



FÍSICA-MATEMÁTICA

RUDI GAELZER (INSTITUTO DE FÍSICA - UFRGS)

Apostila preparada para as disciplinas de Física-Matemática ministradas para os Cursos de Bacharelado em Física do Instituto de Física da Universidade Federal do Rio Grande do Sul, Porto Alegre - RS.



Apostila escrita com:
PROCESSADOR DE DOCUMENTOS
LyX
<http://www.lyx.org/>
<http://wiki.lyx.org/LyX/LyX>

Referências bibliográficas:
SISTEMA biblatex
<http://ctan.org/pkg/biblatex>
<https://github.com/plk/biblatex>

SUMÁRIO

1	Sistemas de Coordenadas Curvilíneas Ortogonais	1
1.1	Coordenadas curvilíneas	1
1.2	Coordenadas curvilíneas ortogonais	6
1.3	Análise vetorial em sistemas de coordenadas curvilíneas ortogonais	7
1.3.1	Álgebra vetorial	7
1.3.2	Integrais de caminho, de superfície e de volume	7
1.4	Operadores vetoriais diferenciais	10
1.4.1	Gradiente	10
1.4.2	Divergente e laplaciano	11
1.4.3	Rotacional	13
1.5	Sistemas de coordenadas curvilíneas ortogonais	14
1.5.1	Coordenadas polares cilíndricas	14
1.5.2	Coordenadas polares esféricas	16
1.5.3	Coordenadas elípticas cilíndricas	18
1.6	Linhas de força e superfícies equipotenciais	20
1.6.1	Linhas de força de um campo vetorial	20
	Referências	22
2	Funções de Uma Variável Complexa	23
2.1	Números e variáveis complexos	24
2.1.1	Representações vetorial e polar	25
2.1.2	Fórmula de Euler	26
2.2	Álgebra de números complexos	26
2.2.1	Fórmula de Moivre	28
2.2.2	Raízes de números complexos	28
2.3	Funções de uma variável complexa	30
2.3.1	Transformações ou mapeamentos	31
2.3.2	Pontos de ramificação, linhas de ramificação e superfícies de Riemann	31
2.3.3	Exemplos de funções unívocas ou plurívocas	32
2.4	O cálculo diferencial de funções de uma variável complexa	33
2.4.1	Limite de uma função complexa	33
2.4.2	Continuidade	35
2.4.3	Derivadas de funções complexas	36
2.4.4	As condições de Cauchy-Riemann	37
2.4.5	Funções analíticas	39
2.4.6	Funções harmônicas	39
2.4.7	Pontos singulares	40
2.5	Integração no plano complexo	40
2.5.1	Integrais de caminho no plano complexo	41
2.5.2	Propriedades matemáticas das integrais de linha	42
2.5.3	Tipos de curvas e domínios no plano complexo	43
2.5.3.1	Tipos de curvas no plano complexo	43
2.5.3.2	Domínios simplesmente ou multiplamente conexos	43
2.5.3.3	Convenção para o percurso de um contorno fechado	44
2.6	O teorema de Cauchy-Goursat	44

2.6.1	O teorema de Green no plano	44
2.6.2	O teorema de Cauchy-Goursat	44
2.7	Fórmulas integrais de Cauchy	48
2.8	Representação em séries de funções analíticas	50
2.8.1	Séries complexas	50
2.8.1.1	Convergência da série	51
2.8.1.2	Convergência absoluta	51
2.8.1.3	Convergência uniforme	52
2.8.2	Testes de convergência	52
2.8.2.1	Testes de convergência absoluta	52
2.8.2.2	Teste de convergência uniforme	52
2.8.3	Séries de potências e séries de Taylor	53
2.8.4	Séries de Taylor de funções elementares	55
2.8.5	Séries de Laurent	55
2.8.6	Teoremas de existência e unicidade	57
2.8.7	Algumas técnicas de construção de séries de Taylor e Laurent	57
2.8.8	Séries de Laurent de funções elementares	59
2.8.9	Classificação de singularidades	60
2.8.9.1	Polos	60
2.8.9.2	Singularidades essenciais	60
2.8.9.3	Singularidades removíveis	60
2.9	Integração no plano complexo pelo método dos resíduos	60
2.9.1	Resíduos	61
2.9.2	Teorema dos resíduos	62
2.9.3	Cálculo de resíduos	62
2.9.3.1	Primeiro método: direto da definição	62
2.9.3.2	Segundo método: polos de ordem m em $z = z_0$	63
2.9.3.3	Terceiro método: resíduo de uma função racional	64
2.9.3.4	Quarto método: pelo desenvolvimento em série de Laurent	64
2.9.4	Cálculo de integrais definidas ou impróprias	65
2.9.4.1	Integrais do tipo I: funções racionais	65
2.9.4.2	Integrais do tipo II: funções racionais de funções trigonométricas	67
2.9.4.3	Integrais do tipo III: integrais de Fourier	68
2.9.4.4	Integrais do tipo IV: integrando com polos no eixo real	69
2.9.4.5	Integrais do tipo V: integração ao longo de linhas de ramificação	72
2.9.4.6	Outros tipos de integrais	74
2.10	Continuação analítica	77
	Referências	80
3	Teoria de Grupos Abstratos	81
3.1	Definições e classificações iniciais	82
3.1.1	Classificação dos grupos	84
3.1.2	Exemplos de grupos	85
3.1.2.1	Grupos infinitos discretos	85
3.1.2.2	Grupos contínuos compactos	85
3.1.2.3	Grupos contínuos não compactos	86
3.1.2.4	Grupos finitos	88
3.2	Grupos finitos	88
3.2.1	Tabela de multiplicação de grupo	88
3.2.2	Grupo cíclico C_n	90
3.2.3	O Grupo simétrico S_n	90
3.2.3.1	Verificação dos axiomas de grupo	92
3.2.3.2	Notação de ciclos	93
3.3	Subgrupos, classes laterais e de conjugação	96
3.3.1	Subgrupos	96
3.3.2	Classes laterais e o teorema de Lagrange	97
3.3.3	Classes de conjugação	99
3.3.4	Subgrupos invariantes e grupo fator	102
3.4	Grupos de simetria	103

3.4.1	Grupos cristalográficos pontuais	106
3.4.2	Projeções estereográficas	110
3.4.3	Grupos cristalográficos espaciais	111
3.5	Produto direto de grupos	113
3.6	Mapeamentos entre grupos	114
3.6.1	Funções e mapeamentos	115
3.6.2	Mapeamento entre grupos e homomorfismo	116
3.7	Estruturas algébricas	121
3.7.1	Estruturas compostas por um conjunto com operações	122
3.7.1.1	Estruturas do tipo grupo	122
3.7.1.2	Estruturas do tipo anel	123
3.7.2	Estruturas do tipo Módulo	125
3.7.3	Estruturas do tipo álgebra	126
	Referências	127
4	Espaços Vetoriais	129
4.1	Espaço vetorial	129
4.2	Subespaços vetoriais e subespaços complementares	130
4.2.1	Subespaço soma	130
4.2.2	Subespaços complementares	130
4.3	Bases de um espaço vetorial	131
4.3.1	Transformação de base	132
4.4	Mapeamentos entre espaços vetoriais	133
4.5	Espaço vetorial dual	136
4.6	Subespaços invariantes	136
4.7	Espaço vetorial normado	136
4.8	Espaço com produto interno	137
4.8.1	Espaço real com produto interno	138
4.8.2	Espaço complexo com produto interno	138
4.8.3	Norma de um vetor induzida pelo produto interno	138
4.8.4	Bases ortonormais em espaços normados com produto interno	139
4.8.5	Operadores lineares em espaços com produto interno	140
4.9	Emprego de matrizes em espaços vetoriais	142
4.9.1	Definições iniciais	142
4.9.2	Álgebra matricial	143
4.9.3	Matrizes quadradas	145
4.9.3.1	Determinante	145
4.9.3.2	Tipos especiais de matrizes quadradas	146
4.9.4	Matrizes associadas a vetores e operadores	148
4.9.4.1	Transformação de base	148
4.9.4.2	Matriz de um operador linear	149
4.10	Autovetores e autovalores de um operador linear	153
4.10.1	O teorema espectral	153
4.11	Espaços tensoriais	155
4.12	Elementos de espaços métricos e topologia	156
4.13	Espaço vetorial métrico	159
4.13.1	Espaço vetorial métrico com norma.	159
4.13.2	Espaço de Banach.	160
4.13.3	Espaço de Hilbert	160
4.13.4	Espaço afim	161
	Referências	165
5	Teoria de Representações de Grupos	167
5.1	Primeiras definições e representações	168
5.1.1	Vetores e funções de base e representações regulares	169
5.1.2	Representação natural	172
5.2	Representações de grupos de transformações lineares	172
5.2.1	Espaços vetoriais e operadores na mecânica quântica	173
5.2.2	Espaços vetoriais e suas representações	174
5.3	Representações equivalentes e caracteres	175

