. Esse fato releva a importância da noção de intervalo na teoria da relatividade: o mesmo representa uma grandeza invariante por transformações de sistemas de referência do tipo descrito acima. Dessa propriedade de invariância extrairemos todas as informações importantes sobre as transformações de Lorentz.

• **Transformações lineares e a estrutura causal**

Vamos aqui provar um teorema de importância central no entendimento da relação entre transformações $L \in \text{Mat}(\mathbb{R}, 4)$ e sua relação com a estrutura causal do espaço-tempo.

Teorema 21.10 *Seja L um elemento inversível de $\text{Mat}(\mathbb{R}, 4)$ que representa uma mudança entre sistemas de referência inerciais que preserva os intervalos de tipo luz. Então,*

$$\eta L^T \eta L = (L^T \eta L)_{00} \mathbb{1} = \pm |\det(L)|^{1/2} \mathbb{1}. \quad (21.190)$$

Se além disso L preserva a estrutura causal, então,

$$\eta L^T \eta L = (L^T \eta L)_{00} \mathbb{1} = |\det(L)|^{1/2} \mathbb{1}. \quad (21.191)$$

Por fim, se L preserva a estrutura causal e não envolve dilatações, então

$$\eta L^T \eta L = \mathbb{1}. \quad (21.192)$$

Uma consequência imediata dessa relação é que $I(Lx, Ly) = I(x, y)$ para todos $x, y \in \mathbb{R}^4$. □

Prova. Para $x \in \mathbb{R}^4$, sejam as formas quadráticas

$$I(x) := \langle x, \eta x \rangle_{\mathbb{R}} \quad \text{e} \quad J(x) := \langle Lx, \eta Lx \rangle_{\mathbb{R}} = \langle x, L^T \eta Lx \rangle_{\mathbb{R}}.$$

É bastante claro que

$$I(x) = (x_0)^2 - \|\vec{x}\|^2 = [x_0 - \|\vec{x}\|] [x_0 + \|\vec{x}\|], \quad (21.193)$$

onde $\vec{x} \equiv (x_1, x_2, x_3)$ e $\|\vec{x}\| := \sqrt{x_1^2 + x_2^2 + x_3^2}$. Por outro lado,

$$J(x) = (L^T \eta L)_{00} (x_0)^2 + a(\vec{x})x_0 + b(\vec{x}), \quad (21.194)$$

onde⁵¹

$$a(\vec{x}) := 2 \sum_{a=1}^3 (L^T \eta L)_{0a} x_a \quad \text{e} \quad b(\vec{x}) := \sum_{a,b=1}^3 (L^T \eta L)_{ab} x_a x_b. \quad (21.195)$$

A partir daqui há dois casos a se considerar⁵²: $(L^T \eta L)_{00} = 0$ e $(L^T \eta L)_{00} \neq 0$.

I. Caso $(L^T \eta L)_{00} = 0$. Sabemos por (21.188)-(21.189) (tomando $y = 0$) que se L preserva intervalos tipo luz, então se tivermos $I(x) = 0$ para algum $x \in \mathbb{R}^4$, valerá também $J(x) = 0$. A condição $I(x) = 0$ se dá caso $x_0 = \|\vec{x}\|$ e se $x_0 = -\|\vec{x}\|$. Nesses casos a condição $J(x) = 0$ significa,

$$a(\vec{x})\|\vec{x}\| + b(\vec{x}) = 0 \quad \text{e} \quad -a(\vec{x})\|\vec{x}\| + b(\vec{x}) = 0,$$

respectivamente. Subtraindo-se a segunda relação da primeira, obtemos que $a(\vec{x})\|\vec{x}\| = 0$, o que implica $a(\vec{x}) = 0$ e $b(\vec{x}) = 0$, sendo \vec{x} arbitrário. Segue de (21.195) que $(L^T \eta L)_{0a} = 0$ para cada $a = 1, 2, 3$. Ora, junto com a hipótese $(L^T \eta L)_{00} = 0$, isso estabeleceu que $(L^T \eta L)_{0a} = 0$ para cada $a = 0, 1, 2, 3$, ou seja, que a primeira linha da matriz $L^T \eta L$ é nula, implicando que $0 = \det(L^T \eta L) = -\det(L)^2$. Isso provou que L não possui inversa, estabelecendo, por contradição com as hipóteses, que o caso $(L^T \eta L)_{00} = 0$ não é possível.

II. Caso $(L^T \eta L)_{00} \neq 0$. Neste caso $J(x)$ é um polinômio de grau dois em x_0 e, devido a (21.194), podemos escrever

$$J(x) = (L^T \eta L)_{00} [x_0 - y_1(\vec{x})] [x_0 - y_2(\vec{x})],$$

sendo $y_1(\vec{x})$ e $y_2(\vec{x})$ tais que

$$-(L^T \eta L)_{00} (y_1(\vec{x}) + y_2(\vec{x})) = a(\vec{x}) \quad \text{e} \quad (L^T \eta L)_{00} y_1(\vec{x})y_2(\vec{x}) = b(\vec{x}).$$

As expressões $y_1(\vec{x})$ e $y_2(\vec{x})$, porém, podem ser determinadas independentemente dessas relações, por serem as raízes de $J(x)$ como função de x_0 . Sabemos por (21.188)-(21.189) (tomando $y = 0$) que se L preserva intervalos tipo luz, então se tivermos $I(x) = 0$ para algum $x \in \mathbb{R}^4$, valerá também $J(x) = 0$. Para \vec{x} fixo qualquer, vemos por (21.193) e (21.194) que

⁵¹Aqui usou-se que $(L^T \eta L)_{0a} = (L^T \eta L)_{a0}$ pois $L^T \eta L$ é simétrica, ou seja, $(L^T \eta L)^T = L^T \eta L$.

⁵²Agradecemos a Bruno Hideki Kimura por apontar-nos a existência de dois casos (e não de apenas um) e pelo tratamento do primeiro.

tanto $I(x)$ quanto $J(x)$ são polinômios de segundo grau em x_0 e, pelo que acabamos de comentar, têm os mesmos zeros. Dessa forma, também por (21.193) e (21.194), podemos escolher $y_1(\vec{x}) = \|\vec{x}\|$ e $y_2(\vec{x}) = -\|\vec{x}\|$.

Com isso e com (21.193), teremos que

$$J(x) = (L^T \eta L)_{00} (x_0 - \|\vec{x}\|)(x_0 + \|\vec{x}\|) = (L^T \eta L)_{00} I(x)$$

para todo $x \in \mathbb{R}^4$. Pela definição de $I(x)$ e $J(x)$, segue disso que

$$\langle Lx, \eta Lx \rangle_{\mathbb{R}} = (L^T \eta L)_{00} \langle x, \eta x \rangle_{\mathbb{R}} \tag{21.196}$$

para todo $x \in \mathbb{R}^4$, ou seja,

$$\langle x, [L^T \eta L - (L^T \eta L)_{00} \eta] x \rangle_{\mathbb{R}} = 0$$

para todo $x \in \mathbb{R}^4$. Como $L^T \eta L - (L^T \eta L)_{00} \eta$ é uma matriz simétrica (verifique!), a Proposição 3.4, página 236, implica $L^T \eta L - (L^T \eta L)_{00} \eta = 0$. Como $\eta^2 = \mathbb{1}$, segue que

$$\eta L^T \eta L = (L^T \eta L)_{00} \mathbb{1}. \tag{21.197}$$

Como $\det(\eta) = -1$ e $\det(L) = \det(L^T)$, obtemos ao tomar o determinante de ambos os lados da igualdade acima que

$$\det(L)^2 = \left[(L^T \eta L)_{00} \right]^4,$$

de onde extraímos que

$$(L^T \eta L)_{00} = \pm |\det(L)|^{1/2}. \tag{21.198}$$

Com (21.197), isso prova (21.190).

Inserindo (21.198) em (21.196) teríamos $\langle Lx, \eta Lx \rangle_{\mathbb{R}} = \pm |\det(L)|^{1/2} \langle x, \eta x \rangle_{\mathbb{R}}$ para todo $x \in \mathbb{R}^4$. Portanto, se L satisfaz (21.191), então L é da forma $L = \lambda L_0$ com $\lambda \in \mathbb{R}_+$ e $L_0 \in \mathcal{L}$. De fato, se $L \neq 0$ satisfaz (21.191), teremos para qualquer $\lambda > 0$ que $\eta(\lambda^{-1}L)^T \eta(\lambda^{-1}L) = \lambda^{-2} |\det(L)|^{1/2} \mathbb{1}$ e, escolhendo $\lambda = |\det(L)|^{1/4}$, concluímos que $\lambda^{-1}L \in \mathcal{L}$.

Assim, se L satisfaz (21.191), L é produto de uma transformação de \mathcal{L} com uma transformação $D(\lambda) = \lambda \mathbb{1}$, $\lambda \in \mathbb{R}_+$, $\lambda \neq 0$. Se L não envolve dilatações, então $L \in \mathcal{L}$. Isso prova (21.192). ■

• **Comentário sobre a linearidade**

Comentário. Como vimos, as transformações lineares de \mathcal{L} têm a propriedade de manterem invariantes os intervalos entre eventos, ou seja, são tais que $I(Lx, Ly) = I(x, y)$ para todos $x, y \in \mathbb{R}^4$. Podemos nos colocar a questão: qual a forma mais geral de uma transformação sobrejetora $A : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ que tenha a mesma propriedade, ou seja, que seja tal que $I(A(x), A(y)) = I(x, y)$ para todos $x, y \in \mathbb{R}^4$, sem que tenhamos de assumir *a priori* que A seja linear? A resposta é dada por uma simples aplicação do Lema 21.3 e do Teorema 21.7, páginas 1034 e 1035, respectivamente.

De fato, notando-se que $\omega(x, y) := \langle x, \eta y \rangle_{\mathbb{R}}$ define uma forma bilinear simétrica e não degenerada em \mathbb{R}^4 , concluímos do Teorema 21.7 que existem um vetor constante c e um operador linear $L : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ satisfazendo $\langle Lx, \eta Ly \rangle_{\mathbb{R}} = \langle x, \eta y \rangle_{\mathbb{R}}$ para todos $x, y \in \mathbb{R}^4$, tais que vale para todo $x \in \mathbb{R}^4$ a relação $A(x) = Lx + c$.

Assim, concluímos que a transformação mais geral no espaço-tempo que preserva os intervalos entre eventos é uma combinação de uma transformação de Lorentz com uma translação. Como discutiremos mais adiante, essas transformações formam um grupo, denominado *grupo de Poincaré*, ou *grupo de Lorentz não homogêneo*.

Vale observar ainda que a condição de se preservar os intervalos entre dois eventos quaisquer x e y é uma condição mais forte que a de simplesmente preservar a estrutura causal, pois essa última consiste apenas em preservar o sinal de $I(x, y)$, mas não necessariamente o valor de $I(x, y)$. ♣

* * *

Como vemos, um papel especial é desempenhado pelas matrizes de \mathcal{L} . Por toda nossa discussão, tais matrizes representam as transformações entre sistemas de referência inerciais que respeitam a imposição física de preservar a estrutura causal e ignoram dilatações e translações. Daqui por diante vamos nos concentrar exclusivamente em tais transformações. Como veremos, o conjunto \mathcal{L} , introduzido acima, tem a estrutura de um grupo, um fato de grande importância. Trata-se do chamado *grupo de Lorentz*, um objeto de importância central na teoria da relatividade especial.

21.7.3 O Grupo de Lorentz

O Teorema 21.10, acima, diz-nos que se $L \in \text{Mat}(\mathbb{R}, 4)$ representa uma transformação linear entre sistemas de referência inerciais que preserva a estrutura causal e não envolve dilatações e translações, então $\eta L^T \eta L = \mathbb{1}$, o que equivale a dizer que $L^{-1} = \eta L^T \eta$. Isso também equivale a dizer que

$$\langle Lx, \eta Ly \rangle_{\mathbb{R}} = \langle x, \eta y \rangle_{\mathbb{R}}$$

para todos $x, y \in \mathbb{R}^4$. Esse fato e a particular forma da matriz η mostram que o conjunto de tais matrizes L coincide com o **grupo** $O(1, 3)$, que previamente definimos (vide página 1039).

Devido à sua grande importância na Física Relativística, o grupo $O(1, 3)$ recebe denominação especial, a saber, é denominado *grupo de Lorentz*⁵³, em honra ao grande físico holandês, pioneiro nos estudos da teoria da relatividade. O grupo $O(1, 3)$ é também denotado pelo símbolo \mathcal{L} . Os elementos de \mathcal{L} são denominados *transformações de Lorentz*.

Equivalentemente, o grupo de Lorentz $\mathcal{L} = O(1, 3)$ é o grupo de todas as matrizes 4×4 que satisfazem

$$L^{-1} = \eta L^T \eta. \tag{21.199}$$

Como todo elemento L do grupo de Lorentz satisfaz $L\eta L^T \eta = \mathbb{1}$, tem-se $\det(L\eta L^T \eta) = 1$, ou seja, $\det(L)^2 = 1$, pois $\det(L\eta L^T \eta) = \det(L) \det(\eta)^2 \det(L^T)$, $\det(\eta) = -1$ e $\det(L) = \det(L^T)$. Assim, $\det(L) = \pm 1$. O subconjunto $SO(1, 3)$ de $O(1, 3)$, formado pelas matrizes L que satisfazem $\det(L) = +1$ é um subgrupo, denotado por \mathcal{L}_+ .

A seguinte proposição sobre o grupo de Lorentz será usada adiante:

Proposição 21.26 *Se $L \in \mathcal{L}$, então $L^T \in \mathcal{L}$.* □

Prova. Sabemos que para qualquer matriz M vale $(M^T)^T = M$ e que para qualquer matriz inversível M vale $(M^T)^{-1} = (M^{-1})^T$. Se $L \in \mathcal{L}$, tem-se, por definição, que $L^{-1} = \eta L^T \eta$. Assim, como $\eta^T = \eta$, segue que $(L^{-1})^T = \eta L \eta$, ou seja, $(L^T)^{-1} = \eta (L^T)^T \eta$, significando que $L^T \in \mathcal{L}$, que é o que se queria provar. ■

• O grupo de Poincaré, ou grupo de Lorentz não homogêneo

Retornemos brevemente às transformações afins gerais que preservam intervalos e que, como vimos, são da forma $Lx + t$, com $t \in \mathbb{R}^4$ sendo uma translação e $L \in \mathcal{L}$. A composição de duas de tais transformações $L'x + t'$ e $Lx + t$, é a transformação $L'(Lx + t) + t' = L'Lx + L't + t'$.

Essa última expressão naturalmente conduz ao seguinte. Seja $\mathcal{P} := \mathcal{L} \times \mathbb{R}^4$ o conjunto de todos os pares ordenados (L, t) com $L \in \mathcal{L}$ e $t \in \mathbb{R}^4$. Então, \mathcal{P} é um grupo com o produto definido por

$$(L', t') \cdot (L, t) := (L'L, L't + t').$$

Como se vê, esse produto faz de \mathcal{P} o produto semidireto $\mathcal{L} \circledast \mathbb{R}^4$. O produto semidireto de grupos foi definido na Seção 2.2.4.2, página 145.

E. 21.88 Exercício. Verifique que o produto acima é de fato associativo. Identifique o elemento neutro e determine a inversa de cada par $(L, t) \in \mathcal{P}$. ✦

Esse grupo, que combina transformações de Lorentz e translações, é denominado *grupo de Poincaré*⁵⁴ (ou também *grupo de Lorentz não homogêneo*) em homenagem ao eminente matemático francês que também foi um dos pioneiros da teoria da relatividade⁵⁵. O grupo de Poincaré é o grupo mais geral de transformações afins do espaço-tempo que mantêm os intervalos invariantes.

Mais adiante (página 1130) vamos retornar ao grupo de Poincaré para analisar sua estrutura enquanto grupo de Lie. Antes, porém, precisamos nos concentrar plenamente no grupo de Lorentz.

⁵³Hendrik Antoon Lorentz (1853–1928).

⁵⁴Jules Henri Poincaré (1854–1912).

⁵⁵Vários historiadores da Ciência apontaram para o fato que Poincaré, assim como Lorentz, antecedeu Einstein em alguns aspectos. Poincaré foi o primeiro (em 1905, o ano da publicação do trabalho seminal de Einstein, mas independente deste) a estudar o caráter de grupo das transformações de Lorentz, tendo provado que toda transformação de Lorentz é combinação de rotações com um “boost”, fato que estabeleceremos no Teorema 21.11, mais adiante.

21.7.4 Alguns Subgrupos do Grupo de Lorentz

Antes de e com o propósito de estudarmos a estrutura do grupo de Lorentz, vamos identificar alguns de seus subgrupos mais importantes.

• **Troca de paridade e reversão temporal**

As seguintes matrizes são elementos do grupo de Lorentz

$$P_1 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_2 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_3 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (21.200)$$

e

$$P := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad T := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (21.201)$$

E. 21.89 *Exercício importante.* Verifique que as cinco matrizes acima são membros do grupo de Lorentz, ou seja, satisfazem $L\eta L^T\eta = \mathbb{1}$. *

As matrizes P, P_1, P_2 e P_3 implementam *trocas de paridade*, ou seja, reversão da orientação dos eixos de coordenadas espaciais de pontos de \mathbb{R}^4 . A matriz T implementa uma *reversão temporal*, ou seja, inversão da coordenada temporal de pontos de \mathbb{R}^4 .

É bastante evidente que $(T)^2 = (P)^2 = (P_1)^2 = (P_2)^2 = (P_3)^2 = \mathbb{1}$ e que $P = P_1 P_2 P_3$. As matrizes $\mathbb{1}, T, P_1, P_2, P_3$ geram um subgrupo do grupo de Lorentz que implementa reversões temporais e de paridade.

• **Os subgrupos Rot e SRot**

Se R é uma matriz 4×4 da forma

$$R := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{} \\ 0 & & & \\ 0 & & & \end{pmatrix},$$

onde r_0 é uma matriz 3×3 pertencente a $O(3)$, então é fácil verificar que R é um elemento do grupo de Lorentz, ou seja, satisfaz $R\eta R^T\eta = \mathbb{1}$.

E. 21.90 *Exercício.* Verifique isso, usando os fatos que $r_0 r_0^T = \mathbb{1}$ e que

$$R^T := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{(r_0)^T} \\ 0 & & & \\ 0 & & & \end{pmatrix} = R^{-1}.$$

*

É fácil constatar que o conjunto das matrizes da forma de R acima forma um subgrupo do grupo de Lorentz. Esse subgrupo será designado aqui⁵⁶ por Rot.

E. 21.91 *Exercício.* Mostre que Rot é isomorfo ao grupo $O(3)$: $\text{Rot} \simeq O(3)$.

*

Se R é da forma acima, é evidente também que $\det(R) = \det(r_0)$. Logo, Rot tem um subgrupo SRot de matrizes R com $\det(R) = 1$ da forma

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{r_0} \\ 0 & & & \\ 0 & & & \end{pmatrix},$$

onde r_0 é uma matriz 3×3 pertencente a $SO(3)$.

E. 21.92 *Exercício.* Mostre que SRot é isomorfo ao grupo $SO(3)$: $\text{SRot} \simeq SO(3)$.

*

E. 21.93 *Exercício.* Mostre que se $R \in \text{Rot}$ mas $R \notin \text{SRot}$, então existe uma matriz $R' \in \text{SRot}$ com $R = PR'$.

*

E. 21.94 *Exercício.* Mostre que se $R \in \text{Rot}$ mas $R \notin \text{SRot}$, então existe uma matriz $R'' \in \text{SRot}$ com $R = P_1 R''$.

*

As matrizes de SRot implementam rotações puras (sem troca de paridade) nas coordenadas espaciais de \mathbb{R}^4 .

• **Os “boosts” de Lorentz**

Um conjunto muito importante de matrizes de Lorentz é formado pelos chamados “boosts”⁵⁷ de Lorentz na direção 1. Tais matrizes são da forma

$$B_1(v) := \begin{pmatrix} \gamma(v) & -v\gamma(v) & 0 & 0 \\ -v\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{21.202}$$

⁵⁶Essa notação não é uniforme na literatura.

⁵⁷Do inglês *to boost*: impulsionar, propelar, impelir, empurrar. Esse vocábulo é geralmente usado em Física para denominar transformações entre sistemas de coordenadas inerciais que envolvam apenas mudanças de velocidades.

onde

$$\gamma(v) := \frac{1}{\sqrt{1-v^2}}$$

e $v \in (-1, 1)$ (lembrar que adotamos a convenção $c = 1$). O estudante deve aperceber-se que as matrizes $B_1(v)$ são o equivalente em $3 + 1$ dimensões dos “boosts” de Lorentz (21.88), em $1 + 1$ dimensões, que encontramos em nosso estudo do grupo $O(1, 1)$ (vide página 1051 e seguintes).

E. 21.95 Exercício muito importante. Verifique que as matrizes $B_1(v)$ acima são membros do grupo de Lorentz, ou seja, satisfazem $B_1(v)\eta B_1(v)^T \eta = \mathbb{1}$ para todo $v \in (-1, 1)$. *

A ação das matrizes $B_1(v)$ sobre um quadrivetor $x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$ é

$$\begin{aligned} (x^0)' &= \gamma(v)(x^0 - vx^1), \\ (x^1)' &= \gamma(v)(x^1 - vx^0), \\ (x^2)' &= x^2, \\ (x^3)' &= x^3, \end{aligned} \tag{21.203}$$

que são as transformações de Lorentz na forma como são usualmente apresentadas em textos de Física.

Outro fato de grande importância é o seguinte: o conjunto de todas as matrizes $B_1(v)$ com $v \in (-1, 1)$ forma um subgrupo do grupo de Lorentz, denominado *subgrupo dos boosts de Lorentz* (na direção 1) e que designaremos aqui por B_1 . Isso decorre do seguinte:

1. Para $v = 0$

$$B_1(0) = \mathbb{1}.$$

2. Para todo $v \in (-1, 1)$

$$B_1(v)^{-1} = B_1(-v).$$

3. Para todos $v, v' \in (-1, 1)$

$$B_1(v')B_1(v) = B_1\left(\frac{v'+v}{1+v'v}\right). \tag{21.204}$$

E. 21.96 Exercício muito importante. Verifique essas três afirmações. *

Observe-se que o item 3, acima, está intimamente associado à regra relativista de composição de velocidades.

Segue também de (21.204) que B_1 é um subgrupo Abeliano: $B_1(v')B_1(v) = B_1(v)B_1(v')$ para todos $v', v \in (-1, 1)$.

E. 21.97 Exercício. Mostre que $\det(B_1(v)) = 1$ para todo $v \in (-1, 1)$ e, portanto, $B_1 \subset \text{SO}(1, 3)$. *

De forma geral, um *boost* de Lorentz é um elemento do grupo de Lorentz que seja da forma $RB_1(v)R^{-1}$ para algum $v \in (-1, 1)$ e algum $R \in \text{SRot}$. O conjunto de todos os *boosts* de Lorentz é

$$B := \left\{ RB_1(v)R^{-1}, v \in (-1, 1) \text{ e } R \in \text{SRot} \right\}. \tag{21.205}$$

Como discutiremos mais adiante, esse conjunto B **não** é um subgrupo do grupo de Lorentz. No entanto, para cada $R \in \text{SRot}$ fixo, o conjunto

$$B_R := \left\{ RB_1(v)R^{-1}, v \in (-1, 1) \right\}$$

é um subgrupo do grupo de Lorentz que contém *boosts* em uma determinada direção, fixada por R . De fato, para cada $R \in \text{SRot}$ fixo, o produto $\left(RB_1(v)R^{-1}\right)\left(RB_1(v')R^{-1}\right) = RB_1\left(\frac{v+v'}{1+vv'}\right)R^{-1}$ é novamente um elemento de B_R e as demais propriedades de um grupo são trivialmente verificáveis.

Analogamente aos *boosts* de Lorentz na direção 1, há os *boosts* de Lorentz nas direções 2 e 3, representados por matrizes como

$$B_2(v) := \begin{pmatrix} \gamma(v) & 0 & -v\gamma(v) & 0 \\ 0 & 1 & 0 & 0 \\ -v\gamma(v) & 0 & \gamma(v) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad B_3(v) := \begin{pmatrix} \gamma(v) & 0 & 0 & -v\gamma(v) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v\gamma(v) & 0 & 0 & \gamma(v) \end{pmatrix}. \quad (21.206)$$

Todas as afirmações feitas sobre as matrizes B_1 têm seu correspondente análogo para as matrizes B_2 e B_3 . Os respectivos subgrupos são aqui denotados por B_2 e B_3 .

Geometricamente as matrizes $B_2(v)$ e $B_1(v)$ estão relacionadas por uma matriz de rotação de SRot que implementa uma rotação de $\pi/2$ em torno do eixo 3:

$$B_2(v) = RB_1(v)R^{-1},$$

onde

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \text{SRot}.$$

E. 21.98 *Exercício.* Verifique! ✱

Analogamente, é possível obter a matriz $B_3(v)$ a partir de $B_1(v)$ ou de $B_2(v)$ através de rotações.

E. 21.99 *Exercício importante.* Uma propriedade de significado profundo na Teoria da Relatividade é o fato que *boosts* de Lorentz em direções distintas não comutam! Mostre explicitamente, por exemplo, que $B_1(v)B_2(v') \neq B_2(v')B_1(v)$, exceto se $v = 0$ ou $v' = 0$.

Esse é um fenômeno puramente relativístico: *boosts* de Galilei em direções distintas comutam! Verifique! ✱

Adiante, em nosso estudo da estrutura geral do grupo de Lorentz, mostraremos o quão importantes os *boosts* de Lorentz são. A saber, mostraremos que toda matriz de Lorentz é obtida por uma sucessão de uma rotação, um *boost* (na direção 1, por exemplo) e eventualmente uma outra rotação. Eventualmente, trocas de paridade e inversões temporais podem ocorrer também. A afirmação precisa está no Teorema 21.11.

21.7.5 Alguns Fatos Sobre a Estrutura do Grupo de Lorentz

Antes de iniciarmos esta seção, sugerimos ao leitor apreciar o estudo do grupo $O(1, 1)$ iniciado à página 1051.