5.4	Soma e produto diretos de matrizes e representações	176
5.4.1	Soma direta de matrizes	177
5.4.2	Soma direta de representações	179
5.4.3	Produto direto de matrizes	180
5.4.4	Produto direto de representações	182
5.5	Representações redutíveis ou irredutíveis	182
5.6	Teoremas fundamentais sobre representações de grupos e caracteres	186
5.6.1	Teoremas sobre representações	187
5.6.2	Interpretação do teorema da ortogonalidade	190
5.6.3	Teoremas sobre caracteres	195
5.6.4	Interpretação do teorema da ortogonalidade dos caracteres	197
5.6.5	Decomposição de uma representação em irreps	198
5.6.6	Construção de uma tabela de caracteres	201
5.7	Bases simetrizadas para representações irredutíveis	203
5.8	Bases para representações de grupos de produto direto	206
5.8.1	Redução da representação do produto direto. Série de Clebsch-Gordan	207
5.8.2	Bases para representações de produtos diretos. Coeficientes de Clebsch-Gordan	207
5.8.3	Representação de um grupo produto direto	209
5.9	Aplicações físicas da teoria de representações de grupo	210
5.9.1	Isomorfismo entre transformações sobre sistemas físicos e transformações sobre espaços funcionais	210
5.9.1.1	Translações espaciais	211
5.9.1.2	Rotações	214
5.9.1.3	Deslocamento temporal	217
5.9.2	O grupo do Hamiltoniano	217
5.9.2.1	Degenerescência normal ou acidental	218
5.9.2.2	Representações de \mathcal{P}_N	219
5.9.2.3	Teoria de grupos e “bons” números quânticos	222
5.9.3	Grupos Abelianos e o teorema de Bloch	224
5.9.3.1	Grupos cíclicos	224
5.9.3.2	O teorema de Bloch	224
5.9.4	Funções de base para representações irredutíveis	226
5.9.5	Perturbações, regras de seleção e simetria	232
5.9.5.1	Perturbações sobre autoestados	232
5.9.5.2	O teorema dos elementos de matriz e regras de seleção	232
5.10	Grupos e álgebras de Lie e suas representações	237
5.10.1	Definições iniciais	237
5.10.2	Transformações infinitesimais do grupo	242
5.10.3	Constantes de estrutura	248
5.10.4	Parametrização das transformações do grupo	252
5.10.5	Álgebras de Lie	254
5.10.6	Representações de grupos de Lie	256
5.10.7	Representações irredutíveis do $SU(2)$	260
5.10.8	O teorema de Casimir	261
	Referências	262
6	Álgebra e Análise Tensoriais	265
6.1	Introdução e definições	265
6.1.1	Convenção de soma de índices e símbolos auxiliares	267
6.1.2	Símbolos auxiliares: Kronecker e Levi-Civita	268
6.2	Propriedades de transformação de escalares, vetores e tensores	269
6.2.1	Rotações	270
6.2.2	Transformações de paridade ou reflexões	274
6.2.3	Reversão temporal	275
6.3	Tensores Cartesianos	275
6.3.1	Espaços funcionais	275
6.3.2	Tensores Cartesianos de postos zero e um	276
6.3.3	Tensores Cartesianos de posto dois ou superior	280

6.4	Álgebra tensorial	283
6.4.1	Adição de tensores	283
6.4.2	Simetria e antissimetria	283
6.4.3	Tensores Hermitianos ou anti-Hermitianos	284
6.4.4	Produto externo de tensores	285
6.4.4.1	Produto externo de dois tensores	285
6.4.4.2	Diádicas	286
6.4.4.3	Gradiente de um vetor	286
6.4.4.4	Produto externo em geral	286
6.4.5	Contração e produto interno	287
6.4.5.1	Contração	287
6.4.5.2	Produto interno	287
6.4.5.3	Produtos com diádicas	288
6.4.6	Regra do quociente	288
6.5	Composição de transformações, rotações infinitesimais e tensores isotrópicos	290
6.5.1	Composição de transformações	291
6.5.2	Rotações infinitesimais	291
6.5.3	Tensores isotrópicos	292
6.6	Rotações impróprias, pseudotensores e tensores duais	294
6.6.1	Rotações impróprias e pseudotensores	295
6.6.2	Tensores duais	300
6.6.3	Tensores irredutíveis	300
6.7	Tensores generalizados	301
6.7.1	Coordenadas curvilíneas generalizadas	301
6.7.2	O espaço de Riemann e o tensor de métrica	302
6.7.2.1	Operação de elevação ou rebaixamento de índice	303
6.7.2.2	Elementos infinitesimais de arco e volume	304
6.8	Transformações generalizadas de coordenadas e tensores generalizados	306
6.9	Tensores relativos	309
6.10	Derivadas dos vetores de base e os símbolos de Christoffel	311
6.11	Diferenciação covariante	313
6.12	Operadores vetoriais na forma tensorial	317
6.12.1	Gradiente de campo escalar	317
6.12.2	Divergente de campo vetorial	317
6.12.3	Laplaciano de um campo escalar	318
6.12.4	Rotacional de um campo vetorial	319
6.13	Diferenciação absoluta e curvas geodésicas	319
6.13.1	Diferenciação absoluta ou intrínseca	319
6.13.2	Curvas Geodésicas	320
6.13.3	Transporte paralelo de campos vetoriais	325
6.14	Os tensores de Riemann, Ricci e Einstein	327
6.14.1	O tensor de curvatura de Riemann-Christoffel	327
6.14.2	O tensor de Ricci e o escalar de curvatura	332
6.14.3	O tensor de Einstein e as equações do campo gravitacional	332
6.15	Aplicações físicas	333
6.15.1	A transformação de Lorentz, o espaço-tempo de Minkowski e a formulação covariante do eletromagnetismo clássico	333
6.15.1.1	A situação anterior a 1900. A transformação de Galileu	334
6.15.1.2	Equações de Maxwell e a transformação de Galileu	335
6.15.1.3	A transformação de Lorentz e os princípios da relatividade restrita	338
6.15.1.4	O espaço-tempo de Minkowski e os quadrivetores	340
6.15.1.5	Formulação covariante do eletromagnetismo clássico	345
6.15.2	A métrica de Schwarzschild na relatividade geral	349
6.15.2.1	Derivação do tensor de métrica	349
6.15.2.2	Consequências e aplicações da métrica de Schwarzschild	351
	Referências	355

A	Distribuições e a “Função” Delta de Dirac	357
A.1	Definição de Distribuições	357
A.1.1	Definição operacional de distribuição	357
A.1.1.1	Exemplos	357
A.1.2	Propriedades de distribuições	358
A.1.2.1	Combinação linear de distribuições	358
A.1.2.2	Produto de duas distribuições	358
A.1.2.3	Séries e integrais de distribuições	358
A.1.2.4	Derivadas de distribuições	359
A.2	Propriedades da “Função” δ	360
A.2.1	Definição da δ	360
A.2.2	Representações da $\delta(x - x_0)$ como o limite de um operador integral	360
A.2.3	Principais propriedades	361
A.2.4	Derivadas da $\delta(x)$	362
A.3	Deltas de Dirac em mais de uma dimensão	362
A.3.1	Coordenadas Curvilíneas Ortogonais	363
A.3.2	Pontos degenerados	363
	Referências	365

SISTEMAS DE COORDENADAS CURVILÍNEAS ORTOGONAIS

A TÉ ESTE PONTO, todas as expressões foram desenvolvidas utilizando um sistema de coordenadas retangulares, também denominado de *Cartesiano*. O sistema Cartesiano de coordenadas oferece a vantagem única de ter os seus três vetores unitários de base, \hat{i} , \hat{j} e \hat{k} , constantes em direção, em qualquer ponto em um espaço vetorial de dimensão 3, doravante denotado por \mathbb{E}^3 e denominado **espaço Euclidiano de dimensão 3**.¹ Algumas propriedades adicionais do espaço Euclidiano são discutidas na seção 4.12 e, posteriormente, no capítulo 6.

Mesmo o vetor unitário $\hat{e}_r = \mathbf{r}/r$ foi intrinsecamente tratado como função das três coordenadas Cartesianas x , y e z . Infelizmente, nem todos os problemas físicos são adequadamente tratados por um sistema de coordenadas Cartesiano. Um exemplo desta situação é a dinâmica de uma partícula sob a ação de uma força central, $\mathbf{F} = F(r)\hat{e}_r$, tal como as forças gravitacional ou eletrostática. Tal dinâmica é, em geral, mais facilmente tratada se for empregado um outro sistema de coordenadas, no caso, o sistema esférico (r, θ, φ) , no qual as coordenadas possuem uma relação unívoca com as coordenadas Cartesianas, mas cujos vetores unitários de base, \hat{e}_r , \hat{e}_θ e \hat{e}_φ , não são constantes em direção.

Existe um número relativamente grande de distintos sistemas de coordenadas ortogonais na literatura. Um determinado sistema deve ser escolhido de forma a explorar adequadamente qualquer simetria ou vínculo no problema físico em estudo. Uma vez escolhido o sistema mais adequado, espera-se que o problema possa ser resolvido mais facilmente que seria caso fosse empregado o sistema Cartesiano. Esta simplificação pode se manifestar, por exemplo, na redução do número de graus de liberdade do problema ou na possibilidade de separação de uma equação diferencial parcial.

Neste capítulo, será desenvolvido inicialmente o formalismo básico de um sistema de coordenadas curvilíneo na seção 1.1. Uma vez definida a forma geral do tensor de métrica, válida para qualquer lei de transformação de coordenadas, a seção 1.2 será voltada exclusivamente para sistemas de coordenadas ortogonais. Em seguida, na seção 1.4, os operadores diferenciais vetoriais gradiente, divergente, rotacional e laplaciano serão derivados para um sistema qualquer, bem como as respectivas formas dos Teoremas de Gauss (ou do divergente) e de Stokes. Na seção 1.5, alguns dos mais importantes e empregados sistemas de coordenadas ortogonais serão discutidos. Finalmente, na seção 1.6 será apresentada uma técnica útil para o traçado de linhas de força de um campo vetorial e suas respectivas superfícies equipotenciais.

1.1 COORDENADAS CURVILÍNEAS

O conceito básico envolvido na definição de um sistema de coordenadas consiste nas famílias de superfícies contidas no \mathbb{E}^3 , sobre as quais uma das três coordenadas permanece constante. Este conceito, neste texto empregado para um espaço de dimensão 3, pode ser facilmente estendido para um número arbitrário de dimensões.

O exemplo mais usual vem do sistema Cartesiano ilustrado na figura 1.1: as coordenadas do sistema Cartesiano são definidas pelos planos mutuamente ortogonais $x = \text{constante}$, $y = \text{constante}$ e $z = \text{constante}$. Um ponto P qualquer ($P \in \mathbb{E}^3$), identificado pela terna de números

¹Na verdade, trata-se de um espaço afim de dimensão 3, no qual é adotada a métrica Euclidiana. Ver exemplo 4.2.

$P = (x, y, z)$, pode ser localizado em um sistema de coordenadas Cartesiano pela intersecção destes três planos.

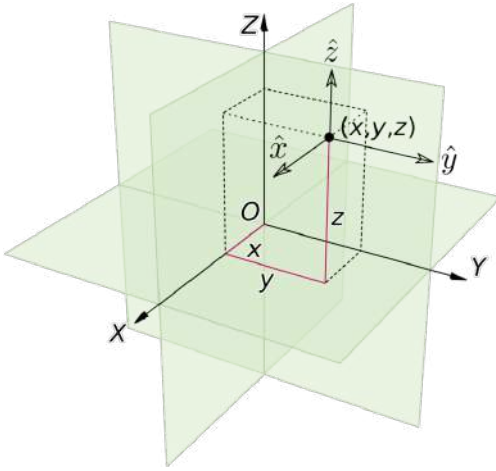


Figura 1.1: Sistema Cartesiano de coordenadas no \mathbb{E}^3 .

Por sua vez, cada coordenada varia ao longo da intersecção dos planos das outras. Por exemplo, a coordenada x , nas vizinhanças do ponto P , varia ao longo da intersecção dos planos $y = \text{cte.}$ e $z = \text{cte.}$ em P .

A figura 1.1 mostra também os vetores da base canônica $\{\hat{x}, \hat{y}, \hat{z}\} \equiv \{\hat{i}, \hat{j}, \hat{k}\}$, os quais são tanto tangentes às retas coordenadas x , y e z quanto ortogonais aos planos coordenados.

Pode-se então imaginar três outras famílias de superfícies contidas no \mathbb{E}^3 , $q_1 = c_1$, $q_2 = c_2$ e $q_3 = c_3$ (c_1 , c_2 e c_3 constantes), denominadas *superfícies coordenadas*, as quais passarão a definir o novo sistema de coordenadas, de modo que agora o mesmo ponto $P \in \mathbb{E}^3$ passa a ser localizado por $P = (q_1, q_2, q_3)$. Esta construção está ilustrada na figura 1.2. Cada par de superfícies coordenadas possui sua intersecção ao longo de uma curva denominada *curva* ou *linha coordenada*. A variação da coordenada oposta às duas superfícies intersectantes ocorre então ao longo desta curva. Por exemplo, na figura 1.2 observa-se que a variação de q_3 ocorre ao longo

da curva coordenada definida pela intersecção das superfícies $q_1 = c_1$ e $q_2 = c_2$.

Ao contrário do sistema Cartesiano, estas três novas famílias de superfícies não necessitam ser mutuamente ortogonais; contudo, é suposto que exista uma relação unívoca entre as novas coordenadas (q_1, q_2, q_3) e as coordenadas Cartesianas (x, y, z) , isto é, deve ser possível escrever tanto as *Leis de Transformação*:

$$x = x(q_1, q_2, q_3) \quad (1.1a)$$

$$y = y(q_1, q_2, q_3) \quad (1.1b)$$

$$z = z(q_1, q_2, q_3), \quad (1.1c)$$

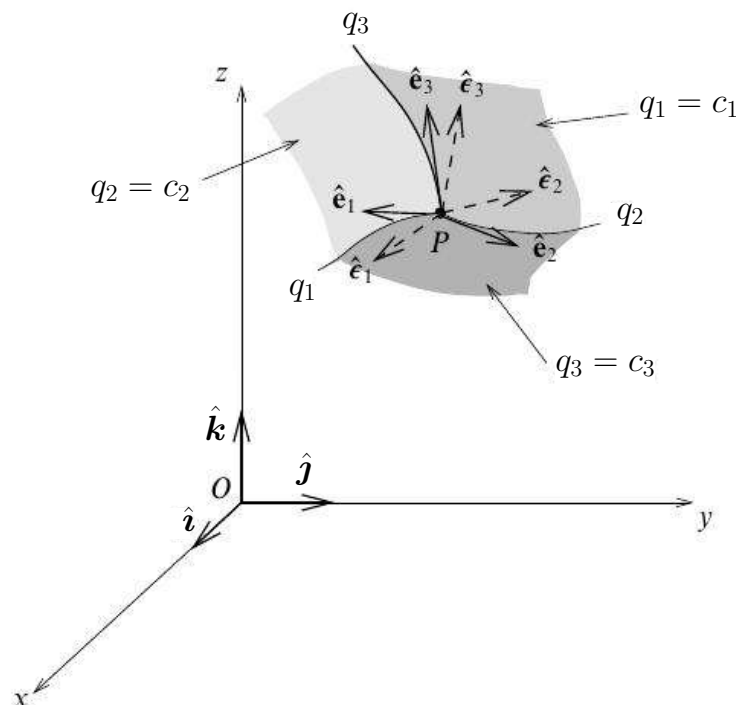


Figura 1.2: Sistema geral de coordenadas curvilíneas.

quanto as suas relações inversas

$$q_1 = q_1(x, y, z) \quad (1.2a)$$

$$q_2 = q_2(x, y, z) \quad (1.2b)$$

$$q_3 = q_3(x, y, z). \quad (1.2c)$$

De acordo com (1.1a), uma variação infinitesimal da coordenada x pode ser escrita

$$dx = \sum_{i=1}^3 \frac{\partial x}{\partial q_i} dq_i = \frac{\partial x}{\partial q_1} dq_1 + \frac{\partial x}{\partial q_2} dq_2 + \frac{\partial x}{\partial q_3} dq_3,$$

existindo expressões equivalentes para as coordenadas y e z . No sistema Cartesiano, o vetor posição r pode ser escrito em termos da base canônica $\{\hat{i}, \hat{j}, \hat{k}\}$, ilustrada nas figuras 1.1 e 1.2. Então, o vetor deslocamento infinitesimal dr pode ser escrito tanto em termos da base canônica, como $dr = dx\hat{i} + dy\hat{j} + dz\hat{k}$, quanto como

$$dr = \sum_{i=1}^3 \frac{\partial r}{\partial q_i} dq_i = \sum_{i=1}^3 \left(\frac{\partial x}{\partial q_i} \hat{i} + \frac{\partial y}{\partial q_i} \hat{j} + \frac{\partial z}{\partial q_i} \hat{k} \right) dq_i. \quad (1.3)$$

De acordo com o Teorema de Pitágoras, o elemento infinitesimal de arco $d\ell^2 = dr \cdot dr$, o qual corresponde à distância entre dois pontos vizinhos no \mathbb{E}^3 , é dado por

$$d\ell^2 = dx^2 + dy^2 + dz^2. \quad (1.4)$$

Para simplificar e condensar a notação empregada, serão feitas as equivalências

$$(x, y, z) \leftrightarrow (x_1, x_2, x_3) \text{ e } \{\hat{i}, \hat{j}, \hat{k}\} \leftrightarrow \{\hat{x}_1, \hat{x}_2, \hat{x}_3\},$$

nas expressões a seguir.

Utilizando as leis de transformação (1.1a-c), deve então ser possível escrever $d\ell^2$ para um sistema qualquer de coordenadas curvilíneas na seguinte forma quadrática:

$$\begin{aligned} d\ell^2 &= g_{11}dq_1^2 + g_{12}dq_1dq_2 + g_{13}dq_1dq_3 \\ &+ g_{21}dq_2dq_1 + g_{22}dq_2^2 + g_{23}dq_2dq_3 \\ &+ g_{31}dq_3dq_1 + g_{32}dq_3dq_2 + g_{33}dq_3^2, \end{aligned} \quad (1.5a)$$

ou

$$d\ell^2 = \sum_{i,j=1}^3 g_{ij}dq_idq_j. \quad (1.5b)$$

Espaços para os quais a definição (1.5) é uma expressão válida para representar $d\ell^2$ são denominados **Riemannianos**. O espaço Euclideano é um tipo particular de um espaço Riemanniano, para o qual o elemento de arco $d\ell^2$ pode sempre ser calculado com o teorema de Pitágoras, independente do sistema de coordenadas adotado.

Inserindo os diferenciais (1.3) em (1.4) e identificando a expressão resultante com (1.5), encontra-se

$$g_{ij} = \frac{\partial x}{\partial q_i} \frac{\partial x}{\partial q_j} + \frac{\partial y}{\partial q_i} \frac{\partial y}{\partial q_j} + \frac{\partial z}{\partial q_i} \frac{\partial z}{\partial q_j} = \sum_{k=1}^3 \frac{\partial x_k}{\partial q_i} \frac{\partial x_k}{\partial q_j} = \frac{\partial r}{\partial q_i} \cdot \frac{\partial r}{\partial q_j}. \quad (1.6)$$

Os coeficientes g_{ij} especificam a natureza do sistema de coordenadas (q_1, q_2, q_3) . Coletivamente, estes coeficientes compõem a *métrica* do espaço \mathbb{E}^3 .² Por sua vez, os g_{ij} são os componentes de um tensor de posto 2, denominado **tensor de métrica**.³ Na relatividade geral, os componentes do tensor de métrica são determinados pelas propriedades da matéria contida no espaço, ou seja, a métrica ou a geometria do espaço passa a ser determinada pelas propriedades físicas deste: geometria se mescla com a física.

²Espaços métricos são definidos na seção 4.12.

³Existem outras condições que são impostas ao tensor de métrica para que a forma quadrática (1.5) forme um espaço Riemanniano. Essas condições são discutidas na seção 6.7.2.