Vamos aqui tentar caracterizar a forma geral de um elemento do grupo de Lorentz $O(1, 3)$. Como já observamos, $O(1, 3)$ possui um subgrupo $\text{SRot} \simeq \text{SO}(3)$ formado por matrizes da forma

$$R := \begin{pmatrix} 1 & & & \\ & \boxed{\begin{matrix} 0 & 0 & 0 \\ & r_0 \end{matrix}} & & \\ & & & \\ & & & \end{pmatrix},$$

onde r_0 é uma matriz 3×3 pertencente a $SO(3)$.

Vamos no que segue demonstrar o seguinte teorema, que nos fornece a forma geral de toda matriz $L \in \mathcal{L}$ e que é de importância em todo estudo detalhado do grupo de Lorentz.

Teorema 21.11 *Seja L um elemento do grupo de Lorentz $O(1, 3)$. Então, vale uma das quatro afirmações seguintes:*

Ia. $\det(L) = +1$, $L_{00} \geq +1$ e L é da forma

$$L = R_a B_1(v) R_b, \tag{21.207}$$

para algum $v \in (-1, 1)$ e para $R_a, R_b \in \text{SRot}$.

Ib. $\det(L) = +1$, $L_{00} \leq -1$ e L é da forma

$$L = T P R_a B_1(v) R_b, \tag{21.208}$$

para algum $v \in (-1, 1)$ e para $R_a, R_b \in \text{SRot}$.

IIa. $\det(L) = -1$, $L_{00} \leq -1$ e L é da forma

$$L = T R_a B_1(v) R_b, \tag{21.209}$$

para algum $v \in (-1, 1)$ e para $R_a, R_b \in \text{SRot}$.

IIb. $\det(L) = -1$, $L_{00} \geq +1$ e L é da forma

$$L = P R_a B_1(v) R_b, \tag{21.210}$$

para algum $v \in (-1, 1)$ e para $R_a, R_b \in \text{SRot}$.

Fazemos notar que as representações (21.207)–(21.210) não são unívocas. □

A demonstração detalhada deste teorema encontra-se na Seção 21.B, página 1151.

• Dois resultados sobre o grupo de Lorentz

Proposição 21.27 *Se L é um elemento do grupo de Lorentz $O(1, 3)$ e L^{-1} é sua inversa, então tem-se que $(L^{-1})_{00} = L_{00}$.* □

Prova. A prova é simples, pois sabemos que $L^{-1} = \eta L^T \eta$. Então, usando-se a representação (21.B.10) e calculando-se explicitamente, tem-se

$$\begin{aligned} L^{-1} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{} \\ 0 & & -\mathbb{1} & \\ 0 & & & \boxed{} \end{pmatrix} \begin{pmatrix} L_{00} & \boxed{a^T} \\ \boxed{b} & \boxed{l^T} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{} \\ 0 & & -\mathbb{1} & \\ 0 & & & \boxed{} \end{pmatrix} \\ &= \begin{pmatrix} L_{00} & \boxed{-a^T} \\ \boxed{-b} & \boxed{l^T} \end{pmatrix}, \end{aligned}$$

o que leva à constatação que $(L^{-1})_{00} = L_{00}$. ■

Proposição 21.28 *Se L e L' são dois elementos quaisquer do grupo de Lorentz $O(1, 3)$, então tem-se que*

$$\text{sinal}((LL')_{00}) = \text{sinal}(L_{00}) \text{sinal}(L'_{00}).$$

□

Prova. Sejam L e L' duas transformações de Lorentz que, como em (21.B.10), representamos na forma de blocos

$$L = \begin{pmatrix} L_{00} & b^T \\ a & l \end{pmatrix}, \quad L' = \begin{pmatrix} L'_{00} & b'^T \\ a' & l' \end{pmatrix}. \quad (21.211)$$

Vamos formar o produto $L'' = LL'$ e estudar o sinal do elemento L''_{00} da matriz resultante. Pela regra de produto de matrizes teremos (verifique!) $L''_{00} = L_{00}L'_{00} + b^T a'$. O produto de matrizes $b^T a'$ é idêntico ao produto escalar $b \cdot a'$ dos vetores b e a' de \mathbb{R}^3 (por quê?). Assim,

$$L''_{00} = L_{00}L'_{00} + b \cdot a'. \quad (21.212)$$

Há dois casos a considerar: o caso em que $\text{sinal}(L_{00}) = \text{sinal}(L'_{00})$ e o caso em que $\text{sinal}(L_{00}) \neq \text{sinal}(L'_{00})$.

1. Caso em que $\text{sinal}(L_{00}) = \text{sinal}(L'_{00})$.

Nessa situação tem-se por (21.212) que

$$L''_{00} \geq L_{00}L'_{00} - |b \cdot a'|.$$

Sabemos que $b \cdot a' = \|b\| \|a'\| \cos \theta$, onde $\|b\|$ é o comprimento de b , $\|a'\|$ é o comprimento de a' e θ é o ângulo que esses dois vetores formam entre si. É óbvio, portanto, que $|b \cdot a'| \leq \|b\| \|a'\|$ (desigualdade de Cauchy). Assim,

$$L''_{00} \geq L_{00}L'_{00} - \|b\| \|a'\|. \quad (21.213)$$

Pela Proposição 21.36, página 1155, valem $\|b\| = \|a\|$ e $\|b'\| = \|a'\|$, assim como $L_{00} = \pm\sqrt{1 + \|a\|^2}$ e $L'_{00} = \pm\sqrt{1 + \|a'\|^2}$. Assim, por (21.213),

$$L''_{00} \geq \sqrt{1 + \|a\|^2} \sqrt{1 + \|a'\|^2} - \|a\| \|a'\| > 0.$$

Portanto, $\text{sinal}(L''_{00}) = +1 = \text{sinal}(L_{00}) \text{sinal}(L'_{00})$, como queríamos provar.

2. Caso em que $\text{sinal}(L_{00}) \neq \text{sinal}(L'_{00})$.

Por (21.212) tem-se

$$L''_{00} \leq L_{00}L'_{00} + |b \cdot a'|.$$

Sabemos que $b \cdot a' = \|b\| \|a'\| \cos \theta$, onde $\|b\|$ é o comprimento de b , $\|a'\|$ é o comprimento de a' e θ é o ângulo que esses dois vetores formam entre si. É óbvio, portanto, que $|b \cdot a'| \leq \|b\| \|a'\|$ (desigualdade de Cauchy). Assim,

$$L''_{00} \leq L_{00}L'_{00} + \|b\| \|a'\|. \quad (21.214)$$

Pela Proposição 21.36, valem $\|b\| = \|a\|$ e $\|b'\| = \|a'\|$, assim como $L_{00} = \pm\sqrt{1 + \|a\|^2}$ e $L'_{00} = \mp\sqrt{1 + \|a'\|^2}$ (pois $\text{sinal}(L_{00}) \neq \text{sinal}(L'_{00})$). Assim, por (21.214),

$$L''_{00} \leq -\sqrt{1 + \|a\|^2} \sqrt{1 + \|a'\|^2} + \|a\| \|a'\| < 0.$$

Portanto, $\text{sinal}(L''_{00}) = -1 = \text{sinal}(L_{00}) \text{sinal}(L'_{00})$, como queríamos provar. ■

• **Os subgrupos próprio, ortócrono e restrito do grupo de Lorentz**

Os conjuntos de transformações de Lorentz que satisfazem as condições Ia, Ib, IIa ou IIb acima são obviamente conjuntos disjuntos. Não é difícil mostrar (mas não o faremos aqui) que cada um é um conjunto conexo. Portanto, o grupo de Lorentz $\mathcal{L} = O(1, 3)$ possui quatro componentes conexas. Seguindo a convenção, denotaremos essas quatro componentes da seguinte forma:

1. $\mathcal{L}_+^\uparrow := \{L \in \mathcal{L} \mid \det(L) = +1 \text{ e sinal}(L_{00}) = +1\}$,
2. $\mathcal{L}_-^\uparrow := \{L \in \mathcal{L} \mid \det(L) = -1 \text{ e sinal}(L_{00}) = +1\}$,
3. $\mathcal{L}_+^\downarrow := \{L \in \mathcal{L} \mid \det(L) = +1 \text{ e sinal}(L_{00}) = -1\}$,
4. $\mathcal{L}_-^\downarrow := \{L \in \mathcal{L} \mid \det(L) = -1 \text{ e sinal}(L_{00}) = -1\}$.

Note-se também que apenas \mathcal{L}_+^\uparrow contém a identidade $\mathbb{1}$. \mathcal{L}_-^\uparrow contém a operação de troca de paridade P . \mathcal{L}_+^\downarrow contém a operação de troca de paridade e inversão temporal PT . \mathcal{L}_-^\downarrow contém a operação de inversão temporal T .

Os conjuntos \mathcal{L}_-^\uparrow , \mathcal{L}_+^\downarrow e \mathcal{L}_-^\downarrow não são subgrupos de \mathcal{L} . Porém, pelas Proposições 21.27 e 21.28, é muito fácil constatar as seguintes afirmações:

1. \mathcal{L}_+^\uparrow é um subgrupo de \mathcal{L} , denominado *grupo de Lorentz próprio ortócrono* ou *grupo de Lorentz restrito*.
2. $\mathcal{L}^\uparrow := \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$ é um subgrupo de \mathcal{L} , denominado *grupo de Lorentz ortócrono*.
3. $\mathcal{L}_+ := \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow$ é um subgrupo de \mathcal{L} , denominado *grupo de Lorentz próprio*.
4. $\mathcal{L}_0 := \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\downarrow$ é um subgrupo de \mathcal{L} , denominado *grupo de Lorentz ortócoro*⁵⁸.

No grupo de Lorentz ortócrono $\mathcal{L}^\uparrow := \mathcal{L}_+^\uparrow \cup \mathcal{L}_-^\uparrow$ não ocorrem reversões temporais⁵⁹. Apenas os elementos de ambos os conjuntos \mathcal{L}_+^\uparrow e \mathcal{L}_-^\uparrow satisfazem $\det(L) = 1$ e, portanto, o grupo de Lorentz próprio $\mathcal{L}_+ := \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow$ coincide com $SO(1, 3)$. O grupo de Lorentz próprio ortócrono \mathcal{L}_+^\uparrow é também por vezes denotado por $SO_0(1, 3)$, pois o subíndice “0” é amiúde usado na literatura matemática para designar a componente conexa de um grupo (no caso, do grupo $SO(1, 3)$) que contém a identidade. Note também que $SRot$ é um subgrupo de \mathcal{L}_+^\uparrow .

• **A importância de \mathcal{L}_+ , \mathcal{L}^\uparrow e \mathcal{L}_+^\uparrow na Física**

É uma crença da Física atual que \mathcal{L}_+^\uparrow representa uma simetria da Natureza (na ausência de campos gravitacionais). Essa crença não se estende aos grupos \mathcal{L}_+ e \mathcal{L}^\uparrow . O problema com esses últimos grupos é que os mesmos envolvem operações de troca de paridade (representada pela matriz P) ou de reversão temporal (representada pela matriz T).

É um fato bem estabelecido experimentalmente que nas chamadas interações fracas da Física das Partículas Elementares a troca de paridade (representada por matrizes como P ou P_1) não é uma transformação de simetria da natureza⁶⁰

⁵⁸Essa denominação, “ortócoro”, é raramente usada.

⁵⁹Esse é o significado da flecha apontando para cima nos símbolos \mathcal{L}_\pm^\uparrow , indicando que o tempo corre na mesma direção nos sistemas de referência inerciais transformados por \mathcal{L}_\pm^\uparrow . Analogamente, a flecha para baixo nos símbolos $\mathcal{L}_\pm^\downarrow$ indica que a direção temporal dos sistemas de referência inerciais transformados por $\mathcal{L}_\pm^\downarrow$ é invertida.

⁶⁰Essa descoberta foi realizada em experimentos de decaimento de neutrons publicados 1957, realizados pela física Chien-Shiung Wu (1912–1997) e colaboradores, baseados em previsão teórica feita por Tsung-Dao Lee (1926–) e Chen Ning Yang (1922–). Esses dois últimos foram agraciados com o Prêmio Nobel de Física de 1957 “for their penetrating investigation of the so-called parity laws which has led to important discoveries regarding the elementary particles”. A Profa. Chien-Shiung Wu foi agraciada com o Prêmio Wolf de Física de 1978.

A referência ao trabalho original de Lee e Yang é: T. D. Lee and C. N. Yang, “Question of Parity Conservation in Weak Interactions”. Phys. Rev. **104** (1), 254–258 (1956). A referência ao trabalho original de Wu e colaboradores é: C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, “Experimental Test of Parity Conservation in Beta Decay”. Phys. Rev. **105** (4), 1413–1415 (1957). A referência [418] contém uma compilação de várias referências originais desses e outros autores sobre o tema.

No contexto da Teoria Quântica de Campos Relativística, é um fato teórico bem estabelecido⁶¹ que a chamada transformação CPT ⁶² é uma transformação de simetria. Violações dessa simetria não foram empiricamente observadas na Física das Partículas Elementares. Por isso, a descoberta experimental, realizada em 1964, que a simetria CP é violada em certos processos de decaimento de partículas⁶³ indica fortemente que a reversão temporal também não seria uma simetria da natureza. Entretanto, evidências experimentais diretas de que a simetria de reversão temporal é violada não foram ainda encontradas, por serem de difícil constatação. Para mais informações a respeito de simetrias e suas violações na Física das Partículas Elementares, vide por exemplo [240], [147], [302], ou outros livros introdutórios sobre o assunto.

• \mathcal{L}_+^\uparrow é um subgrupo normal de \mathcal{L}

Vamos aqui provar a seguinte proposição sobre \mathcal{L}_+^\uparrow :

Proposição 21.29 \mathcal{L}_+^\uparrow é um subgrupo normal do grupo de Lorentz. □

Prova. Tudo o que temos que fazer é provar que se $L \in \mathcal{L}_+^\uparrow$ e $G \in \mathcal{L}$, então $G^{-1}LG \in \mathcal{L}_+^\uparrow$. Isso equivale a provar que $\det(G^{-1}LG) = 1$ e que $\text{sinal}((G^{-1}LG)_{00}) = 1$.

Como $\det(L) = 1$, tem-se obviamente que

$$\det(G^{-1}LG) = \det(G^{-1})\det(L)\det(G) = \det(G^{-1})\det(G) = \det(G^{-1}G) = \det(\mathbb{1}) = 1.$$

Analogamente, pela Proposição 21.28 vale

$$\begin{aligned} \text{sinal}((G^{-1}LG)_{00}) &= \text{sinal}((G^{-1}L)_{00}) \text{sinal}(G_{00}) = \text{sinal}((G^{-1})_{00}) \text{sinal}(L_{00}) \text{sinal}(G_{00}) \\ &= \text{sinal}((G^{-1})_{00}) \text{sinal}(G_{00}) = (\text{sinal}(G_{00}))^2 = 1. \end{aligned}$$

Acima, usamos a Proposição 21.27 na penúltima igualdade. Isso completa a prova. ■

E. 21.100 *Exercício.* Mostre que o grupo quociente $\mathcal{L}/\mathcal{L}_+^\uparrow$ é isomorfo ao grupo gerado por P_1 e T . ✦

21.7.6 Os Geradores Infinitesimais do Grupo de Lorentz

• Os geradores infinitesimais dos *boosts* de Lorentz

Vamos reparametrizar os *boosts* de Lorentz B_1 , B_2 e B_3 , introduzindo um novo parâmetro $z = \text{arctanh } v$, ou seja, $v = \tanh z$, com $-\infty < z < \infty$. Na literatura física, z é por vezes denominado “*rapidez*”. Definindo $\mathcal{B}_a(z) := B_a(\tanh z)$,

⁶¹Vide, e.g., [372] ou [204].

⁶²A chamada *transformação CPT* envolve as operações sucessivas de troca de carga, ou partícula-antipartícula, (denotada por C), de paridade (denotada por P) e de reversão temporal (denotada por T).

⁶³Esta descoberta, realizada em experimentos datados de 1964 sobre o decaimento de mésons K , rendeu o Prêmio Nobel de Física de 1980 a James Watson Cronin (1931–) e Val Logsdon Fitch (1923–), “*for the discovery of violations of fundamental symmetry principles in the decay of neutral K-mesons*”. A referência [418] contém uma compilação de várias referências originais desses e outros autores sobre o tema.

$a = 1, 2, 3$, temos, explicitamente

$$\mathcal{B}_1(z) := \begin{pmatrix} \cosh z & -\sinh z & 0 & 0 \\ -\sinh z & \cosh z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{B}_2(z) := \begin{pmatrix} \cosh z & 0 & -\sinh z & 0 \\ 0 & 1 & 0 & 0 \\ -\sinh z & 0 & \cosh z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (21.215)$$

$$\mathcal{B}_3(z) := \begin{pmatrix} \cosh z & 0 & 0 & -\sinh z \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh z & 0 & 0 & \cosh z \end{pmatrix}. \quad (21.216)$$

As relações de composição (21.204) ficam

$$\mathcal{B}_a(z)\mathcal{B}_a(z') = \mathcal{B}_a(z+z'), \quad a = 1, 2, 3.$$

E. 21.101 *Exercício.* Mostre isso usando (21.204) e a identidade bem conhecida $\tanh(x+y) = \frac{\tanh(x)+\tanh(y)}{1+\tanh(x)\tanh(y)}$. Alternativamente, use a forma explícita das matrizes $\mathcal{B}_a(z)$ dada acima. ✦

Como $\mathcal{B}_a(0) = \mathbb{1}$, constatamos que $\{\mathcal{B}_a(z), -\infty < z < \infty\}$, $a = 1, 2, 3$, são três subgrupos uniparamétricos do grupo de Lorentz. Seus geradores infinitesimais são definidos por

$$\mathcal{M}_a := \left. \frac{d}{dz} \mathcal{B}_a(z) \right|_{z=0}, \quad a = 1, 2, 3,$$

e são explicitamente dados por

$$\mathcal{M}_1 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_3 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \quad (21.217)$$

É também importante notar que

$$\mathcal{B}_a(z) = \exp(z\mathcal{M}_a)$$

para $a = 1, 2, 3$.

E. 21.102 *Exercício.* Verifique isso usando as formas explícitas dos geradores infinitesimais \mathcal{M}_a dadas acima. ✦

• Os geradores infinitesimais de SRot

Além dos *boosts* de Lorentz, consideremos também os três subgrupos uniparamétricos de SRot dados por

$$\mathcal{R}_1(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \phi & -\text{sen} \phi \\ 0 & 0 & \text{sen} \phi & \cos \phi \end{pmatrix}, \quad \mathcal{R}_2(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & 0 & \text{sen} \phi \\ 0 & 0 & 1 & 0 \\ 0 & -\text{sen} \phi & 0 & \cos \phi \end{pmatrix},$$

$$\mathcal{R}_3(\phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & -\text{sen} \phi & 0 \\ 0 & \text{sen} \phi & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

que representam rotações por um ângulo $\phi \in (-\pi, \pi]$ no sentido anti-horário⁶⁴ em torno dos eixos espaciais 1, 2 e 3, respectivamente. Em completa analogia com o grupo SO(3), seus geradores infinitesimais são

$$\mathcal{J}_a := \left. \frac{d}{d\phi} \mathcal{R}_a(\phi) \right|_{\phi=0}, \quad a = 1, 2, 3.$$

É óbvio que

$$\mathcal{J}_a = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \boxed{J_a} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

onde J_a são os geradores infinitesimais de SO(3) dados em (21.92)-(21.94), página 1055. Explicitamente, tem-se

$$\mathcal{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{21.218}$$

E. 21.103 Exercício muito importante.. Mostre que os geradores infinitesimais, \mathcal{M}_a e \mathcal{J}_b , com $a, b = 1, 2, 3$, satisfazem as seguintes

⁶⁴Daí a diferença de sinal nos senos entre $\mathcal{R}_2(\phi)$ e as outras duas matrizes.

relações de comutação:

$$[\mathcal{J}_a, \mathcal{J}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathcal{J}_c, \tag{21.219}$$

$$[\mathcal{M}_a, \mathcal{M}_b] = -\sum_{c=1}^3 \varepsilon_{abc} \mathcal{J}_c, \tag{21.220}$$

$$[\mathcal{J}_a, \mathcal{M}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathcal{M}_c. \tag{21.221}$$

Mostre também que de (21.219)–(21.221) segue que

$$[\vec{\alpha} \cdot \vec{\mathcal{M}} + \vec{\beta} \cdot \vec{\mathcal{J}}, \vec{\gamma} \cdot \vec{\mathcal{M}} + \vec{\delta} \cdot \vec{\mathcal{J}}] = (\vec{\beta} \times \vec{\gamma} - \vec{\alpha} \times \vec{\delta}) \cdot \vec{\mathcal{M}} + (\vec{\beta} \times \vec{\delta} - \vec{\alpha} \times \vec{\gamma}) \cdot \vec{\mathcal{J}}. \tag{21.222}$$

Acima, $\vec{\alpha}$, $\vec{\beta}$, $\vec{\gamma}$ e $\vec{\delta}$ são vetores em \mathbb{R}^3 e expressões como $\vec{\alpha} \cdot \vec{\mathcal{M}}$ e $\vec{\alpha} \cdot \vec{\mathcal{J}}$ denotam $\vec{\alpha} \cdot \vec{\mathcal{M}} \equiv \sum_{c=1}^3 \alpha_c \mathcal{M}_c$ e $\vec{\alpha} \cdot \vec{\mathcal{J}} \equiv \sum_{c=1}^3 \alpha_c \mathcal{J}_c$. *

É claro de (21.219)–(21.221), e especialmente de (21.222), que o espaço vetorial real composto por todas as combinações lineares reais dos seis geradores infinitesimais \mathcal{M}_a e \mathcal{J}_b formam uma álgebra de Lie real, a *álgebra de Lie do grupo de Lorentz* \mathcal{L}_+^\uparrow , a qual é por vezes denotada por ℓ_+^\uparrow . Sabemos que não há mais geradores infinitesimais independentes pois, como provamos, todo elemento do grupo de Lorentz \mathcal{L}_+^\uparrow é produto de *boosts* e rotações.

De (21.220) percebemos o fato notável que os três geradores infinitesimais dos subgrupos de *boost* por si só **não** formam uma álgebra de Lie! Para tal, é preciso incluir os geradores infinitesimais dos subgrupos de rotação! Isso releva uma relação insuspeita, mas profunda, entre os *boosts* (que fisicamente representam transformações entre sistemas de referência inerciais com velocidades relativas não nulas) e as rotações espaciais, pois indica que as rotações espaciais podem ser geradas a partir de *boosts*. Isso é uma característica especial da Física Relativista (vide a comparação com o grupo de Galilei, abaixo) e está relacionada a alguns fenômenos físicos, como a chamada *precessão de Thomas*⁶⁵, importante na discussão do chamado *fator giromagnético* do elétron. Vide qualquer bom livro sobre Mecânica Quântica Relativista (por ex. [325]).

• **Revisitando o Teorema 21.11. Mais sobre *boosts* e rotações**

Como vimos no Teorema 21.11, página 1113, toda $L \in \mathcal{L}_+^\uparrow$ é da forma $L = R_a B_1(v) R_b$, com $R_a, R_b \in \text{SRot}$. Podemos escrever $L = (R_a B_1(v) R_a^{-1}) R$, onde $R := R_a R_b$. Note-se que $B := R_a B_1(v) R_a^{-1}$ é um *boost* de Lorentz em uma certa direção (determinada por v e por R_a), ou seja, é um elemento do conjunto B definido em (21.205), página 1111. De maneira análoga, podemos escrever $L = R_a B_1(v) R_b = R'(R_b^{-1} B_1(v) R_b) = R' B'$ com $R' = R_a R_b \in \text{SRot}$ e $B' = R_b^{-1} B_1(v) R_b \in B$. Em resumo, temos a

Proposição 21.30 *Toda transformação de Lorentz $L \in \mathcal{L}_+^\uparrow$ é da forma $L = BR = R'B'$, para certos $B, B' \in B$ e $R, R' \in \text{SRot}$. Ou seja, toda transformação de Lorentz $L \in \mathcal{L}_+^\uparrow$ é o produto de uma rotação por um boost de Lorentz.* \square

Escrevendo $v \equiv \tanh z$, ficamos com $B = R_a \mathcal{B}_1(z) R_a^{-1}$ ou, usando o gerador infinitesimal \mathcal{M}_1 , $B = R_a \exp(z \mathcal{M}_1) R_a^{-1}$. Isso, por sua vez pode ser reescrito como $B = R_a \exp(z \mathcal{M}_1) R_a^{-1} = \exp(z R_a \mathcal{M}_1 R_a^{-1})$. Vamos agora escrever R_a na forma $R_a = \exp(J)$, onde $J = \sum_{k=1}^3 \alpha_k \mathcal{J}_k \equiv \vec{\alpha} \cdot \vec{\mathcal{J}}$ para certos α_k 's reais. Pela *série de Lie*, expressão (11.60)–(11.61), página 594, teremos

$$R_a \mathcal{M}_1 R_a^{-1} = \exp(J) \mathcal{M}_1 \exp(-J) = \mathcal{M}_1 + [J, \mathcal{M}_1] + \frac{1}{2!} [J, [J, \mathcal{M}_1]] + \frac{1}{3!} [J, [J, [J, \mathcal{M}_1]]] + \dots,$$

sendo a série do lado direito convergente. O fato importante a notar é que, por (21.221), os comutadores múltiplos $[J, \dots [J, \mathcal{M}_1]]$ são combinações lineares de $\mathcal{M}_1, \mathcal{M}_2$ e \mathcal{M}_3 (para mais detalhes, vide a demonstração da Proposição 21.31, página 1120). Concluímos disso que todos os elementos de B são da forma $\exp(\vec{\gamma} \cdot \vec{\mathcal{M}})$ com $\vec{\gamma} \in \mathbb{R}^3$.

⁶⁵Llewellyn Hilleth Thomas (1903–1992). Thomas é também autor, junto com Fermi (Enrico Fermi, 1901–1954.), do chamado *modelo de Thomas-Fermi*, amplamente empregado na Física Atômica.

Provaremos adiante que a recíproca é igualmente verdadeira: todos os elementos da forma $\exp(\vec{\gamma} \cdot \vec{M})$ com $\vec{\gamma} \in \mathbb{R}^3$ são *boosts* de Lorentz. Mais especificamente, mostraremos (no Corolário 21.7) que um elemento de \mathcal{L}_+^\uparrow que seja da forma $\exp(\vec{\gamma} \cdot \vec{M})$, com $\vec{\gamma} \in \mathbb{R}^3$ não nulo, representa um *boost* de Lorentz com “rapidez” $\|\vec{\gamma}\|$ (e, portanto, com velocidade $\tanh(\|\vec{\gamma}\|)$) na direção definida pelo vetor unitário $\vec{\gamma}/\|\vec{\gamma}\|$.