Em (1.6), observa-se que os componentes de g_{ij} podem ser expressos em termos dos **veto- res tangentes** $e_i = e_i(q_1, q_2, q_3) \equiv \partial \mathbf{r} / \partial q_i$. O vetor e_i está na direção da reta tangente à curva coordenada q_i no ponto P e no seu sentido crescente, quando $q_j = \text{cte.}$ ($j \neq i$), conforme pode ser visualizado na figura 1.2. Neste caso, pode-se ver que o conjunto $\{e_1, e_2, e_3\}$ é composto por vetores linearmente independentes e, portanto, forma uma base para o \mathbb{E}^3 , alternativa à base canônica. Nesta base, o vetor deslocamento infinitesimal (1.3), nas vizinhanças do ponto P , pode ser escrito como

$$d\mathbf{r} = \sum_{i=1}^3 \frac{\partial \mathbf{r}}{\partial q_i} dq_i = \sum_{i=1}^3 dq_i e_i.$$

Observa-se aqui também que os vetores tangentes podem ser expressos através da base canônica no \mathbb{E}^3 como

$$e_i = \frac{\partial \mathbf{r}}{\partial q_i} = \sum_{j=1}^3 \frac{\partial x_j}{\partial q_i} \hat{\mathbf{x}}_j = \sum_{j=1}^3 H_{ji} \hat{\mathbf{x}}_j, \text{ onde } H_{ij} = \frac{\partial x_i}{\partial q_j}, \quad (1.7a)$$

sendo que a relação inversa também é útil, isto é,

$$\hat{\mathbf{x}}_i = \sum_{j=1}^3 (H^{-1})_{ji} e_j. \quad (1.7b)$$

A partir deste resultado, calculando o elemento de arco $d\ell^2 = d\mathbf{r} \cdot d\mathbf{r}$ e comparando com (1.6), pode-se identificar os componentes do tensor de métrica como

$$g_{ij} = e_i \cdot e_j = \sum_{k=1}^3 H_{ki} H_{kj}. \quad (1.8)$$

Uma outra base neste sistema de coordenadas pode ser formada, tomando-se os vetores normais às superfícies coordenadas no ponto P . Estes vetores são definidos por $\epsilon_i = \nabla q_i$ e também estão ilustrados na figura 1.2. Desta forma, um vetor $\mathbf{a} \in \mathbb{E}^3$ qualquer pode ser escrito tanto em termos da base $\{e_i\}$ quanto da base $\{\epsilon_i\}$ como

$$\mathbf{a} = \sum_{i=1}^3 \alpha_i e_i = \sum_{i=1}^3 \beta_i \epsilon_i. \quad (1.9)$$

Os componentes $\{\alpha_i\}$ e $\{\beta_i\}$ são denominados, respectivamente, de componentes **contravariantes** e **covariantes** de \mathbf{a} .

Os componentes contravariantes $\{\alpha_i\}$ podem ser expressos em termos da inversa do tensor de métrica,

$$\mathbf{a} \cdot e_j = \sum_{i=1}^3 \alpha_i g_{ij} \implies \alpha_i = \sum_{j=1}^3 (g^{-1})_{ij} (\mathbf{a} \cdot e_j). \quad (1.10a)$$

Por sua vez, as bases $\{e_i\}$ e $\{\epsilon_i\}$ formam um conjunto de vetores *recíprocos*. Ou seja, usando a base canônica para \mathbf{r} e ∇ e a regra de derivação em cadeia, pode-se calcular

$$\begin{aligned} e_i \cdot \epsilon_j &= \frac{\partial \mathbf{r}}{\partial q_i} \cdot \nabla q_j \\ &= \left(\frac{\partial x}{\partial q_i} \hat{\mathbf{i}} + \frac{\partial y}{\partial q_i} \hat{\mathbf{j}} + \frac{\partial z}{\partial q_i} \hat{\mathbf{k}} \right) \cdot \left(\frac{\partial q_j}{\partial x} \hat{\mathbf{i}} + \frac{\partial q_j}{\partial y} \hat{\mathbf{j}} + \frac{\partial q_j}{\partial z} \hat{\mathbf{k}} \right) \\ &= \frac{\partial q_j}{\partial x} \frac{\partial x}{\partial q_i} + \frac{\partial q_j}{\partial y} \frac{\partial y}{\partial q_i} + \frac{\partial q_j}{\partial z} \frac{\partial z}{\partial q_i} = \frac{\partial q_j}{\partial q_i}. \end{aligned}$$

Portanto, $e_i \cdot \epsilon_j = \delta_{ij}$, o que caracteriza dois conjuntos de vetores recíprocos.

Escrevendo os vetores de base $\{\epsilon_i\}$ em termos da base $\{e_i\}$ como

$$\epsilon_i = \sum_{j=1}^3 \gamma_{ij} e_j,$$

o fato destas bases serem recíprocas permite escrever

$$\boldsymbol{\epsilon}_i \cdot \boldsymbol{e}_j = \sum_{k=1}^3 \gamma_{ik} \boldsymbol{e}_k \cdot \boldsymbol{e}_j \implies \delta_{ij} = \sum_{k=1}^3 \gamma_{ik} g_{kj}, \text{ ou seja, } \gamma_{ij} = (g^{-1})_{ij}.$$

Portanto,

$$\boldsymbol{\epsilon}_i = \sum_{j=1}^3 (g^{-1})_{ij} \boldsymbol{e}_j, \quad \boldsymbol{e}_i = \sum_{j=1}^3 g_{ij} \boldsymbol{\epsilon}_j.$$

Desta maneira, pode-se escrever os componentes covariantes $\{\beta_i\}$ de \boldsymbol{a} em (1.9) simplesmente como

$$\boldsymbol{a} = \sum_{i=1}^3 \beta_i \boldsymbol{\epsilon}_i = \sum_{i,j=1}^3 \beta_i (g^{-1})_{ij} \boldsymbol{e}_j \implies \beta_i = \boldsymbol{a} \cdot \boldsymbol{e}_i. \quad (1.10b)$$

Observando as expressões (1.10a,b), pode-se concluir que as relações entre os componentes contravariantes e covariantes do vetor \boldsymbol{a} são determinadas pelo tensor de métrica como

$$\alpha_i = \sum_{j=1}^3 (g^{-1})_{ij} \beta_j \quad (1.10c)$$

$$\beta_i = \sum_{j=1}^3 g_{ij} \alpha_j. \quad (1.10d)$$

Uma discussão mais detalhada destes componentes, no contexto da análise tensorial, será realizada na seção 6.7.

Finalmente, os vetores de base $\{\boldsymbol{e}_i\}$ e $\{\boldsymbol{\epsilon}_i\}$ podem ser normalizados por

$$\hat{\boldsymbol{e}}_i = \frac{1}{h_i} \frac{\partial \boldsymbol{r}}{\partial q_i}, \quad \hat{\boldsymbol{\epsilon}}_i = \frac{\nabla q_i}{|\nabla q_i|},$$

onde

$$h_i \equiv |\boldsymbol{e}_i| = \left| \frac{\partial \boldsymbol{r}}{\partial q_i} \right| = \sqrt{\sum_{j=1}^3 \left(\frac{\partial x_j}{\partial q_i} \right)^2} = \sqrt{\sum_{j=1}^3 H_{ij}^2}, \quad (1.11)$$

é denominado o **fator de escala** da coordenada q_i , formando assim as bases normais $\{\hat{\boldsymbol{e}}_i\}$ e $\{\hat{\boldsymbol{\epsilon}}_i\}$. Estes vetores normalizados estão ilustrados na figura 1.2. A forma matemática dos fatores de escala depende do sistema de coordenadas adotado. Por exemplo, no sistema Cartesiano, $h_i = 1$, simplesmente. Cabe ressaltar também que embora $\boldsymbol{\epsilon}_i$ seja o recíproco de \boldsymbol{e}_i , estes vetores não são necessariamente paralelos entre si.

Pode-se então escrever o vetor deslocamento infinitesimal dado por (1.3), tanto em termos dos fatores de escala quanto em termos da matriz H definida em (1.7), como

$$d\boldsymbol{r} = \sum_{i=1}^3 dq_i \boldsymbol{e}_i = \sum_{i=1}^3 h_i dq_i \hat{\boldsymbol{e}}_i = \sum_{i,j=1}^3 H_{ij} dq_j \hat{\boldsymbol{x}}_i. \quad (1.12)$$

Esta expressão é útil também quando o deslocamento infinitesimal $d\boldsymbol{r}$ é ocasionado pela variação de algum parâmetro contínuo (θ , por exemplo). Neste caso, o vetor posição percorre uma curva no espaço determinada pela forma paramétrica $\boldsymbol{r} = \boldsymbol{r}(\theta)$ e o vetor $d\boldsymbol{r}$ é o vetor tangente a esta curva quando $\theta \rightarrow \theta + d\theta$. Dividindo ambos os lados de (1.12) por $d\theta$ resulta então

$$\frac{d\boldsymbol{r}}{d\theta} = \sum_{i=1}^3 h_i \frac{dq_i}{d\theta} \hat{\boldsymbol{e}}_i,$$

sendo $dq_i/d\theta$ a variação da coordenada q_i com o parâmetro θ .

O resultado acima é particularmente útil na física, onde usualmente o parâmetro livre corresponde ao tempo. Neste caso, $\boldsymbol{v} = d\boldsymbol{r}/dt \doteq \dot{\boldsymbol{r}}$, sendo $\boldsymbol{v} = \boldsymbol{v}(t)$ o vetor velocidade. Portanto,

$$\boldsymbol{v} \equiv \dot{\boldsymbol{r}} = \frac{d\boldsymbol{r}}{dt} = \sum_{i=1}^3 h_i \frac{dq_i}{dt} \hat{\boldsymbol{e}}_i = \sum_{i=1}^3 h_i \dot{q}_i \hat{\boldsymbol{e}}_i, \quad (1.13)$$

onde foi introduzida a notação $\dot{q}_i = dq_i/dt$, comum na mecânica clássica.

1.2 COORDENADAS CURVILÍNEAS ORTOGONAIS

A partir deste ponto, a atenção estará voltada exclusivamente para sistemas de coordenadas *ortogonais*, isto é, para famílias de superfícies mutuamente ortogonais em cada ponto do \mathbb{E}^3 . Neste caso em particular, a base $\{\hat{e}_i\}$ é ortonormal, existindo também uma *regra dextrógi*ra entre os seus vetores, isto é,

$$\hat{e}_i \times \hat{e}_j = \sum_{k=1}^3 \epsilon_{ijk} \hat{e}_k \implies \begin{cases} \hat{e}_1 \times \hat{e}_2 = \hat{e}_3 \\ \hat{e}_3 \times \hat{e}_1 = \hat{e}_2 \\ \hat{e}_2 \times \hat{e}_3 = \hat{e}_1, \end{cases} \quad (1.14)$$

onde ϵ_{ijk} é o tensor de Levi-Civita.⁴ Esta regra se deve ao fato de que o sistema de coordenadas usualmente empregado possui uma orientação dextrógira.

Como $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$, o tensor de métrica (1.6), de acordo com (1.8), possui todos os seus elementos fora da diagonal principal nulos, isto é, $g_{ij} = 0$, para $(i \neq j)$. Pode-se escrever então os componentes g_{ii} não nulos em termos dos fatores de escala, a partir de (1.8), como

$$g_{ij} = h_i^2 \delta_{ij} = \sum_{k=1}^3 H_{ki} H_{kj}, \quad (1.15a)$$

o que permite concluir que

$$\sum_{k=1}^3 H_{ki} H_{kj} = h_i^2 \delta_{ij}. \quad (1.15b)$$

Já as relações (1.7) e (1.15a) mostram que

$$\sum_{k=1}^3 h_k^2 (H^{-1})_{ki} (H^{-1})_{kj} = \delta_{ij}. \quad (1.15c)$$

Além disso, o determinante da matriz g que representa o tensor de métrica é, simplesmente,

$$\det(g) = (h_1 h_2 h_3)^2 = \sum_{i,j,k=1}^3 H_{i1}^2 H_{j2}^2 H_{k3}^2,$$

ao passo que a sua inversa é

$$(g^{-1})_{ij} = h_i^{-2} \delta_{ij}.$$

Neste caso, o elemento infinitesimal de arco $d\ell^2$ em (1.5b) se reduz à seguinte forma quadrática:

$$d\ell^2 = (h_1 dq_1)^2 + (h_2 dq_2)^2 + (h_3 dq_3)^2 = \sum_{i=1}^3 h_i^2 dq_i^2. \quad (1.16)$$

A forma resultante para $d\ell^2$ em (1.16) permite definir a *projeção* de dr ao longo da coordenada q_i como sendo

$$dl_i = h_i dq_i, \quad (1.17)$$

sendo dl_i obtido mantendo-se as outras coordenadas q_j ($j \neq i$) constantes. Porém, o fator de escala h_i pode depender de todas as novas coordenadas, isto é, em geral,

$$h_i = h_i(q_1, q_2, q_3), \quad (i = 1, 2, 3),$$

e pode possuir dimensão de comprimento, em cuja situação q_i não terá dimensão (um ângulo, por exemplo). O vínculo consiste em que dl_i deve possuir dimensão de comprimento.

As bases $\{e_i\}$ e $\{\epsilon_i\}$ serão empregadas novamente na seção 6.7, no contexto da análise tensorial. Para o restante da discussão neste capítulo será empregada somente a base ortonormal $\{\hat{e}_i\}$.

⁴Tensor totalmente antissimétrico. Ver seção 6.1.2.

1.3 ANÁLISE VETORIAL EM SISTEMAS DE COORDENADAS CURVILÍNEAS ORTOGONAIS

Nesta seção as operações de álgebra vetorial, os operadores diferenciais vetoriais e os teoremas integrais para campos escalares e vetoriais serão derivados para um sistema de coordenadas curvilíneas ortogonais em geral.

1.3.1 ÁLGEBRA VETORIAL

Consideram-se dois vetores $\{\mathbf{a}, \mathbf{b}\} \in \mathbb{E}^3$, os quais podem ser escritos tanto em termos da base canônica quanto em termos da base ortonormal $\{\hat{e}_i\}$ como

$$\mathbf{a} = a_x \hat{i} + a_y \hat{j} + a_z \hat{k} = \sum_{i=1}^3 a_i \hat{e}_i \quad \text{e} \quad \mathbf{b} = b_x \hat{i} + b_y \hat{j} + b_z \hat{k} = \sum_{i=1}^3 b_i \hat{e}_i.$$

A adição vetorial se procede como o usual,

$$\mathbf{a} + \mathbf{b} = \sum_{i=1}^3 (a_i + b_i) \hat{e}_i.$$

No caso dos produtos envolvendo vetores, as transformações (1.1) e (1.2) por hipótese são tais que o escalar ou vetor resultantes da operação são invariantes, ou seja, estes objetos geométricos são os mesmos para qualquer sistema de coordenadas curvilíneas. As condições que as leis de transformação (1.1) e (1.2) devem satisfazer para que esta exigência seja satisfeita serão discutidas na seção 6.3.3, dentro do contexto da análise tensorial. Esta condição já foi tacitamente suposta cumprida na seção anterior.

Assumindo que as condições mencionadas acima são satisfeitas, então o resultado do produto escalar entre os vetores \mathbf{a} e \mathbf{b} é, simplesmente,

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^3 a_i b_i.$$

Por sua vez, o resultado do produto vetorial entre os mesmos vetores, de acordo com (1.14), é

$$\mathbf{a} \times \mathbf{b} = \sum_{i,j=1}^3 a_i b_j \hat{e}_i \times \hat{e}_j = \sum_{i,j,k=1}^3 \epsilon_{ijk} a_j b_k \hat{e}_i.$$

Observa-se que o mesmo resultado pode ser obtido pelo cálculo do determinante

$$\mathbf{a} \times \mathbf{b} = \det \begin{pmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}.$$

1.3.2 INTEGRAIS DE CAMINHO, DE SUPERFÍCIE E DE VOLUME

Para o cálculo dos teoremas integrais em coordenadas curvilíneas, é necessário adaptar as integrais em uma, duas ou três dimensões, conforme definidas no sistema Cartesiano, para o sistema de coordenadas geral.

Considerando o campo vetorial $\mathbf{A} = A(\mathbf{r})$,⁵ cujas componentes são

$$\mathbf{A}(\mathbf{r}) = \sum_{i=1}^3 A_i(\mathbf{r}) \hat{e}_i,$$

⁵Uma definição mais rigorosa de um campo vetorial é realizada na seção 6.3.2. A transformação das componentes do campo $\mathbf{A}(\mathbf{r})$ é discutida na seção 6.8.

para calcular a integral de $A(\mathbf{r})$ ao longo de uma curva C contida no \mathbb{E}^3 usando o sistema $\{q_1, q_2, q_3\}$, parte-se do elemento de arco $d\ell$, cujas projeções sobre as coordenadas curvilíneas são dadas por (1.17). Então,

$$\int_C \mathbf{A}(\mathbf{r}) \cdot d\ell = \sum_{i=1}^3 \int_C A_i(\mathbf{r}) h_i dq_i. \quad (1.18)$$

Para o cálculo de uma integral de superfície envolvendo $A(\mathbf{r})$, por exemplo, para se obter o fluxo deste campo através de uma superfície S contida no \mathbb{E}^3 , é necessário, em primeiro lugar, determinar a forma do elemento de superfície no sistema curvilíneo.

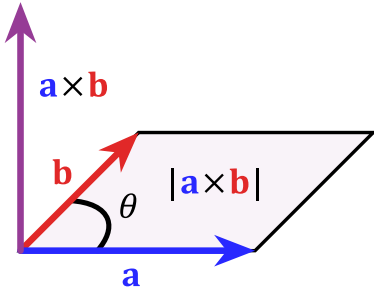


Figura 1.3: A área de um paralelogramo como o módulo de um produto vetorial.

Fazendo uma interpretação geométrica usualmente realizada com o sistema Cartesiano, observa-se, em primeiro lugar, que, dados os vetores \mathbf{a} e \mathbf{b} , a quantidade escalar $|\mathbf{a} \times \mathbf{b}| = ab \sin \theta$ fornece a área do paralelogramo definido por estes, sendo $0 \leq \theta \leq \pi$ o menor ângulo entre os mesmos. Assim, o vetor $\mathbf{s} = \mathbf{a} \times \mathbf{b}$ pode ser interpretado geometricamente como um vetor superfície, cuja área é igual à do paralelogramo mencionado e com a direção e sentido definidos pelo resultado do produto vetorial; ou seja, o vetor \mathbf{s} é sempre ortogonal ao plano definido por \mathbf{a} e \mathbf{b} , com o sentido dado pelo produto de ambos. Este vetor é ilustrado na figura 1.3.

Uma superfície S qualquer contida no \mathbb{E}^3 pode ser subdividida em um número muito grande de superfícies infinitesimais. Com base na interpretação geométrica do produto vetorial, pode-se então definir um *vetor elemento de superfície* $d\sigma$ a partir da área delimitada entre dois elementos de arco $d\ell$ e $d\ell'$ que não estejam paralelos entre si. Com esta notação, obtém-se com a base canônica

$$d\ell \times d\ell' = \sum_{i,j=1}^3 dx_i dx'_j \hat{\mathbf{x}}_i \times \hat{\mathbf{x}}_j = \sum_{i,j,k=1}^3 \epsilon_{ijk} dx_j dx_k \hat{\mathbf{x}}_i.$$

Então, uma definição apropriada para o vetor elemento de superfície de S é

$$d\sigma = \frac{1}{2} \sum_{i,j,k=1}^3 |\epsilon_{ijk}| d\sigma_{jk} \hat{\mathbf{x}}_i = d\sigma_{23} \hat{\mathbf{x}}_1 + d\sigma_{13} \hat{\mathbf{x}}_2 + d\sigma_{12} \hat{\mathbf{x}}_3, \quad (1.19)$$

sendo $d\sigma_{ij} = dx_i dx_j$ o elemento de superfície no plano $x_i - x_j$ ($j \neq i$), perpendicular ao vetor unitário $\hat{\mathbf{x}}_k$ ($k \neq i$ e $k \neq j$).