Os dois resultados a seguir, a Proposição 21.31 e o Corolário 21.7, tornam mais precisa a relação entre *boosts* arbitrários e rotações.

Proposição 21.31 *Seja $r(\theta, \vec{\eta}) \in \text{SO}(3)$ a rotação de θ em torno do vetor unitário $\vec{\eta}$ e seja $R(\theta, \vec{\eta})$ o elemento correspondente de SRot . Seja $\mathcal{B}(\vec{\alpha})$ com $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3$ o boost de Lorentz dado por*

$$\mathcal{B}(\vec{\alpha}) := \exp(\vec{\alpha} \cdot \vec{M}),$$

onde $\vec{\alpha} \cdot \vec{M} := \sum_{k=1}^3 \alpha_k \mathcal{M}_k$, sendo \mathcal{M}_k os geradores infinitesimais dos boosts de \mathcal{L}_+^\uparrow . Então, vale

$$R(\theta, \vec{\eta}) \mathcal{B}(\vec{\alpha}) R(-\theta, \vec{\eta}) = \mathcal{B}(R(\theta, \vec{\eta})\vec{\alpha}). \quad (21.223)$$

□

Prova. Nesta demonstração, seguiremos a convenção de implicitamente somar (de 1 a 3) sobre índices repetidos. Temos,

$$R(\theta, \vec{\eta}) \mathcal{B}(\vec{\alpha}) R(-\theta, \vec{\eta}) = \exp(\theta \vec{\eta} \cdot \vec{\mathcal{J}}) \exp(\vec{\alpha} \cdot \vec{M}) \exp(-\theta \vec{\eta} \cdot \vec{\mathcal{J}}) = \exp\left(\alpha_j \exp(\theta \vec{\eta} \cdot \vec{\mathcal{J}}) \mathcal{M}_j \exp(-\theta \vec{\eta} \cdot \vec{\mathcal{J}})\right).$$

Pela *série de Lie*, expressão (11.60)–(11.61), página 594, temos

$$\exp(\theta \vec{\eta} \cdot \vec{\mathcal{J}}) \mathcal{M}_j \exp(-\theta \vec{\eta} \cdot \vec{\mathcal{J}}) = \mathcal{M}_j + \sum_{a=1}^{\infty} \frac{\theta^a}{a!} \left[\underbrace{[\vec{\eta} \cdot \vec{\mathcal{J}}, [\vec{\eta} \cdot \vec{\mathcal{J}}, \dots, [\vec{\eta} \cdot \vec{\mathcal{J}}, \mathcal{M}_j] \dots]}_a \text{ vezes} \right].$$

Para o primeiro comutador, temos

$$[\vec{\eta} \cdot \vec{\mathcal{J}}, \mathcal{M}_j] = \eta_i [\mathcal{J}_i, \mathcal{M}_j] \stackrel{(21.221)}{=} \eta_i \varepsilon_{ijk} \mathcal{M}_k.$$

Para o segundo comutador, teremos

$$[\vec{\eta} \cdot \vec{\mathcal{J}}, [\vec{\eta} \cdot \vec{\mathcal{J}}, \mathcal{M}_j]] = \eta_l \eta_i \varepsilon_{ijk} [\mathcal{J}_l, \mathcal{M}_k] \stackrel{(21.221)}{=} \eta_l \eta_i \varepsilon_{ijk} \varepsilon_{lkm} \mathcal{M}_m \stackrel{(4.7)}{=} \eta_l \eta_i (\delta_{im} \delta_{jl} - \delta_{il} \delta_{jm}) \mathcal{M}_m = \eta_j (\vec{\eta} \cdot \vec{M}) - \mathcal{M}_j.$$

Assim, o terceiro comutador será igual a $[\vec{\eta} \cdot \vec{\mathcal{J}}, \eta_j (\vec{\eta} \cdot \vec{M}) - \mathcal{M}_j]$. Agora, o primeiro termo é

$$[\vec{\eta} \cdot \vec{\mathcal{J}}, \eta_j (\vec{\eta} \cdot \vec{M})] = \eta_l \eta_j \eta_k [\mathcal{J}_l, \mathcal{M}_k] \stackrel{(21.221)}{=} \eta_l \eta_j \eta_k \varepsilon_{lkn} \mathcal{M}_n = 0,$$

devido à antissimetria de ε_{lkn} . Logo, o terceiro comutador será igual a

$$-[\vec{\eta} \cdot \vec{\mathcal{J}}, \mathcal{M}_j] = -\eta_i \varepsilon_{ijk} \mathcal{M}_k$$

por ser igual ao primeiro comutador, mas com o sinal trocado.

Segue disso uma expressão geral para os comutadores múltiplos:

$$\left[\underbrace{[\vec{\eta} \cdot \vec{\mathcal{J}}, [\vec{\eta} \cdot \vec{\mathcal{J}}, \dots, [\vec{\eta} \cdot \vec{\mathcal{J}}, \mathcal{M}_j] \dots]}_a \text{ vezes} \right] = \begin{cases} (-1)^{a/2} (\eta_j (\vec{\eta} \cdot \vec{M}) - \mathcal{M}_j), & a \text{ par,} \\ (-1)^{(a-1)/2} \eta_i \varepsilon_{ijk} \mathcal{M}_k, & a \text{ ímpar,} \end{cases}$$

a qual pode ser facilmente provada por indução (faça-o!). Assim,

$$\begin{aligned} \exp(\theta \vec{\eta} \cdot \vec{\mathcal{J}}) \mathcal{M}_j \exp(-\theta \vec{\eta} \cdot \vec{\mathcal{J}}) &= \mathcal{M}_j + \left(\sum_{k=0}^{\infty} \frac{(-1)^k \theta^{2k+1}}{(2k+1)!} \right) \eta_i \varepsilon_{ijk} \mathcal{M}_k + \left(\sum_{k=1}^{\infty} \frac{(-1)^k \theta^{2k}}{(2k)!} \right) (\eta_j (\vec{\eta} \cdot \vec{\mathcal{M}}) - \mathcal{M}_j) \\ &= \mathcal{M}_j + \text{sen}(\theta) \eta_i \varepsilon_{ijk} \mathcal{M}_k + (1 - \cos(\theta)) (\eta_j (\vec{\eta} \cdot \vec{\mathcal{M}}) - \mathcal{M}_j) \end{aligned}$$

Logo,

$$\begin{aligned} \exp(\theta \vec{\eta} \cdot \vec{\mathcal{J}}) (\vec{\alpha} \cdot \vec{\mathcal{M}}) \exp(-\theta \vec{\eta} \cdot \vec{\mathcal{J}}) &= \vec{\alpha} \cdot \vec{\mathcal{M}} + \text{sen}(\theta) (\vec{\eta} \times \vec{\alpha}) \cdot \vec{\mathcal{M}} + (1 - \cos(\theta)) ((\vec{\eta} \cdot \vec{\alpha}) (\vec{\eta} \cdot \vec{\mathcal{M}}) - (\vec{\alpha} \cdot \vec{\mathcal{M}})) \\ &= \left(\vec{\alpha} + \text{sen}(\theta) (\vec{\eta} \times \vec{\alpha}) + (1 - \cos(\theta)) ((\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} - \vec{\alpha}) \right) \cdot \vec{\mathcal{M}} \\ &= \left(\cos(\theta) \vec{\alpha} + \text{sen}(\theta) (\vec{\eta} \times \vec{\alpha}) + (1 - \cos(\theta)) (\vec{\eta} \cdot \vec{\alpha}) \vec{\eta} \right) \cdot \vec{\mathcal{M}} \\ &\stackrel{(21.106)}{=} \left(R(\theta, \vec{\eta}) \vec{\alpha} \right) \cdot \vec{\mathcal{M}} \quad , \end{aligned}$$

completando a prova. ■

Da Proposição 21.31, página 1120, e da Proposição 21.8, página 1061, segue a seguinte afirmação, que dispensa demonstração:

Corolário 21.7 *Seja*

$$\mathcal{B}(z, \vec{\xi}) := \exp(z \vec{\xi} \cdot \vec{\mathcal{M}}) \quad ,$$

onde $z \in \mathbb{R}$, $\vec{\xi} = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$ com $\|\vec{\xi}\| = 1$ e $\vec{\xi} \cdot \vec{\mathcal{M}} := \sum_{k=1}^3 \xi_k \mathcal{M}_k$, sendo \mathcal{M}_k os geradores infinitesimais dos boosts de \mathcal{L}_+^\uparrow . Então, existe $R(\theta, \vec{\eta}) \in \text{SRot}$ tal que

$$\mathcal{B}(z, \vec{\xi}) = R(\theta, \vec{\eta}) \mathcal{B}_1(z) R(\theta, \vec{\eta})^{-1} \quad ,$$

onde $\mathcal{B}_1(z)$ é o boost na direção 1 e $R(\theta, \vec{\eta}) \in \text{SRot}$ é o elemento de SRot obtido a partir da rotação $r(\theta, \vec{\eta}) \in \text{SO}(3)$ que leva o vetor unitário na direção 1, \mathbf{e}_1 , no vetor unitário $\vec{\xi}$, ou seja, tal que $r(\theta, \vec{\eta}) \mathbf{e}_1 = \vec{\xi}$.

É claro por isso que um elemento do grupo \mathcal{L}_+^\uparrow que seja da forma $\exp(\vec{\alpha} \cdot \vec{\mathcal{M}})$, com $\vec{\alpha} \in \mathbb{R}^3$ não nulo, representa um boost de Lorentz com “rapidez” $\|\vec{\alpha}\|$ (e, portanto, com velocidade $\tanh(\|\vec{\alpha}\|)$) na direção definida pelo vetor espacial unitário $\vec{\alpha}/\|\vec{\alpha}\|$. □

Dentre as nossas conclusões acima temos também:

Proposição 21.32 *O conjunto de boosts de Lorentz \mathcal{B} coincide com o conjunto das matrizes da forma $\exp\left(\sum_{k=1}^3 \gamma_k \mathcal{M}_k\right)$, com γ_k 's sendo números reais:*

$$\mathcal{B} = \left\{ \exp(\vec{\gamma} \cdot \vec{\mathcal{M}}) \quad , \quad \vec{\gamma} \in \mathbb{R}^3 \right\} .$$

Toda $L \in \mathcal{L}_+^\uparrow$ é da forma $L = \exp(\mathcal{M}) \exp(\mathcal{J})$, onde $\mathcal{J} = \vec{\beta} \cdot \vec{\mathcal{J}}$ e $\mathcal{M} = \vec{\gamma} \cdot \vec{\mathcal{M}}$, sendo que os $\vec{\beta}$'s e $\vec{\gamma}$ são elementos de \mathbb{R}^3 . Alternativamente, podemos escrever $L = \exp(\mathcal{J}') \exp(\mathcal{M}')$, onde $\mathcal{J}' = \vec{\beta}' \cdot \vec{\mathcal{J}}$ e $\mathcal{M}' = \vec{\gamma}' \cdot \vec{\mathcal{M}}$, sendo também $\vec{\beta}'$ e $\vec{\gamma}' \in \mathbb{R}^3$. □

Um outro reconhecimento importante é o de que as matrizes \mathcal{M}_k são simétricas, ou seja, satisfazem $\mathcal{M}_k^T = \mathcal{M}_k$, enquanto que as matrizes \mathcal{J}_k são antissimétricas. Daí segue que as matrizes $\exp(\vec{\alpha} \cdot \vec{\mathcal{M}})$, $\vec{\alpha} \in \mathbb{R}^3$, que implementam boosts de Lorentz, são simétricas, enquanto que as matrizes $\exp(\vec{\beta} \cdot \vec{\mathcal{J}})$, $\vec{\beta} \in \mathbb{R}^3$, $\|\vec{\beta}\| \leq \pi$, que implementam rotações, são matrizes ortogonais.

21.7.6.1 Fórmula de Rodrigues para *Boosts* de Lorentz

Podemos obter para *boosts* de Lorentz uma expressão similar à fórmula de Rodrigues que obtivemos para o grupo $SO(3)$, expressão (21.103), página 1057. Essa expressão é muito útil.

Como vimos acima, os *boosts* de Lorentz são da forma $\exp(\vec{\alpha} \cdot \vec{\mathcal{M}})$, com $\vec{\alpha} \in \mathbb{R}^3$, arbitrário. Usando (21.217), página 1117, é fácil mostrar por um cômputo direto que

$$\vec{\alpha} \cdot \vec{\mathcal{M}} = \begin{pmatrix} 0 & -\alpha_1 & -\alpha_2 & -\alpha_3 \\ -\alpha_1 & 0 & 0 & 0 \\ -\alpha_2 & 0 & 0 & 0 \\ -\alpha_3 & 0 & 0 & 0 \end{pmatrix}, \quad \text{que} \quad (\vec{\alpha} \cdot \vec{\mathcal{M}})^2 = \begin{pmatrix} \|\vec{\alpha}\|^2 & 0 & 0 & 0 \\ 0 & \alpha_1^2 & \alpha_1\alpha_2 & \alpha_1\alpha_3 \\ 0 & \alpha_2\alpha_1 & \alpha_2^2 & \alpha_2\alpha_3 \\ 0 & \alpha_3\alpha_1 & \alpha_3\alpha_2 & \alpha_3^2 \end{pmatrix}$$

e que

$$(\vec{\alpha} \cdot \vec{\mathcal{M}})^3 = \|\vec{\alpha}\|^2 (\vec{\alpha} \cdot \vec{\mathcal{M}}).$$

E. 21.104 *Exercício.* Verifique! *

É agora elementar provar por indução que

$$(\vec{\alpha} \cdot \vec{\mathcal{M}})^{2m} = \|\vec{\alpha}\|^{2m-2} (\vec{\alpha} \cdot \vec{\mathcal{M}})^2, \quad \forall m \in \mathbb{N},$$

$$(\vec{\alpha} \cdot \vec{\mathcal{M}})^{2m+1} = \|\vec{\alpha}\|^{2m} (\vec{\alpha} \cdot \vec{\mathcal{M}}), \quad \forall m \in \mathbb{N}_0.$$

E. 21.105 *Exercício.* Prove isso!! *

Com isso, separando potências pares de ímpares na série de Taylor que define $\exp(\vec{\alpha} \cdot \vec{\mathcal{M}})$, obtemos

$$\begin{aligned} \exp(\vec{\alpha} \cdot \vec{\mathcal{M}}) &= \mathbb{1} + \sum_{n=1}^{\infty} \frac{1}{n!} (\vec{\alpha} \cdot \vec{\mathcal{M}})^n \\ &= \mathbb{1} + \sum_{m=1}^{\infty} \frac{1}{(2m)!} (\vec{\alpha} \cdot \vec{\mathcal{M}})^{2m} + \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} (\vec{\alpha} \cdot \vec{\mathcal{M}})^{2m+1} \\ &= \mathbb{1} + \left(\sum_{m=1}^{\infty} \frac{\|\vec{\alpha}\|^{2m-2}}{(2m)!} \right) (\vec{\alpha} \cdot \vec{\mathcal{M}})^2 + \left(\sum_{m=0}^{\infty} \frac{\|\vec{\alpha}\|^{2m}}{(2m+1)!} \right) (\vec{\alpha} \cdot \vec{\mathcal{M}}). \end{aligned}$$

Concluimos assim que

$$\exp(\vec{\alpha} \cdot \vec{\mathcal{M}}) = \mathbb{1} + \frac{\sinh(\|\vec{\alpha}\|)}{\|\vec{\alpha}\|} (\vec{\alpha} \cdot \vec{\mathcal{M}}) + \left(\frac{\cosh(\|\vec{\alpha}\|) - 1}{\|\vec{\alpha}\|^2} \right) (\vec{\alpha} \cdot \vec{\mathcal{M}})^2, \quad \vec{\alpha} \in \mathbb{R}^3. \quad (21.224)$$

Esta é a *fórmula de Rodrigues para os boosts de Lorentz*.

É conveniente reescrevê-la de forma mais adequada. Assumindo $\vec{\alpha} \neq 0$, definindo $z := \|\vec{\alpha}\|$, definindo $\vec{\xi} := \vec{\alpha}/\|\vec{\alpha}\|$, que é um vetor unitário (*i.e.*, $\|\vec{\xi}\| = 1$) de \mathbb{R}^3 , a expressão (21.224) fica

$$\exp(z\vec{\xi} \cdot \vec{\mathcal{M}}) = \mathbb{1} + \sinh(z) (\vec{\xi} \cdot \vec{\mathcal{M}}) + (\cosh(z) - 1) (\vec{\xi} \cdot \vec{\mathcal{M}})^2. \quad (21.225)$$

E. 21.106 *Exercício.* Verifique que para $\vec{\xi} = (1, 0, 0)$ vale

$$\exp(z\mathcal{M}_1) = \mathbb{1} + \sinh(z)\mathcal{M}_1 + (\cosh(z) - 1)\mathcal{M}_1^2 = \begin{pmatrix} \cosh(z) & -\sinh(z) & 0 & 0 \\ -\sinh(z) & \cosh(z) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \stackrel{(21.215)}{=} \mathcal{B}_1(z),$$

como esperaríamos obter. ✦

• Transformações de Lorentz gerais

Verifiquemos agora a ação de um *boost* $\exp(z\vec{\xi} \cdot \vec{\mathcal{M}})$ sobre um quadrivetor $x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \equiv \begin{pmatrix} x^0 \\ \vec{x} \end{pmatrix}$. Temos que

$$(\vec{\xi} \cdot \vec{\mathcal{M}})x = - \begin{pmatrix} \vec{\xi} \cdot \vec{x} \\ x^0 \vec{\xi} \end{pmatrix} \quad \text{e} \quad (\vec{\xi} \cdot \vec{\mathcal{M}})^2 x = \begin{pmatrix} x^0 \\ (\vec{\xi} \cdot \vec{x}) \vec{\xi} \end{pmatrix}.$$

pois $\|\vec{\xi}\| = 1$. Por isso, temos por (21.225) que

$$x' \equiv \exp(z\vec{\xi} \cdot \vec{\mathcal{M}})x = \begin{pmatrix} x^0 \\ \vec{x} \end{pmatrix} - \sinh(z) \begin{pmatrix} \vec{\xi} \cdot \vec{x} \\ x^0 \vec{\xi} \end{pmatrix} + (\cosh(z) - 1) \begin{pmatrix} x^0 \\ (\vec{\xi} \cdot \vec{x}) \vec{\xi} \end{pmatrix}. \tag{21.226}$$

Definindo como antes $v \equiv \tanh(z)$ e com $\vec{v} := v\vec{\xi}$, teremos $\cosh(z) = \frac{1}{\sqrt{1-v^2}} \equiv \gamma(v)$ e $\sinh(z) = v\gamma(v)$ e, assim,

$$x' = \begin{pmatrix} x^0 \\ \vec{x} \end{pmatrix} - v\gamma(v) \begin{pmatrix} \vec{\xi} \cdot \vec{x} \\ x^0 \vec{\xi} \end{pmatrix} + (\gamma(v) - 1) \begin{pmatrix} x^0 \\ (\vec{\xi} \cdot \vec{x}) \vec{\xi} \end{pmatrix}. \tag{21.227}$$

Escrevendo separadamente as componentes temporal e espaciais, (21.227) fica:

$$(x^0)' = \gamma(v)(x^0 - \vec{v} \cdot \vec{x}), \tag{21.228}$$

$$(\vec{x})' = (\vec{x} - (\vec{\xi} \cdot \vec{x})\vec{\xi}) + \gamma(v)(\vec{\xi} \cdot \vec{x} - vx_0)\vec{\xi}.$$

Separando-se na componente paralela a $\vec{\xi}$ e na componente ortogonal a $\vec{\xi}$ e definindo $x^{\parallel} := \vec{\xi} \cdot \vec{x}$ e $\vec{x}^{\perp} := \vec{x} - (\vec{\xi} \cdot \vec{x})\vec{\xi}$ podemos ainda escrever (21.228) como

$$\begin{aligned} (x^0)' &= \gamma(v)(x^0 - vx^{\parallel}), \\ (x^{\parallel})' &= \gamma(v)(x^{\parallel} - vx^0), \end{aligned} \tag{21.229}$$

$$(\vec{x}^{\perp})' = \vec{x}^{\perp}.$$

Para *boosts* na direção 1, por exemplo, temos $\xi = (1, 0, 0, 0)$ e (21.229) fica

$$\begin{aligned} (x^0)' &= \gamma(v)(x^0 - vx^1), \\ (x^1)' &= \gamma(v)(x^1 - vx^0), \\ (x^2)' &= x^2, \\ (x^3)' &= x^3, \end{aligned} \tag{21.230}$$

que são as transformações de Lorentz na forma usualmente apresentada em textos de Física (vide (21.203), página 1111). As relações (21.227), (21.228) e (21.229) são extensões disso para o caso de *boosts* gerais.

21.7.7 O Grupo de Galilei

Para diversos propósitos é interessante comparar o grupo \mathcal{L}_+^\uparrow com seu correspondente não relativístico: o *grupo de Galilei*⁶⁶. Excluindo-se a reversão temporal e as reflexões espaciais, as bem conhecidas transformações de Galilei da Mecânica Clássica são fornecidas por matrizes reais 4×4 , da forma

$$G(r, \vec{v}) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_1 & \boxed{} & & \\ -v_2 & & & \\ -v_3 & & & \end{pmatrix},$$

onde r é uma matriz 3×3 pertencente a $SO(3)$ e $\vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$ é um vetor com componentes $v_j \in (-\infty, \infty)$, $j = 1, 2, 3$. As matrizes $G(r, \vec{v})$ transformam da seguinte forma um vetor com 4 componentes $\begin{pmatrix} t \\ \vec{x} \end{pmatrix}$:

$$G(r, \vec{v}) \begin{pmatrix} t \\ \vec{x} \end{pmatrix} = \begin{pmatrix} t \\ r\vec{x} - \vec{v}t \end{pmatrix}. \tag{21.231}$$

A interpretação física-geométrica dessa expressão é que ela apresenta uma transformação de sistemas de referência inerciais na qual os eixos espaciais coordenados são rodados de $r \in SO(3)$ e em seguida é aplicado um *boost* não relativístico de velocidade $\vec{v} \in \mathbb{R}^3$, sem alterar intervalos de tempo. Transformações entre sistemas de referência na Mecânica Clássica são estudadas na Seção 44.1, página 2374. Vide, em particular a Proposição 44.2, página 2382.

E. 21.107 *Exercício.* Mostre que o conjunto de matrizes

$$\mathcal{G}_+^\uparrow := \{G(r, \vec{v}), r \in SO(3) \text{ e } \vec{v} \in \mathbb{R}^3\}$$

forma um grupo pelo produto usual de matrizes: mostre que vale a regra de produto

$$G(r_1, \vec{v}_1)G(r_2, \vec{v}_2) = G(r_1r_2, \vec{v}_1 + r_1\vec{v}_2), \tag{21.232}$$

mostre que o elemento neutro é $G(\mathbb{1}_3, \vec{0}) = \mathbb{1}_4$ e mostre que $G(r, \vec{v})^{-1} = G(r^{-1}, -r^{-1}\vec{v})$. Esse grupo \mathcal{G}_+^\uparrow é denominado *Grupo de Galilei* em $(3 + 1)$ -dimensões. *

As matrizes $G(\vec{v}) \equiv G(\mathbb{1}_3, \vec{v})$, ou seja,

$$G(\vec{v}) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ -v_1 & 1 & 0 & 0 \\ -v_2 & 0 & 1 & 0 \\ -v_3 & 0 & 0 & 1 \end{pmatrix},$$

com $\vec{v} = (v_1, v_2, v_3) \in \mathbb{R}^3$, fornecem os chamados *boosts de Galilei* com velocidade \vec{v} . É evidente por (21.232) que

$$G(\vec{v}_1)G(\vec{v}_2) = G(\vec{v}_1 + \vec{v}_2) = G(\vec{v}_2)G(\vec{v}_1).$$

Disso, vê-se que $\{G(\vec{v}), \vec{v} \in \mathbb{R}^3\}$ é um subgrupo Abeliano do grupo de Galilei, denominado *grupo dos boosts de Galilei*. Ao contrário do que ocorre com os *boosts* de Lorentz, os *boosts* de Galilei formam por si só um grupo e esse grupo é Abeliano (e isomorfo ao grupo aditivo \mathbb{R}^3). Esses fatos compõem uma marcante diferença entre \mathcal{L}_+^\uparrow e \mathcal{G}_+^\uparrow .

⁶⁶Galileo Galilei (1564–1642).

De (21.232) é fácil constatar que o grupo de Galilei \mathcal{G}_+^\uparrow é isomorfo ao grupo obtido pelo produto semidireto $SO(3)\mathbb{S}\mathbb{R}^3$, com $SO(3)$ representando rotações e \mathbb{R}^3 representando *boosts* de Galilei.

O conjunto de matrizes $\text{SRot} = \{G(r, \vec{0}), r \in SO(3)\}$ também compõe um subgrupo de \mathcal{G}_+^\uparrow , o subgrupo das rotações. Usando (21.232), é fácil ver que para todos $\vec{v} \in \mathbb{R}^3$ e $r \in SO(3)$ vale

$$G(r, \vec{v}) = G(\vec{v})G(r, \vec{0}) = G(r, \vec{0})G(r^{-1}\vec{v}), \tag{21.233}$$

ou seja, todo elemento de \mathcal{G}_+^\uparrow pode ser escrito como o produto de uma rotação seguida de um *boost* de Galilei (ou na ordem oposta). Compare-se com a afirmativa da Proposição 21.30, página 1119, e comentários que se lhe seguem.

Vamos agora estudar os geradores infinitesimais do grupo de Galilei.

E. 21.108 *Exercício.* Mostre que os geradores infinitesimais dos *boosts* de Galilei são as matrizes

$$\mathcal{M}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{21.234}$$

Compare-os aos geradores infinitesimais dos *boosts* do grupo \mathcal{L}_+^\uparrow dados em (21.217), página 1117. Mostre que os geradores infinitesimais das rotações no grupo de Galilei são as matrizes

$$\mathcal{J}_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathcal{J}_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \mathcal{J}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{21.235}$$

Constata-se que (sem nenhuma surpresa) eles são idênticos aos geradores infinitesimais do subgrupo de rotações do grupo \mathcal{L}_+^\uparrow dados em (21.218), página 1118.

Mostre que os geradores infinitesimais \mathcal{M}_a e \mathcal{J}_b , com $a, b = 1, 2, 3$, acima, satisfazem as seguintes relações de comutação:

$$[\mathcal{J}_a, \mathcal{J}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathcal{J}_c, \tag{21.236}$$

$$[\mathcal{M}_a, \mathcal{M}_b] = 0, \tag{21.237}$$

$$[\mathcal{J}_a, \mathcal{M}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathcal{M}_c. \tag{21.238}$$

Mostre também que de (21.236)–(21.238) segue que

$$[\vec{\alpha} \cdot \vec{\mathcal{M}} + \vec{\beta} \cdot \vec{\mathcal{J}}, \vec{\gamma} \cdot \vec{\mathcal{M}} + \vec{\delta} \cdot \vec{\mathcal{J}}] = (\vec{\beta} \times \vec{\gamma} - \vec{\alpha} \times \vec{\delta}) \cdot \vec{\mathcal{M}} + (\vec{\beta} \times \vec{\delta}) \cdot \vec{\mathcal{J}}. \tag{21.239}$$

Acima, $\vec{\alpha}, \vec{\beta}, \vec{\gamma}$ e $\vec{\delta}$ são vetores em \mathbb{R}^3 . É claro de (21.236)–(21.238), e especialmente de (21.239), que o espaço vetorial real composto por todas as combinações lineares reais dos seis geradores infinitesimais \mathcal{M}_a e \mathcal{J}_b formam uma álgebra de Lie real, a *álgebra de Lie do grupo de Galilei* \mathcal{G}_+^\uparrow . Compare essa álgebra de Lie com a álgebra de Lie do grupo \mathcal{L}_+^\uparrow , tal como exibida em (21.219)–(21.221) e (21.222), página 1119.

Mostre também que $G(\vec{v}) = \exp(\vec{v} \cdot \vec{\mathcal{M}})$, para todo $\vec{v} \in \mathbb{R}^3$, calculando explicitamente a exponencial do lado direito (isso é fácil, pois as matrizes $\vec{v} \cdot \vec{\mathcal{M}}$ são nilpotentes).

Juntando isso à afirmação de (21.233) e ao fato de todo elemento $r \in \text{SO}(3)$ ser da forma $\exp(\theta \vec{\eta} \cdot \vec{J})$, com \vec{J} sendo os geradores infinitesimais de $\text{SO}(3)$, $\theta \in [0, \pi]$ e $\vec{\eta} \in \mathbb{R}^3$ com $\|\vec{\eta}\| = 1$ (vide Seção 21.4.2, página 1053), conclua que todo elemento de \mathcal{G}_+^\uparrow é da forma

$$G(\exp(\theta \vec{\eta} \cdot \vec{J}), \vec{v}) = \exp(\vec{v} \cdot \vec{M}) \exp(\theta \vec{\eta} \cdot \vec{J}).$$

Compare com a afirmação da Proposição 21.32, página 1121, sobre o grupo de Lorentz. ✦

• **O grupo de Galilei não homogêneo**

O chamado *grupo de Galilei não homogêneo* é o grupo obtido do produto semidireto⁶⁷ de \mathcal{G}_+^\uparrow com o grupo aditivo de translações $(\mathbb{R}^4, +)$, ou seja, $\mathcal{G}_+^\uparrow \otimes (\mathbb{R}^4, +)$. Seus elementos são, portanto, pares ordenados $(G(r, \vec{v}), a)$ com $G(r, \vec{v}) \in \mathcal{G}_+^\uparrow$ e $a \in \mathbb{R}^4$, sendo o produto dado por

$$(G(r, \vec{v}), a) \cdot (G(r', \vec{v}'), a') = (G(r, \vec{v})G(r', \vec{v}'), G(r, \vec{v})a' + a).$$

A ação de um elemento $(G(r, \vec{v}), a)$, do grupo de Galilei não homogêneo, no espaço-tempo \mathbb{R}^4 é interpretada como uma transformação de Galilei $G(r, \vec{v})$ seguida de uma translação por $a \in \mathbb{R}^4$.

A versão relativística do grupo de Galilei não homogêneo é o grupo de Poincaré, apresentado na Seção 21.8, página 1129.



Nota. Para um estudo mais aprofundado do grupo de Galilei, de suas representações irredutíveis, da álgebra de Lie dos seus geradores infinitesimais e das aplicações desses temas à Mecânica Quântica Não Relativística, vide:

1. Jean–Marc Lévy–Leblond, “Galilei group and Galilean invariance”, em [253], vol. 2. páginas 221–296.
2. Jean–Marc Lévy–Leblond, “Galilei Group and Nonrelativistic Quantum Mechanics”, J. Math. Phys. **4**, 776–789 (1963). **Doi:** 10.1063/1.1724319.
3. H. Bacry and J.–M. Lévy–Leblond, “Possible kinematics”, J. Math. Phys. **9**, 1605–1614 (1968).
4. J.–M. Lévy–Leblond, “Une nouvelle limite non-relativiste du groupe de Poincaré”, Ann. Inst. H. Poincaré, Sect. A **3**, 1–12 (1965).
Vide também:
5. E. İnönü, E. P. Wigner, “On the Contraction of Groups and Their Representations”. Proc. Nat. Acad. Sci. **39** (6), 510–524 (1953).
Doi: 10.1073/pnas.39.6.510.

Para uma leitura extremamente interessante e que contém comentários sobre as contribuições supralistadas de Lévy–Leblond⁶⁸, İnönü⁶⁹ e Wigner⁷⁰, vide: Freeman J. Dyson⁷¹, “Missed opportunities”. Bull. Amer. Math. Soc. **78**, Number 5, 635–652 (1972). ♣

21.7.7.1 Comparação Entre os Grupos de Galilei e Lorentz. Contrações

Nesta seção discutiremos de que forma podemos obter o grupo de Galilei como um limite do grupo de Lorentz. Essa questão é relevante devida à crença de a Física não relativística ser um limite da Física relativística quando a velocidade da luz for formalmente tomada como infinita. Como veremos, essa tomada de limite pode ser formulada de maneira precisa. Começamos com uma observação que aponta para certas sutilizas dessa questão.

• **Uma transformação de Galilei não é uma transformação de Lorentz lenta**

Uma transformação de Galilei que não envolva rotações, mas apenas um *boost* com velocidade na direção 1 é da forma

$$t' = t, \tag{21.240}$$

$$x' = x - vt, \tag{21.241}$$

$$y' = y, \tag{21.242}$$

$$z' = z. \tag{21.243}$$

⁶⁷A noção geral de produto semidireto de grupos encontra-se definida na Seção 2.2.4.2, página 145.

⁶⁸Jean–Marc Lévy–Leblond (1940–).

⁶⁹Erdal İnönü (1926–2007).

⁷⁰Eugene Paul Wigner (1902–1995).

⁷¹Freeman John Dyson (1923–).

Sob as mesmas restrições, uma transformação de Lorentz é da forma (vide (21.203))

$$t' = \gamma(v) \left(t - \frac{v}{c} x \right), \quad (21.244)$$

$$x' = \gamma(v) (x - vt), \quad (21.245)$$

$$y' = y, \quad (21.246)$$

$$z' = z, \quad (21.247)$$

com $\gamma(v) = 1/\sqrt{1 - (v/c)^2}$. Se desprezarmos termos da ordem $(v/c)^2$, aproximando, portanto, $\gamma(v)$ por 1, essas expressões ficam

$$t' = t - \frac{v}{c} x, \quad (21.248)$$

$$x' = x - vt, \quad (21.249)$$

$$y' = y, \quad (21.250)$$

$$z' = z. \quad (21.251)$$

Comparando-se (21.248)-(21.251) a (21.240)-(21.243), vemos que o lado direito de (21.240) e de (21.248) diferem. Com isso, adquirimos o importante entendimento de que uma transformação de Galilei **não** é meramente uma transformação de Lorentz *lenta*, ou seja, com v/c pequeno, de sorte que possamos desprezar termos da ordem de $(v/c)^2$.

A diferença entre que o lado direito de (21.240) e de (21.248) é o termo $-\frac{v}{c}x$, que se origina na necessidade de, em uma teoria relativística, se considerar tempos de retardo no processo de sincronização de relógios em um sistema de referência inercial, devido à finitude de propagação de sinais luminosos.

Coloca-se, assim, a questão de como exatamente devemos proceder se desejamos obter o grupo de Galilei como algum limite do grupo de Lorentz. Como veremos no que segue, a resposta não é surpreendente: devemos tomar de forma apropriada o limite $c \rightarrow \infty$ e não meramente desprezar termos em $(v/c)^2$. Como exatamente fazer isso é o que discutiremos no que segue.

• Comparação entre os Grupos de Galilei e Lorentz. Contração

As expressões (21.215)-(21.216) para os *boosts* de Lorentz foram obtidas adotando-se, por mera conveniência, $c = 1$. Para uma comparação com o grupo de Galilei é conveniente expressarmos os *boosts* de Lorentz adotando-se $c \neq 1$, supondo, além disso, que os pontos do espaço-tempo sejam representados por coordenadas $\begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$, com t tendo a dimensão de “tempo” e com as coordenadas x_k tendo dimensão de “espaço”. A passagem das coordenadas $\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$, com $x_0 = ct$, para as coordenadas $\begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$ é implementada pela matriz

$$C := \begin{pmatrix} c & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{através de} \quad \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = C \begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Assim, a ação dos *boosts* é dada por

$$\mathcal{B}_k(z) \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathcal{B}_k^c(z) \begin{pmatrix} t \\ x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

$k = 1, 2, 3$, onde

$$\mathcal{B}_k^c(z) := C^{-1} \mathcal{B}_k(z) C,$$

ou seja,

$$\mathcal{B}_1^c(z) := \begin{pmatrix} \cosh z & -\frac{1}{c} \sinh z & 0 & 0 \\ -c \sinh z & \cosh z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{B}_2^c(z) := \begin{pmatrix} \cosh z & 0 & -\frac{1}{c} \sinh z & 0 \\ 0 & 1 & 0 & 0 \\ -c \sinh z & 0 & \cosh z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (21.252)$$

$$\mathcal{B}_3^c(z) := \begin{pmatrix} \cosh z & 0 & 0 & -\frac{1}{c} \sinh z \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -c \sinh z & 0 & 0 & \cosh z \end{pmatrix}. \quad (21.253)$$

Com $c \neq 1$ a “rapidez” z se expressa em termos da velocidade $v \in (-c, c)$ por $z = \tanh^{-1}(v/c)$.

Os respectivos geradores infinitesimais dos *boosts* $\mathcal{B}_k^c(z)$ são agora expressos, não como em (21.217), mas como

$$\mathcal{M}_1^c = \begin{pmatrix} 0 & -c^{-1} & 0 & 0 \\ -c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_2^c = \begin{pmatrix} 0 & 0 & -c^{-1} & 0 \\ 0 & 0 & 0 & 0 \\ -c & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_3^c = \begin{pmatrix} 0 & 0 & 0 & -c^{-1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -c & 0 & 0 & 0 \end{pmatrix}. \quad (21.254)$$

Naturalmente, temos

$$\mathcal{B}_k^c(z) = \exp(z\mathcal{M}_k^c), \quad k = 1, 2, 3. \quad (21.255)$$

Observe-se que vale

$$\mathcal{M}_k^c = C^{-1}\mathcal{M}_k C, \quad k = 1, 2, 3, \quad (21.256)$$

onde $\mathcal{M}_k \equiv \mathcal{M}_k^1$ são as matrizes de (21.217).

Os geradores infinitesimais das rotações não se modificam. Quanto às relações de comutação (21.219)-(21.221), usando-se (21.256) e o fato que C e \mathcal{J}_k comutam para cada $k = 1, 2, 3$, obtemos

$$[\mathcal{J}_i, \mathcal{J}_j] = \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{J}_k, \quad (21.257)$$

$$[\mathcal{M}_i^c, \mathcal{M}_j^c] = -\sum_{k=1}^3 \varepsilon_{ijk} \mathcal{J}_k, \quad (21.258)$$

$$[\mathcal{J}_i, \mathcal{M}_j^c] = \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{M}_k^c. \quad (21.259)$$

Verifique! Observe-se que são as mesmas expressões que em (21.219)-(21.221), e constatamos assim que a forma das mesmas independe de c .

Definamos agora $\mathcal{M}_k^{G,c} := c^{-1}\mathcal{M}_k^c$, $k = 1, 2, 3$, ou seja,

$$\mathcal{M}_1^{G,c} = \begin{pmatrix} 0 & -c^{-2} & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_2^{G,c} = \begin{pmatrix} 0 & 0 & -c^{-2} & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_3^{G,c} = \begin{pmatrix} 0 & 0 & 0 & -c^{-2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \quad (21.260)$$

Com isso, (21.255) fica

$$\mathcal{B}_k^c(z) = \exp((cz)\mathcal{M}_k^{G,c}), \quad k = 1, 2, 3. \quad (21.261)$$

As relações de comutação (21.257)-(21.259) tornam-se

$$[\mathcal{J}_i, \mathcal{J}_j] = \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{J}_k, \tag{21.262}$$

$$[\mathcal{M}_i^{G,c}, \mathcal{M}_j^{G,c}] = -\frac{1}{c^2} \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{J}_k, \tag{21.263}$$

$$[\mathcal{J}_i, \mathcal{M}_j^{G,c}] = \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{M}_k^{G,c}. \tag{21.264}$$

Tomando-se nas expressões (21.262)-(21.264) o limite $c \rightarrow \infty$, obtemos

$$[\mathcal{J}_i, \mathcal{J}_j] = \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{J}_k, \tag{21.265}$$

$$[\mathcal{M}_i^G, \mathcal{M}_j^G] = 0, \tag{21.266}$$

$$[\mathcal{J}_i, \mathcal{M}_j^G] = \sum_{k=1}^3 \varepsilon_{ijk} \mathcal{M}_k^G, \tag{21.267}$$

onde $\mathcal{M}_k^G := \lim_{c \rightarrow \infty} \mathcal{M}_k^{G,c}$, $k = 1, 2, 3$, ou seja,

$$\mathcal{M}_1^G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_2^G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_3^G = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}. \tag{21.268}$$

Esses são precisamente os geradores infinitesimais dos *boosts* de Galilei, indicados em (21.234), e as relações (21.265)-(21.267) são precisamente as relações (21.236)-(21.238), que definem a álgebra de Lie do grupo de Galilei.

Voltando ao lado direito de (21.261), observamos que $cz = c \tanh^{-1}(v/c)$. Para $c \rightarrow \infty$ isso converge a v e, portanto, nesse limite os *boosts* $\mathcal{B}_k^c(z)$ convergem a $\exp(v\mathcal{M}_k^G)$, agora com $v \in \mathbb{R}$, que são os *boosts* de Galilei na direção k .

O processo acima descrito, de transformação da álgebra de Lie do grupo de Lorentz (descrita pelas relações (21.219)-(21.221) ou (21.257)-(21.259)) na álgebra de Lie do grupo de Galilei (descrita pelas relações (21.236)-(21.238)) tomando-se o limite em que um parâmetro (no caso, c) vai a infinito junto com alguns reescalamentos adequados, é denominado *contração de álgebras de Lie*, tendo sido estudado por İnönü⁷² e Wigner⁷³, especialmente em suas implicações sobre a teoria das representações das álgebras de Lie. Para o trabalho original, vide [193]. Para estimulantes notas históricas, vide [194].

21.8 O Grupo de Poincaré

O chamado grupo de Poincaré⁷⁴ (em 3+1 dimensões), também conhecido como *grupo de Lorentz não homogêneo*, é definido como o produto semidireto⁷⁵ de $O(1, 3)$ e do grupo aditivo de translações \mathbb{R}^4 , ou seja, $\mathcal{P} := O(1, 3) \otimes \mathbb{R}^4$. Seus elementos são, portanto, pares ordenados (L, a) com $L \in O(1, 3)$ e $a \in \mathbb{R}^4$, sendo o produto dado por $(L, a) \cdot (L', a') = (LL', La' + a)$. Sua ação no espaço-tempo \mathbb{R}^4 é interpretada como uma transformação de Lorentz L seguida de uma translação por $a \in \mathbb{R}^4$.

O grupo de Poincaré é também denominado *grupo de Lorentz não homogêneo*.

⁷²Erdal İnönü (1926–2007).

⁷³Eugene Paul Wigner (1902–1995).

⁷⁴Jules Henri Poincaré (1854–1912).

⁷⁵A noção geral de produto semidireto de grupos encontra-se definida na Seção 2.2.4.2, página 145.

Há um subgrupo de $GL(\mathbb{R}, 5)$ que é isomorfo a \mathcal{P} . Sejam as matrizes reais 5×5 da forma

$$P(L, a) := \begin{pmatrix} \boxed{L} & \boxed{a} \\ \boxed{0} & \boxed{1} \end{pmatrix}, \quad \text{com } L \in O(1, 3) \text{ e } a \in \mathbb{R}^4.$$

Então, tem-se

$$P(L, a) P(L', a') = P(LL', La' + a).$$

E. 21.109 Exercício importante. Verifique isso e verifique que o conjunto de matrizes

$$\{P(L, a) \in GL(\mathbb{R}, 5), \text{ com } L \in O(1, 3) \text{ e } a \in \mathbb{R}^4\}$$

forma um subgrupo de $GL(\mathbb{R}, 5)$. Mostre que a aplicação $\mathcal{P} \ni (L, a) \mapsto P(L, a)$ é um isomorfismo. *

Assim, o conjunto de matrizes $\{P(L, a) \in GL(\mathbb{R}, 5), \text{ com } L \in O(1, 3) \text{ e } a \in \mathbb{R}^4\}$ forma um subgrupo de $GL(\mathbb{R}, 5)$ que é isomorfo a \mathcal{P} . Também denotaremos esse grupo por \mathcal{P} .

A representação do grupo de Poincaré pelas matrizes $P(L, a)$ serve basicamente à visualização de sua ação. Para tal, pontos do espaço tempo com coordenadas $x \equiv \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^4$ são representados em \mathbb{R}^5 pelos pontos $\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ 1 \end{pmatrix} \equiv \begin{pmatrix} \boxed{x} \\ 1 \end{pmatrix} \in \mathbb{R}^5$.

E. 21.110 Exercício. Mostre que

$$P(L, a) \begin{pmatrix} \boxed{x} \\ 1 \end{pmatrix} = \begin{pmatrix} \boxed{Lx+a} \\ 1 \end{pmatrix}.$$

*

• Os geradores infinitesimais do grupo de Poincaré

O chamado *grupo de Poincaré próprio ortócrono*, denotado por \mathcal{P}_+^\uparrow é o grupo $\mathcal{P}_+^\uparrow := \mathcal{L}_+^\uparrow \otimes \mathbb{R}^4$.

De maneira totalmente análoga ao que fizemos no grupo Euclidiano, podemos determinar os geradores infinitesimais do grupo \mathcal{P}_+^\uparrow . Este possui 10 geradores infinitesimais. Seis da forma

$$\mathbf{m}_k := \begin{pmatrix} \boxed{\mathcal{M}_k} & \boxed{0} \\ \boxed{0} & \boxed{0} \end{pmatrix} \quad \text{ou} \quad \mathbf{j}_k := \begin{pmatrix} \boxed{\mathcal{J}_k} & \boxed{0} \\ \boxed{0} & \boxed{0} \end{pmatrix}, \quad \text{com } k = 1, 2, 3,$$

onde \mathcal{M}_k e \mathcal{J}_k são as matrizes 4×4 definidas em (21.217) e (21.218), respectivamente, e quatro da forma

$$\mathbf{p}_k := \left(\begin{array}{c|c} \boxed{0} & \boxed{x_k} \\ \hline \boxed{0} & \boxed{0} \end{array} \right) \text{ com } k = 0, \dots, 3,$$

onde

$$\mathbf{x}_0 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}_1 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{x}_3 := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

As relações de comutação associadas ao grupo de Poincaré são:

$$[\mathbf{j}_a, \mathbf{j}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathbf{j}_c, \tag{21.269}$$

$$[\mathbf{m}_a, \mathbf{m}_b] = -\sum_{c=1}^3 \varepsilon_{abc} \mathbf{j}_c, \tag{21.270}$$

$$[\mathbf{j}_a, \mathbf{m}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathbf{m}_c, \tag{21.271}$$

$$[\mathbf{p}_a, \mathbf{p}_b] = 0, \tag{21.272}$$

$$[\mathbf{j}_a, \mathbf{p}_b] = (1 - \delta_{b0}) \sum_{c=1}^3 \varepsilon_{abc} \mathbf{p}_c, \tag{21.273}$$

$$[\mathbf{m}_a, \mathbf{p}_b] = -(\delta_{ab} \mathbf{p}_0 + \delta_{b0} \mathbf{p}_a). \tag{21.274}$$

Aqui, os índices dos \mathbf{m} 's e \mathbf{j} 's variam de 1 a 3 e os índices dos \mathbf{p} 's variam de 0 a 3.

E. 21.111 Exercício importante. Todo estudante deve fazê-lo uma vez na vida. Verifique isso. ✱

As três primeiras relações acima seguem de (21.219)-(21.221), página 1119. A relação (21.273) diz que os \mathbf{j} 's comutam com \mathbf{p}_0 e, nos demais casos, tem-se a última relação de (21.183).

Novamente constatamos que a subálgebra gerada pelos \mathbf{p} 's é um ideal de álgebra de Lie do grupo de Poincaré.

• O grupo \mathcal{P}_+^\uparrow em 1+1-dimensões

Com base nas considerações acima e no nosso estudo do grupo $O(1, 1)$ (vide Seção 21.4.1, em especial, página 1051),

sabemos que o grupo \mathcal{P}_+^\uparrow em 1+1-dimensões é isomorfo ao grupo de matrizes da forma

$$\begin{pmatrix} \cosh z & -\sinh z & a_0 \\ -\sinh z & \cosh z & a_1 \\ 0 & 0 & 1 \end{pmatrix},$$

com $z, a_0, a_1 \in \mathbb{R}$. Seus geradores infinitesimais serão

$$m_1 := \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_0 := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad p_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Como é fácil de verificar, as relações de comutação entre esses geradores infinitesimais são

$$[m_1, p_0] = -p_1, \quad [m_1, p_1] = -p_0, \quad [p_0, p_1] = 0.$$

Um elemento genérico dessa álgebra de Lie é da forma

$$I(\mathcal{M}, t) := \begin{pmatrix} \boxed{\mathcal{M}} & \boxed{t} \\ 0 & 0 & 0 \end{pmatrix}$$

onde

$$\mathcal{M} = zm_1 = \begin{pmatrix} 0 & -z \\ -z & 0 \end{pmatrix} \quad \text{e} \quad t = t_0p_0 + t_1p_1 = \begin{pmatrix} t_0 \\ t_1 \end{pmatrix}$$

com $z, t_0, t_1 \in \mathbb{R}$. É um exercício fácil (faça-o!) constatar que para todo $k \in \mathbb{N}$, tem-se

$$I(\mathcal{M}, t)^k = I(\mathcal{M}^k, \mathcal{M}^{k-1}t).$$

Consequentemente, vale que

$$\exp(I(\mathcal{M}, t)) = \mathbb{1} + \sum_{k=1}^{\infty} \frac{1}{k!} I(\mathcal{M}, t)^k = \mathbb{1} + \sum_{k=1}^{\infty} \frac{1}{k!} I(\mathcal{M}^k, \mathcal{M}^{k-1}t) = \begin{pmatrix} \boxed{L} & \boxed{t'} \\ 0 & 0 & 1 \end{pmatrix},$$

onde

$$L := e^{\mathcal{M}} = \begin{pmatrix} \cosh z & -\sinh z \\ -\sinh z & \cosh z \end{pmatrix} \quad \text{e} \quad t' = f(\mathcal{M})t,$$

sendo f a função analítica inteira definida em (21.184), página 1098. A matriz $f(\mathcal{M})$ pode ser calculada facilmente usando-se o fato que

$$\mathcal{M}^{2k} = \mathbb{1} \quad \text{e} \quad \mathcal{M}^{2k+1} = \mathcal{M}, \quad k \in \mathbb{N}_0,$$

de onde se extrai

$$\begin{aligned} f(\mathcal{M}) &:= \mathbb{1} + \sum_{k=2}^{\infty} \frac{1}{k!} \mathcal{M}^{k-1} = \mathbb{1} + \sum_{m=1}^{\infty} \frac{1}{(2m)!} \mathcal{M}^{2m-1} + \sum_{m=1}^{\infty} \frac{1}{(2m+1)!} \mathcal{M}^{2m} \\ &= \left(\sum_{m=1}^{\infty} \frac{z^{2m-1}}{(2m)!} \right) \mathcal{M} + \left(\sum_{m=0}^{\infty} \frac{z^{2m}}{(2m+1)!} \right) \mathbb{1} \\ &= \frac{\cosh z - 1}{z} \mathcal{M} + \frac{\sinh z}{z} \mathbb{1} \\ &= \begin{pmatrix} \frac{\sinh z}{z} & -\frac{\cosh z - 1}{z} \\ -\frac{\cosh z - 1}{z} & \frac{\sinh z}{z} \end{pmatrix}. \end{aligned}$$

Notemos que

$$\det(f(\mathcal{M})) = 2 \left(\frac{\cosh z - 1}{z^2} \right) \neq 0$$

para $z \in \mathbb{R}$. Assim, $f(\mathcal{M})$ é inversível e se escolhermos $t = f(\mathcal{M})^{-1}a$, para qualquer $a = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \in \mathbb{R}^2$, teremos

$$\exp\left(I(\mathcal{M}, f(\mathcal{M})^{-1}a)\right) = \begin{pmatrix} \boxed{L} & \boxed{a} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cosh z & -\sinh z & a_0 \\ -\sinh z & \cosh z & a_1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Isso prova que todo elemento do grupo \mathcal{P}_+^\uparrow em 1+1 dimensões pode ser escrito como exponencial de um elemento da sua própria álgebra de Lie.

21.9 Mais Sobre Grupos Simpléticos

Vamos aqui dar continuidade à nossa discussão sobre propriedades dos grupos simpléticos, que introduzimos na Seção 21.3.3.3, página 1041.