Nota-se que esta definição satisfaz a interpretação geométrica de $d\sigma$. Por exemplo, se a superfície S estiver contida totalmente no plano $x_2 - x_3 \leftrightarrow y - z$, então $dx_1 = dx = 0$ e

$$d\sigma = d\sigma_{23} \hat{\mathbf{x}}_1 = dx_2 dx_3 \hat{\mathbf{x}}_1 = dy dz \hat{\mathbf{i}},$$

como era o esperado.

Passando agora para o sistema curvilíneo $\{q_1, q_2, q_3\}$, o elemento de arco $d\ell$ é dado por (1.12) e daí o elemento de superfície $d\sigma$ pode ser definido, por analogia a (1.19), como

$$d\sigma = \frac{1}{2} \sum_{i,j,k=1}^3 |\epsilon_{ijk}| d\sigma_{jk} \hat{\mathbf{e}}_i = d\sigma_{23} \hat{\mathbf{e}}_1 + d\sigma_{13} \hat{\mathbf{e}}_2 + d\sigma_{12} \hat{\mathbf{e}}_3, \quad (1.20a)$$

sendo agora $d\sigma_{ij} \equiv d\ell_i d\ell_j = h_i h_j dq_i dq_j$ ($j \neq i$).

Portanto, a integral de fluxo de $\mathbf{A}(\mathbf{r})$ através de S resulta

$$\begin{aligned} \int_S \mathbf{A} \cdot d\sigma &= \frac{1}{2} \sum_{i,j,k=1}^3 |\epsilon_{ijk}| \int_S A_i d\sigma_{jk} = \frac{1}{2} \sum_{i,j,k=1}^3 |\epsilon_{ijk}| \int_S A_i h_j h_k dq_j dq_k, \\ \int_S \mathbf{A} \cdot d\sigma &= \int_S A_1 h_2 h_3 dq_2 dq_3 + \int_S A_2 h_1 h_3 dq_1 dq_3 + \int_S A_3 h_1 h_2 dq_1 dq_2. \end{aligned} \quad (1.20b)$$

Finalmente, para o cálculo de uma integral de volume é necessário o elemento infinitesimal de volume d^3r no \mathbb{E}^3 . Novamente realizando uma interpretação geométrica, ilustrada na figura 1.4, dados os vetores \mathbf{a} , \mathbf{b} e \mathbf{c} , o produto misto

$$V_p = |\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})| = abc \operatorname{sen} \theta \cos \alpha$$

fornece o volume do paralelepípedo definido pelos vetores.

Então, no sistema Cartesiano, o elemento de volume é escrito simplesmente como

$$d^3r = dx dy dz \left| \hat{i} \cdot (\hat{j} \times \hat{k}) \right| = dx dy dz.$$

Por consequência, se $d\ell_i = h_i dq_i \hat{e}_i$ é o vetor elemento de arco na direção i , então o correspondente elemento de volume pode ser calculado como

$$d^3r = |d\ell_1 \cdot (d\ell_2 \times d\ell_3)| = h_1 h_2 h_3 dq_1 dq_2 dq_3, \quad (1.21a)$$

uma vez que $|\hat{e}_1 \cdot (\hat{e}_2 \times \hat{e}_3)| = 1$ para um sistema de coordenadas ortogonais.

Uma expressão equivalente para d^3r é obtida escrevendo-se, inicialmente, $d\ell_i = dq_i \mathbf{e}_i$. Então, de (1.7a), resulta

$$d\ell_i = \sum_{j=1}^3 H_{ji} dq_j \hat{\mathbf{x}}_j.$$

Percebe-se então que o elemento de volume fica

$$\begin{aligned} d^3r &= \left| \sum_{i,j,k=1}^3 H_{i1} H_{j2} H_{k3} \hat{\mathbf{x}}_i \cdot (\hat{\mathbf{x}}_j \times \hat{\mathbf{x}}_k) \right| dq_1 dq_2 dq_3 \\ &= \left| \sum_{i,j,k=1}^3 \epsilon_{ijk} H_{i1} H_{j2} H_{k3} \right| dq_1 dq_2 dq_3. \end{aligned}$$

Porém, uma expressão conhecida envolvendo o tensor de Levi-Civita e o determinante de uma matriz A de ordem 3 é

$$\det(A) \epsilon_{abc} = \sum_{i,j,k=1}^3 \epsilon_{ijk} A_{ai} A_{bj} A_{ck} = \sum_{i,j,k=1}^3 \epsilon_{ijk} A_{ia} A_{jb} A_{kc}.$$

Portanto,

$$d^3r = |\det(H)| dq_1 dq_2 dq_3. \quad (1.21b)$$

Mas

$$J \equiv \det(H) = \det \begin{pmatrix} \frac{\partial x_1}{\partial q_1} & \frac{\partial x_1}{\partial q_2} & \frac{\partial x_1}{\partial q_3} \\ \frac{\partial x_2}{\partial q_1} & \frac{\partial x_2}{\partial q_2} & \frac{\partial x_2}{\partial q_3} \\ \frac{\partial x_3}{\partial q_1} & \frac{\partial x_3}{\partial q_2} & \frac{\partial x_3}{\partial q_3} \end{pmatrix} \equiv \frac{\partial(x_1, x_2, x_3)}{\partial(q_1, q_2, q_3)},$$

o qual é justamente o Jacobiano da transformação $\{x_1, x_2, x_3\} \rightarrow \{q_1, q_2, q_3\}$. Ou seja,

$$d^3r = |J| dq_1 dq_2 dq_3 = \left| \frac{\partial(x_1, x_2, x_3)}{\partial(q_1, q_2, q_3)} \right| dq_1 dq_2 dq_3. \quad (1.21c)$$

Portanto, para sistemas de coordenadas curvilíneas ortogonais, infere-se que $|J| = J = h_1 h_2 h_3$. Com este elemento de volume, pode-se agora calcular as integrais de volume no sistema curvilíneo.

Exercício 1.1. Mostre que para sistemas de coordenadas curvilíneas ortogonais o Jacobiano da transformação está relacionado aos fatores de escala por $J = h_1 h_2 h_3$.

Solução. Para demonstrar este resultado, é mais fácil elevar ambos os lados ao quadrado. Então, de (1.21c) e (1.11), deseja-se demonstrar que

$$J^2 = \left(\sum_{i,j,k=1}^3 \epsilon_{ijk} H_{i1} H_{j2} H_{k3} \right)^2 = (h_1 h_2 h_3)^2,$$

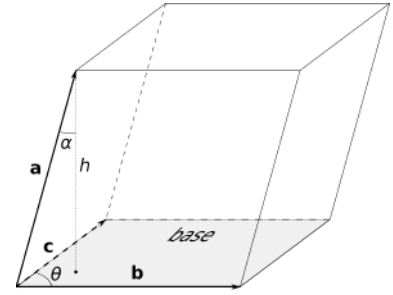


Figura 1.4: Volume de um paralelepípedo como o módulo do produto misto.

$$\left(\sum_{i,j,k=1}^3 \epsilon_{ijk} H_{i1} H_{j2} H_{k3} \right)^2 = \sum_{i,j,k=1}^3 H_{i1}^2 H_{j2}^2 H_{k3}^2.$$

O lado esquerdo pode ser escrito como

$$\left(\sum_{i,j,k=1}^3 \epsilon_{ijk} H_{i1} H_{j2} H_{k3} \right)^2 = \sum_{i,j,k=1}^3 \sum_{\ell,m,n=1}^3 \epsilon_{ijk} \epsilon_{lmn} H_{i1} H_{j2} H_{k3} H_{\ell 1} H_{m2} H_{n3}.$$

Usando a propriedade (6.3),

$$\epsilon_{ijk} \epsilon_{lmn} = \delta_{i\ell} (\delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}) + \delta_{im} (\delta_{jn} \delta_{k\ell} - \delta_{j\ell} \delta_{kn}) + \delta_{in} (\delta_{j\ell} \delta_{km} - \delta_{jm} \delta_{k\ell}),$$

resulta

$$\begin{aligned} \left(\sum_{i,j,k=1}^3 \epsilon_{ijk} H_{i1} H_{j2} H_{k3} \right)^2 &= \sum_{i,j,k=1}^3 [H_{i1} H_{j2} H_{k3} - H_{i1} H_{k2} H_{j3} \\ &\quad + H_{k1} H_{i2} H_{j3} - H_{j1} H_{i2} H_{k3} + H_{j1} H_{k2} H_{i3} - H_{k1} H_{j2} H_{i3}] H_{i1} H_{j2} H_{k3}. \end{aligned}$$

Agora, de acordo com (1.15b), somente o primeiro termo permanece não nulo na expressão acima, resultando então

$$J^2 = \sum_{i,j,k=1}^3 H_{i1}^2 H_{j2}^2 H_{k3}^2 = (h_1 h_2 h_3)^2,$$

como se queria demonstrar.

1.4 OPERADORES VETORIAIS DIFERENCIAIS

Nesta seção serão obtidas as expressões gerais para os operadores gradiente, divergente e rotacional, válidas para qualquer sistema de coordenadas ortogonais.

1.4.1 GRADIENTE

O ponto de partida para a obtenção dos operadores vetoriais diferenciais em qualquer sistema de coordenadas ortogonal consiste na interpretação geométrica do gradiente como aquele vetor que possui o módulo, direção e sentido da máxima taxa de variação de um determinado campo escalar sobre uma de suas superfícies equipotenciais.

Sendo $\psi = \psi(q_1, q_2, q_3)$ um campo escalar, função das coordenadas curvilíneas ortogonais, a taxa máxima de variação deste campo sobre a superfície $q_1 = \text{cte.}$, por exemplo, será dada pelo maior valor da derivada direcional em qualquer ponto sobre esta superfície. A variação de ψ , de uma forma geral, é dada por

$$d\psi = \sum_{i=1}^3 \frac{\partial \psi}{\partial q_i} dq_i, \quad (1.22)$$

ao passo que um arco infinitesimal $d\ell$ é dado por (1.12),

$$d\ell = \sum_{i=1}^3 h_i dq_i \hat{e}_i.$$

Portanto, a variação do campo ψ ao longo de $d\ell$ será dado pela derivada direcional

$$d\psi|_{d\ell} = \nabla \psi \cdot d\ell,$$

sendo $\nabla \psi$ a expressão procurada para o gradiente do campo. Então,

$$d\psi|_{d\ell} = \sum_{i=1}^3 (\nabla \psi)_i h_i dq_i,$$

sendo $(\nabla\psi)_i$ a expressão procurada. Comparando com (1.22), resulta que

$$h_i (\nabla\psi)_i = \frac{\partial\psi}{\partial q_i} \implies (\nabla\psi)_i = \frac{1}{h_i} \frac{\partial\psi}{\partial q_i}.$$

Portanto,

$$\nabla\psi = \sum_{i=1}^3 \frac{1}{h_i} \frac{\partial\psi}{\partial q_i} \hat{e}_i. \quad (1.23)$$

Pode-se verificar facilmente a interpretação geométrica do gradiente. Sejam P e Q dois pontos infinitesimalmente próximos sobre a superfície equipotencial $\psi = \text{cte}$. Ao longo do deslocamento $P \rightarrow Q$, a taxa de variação do campo é, portanto, $d\psi = \nabla\psi \cdot d\ell = 0$. Ou seja, o gradiente é perpendicular à superfície equipotencial em P . Agora, sejam $S_1: \psi = C_1$ e $S_2: \psi' = C_2$ duas superfícies infinitesimalmente próximas entre si e seja P um ponto sobre S_1 e Q um ponto sobre S_2 . Ao longo do deslocamento $P \rightarrow Q$, a variação do campo será agora $d\psi = \psi' - \psi = C_2 - C_1 = \nabla\psi \cdot d\ell$. Dentre todas as posições relativas entre P e Q , a taxa de variação do campo, para um dado $|d\ell|$, será maximizada quando o ângulo entre $\nabla\psi$ e $d\ell$ for nulo. Ou seja, o gradiente está na direção de máxima variação do campo.

Exercício 1.2. Mostre que para coordenadas curvilíneas ortogonais, $\nabla q_i = \hat{e}_i/h_i$. Como consequência, mostre que os dois conjuntos de vetores recíprocos $\{\hat{e}_i\}$ e $\{\hat{\epsilon}_i\}$ são idênticos neste caso.

Solução. Fazendo $\psi = q_i$ em (1.23), resulta

$$\nabla q_i = \sum_{j=1}^3 \frac{1}{h_j} \frac{\partial q_i}{\partial q_j} \hat{e}_j = \frac{\hat{e}_i}{h_i} = \epsilon_i.$$

Portanto, como $|\nabla q_i| = h_i^{-1}$, resulta que

$$\hat{\epsilon}_i = \frac{\nabla q_i}{|\nabla q_i|} = \hat{e}_i,$$

como se desejava demonstrar.

1.4.2 DIVERGENTE E LAPLACIANO

O operador divergente pode ser obtido a partir de sua definição relacionada ao teorema do divergente. Sendo $\mathbf{A} = \mathbf{A}(q_1, q_2, q_3)$ um campo vetorial definido em um volume elementar d^3r do \mathbb{E}^3 , centrado em torno do ponto $P = (q_1, q_2, q_3)$ e delimitado pela superfície Gaussiana S , então o divergente do campo em P é dado por

$$\nabla \cdot \mathbf{A}(q_1, q_2, q_3) = \lim_{\int d^3r \rightarrow 0} \frac{\oint_S \mathbf{A} \cdot d\boldsymbol{\sigma}}{\int d^3r}.$$

A figura 1.5 ilustra o elemento de volume d^3r em questão, sendo que os lados do elemento de volume são os elementos infinitesimais de arco $dl_i = h_i dq_i$, ao longo de cada coordenada curvilínea. Consideram-se agora as integrais de superfície através de cada lado do elemento de volume.

Por exemplo, a integral de fluxo de \mathbf{A} através de uma das superfícies $S_1: q_1 = \text{cte.}$, correspondentes à face 3 e sua oposta na figura 1.5. A partir de (1.20a,b), esta integral é dada simplesmente por

$$\Phi_1 = \int_{S_1} \mathbf{A} \cdot d\boldsymbol{\sigma} = \pm \int_S A_1 h_2 h_3 dq_2 dq_3 \approx \pm A_1(q_1, q_2, q_3) h_2 h_3 dq_2 dq_3,$$

uma vez que a área de S_1 é infinitesimal.

O sinal será tomado a partir de uma convenção para fluxo positivo. Como o sistema de coordenadas é dextrógiro, isto é, $\hat{e}_2 \times \hat{e}_3 = \hat{e}_1$, o fluxo de \mathbf{A} através da face 3 ou sua oposta

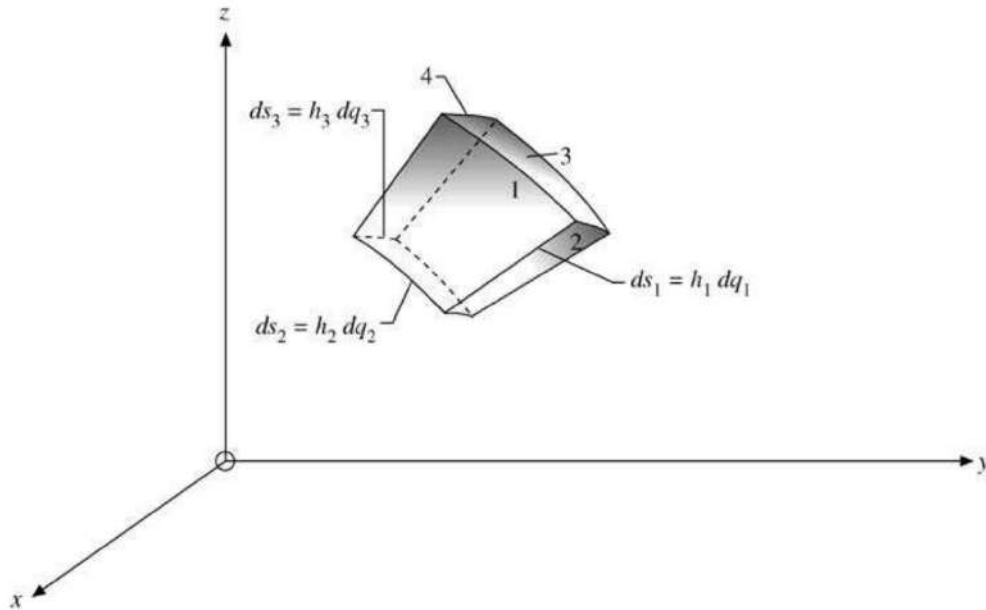


Figura 1.5: Um elemento de volume em termos das coordenadas curvilíneas ortogonais.

será tomado positivo quando este ocorrer no sentido de \hat{e}_1 , ou seja, através da face em $q_1 + dq_1$. Nesta,⁶

$$\Phi'_1 = \int_{S_1} \mathbf{A} \cdot d\boldsymbol{\sigma} \approx A_1(q_1 + dq_1, q_2, q_3) h_2 h_3 dq_2 dq_3 \approx \left[A_1(q_1, q_2, q_3) h_2 h_3 + \frac{\partial}{\partial q_1} (A_1 h_2 h_3) dq_1 \right] dq_2 dq_3.$$

Portanto, o fluxo líquido através da superfície 3 e de sua oposta será

$$\Delta\Phi_1 = \Phi'_1 - \Phi_1 \approx \frac{\partial}{\partial q_1} (A_1 h_2 h_3) dq_1 dq_2 dq_3.$$

Procedendo da mesma maneira nas superfícies 1 e oposta, 2 e 4 da figura 1.5, obtém-se que o fluxo total através da superfície gaussiana elementar é

$$\oint_S \mathbf{A} \cdot d\boldsymbol{\sigma} \approx \left[\frac{\partial}{\partial q_1} (A_1 h_2 h_3) + \frac{\partial}{\partial q_2} (A_2 h_1 h_3) + \frac{\partial}{\partial q_3} (A_3 h_1 h_2) \right] dq_1 dq_2 dq_3.$$

Por sua vez, a integral de volume é, de acordo com (1.21), simplesmente,

$$\int d^3r \approx h_1 h_2 h_3 dq_1 dq_2 dq_3.$$

Tomando o limite $d^3r \rightarrow 0$, resulta a expressão desejada para o divergente do campo vetorial,

$$\nabla \cdot \mathbf{A}(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1} (A_1 h_2 h_3) + \frac{\partial}{\partial q_2} (A_2 h_1 h_3) + \frac{\partial}{\partial q_3} (A_3 h_1 h_2) \right]. \quad (1.24)$$

O operador laplaciano em coordenadas curvilíneas ortogonais é obtido simplesmente aplicando-se o divergente (1.24) sobre o gradiente (1.23) de um campo escalar ψ ,

$$\nabla^2 \psi(q_1, q_2, q_3) = \nabla \cdot (\nabla \psi),$$

ou seja,

$$\nabla^2 \psi(q_1, q_2, q_3) = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^3 \frac{\partial}{\partial q_i} \left(\frac{h_1 h_2 h_3}{h_i^2} \frac{\partial \psi}{\partial q_i} \right). \quad (1.25)$$

⁶Lembrando que $f(\mathbf{a} + \boldsymbol{\alpha}) \approx f(\mathbf{a}) + \boldsymbol{\alpha} \cdot (\partial f / \partial \mathbf{a})$.