Nosso primeiro resultado aqui diz respeito a propriedades espectrais de matrizes do grupos $\text{Sp}(2n, \mathbb{C})$ e $\text{Sp}(2n, \mathbb{R})$.

Proposição 21.33 *Se $A \in \text{Sp}(2n, \mathbb{C})$, então os polinômios característicos de A e de A^{-1} são iguais: $q_A(x) = q_{A^{-1}}(x)$ para todo $x \in \mathbb{C}$. Consequentemente, A e A^{-1} têm o mesmo espectro de autovalores (incluindo multiplicidades) o que implica que se λ é autovalor de A , λ^{-1} também o é. Como consequência $\det A = 1$.*

Disso segue também que $q_A(x) = x^{2n} q_A(x^{-1})$ para todo $x \in \mathbb{C}$, $x \neq 0$. Um polinômio com essa propriedade é dito reflexivo (vide Comentário após a demonstração).

Como $\text{Sp}(2n, \mathbb{R}) \subset \text{Sp}(2n, \mathbb{C})$ essas afirmações valem também para $A \in \text{Sp}(2n, \mathbb{R})$, mas como as matrizes desse grupo são reais, vale adicionalmente que se λ é autovalor de $A \in \text{Sp}(2n, \mathbb{R})$, então $\bar{\lambda}$ também o é. Assim, se $A \in \text{Sp}(2n, \mathbb{R})$ e λ é um seu autovalor, concluímos que λ^{-1} , $\bar{\lambda}$ e $\overline{\lambda^{-1}}$ também são autovalores de A . \square

Prova da Proposição 21.33. Por simplicidade, denotaremos aqui J_{2n} simplesmente por J e $\mathbb{1}_{2n}$ por $\mathbb{1}$. Seja $q_M(x) = \det(x\mathbb{1} - M)$, $x \in \mathbb{C}$, o polinômio característico de uma matriz $M \in \text{Mat}(\mathbb{C}, 2n)$. Se $A \in \text{Sp}(2n, \mathbb{C})$, temos, usando o fato que $\det J = 1$, usando que $J^2 = -\mathbb{1}$, usando (21.52), usando que um determinante é invariante por transposições e usando que a dimensão do espaço é par,

$$\begin{aligned} q_A(x) &= \det(x\mathbb{1} - A) = \det\left(J(x\mathbb{1} - A)J\right) = \det\left(-x\mathbb{1} - JAJ\right) = \det\left(-x\mathbb{1} + (A^{-1})^T\right) \\ &= \det\left(-x\mathbb{1} + A^{-1}\right) = \det\left(x\mathbb{1} - A^{-1}\right) = q_{A^{-1}}(x). \end{aligned}$$

Isso mostra que A e A^{-1} têm o mesmo conjunto de autovalores, incluindo multiplicidades. Se $\{\lambda_1, \dots, \lambda_{2n}\}$ são os autovalores de A (repetindo-se multiplicidades) sabemos que $\{\lambda_1^{-1}, \dots, \lambda_{2n}^{-1}\}$ são os autovalores de A^{-1} (repetindo-se multiplicidades). Dessa forma, a igualdade $q_A(x) = q_{A^{-1}}(x)$ entre os polinômios característicos significa que $\prod_{k=1}^{2n} (x - \lambda_k) = \prod_{k=1}^{2n} (x - \lambda_k^{-1})$, o que implica que se λ é autovalor de A , λ^{-1} também o é. Concluimos disso que $\det(A) = \prod_{k=1}^{2n} \lambda_k = 1$.

Observe-se ainda que, como $\det(A) = 1$, podemos escrever, para $x \neq 0$,

$$q_{A^{-1}}(x) = \det\left(x\mathbb{1} - A^{-1}\right) = \det\left(xA - \mathbb{1}\right) = x^{2n} \det\left(A - x^{-1}\mathbb{1}\right) = (-1)^{2n} x^{2n} \det\left(x^{-1}\mathbb{1} - A\right) = x^{2n} q_A(x^{-1}),$$

e portanto, estabelecemos que, para $x \neq 0$,

$$q_A(x) = x^{2n} q_A(x^{-1}).$$

Um polinômio com essa propriedade é dito reflexivo (vide Comentário adiante).

Para qualquer matriz real $M \in \text{Mat}(\mathbb{R}, 2n)$ vale ainda que $\overline{q_M(\bar{x})} = \det(\overline{x\mathbb{1} - M}) = \det(x\mathbb{1} - M) = q_M(x)$. Se os autovalores que M são $\{\mu_1, \dots, \mu_{2n}\}$, essa igualdade significa que $\prod_{k=1}^{2n} (x - \mu_k) = \prod_{k=1}^{2n} (x - \bar{\mu}_k)$, implicando que se μ é autovalor de A , $\bar{\mu}$ também o é. Dessa forma, se $A \in \text{Sp}(2n, \mathbb{R})$ e λ é um seu autovalor, concluimos que λ^{-1} , $\bar{\lambda}$ e $\overline{\lambda^{-1}}$ também são autovalores de A . ■

Comentário. Polinômios reflexivos. Um polinômio de grau m , $p(x) = a_0 + a_1x + \dots + a_{m-1}x^{m-1} + a_mx^m$, é dito ser um *polinômio reflexivo* se seus coeficientes forem simétricos por reflexão em relação a seus índices, ou seja, se valer $a_k = a_{m-k}$ para todo $k \in \{0, \dots, m\}$. De forma explícita isso significa que $a_0 = a_m$, $a_1 = a_{m-1}$, $a_2 = a_{m-2}$ etc,

É elementar constatar (faça-o!) que um polinômio $p(x)$ de grau m é reflexivo se e somente se $p(x) = x^m p(x^{-1})$ para todo $x \neq 0$. ♣

21.9.1 O Subgrupo Simplético Real Ortogonal

Vamos agora estudar com mais detalhe os grupos simpléticos reais compactos $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ que introduzimos à página 1045. Seu estudo requer compreender a relação entre os espaços vetoriais reais \mathbb{R}^{2n} e os espaços vetoriais complexos \mathbb{C}^n . Nosso objetivo é apresentar uma demonstração alternativa do Teorema 21.8, página 1046, que afirma que os grupos $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ e $\text{U}(n)$ são isomorfos.

- **Grupos simpléticos e a estrutura complexa induzida por J_{2n}**

Há um isomorfismo \mathcal{I}_n entre o espaço vetorial complexo \mathbb{C}^n e o espaço vetorial real \mathbb{R}^{2n} , o qual é dado da seguinte

forma: $\mathcal{I}_n : \mathbb{C}^n \rightarrow \mathbb{R}^{2n}$ é definido por

$$\mathcal{I}_n \begin{pmatrix} u_1 + iu_{n+1} \\ \vdots \\ u_n + iu_{2n} \end{pmatrix} := \begin{pmatrix} u_1 \\ \vdots \\ u_n \\ u_{n+1} \\ \vdots \\ u_{2n} \end{pmatrix}, \quad \text{sendo} \quad \mathcal{I}_n^{-1} \begin{pmatrix} u_1 \\ \vdots \\ u_n \\ u_{n+1} \\ \vdots \\ u_{2n} \end{pmatrix} = \begin{pmatrix} u_1 + iu_{n+1} \\ \vdots \\ u_n + iu_{2n} \end{pmatrix},$$

sendo $u_k \in \mathbb{R}$ para todo $k = 1, \dots, 2n$.

É fácil constatar que \mathcal{I}_n^{-1} pode ser representada por uma matriz $n \times (2n)$:

$$\mathcal{I}_n^{-1} = \begin{pmatrix} \mathbb{1}_n & i\mathbb{1}_n \end{pmatrix} \in \text{Mat}(n, \mathbb{C}, 2n). \tag{21.275}$$

Para a notação, vide página 448 e seguintes.

Para números complexos da forma $x + iy$, com $x, y \in \mathbb{R}$, temos

$$\mathcal{I}_n \left((x + iy) \begin{pmatrix} u_1 + iu_{n+1} \\ \vdots \\ u_n + iu_{2n} \end{pmatrix} \right) = \mathcal{I}_n \begin{pmatrix} (xu_1 - yu_{n+1}) + i(xu_{n+1} + yu_1) \\ \vdots \\ (xu_n - yu_{2n}) + i(xu_{2n} + yu_n) \end{pmatrix} = (x\mathbb{1}_{2n} - yJ_{2n}) \begin{pmatrix} u_1 \\ \vdots \\ u_n \\ u_{n+1} \\ \vdots \\ u_{2n} \end{pmatrix}.$$

Verifique! E veja também a discussão da Seção 3.5, página 245. Assim, vemos que para a multiplicação por escalares, temos a regra

$$\mathcal{I}_n(z\mathcal{I}_n^{-1}(u)) := (x\mathbb{1}_{2n} - yJ_{2n})u,$$

onde $z = x + iy \in \mathbb{C}$, com $x, y \in \mathbb{R}$, e $u \in \mathbb{R}^{2n}$, ou seja,

$$\mathcal{I}_n z \mathcal{I}_n^{-1} = x\mathbb{1}_{2n} - yJ_{2n}.$$

Em particular, para $z = i$ isso nos diz que $\mathcal{I}_n i \mathcal{I}_n^{-1} = -J_{2n}$ e, portanto,

$$\mathcal{I}_n^{-1} J_{2n} \mathcal{I}_n = -i. \tag{21.276}$$

resultado que usaremos adiante.

Precisaremos também do seguinte resultado:

Lema 21.6 *Seja $A \in \text{Sp}(2n, \mathbb{R})$. Então, $A \in \text{O}(2n)$ se e somente se $AJ_{2n} = J_{2n}A$. Em outras palavras, $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n) = \mathbb{C}(\{J_{2n}\}, \text{Sp}(2n, \mathbb{R}))$, o centralizador⁷⁶ de $\{J_{2n}\}$ em $\text{Sp}(2n, \mathbb{R})$. \square*

⁷⁶A noção de “centralizador” foi introduzida à página 138

Prova. A pertence a $\text{Sp}(2n, \mathbb{R})$ se e somente se $A^T J_{2n} A = J_{2n}$, ou seja, se e somente se $J_{2n} A^T J_{2n} = -A^{-1}$. Assim, se $A^T = A^{-1}$ tem-se $J_{2n} A^T J_{2n} = -A^T$, o que implica $J_{2n} A^T = A^T J_{2n}$, que por sua vez implica $J_{2n} A = A J_{2n}$, já que $J_{2n}^T = -J_{2n}$. Reciprocamente, se $A^T J_{2n} A = J_{2n}$ e $J_{2n} A = A J_{2n}$, temos $A^T A J_{2n} = J_{2n}$ e, portanto, $A^T A = \mathbb{1}_{2n}$, o que diz-nos que $A^T = A^{-1}$. ■

• **Mais alguns resultados preparatórios**

O lema a seguir contém um resultado técnico que empregaremos adiante.

Lema 21.7 *Com as definições de acima, tem-se*

$$\langle u, \mathcal{I}_n w \rangle_{\mathbb{R}} = \text{Re} \left(\langle \mathcal{I}_n^{-1} u, w \rangle_{\mathbb{C}} \right) \tag{21.277}$$

e

$$\text{Re} \left(\langle w, \mathcal{I}_n^{-1} u \rangle_{\mathbb{C}} \right) = \langle u, \mathcal{I}_n w \rangle_{\mathbb{R}}, \tag{21.278}$$

para todos $u \in \mathbb{R}^{2n}$ e $w \in \mathbb{C}^n$. □

Prova. Escrevendo $w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$, temos $\mathcal{I}_n w = \begin{pmatrix} \text{Re}(w_1) \\ \vdots \\ \text{Re}(w_n) \\ \text{Im}(w_1) \\ \vdots \\ \text{Im}(w_n) \end{pmatrix}$. Assim,

$$\begin{aligned} \langle u, \mathcal{I}_n w \rangle_{\mathbb{R}} &= \sum_{k=1}^n u_k \text{Re}(w_k) + \sum_{k=1}^n u_{n+k} \text{Im}(w_k) = \sum_{k=1}^n (u_k \text{Re}(w_k) + u_{n+k} \text{Im}(w_k)) \\ &= \sum_{k=1}^n \text{Re} \left((u_k - i u_{n+k}) w_k \right) = \text{Re} \left(\sum_{k=1}^n \overline{(u_k + i u_{n+k})} w_k \right) = \text{Re} \left(\langle \mathcal{I}_n^{-1} u, w \rangle_{\mathbb{C}} \right). \end{aligned}$$

A relação (21.277) é imediata a partir de (21.278), pois $\langle w, \mathcal{I}_n^{-1} u \rangle_{\mathbb{C}}$ é o complexo conjugado de $\langle \mathcal{I}_n^{-1} u, w \rangle_{\mathbb{C}}$. ■

Faremos uso também do seguinte resultado elementar:

Lema 21.8 *Seja $M \in \text{Mat}(n, \mathbb{C})$, Então, a aplicação linear real $\mathcal{I}_n M \mathcal{I}_n^{-1} \in \text{Mat}(\mathbb{R}, 2n)$ satisfaz*

$$(\mathcal{I}_n M \mathcal{I}_n^{-1})^T = \mathcal{I}_n M^* \mathcal{I}_n^{-1}. \tag{21.279}$$

□

Prova. Sejam $u, v \in \mathbb{R}^{2n}$. Então,

$$\begin{aligned} \langle u, \mathcal{I}_n M \mathcal{I}_n^{-1} v \rangle_{\mathbb{R}} &\stackrel{(21.277)}{=} \text{Re} \left(\langle \mathcal{I}_n^{-1} u, M \mathcal{I}_n^{-1} v \rangle_{\mathbb{C}} \right) = \text{Re} \left(\langle M^* \mathcal{I}_n^{-1} u, \mathcal{I}_n^{-1} v \rangle_{\mathbb{C}} \right) \\ &\stackrel{(21.278)}{=} \langle v, \mathcal{I}_n M^* \mathcal{I}_n^{-1} u \rangle_{\mathbb{R}} = \langle \mathcal{I}_n M^* \mathcal{I}_n^{-1} u, v \rangle_{\mathbb{R}}, \end{aligned}$$

estabelecendo que $(\mathcal{I}_n M \mathcal{I}_n^{-1})^T = \mathcal{I}_n M^* \mathcal{I}_n^{-1}$. ■

• **Provando o Lema 21.8 de um outro ponto de vista**

A relação (21.279) talvez possa ser melhor entendida se procurarmos expressar adequadamente a matriz $\mathcal{I}_n M \mathcal{I}_n^{-1} \in \text{Mat}(\mathbb{R}, 2n)$ para $M \in \text{Mat}(n, \mathbb{C})$. Se separarmos cada elemento de matriz de M em sua parte real e imaginária,

podemos escrever

$$M = M_R + iM_I .$$

onde $M_R, M_I \in \text{Mat}(n, \mathbb{R})$, sendo seus elementos de matriz dados por $(M_R)_{ij} = \text{Re}(M_{ij})$ e $(M_I)_{ij} = \text{Im}(M_{ij})$. Afirmamos que $\mathcal{I}_n M \mathcal{I}_n^{-1}$ pode ser representada em blocos da seguinte forma:

$$\mathcal{I}_n M \mathcal{I}_n^{-1} = \begin{pmatrix} M_R & -M_I \\ M_I & M_R \end{pmatrix} \in \text{Mat}(\mathbb{R}, 2n) . \quad (21.280)$$

De fato, escrevendo $u \in \mathbb{R}^{2n}$ na forma $u = \begin{pmatrix} U_1 \\ U_2 \end{pmatrix}$, onde $U_1, U_2 \in \mathbb{R}^n$ são dados por $U_1 = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ e $U_2 = \begin{pmatrix} u_{n+1} \\ \vdots \\ u_{2n} \end{pmatrix}$,

teremos

$$\begin{aligned} \mathcal{I}_n M \mathcal{I}_n^{-1} u &= \mathcal{I}_n M (U_1 + iU_2) = \mathcal{I}_n (M_R U_1 - M_I U_2 + i(M_I U_1 + M_R U_2)) \\ &= \begin{pmatrix} M_R U_1 - M_I U_2 \\ M_I U_1 + M_R U_2 \end{pmatrix} = \begin{pmatrix} M_R & -M_I \\ M_I & M_R \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} M_R & -M_I \\ M_I & M_R \end{pmatrix} u , \end{aligned}$$

provando (21.280).

Vê-se com isso que

$$(\mathcal{I}_n M \mathcal{I}_n^{-1})^T = \begin{pmatrix} M_R^T & M_I^T \\ -M_I^T & M_R^T \end{pmatrix} . \quad (21.281)$$

Como, porém, $M^* = M_R^T - iM_I^T$, vemos que o lado direito de (21.281) coincide com $\mathcal{I}_n M^* \mathcal{I}_n^{-1}$, estabelecendo novamente (21.279).

• **O isomorfismo entre $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ e $\text{U}(n)$**

Chegamos agora a um resultado já anunciado sobre o grupo $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ (e demonstrado por outros meios o Teorema 21.8, página 1046).

Proposição 21.34 *Os grupos $\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$ e $\text{U}(n)$ são isomorfos.* □

Prova. Se $A \in \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$, então A respeita a estrutura linear complexa que pode ser definida em \mathbb{R}^{2n} . Essa estrutura é isomorfa à de \mathbb{C}^n pelo isomorfismo $\mathcal{I}_n : \mathbb{C}^n \rightarrow \mathbb{R}^{2n}$. Assim, $\mathcal{I}_n^{-1} A \mathcal{I}_n$ é um operador linear em \mathbb{C}^n , e pode ser identificado como um elemento de $\text{Mat}(n, \mathbb{C})$.

Com isso, podemos aplicar o Lema 21.8, adotando $M = \mathcal{I}_n^{-1} A \mathcal{I}_n$ e obtemos de (21.279) que $A^T = \mathcal{I}_n (\mathcal{I}_n^{-1} A \mathcal{I}_n)^* \mathcal{I}_n^{-1}$, implicando

$$(\mathcal{I}_n^{-1} A \mathcal{I}_n)^* = \mathcal{I}_n^{-1} A^T \mathcal{I}_n . \quad (21.282)$$

Afirmamos que $\mathcal{I}_n^{-1} A \mathcal{I}_n \in \text{U}(n)$, o grupo das matrizes unitárias complexas $n \times n$, introduzido à página 1039. De fato,

$$(\mathcal{I}_n^{-1} A \mathcal{I}_n)^* (\mathcal{I}_n^{-1} A \mathcal{I}_n) \stackrel{(21.282)}{=} (\mathcal{I}_n^{-1} A^T \mathcal{I}_n) (\mathcal{I}_n^{-1} A \mathcal{I}_n) = \mathcal{I}_n^{-1} A^T A \mathcal{I}_n = \mathcal{I}_n^{-1} \mathbb{1}_{2n} \mathcal{I}_n = \mathbb{1}_n ,$$

o que estabelece que $\mathcal{I}_n^{-1} A \mathcal{I}_n$ é uma matriz unitária complexa, ou seja, é um elemento do grupo $\text{U}(n)$.

Vamos agora supor que $K \in \text{U}(n)$ e analisemos $\mathcal{I}_n K \mathcal{I}_n^{-1}$. Trata-se de um operador linear real, pois \mathcal{I}_n e sua inversa são lineares reais. Assim, podemos afirmar que $\mathcal{I}_n K \mathcal{I}_n^{-1} \in \text{Mat}(\mathbb{R}, 2n)$.

Temos que

$$\left(\mathcal{I}_n K \mathcal{I}_n^{-1}\right)^T J_{2n} \left(\mathcal{I}_n K \mathcal{I}_n^{-1}\right) \stackrel{(21.279)}{=} \mathcal{I}_n K^* \left(\mathcal{I}_n^{-1} J_{2n} \mathcal{I}_n\right) K \mathcal{I}_n^{-1} = \mathcal{I}_n(-i) K^* K \mathcal{I}_n^{-1} = \mathcal{I}_n(-1) \mathcal{I}_n^{-1} = J_{2n},$$

estabelecendo que $\mathcal{I}_n K \mathcal{I}_n^{-1} \in \text{Sp}(2n, \mathbb{R})$. Fora isso, tem-se também e

$$\left(\mathcal{I}_n K \mathcal{I}_n^{-1}\right)^T \left(\mathcal{I}_n K \mathcal{I}_n^{-1}\right) \stackrel{(21.279)}{=} \mathcal{I}_n K^* K \mathcal{I}_n^{-1} = \mathcal{I}_n \mathbb{1}_n \mathcal{I}_n^{-1} = \mathbb{1}_{2n},$$

provando que $\mathcal{I}_n K \mathcal{I}_n^{-1} \in \text{O}(2n)$. Estabelecemos, portanto, que $\mathcal{I}_n K \mathcal{I}_n^{-1} \in \text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n)$. Que

$$\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n) \ni A \mapsto \mathcal{I}_n^{-1} A \mathcal{I}_n \in \text{U}(n)$$

é um homomorfismo é bastante evidente, pois $(\mathcal{I}_n^{-1} A \mathcal{I}_n)(\mathcal{I}_n^{-1} A' \mathcal{I}_n) = \mathcal{I}_n^{-1} (AA') \mathcal{I}_n$. Isso completa a demonstração que

$$\text{Sp}(2n, \mathbb{R}) \cap \text{O}(2n) \simeq \text{U}(n),$$

com o isomorfismo sendo definido por \mathcal{I}_n . ■

21.9.2 Grupos Simpléticos. Álgebras de Lie e Parametrização de Cayley

Na Seção 21.4.2.3, página 1069, apresentamos a chamada *parametrização de Cayley* dos grupos $\text{SO}(n)$, tratando com algum detalhe dos casos particulares de $\text{SO}(2)$ e $\text{SO}(3)$. Nesta seção mostraremos que os grupos simpléticos também admitem uma parametrização similar.

• A álgebra de Lie de $\text{sp}(2n, \mathbb{C})$ e $\text{sp}(2n, \mathbb{R})$

Já demonstramos na Proposição 22.6, página 1170, que os grupos simpléticos são grupo de Lie. Um subgrupo uniparamétrico de $\text{Sp}(2n, \mathbb{C})$ possui a forma exponencial $A_t = e^{tX}$, $t \in \mathbb{R}$, com $X \in \text{Mat}(\mathbb{C}, 2n)$ sendo seu gerador infinitesimais (vide Proposição 22.7, página 1171). A condição de pertencimento a $\text{Sp}(2n, \mathbb{C})$ significa $(e^{tX})^T J_{2n} e^{tX} = J_{2n}$, ou seja $e^{tX^T} J_{2n} e^{tX} = J_{2n}$. Diferenciando-se essa igualdade em relação a t e tomando-se $t = 0$, obtemos

$$X^T J_{2n} + J_{2n} X = 0.$$

No caso do grupo $\text{sp}(2n, \mathbb{R})$ temos a mesma conclusão, apenas que X é uma matriz real. Definimos

$$\text{sp}(2n, \mathbb{R}) := \left\{ X \in \text{Mat}(\mathbb{R}, 2n) \mid X^T J_{2n} = -J_{2n} X \right\} \quad \text{e}$$

$$\text{sp}(2n, \mathbb{C}) := \left\{ X \in \text{Mat}(\mathbb{C}, 2n) \mid X^T J_{2n} = -J_{2n} X \right\}.$$

É claro que $\text{sp}(2n, \mathbb{R}) \subset \text{sp}(2n, \mathbb{C})$. Como é esperado, esses dois conjuntos compõem álgebras de Lie em relação ao produto dado pelo comutador, pois se $X, Y \in \text{sp}(2n, \mathbb{C})$ temos

$$[X, Y]^T J_{2n} = Y^T X^T J_{2n} - X^T Y^T J_{2n} = J_{2n} (YX - XY) = -J_{2n} [X, Y].$$

Analogamente para $\text{sp}(2n, \mathbb{R})$.

A álgebra de Lie $\text{sp}(2n, \mathbb{C})$ é uma álgebra de Lie complexa e $\text{sp}(2n, \mathbb{R})$ é uma álgebra de Lie real. Elas são denominadas *álgebras de Lie associadas aos grupos simpléticos* $\text{sp}(2n, \mathbb{C})$ e $\text{sp}(2n, \mathbb{R})$, respectivamente.

Multiplicando-se a relação $X^T J_{2n} = -J_{2n} X$ à esquerda e à direita por J_{2n} , tem-se $J_{2n} X^T = -X J_{2n}$, ou seja, $(X^T)^T J_{2n} = -J_{2n} X^T$. Conclui-se disso que se $X \in \text{sp}(2n, \mathbb{C})$, então $X^T \in \text{sp}(2n, \mathbb{C})$. O mesmo vale para $\text{sp}(2n, \mathbb{R})$. No caso de $\text{sp}(2n, \mathbb{C})$ vale ainda que se $X \in \text{sp}(2n, \mathbb{C})$, então $\overline{X} \in \text{sp}(2n, \mathbb{C})$. Portanto, se $X \in \text{sp}(2n, \mathbb{C})$, então $X^* \in \text{sp}(2n, \mathbb{C})$.

Não é difícil ver que o mapa exponencial $\mathfrak{sp}(2n, \mathbb{C}) \ni X \mapsto e^{tX}$, $t \in \mathbb{R}$, mapeia $\mathfrak{sp}(2n, \mathbb{C})$ em $\mathrm{Sp}(2n, \mathbb{C})$ (e, analogamente, mapeia $\mathfrak{sp}(2n, \mathbb{R})$ em $\mathrm{Sp}(2n, \mathbb{R})$). De fato, para $X \in \mathfrak{sp}(2n, \mathbb{C})$, tem-se $J_{2n}X^k = (-1)^k (X^T)^k J_{2n}$, $k \in \mathbb{N}$ e, portanto,

$$J_{2n}e^{tX} = J_{2n} \left(\mathbb{1} + \sum_{k=1}^{\infty} \frac{t^k}{k!} X^k \right) = \left(\mathbb{1} + \sum_{k=1}^{\infty} \frac{(-1)^k t^k}{k!} (X^T)^k \right) J_{2n} = e^{-tX^T} J_{2n},$$

o que conduz a $(e^{tX})^T J_{2n} e^{tX} = e^{tX^T} J_{2n} e^{tX} = J_{2n}$, provando que $e^{tX} \in \mathrm{Sp}(2n, \mathbb{C})$. Para $\mathfrak{sp}(2n, \mathbb{R})$ o resultado é análogo.

É importante ainda frisar que o mapa exponencial $\mathfrak{sp}(2n, \mathbb{C}) \ni X \mapsto e^{tX} \in \mathrm{Sp}(2n, \mathbb{C})$, $t \in \mathbb{R}$, nem sempre é injetor e/ou sobrejetor.

E. 21.112 Exercício. Mostre que $\mathfrak{sp}(2, \mathbb{C})$ coincide com a álgebra de Lie complexa gerada pelas matrizes de Pauli, que é $\mathfrak{sl}(2, \mathbb{C})$. Use para tal que $J_2 = -i\sigma_2$ e verifique que as três matrizes σ_1, σ_2 e σ_3 satisfazem a relação $(\sigma_k)^T J_2 = -J_2 \sigma_k$, $k = 1, 2, 3$. Nesse caso o mapa exponencial é injetor e sobrejetor. Vide a discussão sobre o grupo $\mathrm{SL}(2, \mathbb{C})$ na Seção 21.4.6, página 1085. \spadesuit

Escrevendo os elementos de $\mathfrak{sp}(2n, \mathbb{C})$ na forma de blocos $X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, com a, b, c e $d \in \mathrm{Mat}(n, \mathbb{C})$, a condição $X^T J_{2n} = -J_{2n} X$ implica $d = -a^T$, $b = b^T$ e $c = c^T$. Assim, $X \in \mathfrak{sp}(2n, \mathbb{C})$ se e somente se possuir a forma de blocos

$$X = \begin{pmatrix} a & b \\ c & -a^T \end{pmatrix}, \quad \text{com} \quad b = b^T \quad \text{e} \quad c = c^T, \quad a, b, c \in \mathrm{Mat}(n, \mathbb{C}). \quad (21.283)$$

O mesmo se dá para os elementos de $\mathfrak{sp}(2n, \mathbb{R})$, agora com a, b e $c \in \mathrm{Mat}(n, \mathbb{R})$.

A forma (21.283) das matrizes de $\mathfrak{sp}(2n, \mathbb{C})$ permite-nos calcular a dimensão desse espaço vetorial complexo. Há n^2 parâmetros complexos fixando as matrizes a ; há $n(n+1)/2$ parâmetros complexos fixando as matrizes simétricas b e mais $n(n+1)/2$ parâmetros complexos fixando as matrizes simétricas c . Assim, a dimensão da álgebra de Lie $\mathfrak{sp}(2n, \mathbb{C})$ como espaço vetorial complexo é

$$\dim(\mathfrak{sp}(2n, \mathbb{C})) = 2n^2 + n.$$

Analogamente, a dimensão da álgebra de Lie $\mathfrak{sp}(2n, \mathbb{R})$ como espaço vetorial real é

$$\dim(\mathfrak{sp}(2n, \mathbb{R})) = 2n^2 + n.$$

• **A álgebra de Lie de $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$**

Os geradores infinitesimais do grupo $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ têm de satisfazer a condição adicional de serem antissimétricos: $X = -X^T$. Dessa forma, a álgebra de Lie correspondente é $\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$ que é

$$\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n) := \left\{ X \in \mathrm{Mat}(\mathbb{R}, 2n) \mid X J_{2n} = J_{2n} X \text{ e } X = -X^T \right\}. \quad (21.284)$$

Observe o leitor que isso é consistente com (21.65), página 1046.

E. 21.113 Exercício. Mostre explicitamente que o conjunto do lado direito de (21.284) é uma álgebra de Lie real para o produto definido pelo comutador de matrizes. \spadesuit

O mapa exponencial é sobrejetor para essa álgebra de Lie, pois o é para o grupo $\mathrm{SO}(2n)$ (lembrar aqui que $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) = \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{SO}(2n)$).

Se retornarmos a (21.283), a imposição adicional de antissimetria implica que $X \in \mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$ se e somente se possuir a forma de blocos

$$X = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad \text{com} \quad b = b^T \quad \text{e} \quad a = -a^T, \quad a, b \in \mathrm{Mat}(n, \mathbb{R}). \quad (21.285)$$

E. 21.114 Exercício importante. Verifique a validade dessa afirmação! *

A forma (21.285) das matrizes de $\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$ permite-nos calcular a dimensão desse espaço vetorial complexo. Há $n(n-1)/2$ parâmetros reais fixando as matrizes antissimétricas a ; há $n(n+1)/2$ parâmetros reais fixando as matrizes simétricas b . Assim, a dimensão da álgebra de Lie $\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$ como espaço vetorial real é

$$\dim(\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)) = n^2.$$

• **A álgebra de Lie de $\mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n)$**

Os geradores infinitesimais do grupo $\mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n)$ têm de satisfazer a condição adicional de serem anti-autoadjuntos: $X = -X^*$. Dessa forma, a álgebra de Lie correspondente é $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ que é

$$\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n) := \left\{ X \in \mathrm{Mat}(\mathbb{C}, 2n) \mid X^T J_{2n} = -J_{2n} X \text{ e } X = -X^* \right\}. \quad (21.286)$$

Como $X = -X^*$, a condição $X^T J_{2n} = -J_{2n} X$ é equivalente a $\overline{X} J_{2n} = J_{2n} X$.

E. 21.115 Exercício. Mostre explicitamente que o conjunto do lado direito de (21.286) é uma álgebra de Lie real para o produto definido pelo comutador de matrizes. *

O mapa exponencial é sobrejetor para essa álgebra de Lie, pois o é para o grupo $\mathrm{SU}(2n)$ (lembrar aqui que $\mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n) = \mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{SU}(2n)$).

E. 21.116 Exercício. Mostre que $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ coincide com a álgebra de Lie real gerada pelas matrizes de Pauli, que é $\mathfrak{su}(2)$. Use para tal que $J_2 = -i\sigma_2$ e verifique que as três matrizes σ_1, σ_2 e σ_3 satisfazem a relação $(\sigma_k)^T J_2 = -J_2 \sigma_k, k = 1, 2, 3$. Nesse caso o mapa exponencial é injetor e sobrejetor. Vide a discussão sobre o grupo $\mathrm{SU}(2)$ na Seção 21.4.4, página 1076. *

Se retornarmos a (21.283), a imposição adicional de antiautoadjunção implica que $X \in \mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ se e somente se possuir a forma de blocos

$$X = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}, \quad \text{com} \quad b = b^T \quad \text{e} \quad a^* = -a, \quad a, b \in \mathrm{Mat}(n, \mathbb{C}). \quad (21.287)$$

E. 21.117 Exercício importante. Verifique a validade dessa afirmação! *

A forma (21.287) das matrizes de $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ permite-nos calcular a dimensão desse espaço vetorial real. Há $n(n-1) + n = n^2$ parâmetros reais fixando as matrizes anti-autoadjuntas a ; há $n(n+1)$ parâmetros reais fixando as matrizes simétricas b . Assim, a dimensão da álgebra de Lie $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ como espaço vetorial real é

$$\dim(\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)) = 2n^2 + n.$$

• **As transformações de Cayley nos grupos simpléticos**

Inspirados nos resultados da Seção 21.4.2.3, página 1069, vamos demonstrar primeiro que a transformada de Cayley mapeia bijetivamente certos subconjuntos $\mathfrak{sp}(2n, \mathbb{C})$ e $\mathrm{Sp}(2n, \mathbb{C})$ (e os conjuntos $\mathfrak{sp}(2n, \mathbb{R})$ e $\mathrm{Sp}(2n, \mathbb{R})$).

No contexto das transformadas de Cayley, da forma como a definimos, uma matriz A é dita ser uma *matriz excepcional* se $\det(\mathbb{1} + A) = 0$, ou seja, se $\mathbb{1} + A$ não possuir inversa. Denotamos por $F(n, \mathbb{R})$ (por $F(n, \mathbb{C})$) o conjunto das matrizes não excepcionais reais (respectivamente, complexas) $n \times n$.

A transformada de Cayley encontra-se definida em (21.138), página 1069. Lá é provado que mapeia bijetivamente matrizes não excepcionais em matrizes não excepcionais.

Se $A \in \mathfrak{sp}(2n, \mathbb{R})$ ou $A \in \mathfrak{sp}(2n, \mathbb{C})$, vale $A^T J_{2n} = -A J_{2n}$ e disso segue que

$$(\mathbb{1} - A^T) J_{2n} = J_{2n} (\mathbb{1} + A) \quad (21.288)$$

e que $J_{2n}(\mathbb{1} + A^T) = (\mathbb{1} - A)J_{2n}$, o que implica

$$(\mathbb{1} + A^T)^{-1} J_{2n} = J_{2n}(\mathbb{1} - A)^{-1}. \quad (21.289)$$

Assim, se $A \in \mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{F}(n, \mathbb{R})$, ou $A \in \mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{F}(n, \mathbb{C})$,

$$\begin{aligned} \mathcal{C}(A)^T J_{2n} \mathcal{C}(A) &= (\mathbb{1} + A^T)^{-1} (\mathbb{1} - A^T) J_{2n} (\mathbb{1} - A) (\mathbb{1} + A)^{-1} \stackrel{(21.288)}{=} (\mathbb{1} + A^T)^{-1} J_{2n} (\mathbb{1} + A) (\mathbb{1} - A) (\mathbb{1} + A)^{-1} \\ &\stackrel{(21.289)}{=} J_{2n} (\mathbb{1} - A)^{-1} (\mathbb{1} + A) (\mathbb{1} - A) (\mathbb{1} + A)^{-1} = J_{2n}. \end{aligned}$$

Tem-se portanto, $\mathcal{C}(A) \in \mathrm{Sp}(2n, \mathbb{R})$ (ou $\mathcal{C}(A) \in \mathrm{Sp}(2n, \mathbb{C})$), caso $A \in \mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{F}(n, \mathbb{R})$, (respectivamente, caso $A \in \mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{F}(n, \mathbb{C})$).

Além disso, se $A \in \mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$ for não excepcional, tem-se adicionalmente $A^T = -A$ e, portanto,

$$\mathcal{C}(A)^T \mathcal{C}(A) = (\mathbb{1} + A^T)^{-1} (\mathbb{1} - A^T) (\mathbb{1} - A) (\mathbb{1} + A)^{-1} = (\mathbb{1} - A)^{-1} (\mathbb{1} + A) (\mathbb{1} - A) (\mathbb{1} + A)^{-1} = \mathbb{1}_{2n}.$$

Assim, $\mathcal{C}(A) \in \mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n)$ caso $A \in \mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n) \cap \mathfrak{F}(\mathbb{R}, n)$.

Conforme já argumentamos na Seção 21.4.2.3, página 1069, os elementos de $\mathfrak{o}(2n)$ são sempre não excepcionais (por serem matrizes antissimétricas), de sorte que $\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n) \cap \mathfrak{F}(\mathbb{R}, n) = \mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$.

Se agora $A \in \mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ for não excepcional, tem-se adicionalmente $A^* = -A$ e, portanto,

$$\mathcal{C}(A)^* \mathcal{C}(A) = (\mathbb{1} + A^*)^{-1} (\mathbb{1} - A^*) (\mathbb{1} - A) (\mathbb{1} + A)^{-1} = (\mathbb{1} - A)^{-1} (\mathbb{1} + A) (\mathbb{1} - A) (\mathbb{1} + A)^{-1} = \mathbb{1}_{2n}.$$

Assim, $\mathcal{C}(A) \in \mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n)$ caso $A \in \mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n) \cap \mathfrak{F}(\mathbb{C}, n)$, com A não excepcional. Matrizes anti-autoadjuntas, porém, também são não excepcionais. De fato, se $A^* = -A$, então iA é autoadjunta e, portanto, possui espectro real, o que significa que o espectro de A é imaginário puro. Assim, $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n) \cap \mathfrak{F}(\mathbb{C}, n) = \mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$.

Reunindo os resultados acima, estabelecemos o seguinte:

Proposição 21.35 *A transformada de Cayley \mathcal{C} mapeia continuamente e bijectivamente*

1. $\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{F}(\mathbb{R}, n)$ e $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathfrak{F}(\mathbb{R}, n)$,
2. $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{F}(\mathbb{C}, n)$ e $\mathrm{Sp}(2n, \mathbb{C}) \cap \mathfrak{F}(\mathbb{C}, n)$,
3. a álgebra de Lie real $\mathfrak{sp}(2n, \mathbb{R}) \cap \mathfrak{o}(2n)$ e $\mathrm{Sp}(2n, \mathbb{R}) \cap \mathrm{O}(2n) \cap \mathfrak{F}(\mathbb{R}, n)$,
4. a álgebra de Lie real $\mathfrak{sp}(2n, \mathbb{C}) \cap \mathfrak{u}(2n)$ e $\mathrm{Sp}(2n, \mathbb{C}) \cap \mathrm{U}(2n) \cap \mathfrak{F}(\mathbb{C}, n)$.

A transformada de Cayley \mathcal{C} é definida em (21.138), página 1069. □

21.9.3 O Teorema de Williamson

Matrizes do grupo simplético podem ser utilizadas para transformar matrizes reais simétricas e positivas de ordem par em uma forma diagonal específica através de uma transformação de congruência (não necessariamente de similaridade!). Esse resultado, conhecido como *Teorema de Williamson*⁷⁷, é relevante na Mecânica Quântica e na Geometria Simplética (i.e., na Mecânica Clássica). Vide e.g., [142].

Mais especificamente, temos:

⁷⁷John Williamson (1901–1949). O trabalho original é: John Williamson, “On the Algebraic Problem Concerning the Normal Forms of Linear Dynamical Systems”. American Journal of Mathematics Vol. 58, N. 1, 141–163 (1936). doi:10.2307/2371062.

Teorema 21.12 (Teorema de Williamson) Para $m \in \mathbb{N}$, seja $M \in \text{Mat}(\mathbb{R}, 2m)$ uma matriz real simétrica positiva. Então, existe uma matriz simplética $S \in \text{Sp}(\mathbb{R}, 2m)$ e uma matriz $m \times m$ diagonal $D \in \text{Mat}(\mathbb{R}, m)$ tais que

$$M = S^T \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} S. \tag{21.290}$$

Acima, D é a matriz diagonal $D = \text{diag}\{\sigma_1, \dots, \sigma_m\}$, onde $(\sigma_1, \dots, \sigma_m)$ é a lista canônica dos valores singulares (definida à página 541) da matriz antissimétrica de ordem par $A := M^{1/2} J_{2n} M^{1/2}$.

Assim, toda matriz simétrica positiva de ordem par pode ser diagonalizada por uma transformação de congruência gerada por uma matriz do grupo simplético.

As quantidades $(\sigma_1, \dots, \sigma_m)$ são frequentemente denominados autovalores simpléticos de M . □

A demonstração desse teorema é apresentada na Seção 10.8.6, página 545, fazendo uso de resultados e definições da Seção 10.8.5, página 538.

21.10 Operadores Diferenciais como Geradores Infinitesimais sobre Funções

Nosso estudo sobre usos de grupos contínuos e suas representações, em Física, não poderia estar completo sem que gastássemos algumas palavras apresentando alguns fatos elementares ligados à ação de certos grupos de maior interesse (como o grupo de translações e o de rotações) em espaços de funções. Esse tema é relevante na Mecânica Quântica, onde essas ações são amplamente usadas, mas também na Mecânica Clássica, na discussão da ação de grupos de simetria e da evolução temporal no espaço de fase.

Consideremos uma situação na qual temos uma função $f : \mathbb{R}^n \rightarrow \mathbb{C}$ ao menos uma vez diferenciável. Seja V_t , $t \in \mathbb{R}$ um grupo uniparamétrico contínuo de matrizes inversíveis agindo em \mathbb{R}^n , satisfazendo, portanto, $V_0 = \mathbb{1}_n$ e $V_t V_{t'} = V_{t+t'}$ para todos $t, t' \in \mathbb{R}$. Seja $J := \left. \frac{d}{dt} V_t \right|_{t=0}$ o gerador infinitesimal desse grupo uniparamétrico, de sorte que temos $V_t = e^{tJ}$ para todo $t \in \mathbb{R}$, como já esclarecemos. A expressão

$$(\mathcal{V}_t f)(x) := f(V_{-t}x), \quad x \in \mathbb{R}^n, \quad t \in \mathbb{R} \tag{21.291}$$

define uma ação do grupo uniparamétrico sobre a função f . Temos, pela regra da cadeia,

$$\frac{d}{dt} (\mathcal{V}_t f)(x) = \sum_{k=1}^n \left[\left(\frac{d}{dt} V_{-t} \right)_k x \right] \frac{\partial f}{\partial x^k} (V_{-t}x). \tag{21.292}$$

Tomando-se $t = 0$, temos

$$\left. \frac{d}{dt} (\mathcal{V}_t f)(x) \right|_{t=0} = - \sum_{k=1}^n (Jx)_k \frac{\partial f}{\partial x^k} (x). \tag{21.293}$$

Dessa forma, poderemos identificar o operador diferencial

$$\mathcal{J} := - \sum_{k=1}^n (Jx)_k \frac{\partial}{\partial x^k} = -(Jx) \cdot \vec{\nabla} \tag{21.294}$$

como sendo o *gerador infinitesimal* da representação \mathcal{V} no espaço das funções diferenciáveis em \mathbb{R}^n , com o que temos, em um sentido que pode ser formalizado

$$\mathcal{V}_t = \exp(t\mathcal{J}), \quad t \in \mathbb{R}. \tag{21.295}$$

Não vamos aqui atribuir um significado preciso à exponenciação do operador diferencial \mathcal{J} , o que demandaria mais restrições sobre a função f e um mergulho na teoria dos operadores definidos em espaços de Fréchet ou de Banach. O

ponto que queremos destacar é que a relação acima é formalmente consistente com a noção de que o operador diferencial \mathcal{J} é o gerador infinitesimal do grupo uniparamétrico \mathcal{V}_t , $t \in \mathbb{R}$ sobre funções como f .

• O grupo de translações em \mathbb{R} e \mathbb{R}^n

Seja f uma função definida em \mathbb{R} assumindo valores em \mathbb{R} ou \mathbb{C} . Em \mathbb{R} age naturalmente o grupo de translações $(\mathbb{R}, +)$, o grupo aditivo dos reais. Segundo a definição geral (vide Seção 2.1.9.1, página 117), uma ação desse grupo sobre f é dada por $(T_a f)(x) := f(x - a)$, $a \in \mathbb{R}$. Afirmamos que, ao menos para funções diferenciáveis, o operador $-\frac{d}{dx}$ é o gerador infinitesimal das translações. De fato, $\lim_{a \rightarrow 0} a^{-1} ((T_a f)(x) - f(x)) \Big|_{a=0} = -f'(x)$.

Se f satisfizer condições adequadas, como a de ser real-analítica, ela pode ser expandida em série de Taylor e, assim, $f(x - a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} (-a)^n$. Com isso, podemos escrever de forma simbólica

$$(T_a f)(x) = f(x - a) = \left(\left(\sum_{n=0}^{\infty} \frac{(-a)^n}{n!} \frac{d^n}{dx^n} \right) f \right) (x) = \left(e^{-a \frac{d}{dx}} f \right) (x).$$

Não vamos aqui atribuir um significado preciso à exponenciação do operador de diferenciação $\frac{d}{dx}$, ou seja, não vamos discutir de forma precisa em que sentido a série que define essa exponencial é convergente, em quais tipos de funções ela age etc. Isso pode ser feito de diversas formas, em espaços de funções diferentes, e trataremos desse tipo de questão em outros capítulos. O ponto que queremos destacar é que essa relação é consistente com a noção de que o operador diferencial $-\frac{d}{dx}$ é o gerador infinitesimal das translações no espaço de funções considerado, pois ela nos mostra que T_a é obtido pela exponenciação de a vezes o gerador infinitesimal $-\frac{d}{dx}$.

Essa caracterização de $-\frac{d}{dx}$ como o gerador infinitesimal de translações é de grande relevância, por exemplo, na Mecânica Quântica.

O leitor facilmente se apercebe que a análise acima pode ser estendida para a ação de translações sobre funções diferenciáveis em \mathbb{R}^n , com o gerador infinitesimal sendo agora $-\vec{\nabla} \equiv -\text{grad}$.

Comentário. O estudante pode estar se perguntando como o formalismo desenvolvido entre (21.291) e (21.295) se aplica nesse contexto. Para tal, podemos proceder da seguinte forma. Primeiramente identificamos \mathbb{R} com o subconjunto \mathbf{R} de \mathbb{R}^2 dado por $\mathbf{R} = \left\{ \begin{pmatrix} x \\ 1 \end{pmatrix}, x \in \mathbb{R} \right\}$ (note-se que \mathbf{R} não é um subespaço vetorial de \mathbb{R}^2). Assim, identificamos a função $f\left(\begin{pmatrix} x \\ 1 \end{pmatrix}\right)$ com $f(x)$.

Com essas identificações, a ação das translações é realizada pelas matrizes $V_t = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, $t \in \mathbb{R}$. De fato,

$$V_t \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} x + t \\ 1 \end{pmatrix}, \quad \text{sendo que} \quad V_0 = \mathbb{1}_2 \quad \text{e} \quad V_t V_{t'} = V_{t+t'}, \quad t, t' \in \mathbb{R}.$$

Entenda-se que as matrizes V_t não são operadores lineares em \mathbf{R} , pois \mathbf{R} sequer é um subespaço vetorial de \mathbb{R}^2 ! As matrizes V_t representam, isto sim, ações por matrizes do grupo de translações sobre \mathbf{R} .

O gerador infinitesimal desse grupo uniparamétrico de matrizes é $J = \frac{d}{dt} V_t \Big|_{t=0} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, sendo que $J\left(\begin{pmatrix} x \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Entenda-se novamente que J não é um operador linear em \mathbf{R} , pois \mathbf{R} sequer é um subespaço vetorial de \mathbb{R}^2 !

Com isso, as funções $f : \mathbb{R} \rightarrow \mathbb{C}$ devem ser entendidas como funções de \mathbf{R} em \mathbb{C} e as translações tornam-se

$$\mathcal{V}_t f\left(\begin{pmatrix} x \\ 1 \end{pmatrix}\right) := f\left(V_{-t}\left(\begin{pmatrix} x \\ 1 \end{pmatrix}\right)\right) = f\left(\begin{pmatrix} x-t \\ 1 \end{pmatrix}\right).$$

A expressão (21.294) para o gerador infinitesimal sobre as funções fica simplesmente $\mathcal{J} = -\frac{d}{dx}$, onde $\frac{d}{dx} f\left(\begin{pmatrix} x \\ 1 \end{pmatrix}\right) := f'\left(\begin{pmatrix} x \\ 1 \end{pmatrix}\right)$. ♣

• O grupo de dilatações agindo em funções de \mathbb{R}^n

O grupo multiplicativo dos reais positivos, ou grupo de dilatações, (\mathbb{R}_+, \cdot) , pode ser feito agir em \mathbb{R}^n através da multiplicação de vetores de \mathbb{R}^n por números da forma e^t , $t \in \mathbb{R}$: $\mathbb{R}^n \ni \vec{x} \mapsto e^t \vec{x}$, onde $\vec{x} \equiv (x_1, \dots, x_n) \in \mathbb{R}^n$ and $t \in \mathbb{R}$. Assim, se f é uma função em \mathbb{R}^n , temos $(\mathcal{V}_t f)(\vec{x}) := f(e^{-t} \vec{x})$ e, se f for diferenciável,

$$\frac{d}{dt} (\mathcal{V}_t f)(\vec{x}) \Big|_{t=0} = -\vec{x} \cdot (\vec{\nabla} f)(\vec{x}) \quad \text{e, portanto,} \quad \mathcal{J} \equiv -\vec{x} \cdot \vec{\nabla} = -x_1 \frac{\partial}{\partial x_1} - \dots - x_n \frac{\partial}{\partial x_n} \quad (21.296)$$

é o gerador infinitesimal das dilatações em funções diferenciáveis definidas em \mathbb{R}^n . Podemos, assim, escrever formalmente

$$(\mathcal{V}_t f)(\vec{x}) := f(e^{-t} \vec{x}) = \left(\exp(t\mathcal{J}) f \right) (\vec{x}),$$

com \mathcal{J} dado em (21.296). Como antes, deixamos aberto o sentido preciso em que essa exponenciação de \mathcal{J} é definida, e para quais funções.

• **O grupo $SO(2)$ agindo em funções de \mathbb{R}^2**

Seja f uma função definida em \mathbb{R}^2 assumindo valores em \mathbb{R} ou \mathbb{C} . Denotaremos os vetores de \mathbb{R}^2 por \vec{x} ou diretamente em termos de suas componentes pelo par ordenado $(x, y) \equiv \vec{x}$. O grupo $SO(2)$ age no domínio de definição dessa função e seja um elemento de $SO(2)$ da forma $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. Por convenção, $R(\theta)$ roda cada vetor \vec{x} de \mathbb{R}^2 de um ângulo $\theta \in (-\pi, \pi]$ no sentido anti-horário em torno da origem.

Sua ação na função f é, por definição, dada por $(T_\theta f)(\vec{x}) := f(R(-\theta)\vec{x})$. Escrevendo em termos de componentes, $(T_\theta f)(x, y) = f((\cos \theta)x + (\sin \theta)y, -(\sin \theta)x + (\cos \theta)y)$. O gerador infinitesimal associado à ação T_θ é definido por

$$\begin{aligned} (\mathcal{J}f)(x, y) &:= \left. \frac{d}{d\theta}(T_\theta f)(x, y) \right|_{\theta=0} \\ &= \left[\begin{aligned} &(-(\sin \theta)x + (\cos \theta)y) \frac{\partial f}{\partial x} \left((\cos \theta)x + (\sin \theta)y, -(\sin \theta)x + (\cos \theta)y \right) \\ &+ (-(\cos \theta)x - (\sin \theta)y) \frac{\partial f}{\partial y} \left((\cos \theta)x + (\sin \theta)y, -(\sin \theta)x + (\cos \theta)y \right) \end{aligned} \right] \Big|_{\theta=0} \\ &= \left(\left[y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right] f \right)(x, y), \end{aligned}$$

ou seja, $\mathcal{J} = y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}$ e, por conseguinte,

$$\left(e^{\theta \mathcal{J}} f \right)(\vec{x}) = f(R(-\theta)\vec{x}).$$

Essa relação decorre da afirmação, já utilizada várias vezes, que todo grupo uniparamétrico pode ser escrito como exponencial de seu gerador infinitesimal (vide, por exemplo a Proposição 22.7, página 1171), mas pode também ser obtida formalmente da expansão em série de Taylor de $(T_\theta f)(x, y)$ como função da variável θ em torno de $\theta = 0$.

E. 21.118 *Exercício.* Mostre isso! *

Novamente omitiremos a discussão sobre a convergência da série que define a exponencial do operador diferencial \mathcal{J} .

Fazemos ainda notar que a expressão acima para \mathcal{J} é realmente idêntica à que temos por (21.294). Ali, $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ e, assim, $J \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$. Com isso, o lado direito de (21.294) fica $y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}$, que coincide com nossa expressão para \mathcal{J} .

• **O grupo $SO(3)$ agindo em funções de \mathbb{R}^3**

Seja f uma função definida em \mathbb{R}^3 assumindo valores em \mathbb{R} ou \mathbb{C} . A ação de uma rotação anti-horária de um ângulo θ em torno do eixo definido por um versor $\vec{\eta}$, fixo, é

$$\left(T_{\theta, \vec{\eta}} f \right)(\vec{x}) := f \left(e^{-\theta \vec{\eta} \cdot \vec{J}} \vec{x} \right).$$

Para $\vec{\eta}$ fixo $e^{\theta \vec{\eta} \cdot \vec{J}}$ é da ação de um grupo uniparamétrico em \mathbb{R}^3 , que tem por gerador infinitesimal $J_{\vec{\eta}} = \vec{\eta} \cdot \vec{J}$. Assim, por A expressão (21.294) para o gerador infinitesimal sobre as funções fica simplesmente

$$\mathcal{J}_{\vec{\eta}} = - \sum_{k=1}^3 \left((\vec{\eta} \cdot \vec{J}) \vec{x} \right)_k \frac{\partial}{\partial x_k}$$

onde $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$. Sabemos por (21.112), página 1059, que $(\vec{\eta} \cdot \vec{J}) \vec{x} = \vec{\eta} \times \vec{x}$. Assim, $\mathcal{J}_{\vec{\eta}} = -(\vec{\eta} \times \vec{x}) \cdot \vec{\nabla}$. Com isso,

$$\left(\mathcal{J}_{\vec{\eta}} f \right)(\vec{x}) = -(\vec{\eta} \times \vec{x}) \cdot \left(\vec{\nabla} f \right)(\vec{x}) \stackrel{(4.13)}{=} -\vec{\eta} \cdot \left(\vec{x} \times \vec{\nabla} f \right)(\vec{x}).$$

Notemos *en passant* que $-(\vec{x} \times \vec{\nabla} f)(\vec{x}) = \vec{\nabla} \times (f(\vec{x}) \vec{x})$, pela identidade (4.21), página 261.

Concluimos disso que, para cada versor $\vec{\eta}$, o gerador infinitesimal das rotações é

$$\mathcal{J}_{\vec{\eta}} = -\vec{\eta} \cdot (\vec{x} \times \vec{\nabla}). \tag{21.297}$$

Portanto, formalmente,

$$(e^{\theta \mathcal{J}_{\vec{\eta}}} f)(\vec{x}) = f(e^{-\theta \vec{\eta} \cdot \vec{J}} \vec{x}) = f(R(-\theta, \vec{\eta}) \vec{x}).$$

Na Mecânica Quântica, os operadores $\mathcal{J}_{\vec{\eta}}$ são associados (após multiplicação por $-i$) com a projeção do momento angular na direção $\vec{\eta}$.

Segundo (21.297), os geradores infinitesimais das rotações espaciais em torno das direções 1, 2 e 3 sobre as funções em \mathbb{R}^4 são, respectivamente,

$$\mathcal{J}_1 = x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3}, \quad \mathcal{J}_2 = x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1}, \quad \mathcal{J}_3 = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}.$$

Naturalmente, $\mathcal{J}_{\vec{\eta}} = \sum_{k=1}^3 \eta_k \mathcal{J}_k$.

E. 21.119 *Exercício.* Verifique também que os operadores diferenciais \mathcal{J}_a , $a = 1, 2, 3$, satisfazem as relações de comutação formais

$$[\mathcal{J}_a, \mathcal{J}_b] = \sum_{c=1}^3 \varepsilon_{abc} \mathcal{J}_c,$$

ou, equivalentemente,

$$\left[-\vec{\alpha} \cdot (\vec{x} \times \vec{\nabla}), -\vec{\beta} \cdot (\vec{x} \times \vec{\nabla}) \right] = -(\vec{\alpha} \times \vec{\beta}) \cdot (\vec{x} \times \vec{\nabla}), \quad \text{ou seja,} \quad [\mathcal{J}_{\vec{\alpha}}, \mathcal{J}_{\vec{\beta}}] = \mathcal{J}_{\vec{\alpha} \times \vec{\beta}},$$

onde $\vec{\alpha}$ e $\vec{\beta}$ são vetores unitários constantes de \mathbb{R}^3 . As diversas relações acima caracterizam a álgebra de Lie do grupo $SO(3)$. ♣

Comentário. O leitor que não deseje evocar a expressão (21.294) para o gerador infinitesimal pode seguir a seguinte abordagem direta. Temos,

$$(T_{\theta, \vec{\eta}} f)(\vec{x}) := f(e^{-\theta \vec{\eta} \cdot \vec{J}} \vec{x}) = f\left(\vec{x} + (1 - \cos(\theta))(\vec{\eta} \times (\vec{\eta} \times \vec{x})) - \sin(\theta) \vec{\eta} \times \vec{x}\right)$$

onde, na última igualdade, usamos a fórmula de Rodrigues (21.105), página 1057. Assumindo f diferenciável, a derivada da expressão do lado direito em relação a θ é

$$\left(\sin(\theta)(\vec{\eta} \times (\vec{\eta} \times \vec{x})) - \cos(\theta) \vec{\eta} \times \vec{x} \right) \cdot (\vec{\nabla} f) \left(\vec{x} + (1 - \cos(\theta))(\vec{\eta} \times (\vec{\eta} \times \vec{x})) - \sin(\theta) \vec{\eta} \times \vec{x} \right)$$

e essa expressão, quando calculada em $\theta = 0$, torna-se $-(\vec{\eta} \times \vec{x}) \cdot (\vec{\nabla} f)(\vec{x}) \stackrel{(4.13)}{=} -\vec{\eta} \cdot (\vec{x} \times \vec{\nabla} f)(\vec{x}) = (\mathcal{J}_{\vec{\eta}} f)(\vec{x})$. Notemos *en passant* que $-(\vec{x} \times \vec{\nabla} f)(\vec{x}) = \vec{\nabla} \times (f(\vec{x}) \vec{x})$, pela identidade (4.21), página 261. ♣

• **A ação de boosts de Lorentz e rotações sobre funções em \mathbb{R}^4**

Pelo Corolário 21.7, página 1121, sabemos que o gerador infinitesimal do subgrupo uniparamétrico $\{\mathcal{B}(z, \vec{\xi}), z \in \mathbb{R}\}$ de boosts de Lorentz em uma direção fixada por um vetor unitário $\vec{\xi} = (\xi_1, \xi_2, \xi_3) \in \mathbb{R}^3$ é dado por $\vec{\xi} \cdot \vec{\mathcal{M}}$, sendo que, por (21.217), página 1117,

$$\vec{\xi} \cdot \vec{\mathcal{M}} = \begin{pmatrix} 0 & -\xi_1 & -\xi_2 & -\xi_3 \\ -\xi_1 & 0 & 0 & 0 \\ -\xi_2 & 0 & 0 & 0 \\ -\xi_3 & 0 & 0 & 0 \end{pmatrix}.$$

Agindo sobre uma função $f : \mathbb{R}^4 \rightarrow \mathbb{C}$, os boosts são expressos por um operador, que denotaremos por $\mathcal{B}(z, \vec{\xi})$, da seguinte forma:

$$(\mathcal{B}(z, \vec{\xi}) f)(x) := f(\mathcal{B}(-z, \vec{\xi}) x) = f(\exp(-z \vec{\xi} \cdot \vec{\mathcal{M}}) x),$$

com $z \in \mathbb{R}$ e $x = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_0 \\ \vec{x} \end{pmatrix} \in \mathbb{R}^4$.

De acordo com (21.294), página 1142, o gerador infinitesimal do subgrupo uniparamétrico $B(z, \vec{\xi})$, $z \in \mathbb{R}$, é dado por

$$M_{\vec{\xi}} := - \sum_{k=0}^3 \left[(\vec{\xi} \cdot \vec{M}) x \right]_k \frac{\partial}{\partial x^k}. \quad (21.298)$$

É fácil ver que $(\vec{\xi} \cdot \vec{M}) x = - \left(\frac{\vec{\xi} \cdot \vec{x}}{x_0 \vec{\xi}} \right)$. Portanto,

$$M_{\vec{\xi}} = (\vec{\xi} \cdot \vec{x}) \frac{\partial}{\partial x_0} + x_0 \vec{\xi} \cdot \vec{\nabla}, \quad (21.299)$$

onde, naturalmente, $\vec{\xi} \cdot \vec{\nabla} = \xi_1 \frac{\partial}{\partial x_1} + \xi_2 \frac{\partial}{\partial x_2} + \xi_3 \frac{\partial}{\partial x_3}$.

Por exemplo, tomando-se sucessivamente $\vec{\xi} = (1, 0, 0)$, $(0, 1, 0)$ e $(0, 0, 1)$, temos os geradores infinitesimais

$$M_1 = x_1 \frac{\partial}{\partial x_0} + x_0 \frac{\partial}{\partial x_1}, \quad M_2 = x_2 \frac{\partial}{\partial x_0} + x_0 \frac{\partial}{\partial x_2}, \quad M_3 = x_3 \frac{\partial}{\partial x_0} + x_0 \frac{\partial}{\partial x_3},$$

que geram sobre as funções os *boosts* nas direções 1, 2 e 3, respectivamente. Observe-se ainda que $M_{\vec{\xi}} = \sum_{k=1}^3 \xi_k M_k$.

As expressões acima são empregadas na Mecânica Quântica Relativística e na Teoria Quântica de Campos.

E. 21.120 *Exercício importante.* Mostre que os geradores infinitesimais das rotações espaciais em torno das direções 1, 2 e 3 sobre as funções em \mathbb{R}^4 são, respectivamente,

$$J_1 = x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3}, \quad J_2 = x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1}, \quad J_3 = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}.$$

Verifique também que os operadores diferenciais M_a e J_b , $a, b = 1, 2, 3$, satisfazem as relações de comutação que caracterizam a álgebra de Lie do grupo de Lorentz próprio ortócrono \mathcal{L}_\uparrow^+ :

$$[J_a, J_b] = \sum_{c=1}^3 \varepsilon_{abc} J_c, \quad (21.300)$$

$$[M_a, M_b] = - \sum_{c=1}^3 \varepsilon_{abc} J_c, \quad (21.301)$$

$$[J_a, M_b] = \sum_{c=1}^3 \varepsilon_{abc} M_c. \quad (21.302)$$

Compare a (21.219)-(21.221), página 1119. ✦

E. 21.121 *Exercício.* Procedendo analogamente, obtenha os geradores infinitesimais sobre funções das transformações de Galilei. ✦

21.11 Exercícios Adicionais

E. 21.122 *Exercício.* a. Usando (21.24), página 1030, verifique a validade da fórmula de Lie-Trotter (relação (11.41), página 585):

$$\exp(A + B) = \lim_{m \rightarrow \infty} \left[\exp\left(\frac{1}{m}A\right) \exp\left(\frac{1}{m}B\right) \right]^m$$

para quando A e B são elementos da álgebra de Heisenberg $\mathfrak{gh}_3(\mathbb{C})$, ou seja, $A = h(a, b, c)$ e $B = h(a', b', c')$.

b. Nas mesmas circunstâncias, e também usando (21.24), verifique a fórmula do comutador (relação (11.42), página 585):

$$\exp([A, B]) = \lim_{m \rightarrow \infty} \left[\exp\left(\frac{1}{m}A\right) \exp\left(\frac{1}{m}B\right) \exp\left(-\frac{1}{m}A\right) \exp\left(-\frac{1}{m}B\right) \right]^{m^2}.$$

✱

E. 21.123 *Exercício.* A fórmula de Rodrigues (21.103) também pode ser obtida com o uso do Cálculo Funcional (Teorema Espectral). Seja $M \equiv \vec{\eta} \cdot \vec{J}$, com $\vec{\eta}$ sendo um vetor unitário de \mathbb{R}^3 . Obtenha explicitamente o polinômio característico de M e mostre que seus autovalores são $\alpha_1 = 0$, $\alpha_2 = i$ e $\alpha_3 = -i$. Por serem distintos, conclua que M é diagonalizável e que, portanto, vale para a mesma o Teorema Espectral, Teorema 10.7, página 480. Usando (10.60), página 483, mostre que seus respectivos projetores espectrais são

$$E_1 = M^2 + \mathbb{1}, \quad E_2 = -\frac{1}{2}(M^2 + iM) \quad \text{e} \quad E_3 = -\frac{1}{2}(M^2 - iM).$$

Constatare que $\mathbb{1} = \sum_{k=1}^3 E_k$ e que $M = \sum_{k=1}^3 \alpha_k E_k = i(E_2 - E_3)$. Conclua também que

$$e^{\theta M} = \sum_{k=1}^3 e^{\alpha_k \theta} E_k = E_1 + e^{i\theta} E_2 + e^{-i\theta} E_3 = \mathbb{1} + \text{sen}(\theta)M + (1 - \cos(\theta))M^2,$$

que é a fórmula de Rodrigues (21.103).

✱

E. 21.124 *Exercício.* Usando apenas considerações geométricas, deduza que o efeito de $R(\theta, \vec{\eta})$ (uma rotação de um ângulo θ no sentido anti-horário em torno de um eixo definido por um vetor unitário $\vec{\eta} \in \mathbb{R}^3$) em um vetor qualquer $\vec{\alpha} \in \mathbb{R}^3$ é, de fato, dado por (21.106).

Sugestão: Suponha primeiramente que $\vec{\alpha}$ e $\vec{\eta}$ não sejam paralelos e defina uma base ortogonal em \mathbb{R}^3 constituída pelos vetores $\vec{\eta}$, $\vec{\eta} \times \vec{\alpha}$ e $\vec{\eta} \times (\vec{\eta} \times \vec{\alpha})$. Normalize esses vetores e exprima $\vec{\alpha}$ em termos dessa base ortonormalizada. Efetue a rotação de um ângulo θ do vetor $\vec{\alpha}$ em torno de $\vec{\eta}$ observando que a componente de $\vec{\alpha}$ na direção de $\vec{\eta}$ permanece invariante, enquanto que as componentes ao longo de $\vec{\eta} \times (\vec{\eta} \times \vec{\alpha})$ e $\vec{\eta} \times \vec{\alpha}$ são rodadas de θ . Por fim, convença-se que (21.106) é trivialmente válida caso $\vec{\alpha}$ e $\vec{\eta}$ sejam paralelos. ✱

E. 21.125 *Exercício.* Uma forma alternativa de se escrever a álgebra de Lie do grupo $\text{SO}(3)$ é obtida definindo-se

$$J_{k\ell} := \sum_{m=1}^3 \varepsilon_{k\ell m} J_m = \varepsilon_{k\ell m} J_m$$

como o gerador infinitesimal de rotação no plano k - ℓ . Acima, $\varepsilon_{k\ell m}$ é o símbolo totalmente antissimétrico de Levi-Civita, J_m representa o gerador infinitesimal usual de rotação em torno do eixo m e, na última igualdade, empregou-se a convenção de soma para índices repetidos. Mostre que, em termos dos geradores infinitesimais $J_{k\ell}$, a álgebra de Lie de $\text{SO}(3)$, $[J_k, J_\ell] = \varepsilon_{k\ell m} J_m$, pode ser escrita como

$$[J_{k\ell}, J_{mn}] = \delta_{km} J_{\ell n} - \delta_{kn} J_{\ell m} - \delta_{\ell m} J_{kn} + \delta_{\ell n} J_{km},$$

onde δ_{km} é o delta de Kronecker. (*Sugestão:* É desaconselhável, embora possível, verificar as relações de comutação explicitamente caso-a-caso. O mais prático é explorar as bem conhecidas propriedades (4.4)–(4.10) do símbolo de Levi-Civita $\varepsilon_{k\ell m}$ obtidas na Seção 4.1, página 258.

Observação. A utilidade de se escrever a álgebra de $\text{SO}(3)$ em termos dos geradores infinitesimais $J_{k\ell}$ está no fato de podermos generalizá-la para dimensões maiores de uma forma direta. Por exemplo, a álgebra de Lie do grupo de Lorentz próprio ortócrono $\mathcal{L}_+^\uparrow \subset \text{SO}(1, 3)$ pode ser escrita como

$$[J_{\mu\nu}, J_{\lambda\sigma}] = -\eta_{\mu\lambda} J_{\nu\sigma} + \eta_{\mu\sigma} J_{\nu\lambda} + \eta_{\nu\lambda} J_{\mu\sigma} - \eta_{\nu\sigma} J_{\mu\lambda},$$

com $\mu, \nu, \lambda, \sigma \in \{0, 1, 2, 3\}$, onde $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ é o tensor métrico de Minkowski e onde, adicionalmente, $J_{00} = 0$ e $J_{0k} = -J_{k0} = \mathcal{M}_k$, $k = 1, 2, 3$. Note que, assim, um *boost* na direção k “equivale” à uma “rotação” no plano formado pela direção (espacial) k e pela direção temporal 0. ✱

E. 21.126 *Exercício.* Adaptado de [284]. Este exercício fornece uma alternativa ao homomorfismo dado em (21.168). Já vimos que um elemento genérico de $SU(2)$ é da forma $U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$, sendo $a, b \in \mathbb{C}$ com $|a|^2 + |b|^2 = 1$. Mostre que a aplicação ϕ' definida por

$$\phi' \left(\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \right) := \begin{pmatrix} \frac{1}{2}(\bar{a}^2 - \bar{b}^2 + a^2 - b^2) & \frac{i}{2}(\bar{a}^2 + \bar{b}^2 - a^2 - b^2) & -\bar{a}\bar{b} - ab \\ \frac{i}{2}(-a^2 + b^2 + \bar{a}^2 - \bar{b}^2) & \frac{1}{2}(\bar{a}^2 + \bar{b}^2 + a^2 + b^2) & i(ab - \bar{a}\bar{b}) \\ \bar{a}\bar{b} + ab & i(\bar{a}\bar{b} - ab) & \bar{a}\bar{a} - b\bar{b} \end{pmatrix}$$

é um epimorfismo (um homomorfismo sobrejetor) de $SU(2)$ em $SO(3)$ e mostre que $\text{Ker } \phi' = \{\mathbb{1}, -\mathbb{1}\}$. ✦

E. 21.127 *Exercício.* Sejam $\mathcal{B} := \{B(t) \in GL(\mathbb{C}, n), t \in \mathbb{R}\}$ e $\mathcal{C} := \{C(s) \in GL(n, \mathbb{C}), s \in \mathbb{R}\}$ dois subgrupos uniparamétricos de matrizes $n \times n$ e suponha que valha

$$B(t)C(s) = C(e^{\lambda t}s)B(t), \tag{21.303}$$

para todos $t, s \in \mathbb{R}$, com $\lambda \in \mathbb{R}$ sendo uma constante. Mostre que o conjunto $\mathcal{A} := \{B(t)C(s), t, s \in \mathbb{R}\}$, composto por todos os possíveis produtos de elementos de \mathcal{B} e de \mathcal{C} (nesta ordem), forma um subgrupo de $GL(n, \mathbb{C})$.

O grupo \mathcal{A} , acima, é denominado *grupo de Anosov*⁷⁸ e está relacionado a uma classe importante de sistemas dinâmicos hiperbólicos, denominados *sistemas de Anosov* ou *fluxos de Anosov*. A relação (21.303) é por vezes denominada *relação de Anosov*. A constante λ é denominada *constante de Lyapunov*⁷⁹, nesse contexto.

Um exemplo concreto (com $\lambda = 1$) é fornecido pelas matrizes 3×3

$$B(t) = \begin{pmatrix} \cosh t & 0 & -\sinh t \\ 0 & 1 & 0 \\ -\sinh t & 0 & \cosh t \end{pmatrix} \quad \text{e} \quad C(s) = \begin{pmatrix} 1 + \frac{s^2}{2} & s & \frac{s^2}{2} \\ s & 1 & s \\ -\frac{s^2}{2} & -s & 1 - \frac{s^2}{2} \end{pmatrix}.$$

Verifique que $B(t)B(t') = B(t+t')$, que $C(s)C(s') = C(s+s')$, que $B(0) = C(0) = \mathbb{1}$ e, mais importante, verifique que a propriedade (21.303) é satisfeita nesse caso, na forma $B(t)C(s) = C(e^t s)B(t)$, para todos $t, s \in \mathbb{R}$.

As matrizes $B(t)$ e $C(s)$, acima, são elementos do grupo de Lorentz em $2+1$ dimensões. As matrizes $B(t)$ implementam um “boost” de Lorentz na direção 2 e as matrizes $C(s)$ implementam as chamadas *translações horosféricas*, que são transformações de Lorentz que mantêm invariante um raio de luz, no caso, o raio de luz que aponta na direção $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Vide também o Exercício E. 11.28, página 610. Naquele exercício as translações horosféricas mantêm invariante o raio de luz que aponta na direção $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

Um outro exemplo com matrizes 2×2 , também com $\lambda = 1$, é $B(t) = \begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{pmatrix}$ e $C(s) = \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}$. Verifique! ✦

⁷⁸Dmitri Victorovich Anosov (1936–2014).

⁷⁹Aleksandr Mikhailovich Lyapunov (1857–1918).

Apêndices

21.A Extensão do Lema 21.3 e do Teorema 21.7 ao Caso Complexo

O Lema 21.3 e o Teorema 21.7, páginas 1034 e 1035, respectivamente, possuem versões para formas sesquilineares não degeneradas em espaços vetoriais complexos. Trataremos disso na seção corrente.

Lema 21.9 *Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \rightarrow \mathbb{C}$. Seja q a forma quadrática associada a ω , definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $R : V \rightarrow V$ uma função sobrejetora satisfazendo $R(iv) = \pm iR(v)$. Então, R satisfaz as condições*

$$R(0) = 0 \quad e \quad q(R(u) - R(v)) = q(u - v) \quad \text{para todos } u, v \in V \quad (21.A.1)$$

se e somente se satisfizer a condição (caso “+”)

$$\omega(R(u), R(v)) = \omega(u, v) \quad (21.A.2)$$

para todos $u, v \in V$, ou a condição (caso “-”)

$$\omega(R(u), R(v)) = \omega(v, u), \quad (21.A.3)$$

também para todos $u, v \in V$. □

Prova do Lema 21.9. A prova é dividida em duas partes

I. Vamos supor a validade de qualquer um dos dois casos (21.A.2) ou (21.A.3). Tomando-se nela $v = 0$, temos que $\omega(R(u), R(0)) = 0$ para todo $u \in V$, o que implica $R(0) = 0$ pois ω é não degenerada e R é sobrejetora. Além disso,

$$\begin{aligned} q(R(u) - R(v)) &= \omega(R(u) - R(v), R(u) - R(v)) = \omega(R(u), R(u)) + \omega(R(v), R(v)) - \omega(R(u), R(v)) - \omega(R(v), R(u)) \\ &\stackrel{(21.A.2) \text{ ou } (21.A.3)}{=} \omega(u, u) + \omega(v, v) - \omega(u, v) - \omega(v, u) = \omega(u - v, u - v) = q(u - v), \end{aligned} \quad (21.A.4)$$

que estabelece (21.A.1).

II. Vamos supor a validade de (21.A.1). Tomando-se $v = 0$ na segunda igualdade de (21.A.1) e usando a primeira ($R(0) = 0$), tem-se $q(R(u)) = q(u)$ para todo u . Agora, por um lado

$$q(R(u)) + q(R(v)) - q(R(u) - R(v)) = q(u) + q(v) - q(u - v) = \omega(u, u) + \omega(v, v) - \omega(u - v, u - v) = \omega(u, v) + \omega(v, u).$$

Por outro lado,

$$\begin{aligned} q(R(u)) + q(R(v)) - q(R(u) - R(v)) &= \omega(R(u), R(u)) + \omega(R(v), R(v)) - \omega(R(u) - R(v), R(u) - R(v)) \\ &= \omega(R(u), R(v)) + \omega(R(v), R(u)). \end{aligned}$$

Isso estabeleceu que

$$\omega(R(u), R(v)) + \omega(R(v), R(u)) = \omega(u, v) + \omega(v, u). \quad (21.A.5)$$

Trocando-se v por iv em (21.A.5) e adicionando-se o resultado à própria (21.A.5), obtemos

$$2\omega(u, v) = \omega(R(u), R(v)) + \omega(R(v), R(u)) - i[\omega(R(u), R(iv)) + \omega(R(iv), R(u))] \quad (21.A.6)$$

Se tivermos válida a condição $R(iv) = +iR(v)$, a relação (21.A.6) e a sesquilinearidade de ω implicam

$$\omega(u, v) = \omega(R(u), R(v)).$$

Se tivermos válida a condição $R(iv) = -iR(v)$, a relação (21.A.6) e a sesquilinearidade de ω implicam

$$\omega(u, v) = \omega(R(v), R(u)) .$$

Isso provou a validade de (21.A.2) ou de (21.A.3), completando a prova do Lema 21.9. ■

Teorema 21.13 *Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \rightarrow \mathbb{C}$. Seja q a forma quadrática associada a ω , definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \rightarrow V$ uma função sobrejetora definida em V e tal que $q(A(u) - A(v)) = q(u - v)$ para todos $u, v \in V$ e satisfazendo $A(iu) - A(iv) = \pm i(A(u) - A(v))$ também para todos $u, v \in V$. Então, existem: 1^o um vetor constante $c \in V$ e, 2^o um operador sobrejetor $R : V \rightarrow V$ satisfazendo*

- (caso “+”): $R : V \rightarrow V$ é linear e $\omega(Ru, Rv) = \omega(u, v)$ para todos $u, v \in V$;
- (caso “-”): $R : V \rightarrow V$ é antilinear e $\omega(Ru, Rv) = \omega(v, u)$ para todos $u, v \in V$.

e tais que

$$A(u) = Ru + c$$

para todo $u \in V$. O operador R e a constante c são univocamente determinados. □

Prova do Teorema 21.13. Seja V um espaço vetorial complexo dotado de uma forma sesquilinear e não degenerada $\omega : V \times V \rightarrow \mathbb{C}$. Seja q a forma quadrática associada a ω , definida por $q(u) := \omega(u, u)$, para todo $u \in V$. Seja $A : V \rightarrow V$ uma função sobrejetora definida em V tal que $q(A(u) - A(v)) = q(u - v)$ para todos $u, v \in V$.

Defina-se $c := A(0) \in V$ e defina-se uma nova função $R : V \rightarrow V$ dada por $R(v) := A(v) - c = A(v) - A(0)$ para todo $v \in V$. É claro que R é sobrejetora, por A o ser. É claro também que $R(0) = 0$ e que $q(R(u) - R(v)) = q(u - v)$ para todos $u, v \in V$. Além disso, $R(iu) = A(iu) - A(i0) = \pm i(A(u) - A(0)) = \pm iR(u)$, para todo $u \in V$.

Vamos agora mostrar que se $R : V \rightarrow V$ satisfaz (21.A.1) (e, portanto, (21.A.2) ou (21.A.3), pelo Lema 21.9) e as demais hipóteses então R é linear ou antilinear. Temos que, para $w \in V$, arbitrário, e assumindo o caso $R(iv) = +iR(v)$,

$$\begin{aligned} & \omega\left(R(w), R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2)\right) \\ &= \omega\left(R(w), R(\alpha_1 u_1 + \alpha_2 u_2)\right) - \alpha_1 \omega\left(R(w), R(u_1)\right) - \alpha_2 \omega\left(R(w), R(u_2)\right) \\ & \stackrel{(21.A.2)}{=} \omega\left(w, \underbrace{(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2}_{=0}\right) = 0. \end{aligned} \quad (21.A.7)$$

Como R é sobrejetora e w é arbitrário, a não degenerescência de ω implica que $R(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 R(u_1) - \alpha_2 R(u_2) = 0$ para todos $\alpha_1, \alpha_2 \in \mathbb{C}$ e todos $u_1, u_2 \in V$, significando que $R : V \rightarrow V$ é linear.

Analogamente, assumindo o caso $R(iv) = -iR(v)$,

$$\begin{aligned} & \omega\left(R(w), R(\alpha_1 u_1 + \alpha_2 u_2) - \overline{\alpha_1}R(u_1) - \overline{\alpha_2}R(u_2)\right) \\ &= \omega\left(R(w), R(\alpha_1 u_1 + \alpha_2 u_2)\right) - \overline{\alpha_1}\omega\left(R(w), R(u_1)\right) - \overline{\alpha_2}\omega\left(R(w), R(u_2)\right) \\ & \stackrel{(21.A.3)}{=} \omega\left(\alpha_1 u_1 + \alpha_2 u_2, w\right) - \overline{\alpha_1}\omega(u_1, w) - \overline{\alpha_2}\omega(u_2, w) \\ &= \omega\left(\underbrace{(\alpha_1 u_1 + \alpha_2 u_2) - \alpha_1 u_1 - \alpha_2 u_2}_{=0}, w\right) = 0. \end{aligned} \quad (21.A.8)$$

Como R é sobrejetora e w é arbitrário, a não degenerescência de ω implica que $R(\alpha_1 u_1 + \alpha_2 u_2) - \overline{\alpha_1}R(u_1) - \overline{\alpha_2}R(u_2) = 0$ para todos $\alpha_1, \alpha_2 \in \mathbb{C}$ e todos $u_1, u_2 \in V$, significando que $R : V \rightarrow V$ é antilinear.

Resta apenas demonstrar a unicidade de R e c . Vamos supor que para todo $u \in V$ tenhamos $Ru + c = R'u + c'$. Tomando-se $u = 0$, concluiríamos que $c = c'$. A relação restante $Ru = R'u$, válida para todo $u \in V$, significa que $R = R'$. Isso completa a prova do Teorema 21.13. ■

21.B Prova do Teorema 21.11

Nesta seção será apresentada a demonstração do Teorema 21.11, página 1113.

Seja L um elemento do grupo de Lorentz $O(1, 3)$, representada como matriz 4×4 na forma

$$L = \begin{pmatrix} L_{00} & L_{01} & L_{02} & L_{03} \\ L_{10} & L_{11} & L_{12} & L_{13} \\ L_{20} & L_{21} & L_{22} & L_{23} \\ L_{30} & L_{31} & L_{32} & L_{33} \end{pmatrix}. \quad (21.B.9)$$

Vamos definir vetores coluna (ou seja, matrizes 3×1) a e b por

$$a := \begin{pmatrix} L_{10} \\ L_{20} \\ L_{30} \end{pmatrix}, \quad b := \begin{pmatrix} L_{01} \\ L_{02} \\ L_{03} \end{pmatrix}.$$

É evidente que podemos representar L na forma de blocos

$$L = \begin{pmatrix} L_{00} & \boxed{b^T} \\ \boxed{a} & \boxed{l} \end{pmatrix}, \quad (21.B.10)$$

onde b^T , a transposta de b , é o vetor linha (matriz 1×3) dado por $b^T = (L_{01}, L_{02}, L_{03})$ e l é a matriz 3×3 dada por

$$l := \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}.$$

Vamos agora considerar duas matrizes R_a e R_b pertencentes a SRot, ou seja,

$$R_a := \begin{pmatrix} 1 & 0 & 0 & 0 \\ \boxed{} & & & \\ 0 & & & \\ 0 & & & \\ 0 & & & \end{pmatrix}, \quad R_b := \begin{pmatrix} 1 & 0 & 0 & 0 \\ \boxed{} & & & \\ 0 & & & \\ 0 & & & \\ 0 & & & \end{pmatrix},$$

com r_a e r_b matrizes 3×3 pertencentes a $SO(3)$. Precisamos estudar a forma da matriz $R_a L R_b^T$. A regra de produto de matrizes diz-nos que

$$R_a L R_b^T = \begin{pmatrix} L_{00} & \boxed{(r_b b)^T} \\ \boxed{r_a a} & \boxed{r_a l r_b^T} \end{pmatrix}. \tag{21.B.11}$$

E. 21.128 Exercício importante. Verifique! *

A expressão acima é válida de modo bastante geral, para quaisquer que sejam as matrizes de rotação r_a e r_b . Vamos agora, porém, considerar matrizes de rotação r_a e r_b particulares. Escolhemos r_a da forma $r_a = s^a t^a$, onde $t^a \in SO(3)$ é a matriz de rotação que roda o vetor a de modo que apenas a primeira componente do vetor resultante seja não nula:

$$t^a a = \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}. \tag{21.B.12}$$

A matriz $s^a \in SO(3)$, por sua vez, é uma matriz de rotação em torno do eixo 1, e que, portanto, deixa o vetor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ invariante. s^a é da forma

$$s^a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & s_{22}^a & s_{23}^a \\ 0 & s_{32}^a & s_{33}^a \end{pmatrix} =: \begin{pmatrix} 1 & 0 & 0 \\ 0 & \boxed{s^{a'}} & \\ 0 & & \end{pmatrix}, \tag{21.B.13}$$

com

$$s^{a'} := \begin{pmatrix} s_{22}^a & s_{23}^a \\ s_{32}^a & s_{33}^a \end{pmatrix} \in \text{SO}(2).$$

Assim, temos também

$$s^a t^a a = \begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}.$$

Analogamente, escolhemos r_b da forma $r_b = s^b t^b$, onde $t^b \in \text{SO}(3)$ é a matriz de rotação que roda o vetor b de modo que apenas a primeira componente do vetor resultante seja não nula:

$$t^b b = \begin{pmatrix} \beta \\ 0 \\ 0 \end{pmatrix}. \tag{21.B.14}$$

A matriz $s^b \in \text{SO}(3)$, por sua vez, é uma matriz de rotação em torno do eixo 1, e que, portanto, deixa o vetor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ invariante. s^b é da forma

$$s^b = \begin{pmatrix} 1 & 0 & 0 \\ 0 & s_{22}^b & s_{23}^b \\ 0 & s_{32}^b & s_{33}^b \end{pmatrix} =: \begin{pmatrix} 1 & 0 & 0 \\ 0 & \boxed{s^{b'}} \\ 0 & & \end{pmatrix}, \tag{21.B.15}$$

com

$$s^{b'} := \begin{pmatrix} s_{22}^b & s_{23}^b \\ s_{32}^b & s_{33}^b \end{pmatrix} \in \text{SO}(2).$$

Pela definição de s^b acima, também temos

$$s^b t^b b = \begin{pmatrix} \beta \\ 0 \\ 0 \end{pmatrix}.$$

Daqui por diante as matrizes t^a e t^b estarão fixas. As matrizes s^a e s^b são ainda arbitrárias, mas serão fixadas mais adiante. Com essas escolhas temos agora

$$R_a L R_b^T = \begin{pmatrix} L_{00} & \beta & 0 & 0 \\ \alpha & \boxed{s^a t_t (s^b)^T} \\ 0 & & & \\ 0 & & & \end{pmatrix}, \tag{21.B.16}$$

onde $l_t := t^a l(t^b)^T$.

A matriz $L' := R_a L R_b^T$ é certamente um elemento do grupo de Lorentz $O(1, 3)$, pois R_a , L e R_b^T o são. Assim, L' satisfaz $L' \eta (L')^T \eta = \mathbb{1}$. Calculemos o lado esquerdo dessa igualdade:

$$\begin{aligned}
 L' \eta (L')^T \eta &= \begin{pmatrix} L_{00} & \beta & 0 & 0 \\ \alpha & \boxed{} & & \\ 0 & s^a l_t (s^b)^T & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{} & & \\ 0 & -\mathbb{1} & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} L_{00} & \alpha & 0 & 0 \\ \beta & \boxed{} & & \\ 0 & s^b l_t^T (s^a)^T & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \boxed{} & & \\ 0 & & & -\mathbb{1} \\ 0 & & & \end{pmatrix} \\
 &= \begin{pmatrix} L_{00} & -\beta & 0 & 0 \\ \alpha & \boxed{} & & \\ 0 & -s^a l_t (s^b)^T & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} L_{00} & -\alpha & 0 & 0 \\ \beta & \boxed{} & & \\ 0 & -s^b l_t^T (s^a)^T & & \\ 0 & & & \end{pmatrix} \\
 &= \begin{pmatrix} L_{00}^2 - \beta^2 & \boxed{-g^T} \\ \boxed{g} & \boxed{f} \end{pmatrix},
 \end{aligned}$$

onde

$$f := s^a l_t (l_t)^T (s^a)^T - \alpha^2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{e} \quad g := -\beta s^a l_t (s^b)^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + L_{00} \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Na expressão para f usamos o fato que $(s^b)^T s^b = \mathbb{1}$, pois s^b é uma matriz ortogonal.

E. 21.129 *Exercício importante.* Verifique as expressões acima. Sugestão: exerça a virtude da Paciência. ✦

Como mencionamos, $L' \eta (L')^T \eta = \mathbb{1}$. Portanto, devemos ter

$$f = \mathbb{1}, \tag{21.B.17}$$

$$g = 0 \quad \text{e} \tag{21.B.18}$$

$$L_{00}^2 - \beta^2 = 1 \tag{21.B.19}$$

Logo,

$$s^a l_t (l_t)^T (s^a)^T = \begin{pmatrix} 1 + \alpha^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{21.B.20}$$

e

$$\beta s^a l_t (s^b)^T = L_{00} \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{21.B.21}$$

Devido à forma de s^a e s^b em (21.B.13) e (21.B.15) essas relações implicam

$$l_t (l_t)^T = \begin{pmatrix} 1 + \alpha^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{21.B.22}$$

e

$$\beta l_t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = L_{00} \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{21.B.23}$$

respectivamente

E. 21.130 *Exercício.* Verifique! ✦

Com as relações acima temos em mãos algumas conclusões sobre a estrutura do grupo de Lorentz. A primeira é a seguinte proposição:

Proposição 21.36 *Para qualquer transformação de Lorentz L vale*

$$L_{00}^2 - \beta^2 = 1, \tag{21.B.24}$$

$$L_{00}^2 - \alpha^2 = 1 \tag{21.B.25}$$

e, conseqüentemente,

$$\alpha^2 = \beta^2. \tag{21.B.26}$$

Fora isso,

$$\|a\|^2 = \alpha^2 = \beta^2 = \|b\|^2,$$

onde $\|a\|$ e $\|b\|$ são os módulos dos vetores a e b , respectivamente, ou seja,

$$\|a\|^2 = (L_{10})^2 + (L_{20})^2 + (L_{30})^2 \quad e \quad \|b\|^2 = (L_{01})^2 + (L_{02})^2 + (L_{03})^2.$$

Portanto,

$$L_{00}^2 = 1 + (L_{10})^2 + (L_{20})^2 + (L_{30})^2 = 1 + (L_{01})^2 + (L_{02})^2 + (L_{03})^2.$$

□

Prova. (21.B.24) é o mesmo que (21.B.19). Para provar (21.B.25), notemos que, pela Proposição 21.26, L^T é também uma transformação de Lorentz. Logo, para L^T a relação (21.B.24) significa $L_{00}^2 - \alpha^2 = 1$, pois ao passarmos de L para L^T o elemento L_{00} não muda, mas ocorre a troca $\alpha \leftrightarrow \beta$. (21.B.26) segue de (21.B.24) e (21.B.25). Para provar que $\|a\|^2 = \alpha^2$, notemos que, por (21.B.12), o vetor $\begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}$ é obtido de a por uma rotação $t^a \in \text{SO}(3)$, que não altera o comprimento de vetores. De modo análogo prova-se que $\|b\|^2 = \beta^2$. ■

Segue dessa proposição que, para prosseguirmos, teremos que considerar dois casos: o caso $\alpha = \beta = 0$ e o caso em que $\alpha \neq 0$ e $\beta \neq 0$.

Caso $\alpha = \beta = 0$.

Como comentamos, nesse caso temos $a = b = 0$. Podemos adotar $s^a = s^b = t^a = t^b = \mathbb{1}$ e, portanto, L é simplesmente da forma

$$L = \begin{pmatrix} L_{00} & 0 & 0 & 0 \\ 0 & \boxed{} \\ 0 & & & \\ 0 & & & \end{pmatrix}.$$

Com $\alpha = 0$ e $s^a = s^b = t^a = t^b = \mathbb{1}$, a relação (21.B.22) reduz-se a $l^T = \mathbb{1}$, ou seja, $l \in \text{O}(3)$. Como $\det(L) = \pm 1$ e $\det(l) = \pm 1$ há quatro situações a considerar:

Ia. $\det(L) = 1$ e $\det(l) = 1$.

Nessa situação tem-se $l \in \text{SO}(3)$ e $L_{00} = 1$. Portanto, $L \in \text{SRot}$.

Ib. $\det(L) = 1$ e $\det(l) = -1$.

Nessa situação $l \in \text{O}(3)$ mas $l \notin \text{SO}(3)$ e $L_{00} = -1$. Assim L é da forma $L = P_1 T R$ com $R \in \text{SRot}$.

IIa. $\det(L) = -1$ e $\det(l) = 1$.

Nessa situação $l \in \text{SO}(3)$ e $L_{00} = -1$. Assim L é da forma $L = T R$ com $R \in \text{SRot}$.

IIb. $\det(L) = -1$ e $\det(l) = -1$.

Nessa situação $l \in \text{O}(3)$ mas $l \notin \text{SO}(3)$ e $L_{00} = 1$. Assim L é da forma $L = P_1 R$ com $R \in \text{SRot}$.

E. 21.131 *Exercício.* Justifique as afirmações acima. ★

Resumindo, vimos para o caso $a = b = 0$ que nas quatro situações possíveis L consiste apenas numa simples rotação, seguida eventualmente de uma inversão de paridade (Ib e IIb) e/ou de uma reversão temporal (Ib e IIa.). Como veremos, o caso $\alpha \neq 0$ e $\beta \neq 0$ envolve também um “boost de Lorentz”, ou seja, uma mudança entre dois sistemas de referência inerciais com uma velocidade relativa eventualmente não nula.

Caso $\alpha \neq 0$ e $\beta \neq 0$.

Como $\beta \neq 0$, (21.B.23) pode ser escrita como

$$l_t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{L_{00}\alpha}{\beta} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \tag{21.B.27}$$

ou seja, $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ é um autovetor de l_t com autovalor $\omega := \frac{L_{00}\alpha}{\beta}$. Escrevendo l_t na sua forma explícita como

$$l_t = \begin{pmatrix} (l_t)_{11} & (l_t)_{12} & (l_t)_{13} \\ (l_t)_{21} & (l_t)_{22} & (l_t)_{23} \\ (l_t)_{31} & (l_t)_{32} & (l_t)_{33} \end{pmatrix},$$

a relação (21.B.27) fica

$$\begin{pmatrix} (l_t)_{11} & (l_t)_{12} & (l_t)_{13} \\ (l_t)_{21} & (l_t)_{22} & (l_t)_{23} \\ (l_t)_{31} & (l_t)_{32} & (l_t)_{33} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \omega \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

o que implica $(l_t)_{11} = \omega$ e $(l_t)_{21} = (l_t)_{31} = 0$. Assim,

$$l_t = \begin{pmatrix} \omega & (l_t)_{12} & (l_t)_{13} \\ 0 & (l_t)_{22} & (l_t)_{23} \\ 0 & (l_t)_{32} & (l_t)_{33} \end{pmatrix} = \begin{pmatrix} \omega & \boxed{\phi^T} \\ 0 & \boxed{l'_t} \\ 0 & \end{pmatrix},$$

onde ϕ é o vetor coluna $\phi = \begin{pmatrix} (l_t)_{12} \\ (l_t)_{13} \end{pmatrix}$ e l'_t é a matriz 2×2 dada por $l'_t := \begin{pmatrix} (l_t)_{22} & (l_t)_{23} \\ (l_t)_{32} & (l_t)_{33} \end{pmatrix}$. Ocorre que também

vale que $\phi = 0$. Para ver isso, notemos que (21.B.22) diz-nos que

$$l_t(l_t)^T = \begin{pmatrix} \omega & \boxed{\phi^T} \\ 0 & \boxed{l'^T_t} \\ 0 & \end{pmatrix} \begin{pmatrix} \omega & \boxed{0 \ 0} \\ \boxed{\phi} & \boxed{l'_t} \end{pmatrix} = \begin{pmatrix} 1 + \alpha^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

ou seja,

$$\begin{pmatrix} \omega^2 + \phi^T \phi & \boxed{(l'_t \phi)^T} \\ \boxed{l'_t \phi} & \boxed{l'_t (l'_t)^T} \end{pmatrix} = \begin{pmatrix} 1 + \alpha^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Logo,

$$l'_t (l'_t)^T = \mathbb{1}, \tag{21.B.28}$$

$$l'_t \phi = 0 \tag{21.B.29}$$

e

$$\omega^2 + \phi^T \phi = 1 + \alpha^2. \tag{21.B.30}$$

Agora, (21.B.28) afirma que l'_t é uma matriz ortogonal: $(l'_t)^{-1} = (l'_t)^T$. Aplicando, portanto, $(l'_t)^{-1}$ à esquerda em (21.B.29) segue que $\phi = 0$. Chegamos assim à conclusão que

$$l_t = \begin{pmatrix} \omega & 0 & 0 \\ 0 & (l_t)_{22} & (l_t)_{23} \\ 0 & (l_t)_{32} & (l_t)_{33} \end{pmatrix} = \begin{pmatrix} \omega & 0 & 0 \\ 0 & \boxed{l'_t} \\ 0 & & \end{pmatrix},$$

com $\omega^2 = 1 + \alpha^2$ (por (21.B.30)). Segue daí que

$$s^a l_t (s^b)^T = \begin{pmatrix} \omega & 0 & 0 \\ 0 & \boxed{s^{a'} l'_t (s^{b'})^T} \\ 0 & & \end{pmatrix},$$

($s^{a'}$ e $s^{b'}$ estão definidos em (21.B.13) e (21.B.15)). Neste momento vamos fixar s^a e s^b , adotando

$$s^{a'} = s^{b'} (l'_t)^{-1} = s^{b'} (l'_t)^T.$$

Com isso, obviamente

$$s^{a'} l'_t (s^{b'})^T = \mathbb{1}.$$

Logo,

$$s^a l_t (s^b)^T = \begin{pmatrix} \omega & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Retornando a (21.B.16)

$$R_a L R_b^T = \begin{pmatrix} L_{00} & \beta & 0 & 0 \\ \alpha & \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{21.B.31}$$

onde, recordando,

$$\omega = \frac{L_{00}\alpha}{\beta} \quad \text{e} \quad \omega^2 = 1 + \alpha^2. \tag{21.B.32}$$

Resta-nos mostrar que a matriz do lado direito de (21.B.31) tem a forma de um *boost* de Lorentz, acompanhado eventualmente de uma operação de troca de paridade e/ou reversão temporal. É o que faremos agora.

Como $R_a L R_b^T$ é um elemento do grupo de Lorentz $O(1, 3)$, tem-se que $\det(R_a L R_b^T) = \pm 1$. Calculando o determinante da matriz do lado direito (21.B.31) tem-se, então, $\omega L_{00} - \alpha\beta = \pm 1$. Multiplicando-se por α/β teremos $\omega \frac{L_{00}\alpha}{\beta} - \alpha^2 = \pm \frac{\alpha}{\beta}$, ou seja,

$$\omega^2 - \alpha^2 = \pm \frac{\alpha}{\beta}.$$

Pela segunda equação em (21.B.32) isso implica

$$\beta = \pm\alpha \quad \text{e} \quad L_{00} = \pm\omega,$$

os dois sinais \pm acima sendo iguais ao sinal de $\det(R_a L R_b^T)$. Porém, ω é dado por $\pm\sqrt{1+\alpha^2}$ (vide (21.B.32)), mas a escolha do sinal dessa raiz quadrada é independente do sinal de $\det(R_a L R_b^T)$. Há, portanto, quatro situações possíveis que deveremos considerar separadamente:

Ia. Escolhendo $\det(R_a L R_b^T) = +1$ e $\omega = +\sqrt{1+\alpha^2}$, (21.B.31) fica

$$L = (R_a)^T \begin{pmatrix} \sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} R_b. \quad (21.B.33)$$

R_a e R_b são elementos de $\text{SRot} \simeq \text{SO}(3)$, temos $\det(R_a) = \det(R_b) = 1$. Logo, neste caso temos $\det(L) = 1$. Fora isso $L_{00} \geq 1$.

É conveniente escrever (21.B.33) de outra forma. Como α é um número real arbitrário, vamos definir $v \in (-1, 1)$ por

$$v := -\frac{\alpha}{\sqrt{1+\alpha^2}}, \quad \text{de modo que} \quad \alpha = -\frac{v}{\sqrt{1-v^2}}. \quad (21.B.34)$$

Teremos

$$\begin{pmatrix} \sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \gamma(v) & -v\gamma(v) & 0 & 0 \\ -v\gamma(v) & \gamma(v) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =: B_1(v),$$

onde

$$\gamma(v) := \frac{1}{\sqrt{1-v^2}}.$$

Como se vê, chegamos dessa forma aos *boosts* de Lorentz $B_1(v)$ utilizando apenas as propriedades definidoras do grupo de Lorentz. Compare com o estudo do grupo $O(1, 1)$, página 1051.

Com essa parametrização, (21.B.33) fica

$$L = (R_a)^T B_1(v) R_b, \quad (21.B.35)$$

para $R_a, R_b \in \text{SRot}$.

Ib. Escolhendo $\det(R_a L R_b^T) = +1$ e $\omega = -\sqrt{1+\alpha^2}$, (21.B.31) fica

$$R_a L R_b^T = \begin{pmatrix} -\sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & -\sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (21.B.36)$$

Logo, usando-se as matrizes P_1 e T definidas em (21.200) e (21.201), segue

$$P_1 R_a L R_b^T T = \begin{pmatrix} \sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (21.B.37)$$

como facilmente se verifica. Daí, notando que T e R_b comutam (por que?), conclui-se que nesse caso temos

$$L = (P_1 R_a)^T \begin{pmatrix} \sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} R_b T. \quad (21.B.38)$$

Assim, com a parametrização (21.B.34),

$$L = (P_1 R_a)^T B_1(v) R_b T, \quad (21.B.39)$$

para $R_a, R_b \in \text{SRot}$. Note-se que neste caso temos $\det(L) = 1$ com $L_{00} \leq -1$.

IIa. Escolhendo $\det(R_a L R_b^T) = -1$ e $\omega = +\sqrt{1+\alpha^2}$, (21.B.31) fica

$$R_a L R_b^T = \begin{pmatrix} -\sqrt{1+\alpha^2} & -\alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (21.B.40)$$

Assim,

$$T R_a L R_b^T = \begin{pmatrix} \sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (21.B.41)$$

como facilmente se verifica. Nesse caso, então,

$$L = T(R_a)^T \begin{pmatrix} \sqrt{1+\alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1+\alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} R_b. \quad (21.B.42)$$

Assim, com a parametrização (21.B.34),

$$L = T(R_a)^T B_1(v) R_b, \tag{21.B.43}$$

para $R_a, R_b \in \text{SRot}$. Note-se que neste caso temos $\det(L) = -1$ com $L_{00} \leq -1$.

IIb. Escolhendo $\det(R_a L R_b^T) = -1$ e $\omega = -\sqrt{1 + \alpha^2}$, (21.B.31) fica

$$R_a L R_b^T = \begin{pmatrix} \sqrt{1 + \alpha^2} & -\alpha & 0 & 0 \\ \alpha & -\sqrt{1 + \alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{21.B.44}$$

Assim,

$$R_a L R_b^T P_1 = \begin{pmatrix} \sqrt{1 + \alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1 + \alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{21.B.45}$$

como facilmente se verifica. Nesse caso, então,

$$L = (R_a)^T \begin{pmatrix} \sqrt{1 + \alpha^2} & \alpha & 0 & 0 \\ \alpha & \sqrt{1 + \alpha^2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} P_1 R_b. \tag{21.B.46}$$

Assim, com a parametrização (21.B.34),

$$L = (R_a)^T B_1(v) P_1 R_b, \tag{21.B.47}$$

para $R_a, R_b \in \text{SRot}$. Note-se que neste caso temos $\det(L) = -1$ e $L_{00} \geq 1$.

A demonstração do Teorema 21.11 está assim completa. ■