1.4.3 ROTACIONAL

Finalmente, dado novamente o campo vetorial \mathbf{A} , o seu rotacional no ponto P será obtido a partir de sua definição baseada no teorema de Stokes. Em \mathbb{E}^3 , se S é uma superfície delimitada pela curva fechada C , então

$$(\nabla \times \mathbf{A}) \cdot \hat{\mathbf{n}} = \lim_{S \rightarrow 0} \left(\frac{1}{S} \oint_C \mathbf{A} \cdot d\ell \right),$$

onde $\hat{\mathbf{n}}$ é o vetor unitário normal a S .

Tomando como guia geométrico agora a figura 1.6, a qual ilustra o elemento de superfície $d\sigma_{23}$, orientado sobre a superfície $S_1: q_1 = \text{cte.}$, calcula-se então a circulação do campo \mathbf{A} ao longo do caminho fechado ℓ_{14} : $1 \rightarrow 4$ que delimita $d\sigma_{23}$. De acordo com o teorema da média e de (1.20b), pode-se escrever

$$\int_{S_1} \nabla \times \mathbf{A} \cdot d\sigma = \hat{\mathbf{e}}_1 \cdot [\nabla \times \mathbf{A}(q_1, q_2^*, q_3^*)] d\sigma_{23} = \hat{\mathbf{e}}_1 \cdot [\nabla \times \mathbf{A}(q_1, q_2^*, q_3^*)] h_2 h_3 dq_2 dq_3,$$

sendo (q_1, q_2^*, q_3^*) um ponto sobre S_1 , contido dentro do caminho fechado ℓ_{14} . De acordo com o teorema de Stokes, o resultado anterior é dado por

$$\hat{\mathbf{e}}_1 \cdot [\nabla \times \mathbf{A}(q_1, q_2^*, q_3^*)] h_2 h_3 dq_2 dq_3 = \oint_{\ell_{14}} \mathbf{A} \cdot d\ell.$$

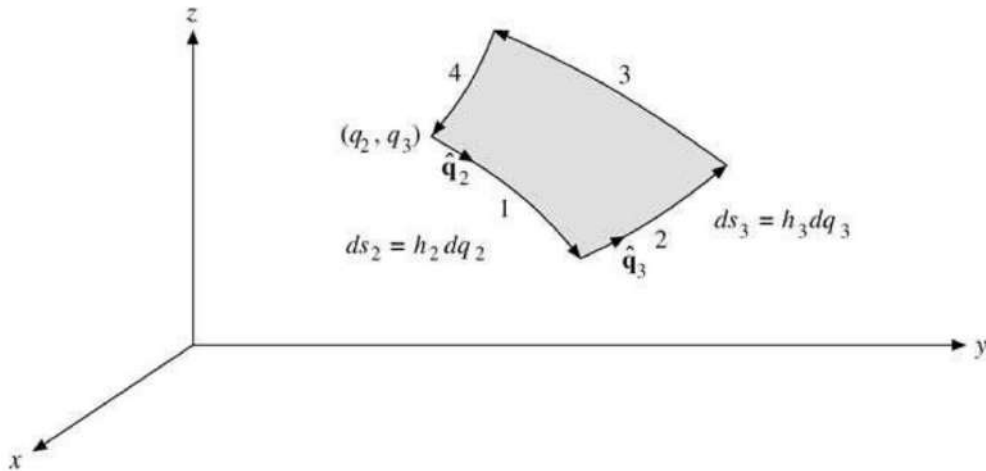


Figura 1.6: Elemento $d\sigma_{23}$, sobre uma superfície $S_1: q_1 = \text{cte.}$

Seguindo ao longo do contorno fechado ℓ_{14} , que delimita S_1 , a integral de caminho acima pode ser escrita, usando (1.17) e (1.18) e cuidando novamente os sentidos dos vetores unitários $\{\hat{\mathbf{e}}_i\}$, como

$$\begin{aligned} \oint_{\ell_{14}} \mathbf{A} \cdot d\ell \approx & A_2(q_1, q_2, q_3) h_2 dq_2 + A_3(q_1, q_2 + dq_2, q_3) h_3 dq_3 \\ & - A_2(q_1, q_2, q_3 + dq_3) h_2 dq_2 - A_3(q_1, q_2, q_3) h_3 dq_3, \end{aligned}$$

a qual pode ser ainda escrita como

$$\oint_{\ell_{14}} \mathbf{A} \cdot d\ell \approx A_2 h_2 dq_2 + \left[A_3 h_3 + \frac{\partial}{\partial q_2} (A_3 h_3) dq_2 \right] dq_3 - \left[A_2 h_2 + \frac{\partial}{\partial q_3} (A_2 h_2) dq_3 \right] dq_2 - A_3 h_3 dq_3,$$

ou seja,

$$\hat{\mathbf{e}}_1 \cdot [\nabla \times \mathbf{A}(q_1, q_2^*, q_3^*)] h_2 h_3 dq_2 dq_3 \approx \frac{\partial}{\partial q_2} (A_3 h_3) dq_2 dq_3 - \frac{\partial}{\partial q_3} (A_2 h_2) dq_2 dq_3.$$

No limite $dq_2 dq_3 \rightarrow 0$, obtém-se

$$\hat{\mathbf{e}}_1 \cdot [\nabla \times \mathbf{A}(q_1, q_2, q_3)] = \frac{1}{h_2 h_3} \left[\frac{\partial}{\partial q_2} (A_3 h_3) - \frac{\partial}{\partial q_3} (A_2 h_2) \right].$$

Procedendo da mesma maneira nos contornos das superfícies S_2 e S_3 , obtém-se as outras componentes de $\nabla \times \mathbf{A}$ pela permutação cíclica de índices do resultado acima. Desta maneira, pode-se escrever o rotacional do campo \mathbf{A} como o determinante

$$\nabla \times \mathbf{A} = \frac{1}{h_1 h_2 h_3} \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial/\partial q_1 & \partial/\partial q_2 & \partial/\partial q_3 \\ h_1 A_1 & h_2 A_2 & h_3 A_3 \end{vmatrix}. \quad (1.26)$$

Devido à presença dos operadores diferenciais, este determinante deve ser desenvolvido a partir da primeira linha.

1.5 SISTEMAS DE COORDENADAS CURVILÍNEAS ORTOGONAIS

Serão obtidas agora as expressões desenvolvidas nas seções 1.2 e 1.4 para alguns dos sistemas de coordenadas curvilíneas ortogonais mais empregados na literatura.

1.5.1 COORDENADAS POLARES CILÍNDRICAS

Este sistema de coordenadas é usualmente empregado em problemas com simetria cilíndrica. As coordenadas polares cilíndricas são identificadas pelo conjunto $\{q_1, q_2, q_3\} \rightarrow \{\rho, \phi, x_3\}$, sendo identificadas na figura 1.7a.

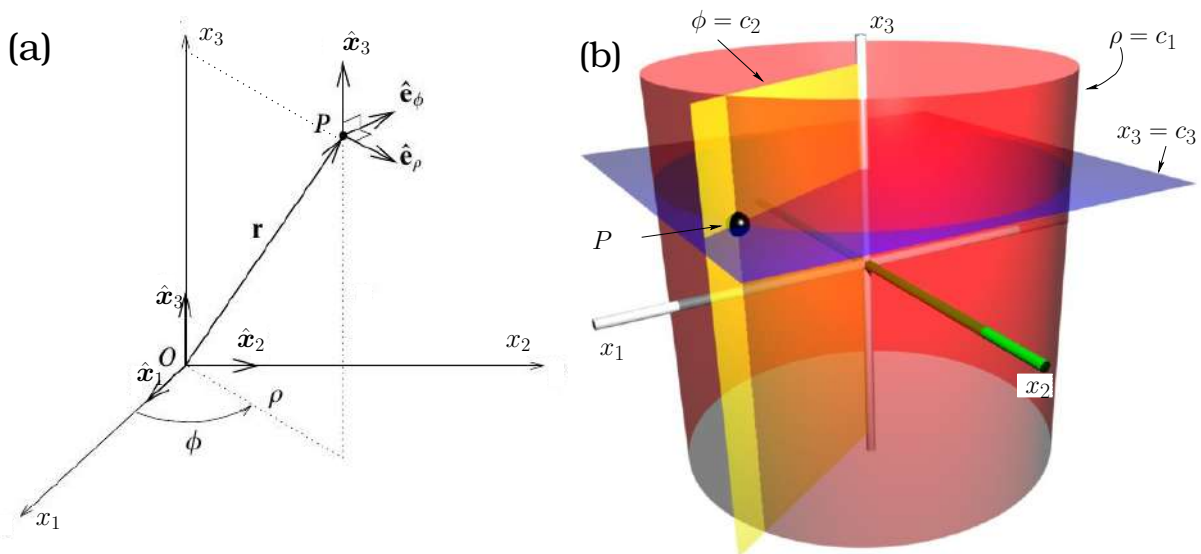


Figura 1.7: (a) Coordenadas cilíndricas e seus vetores unitários. (b) Superfícies coordenadas. O ponto P está localizado na intersecção das superfícies.

As leis de transformação para este sistema são as seguintes,

$$\begin{cases} \rho = \sqrt{x_1^2 + x_2^2} \\ \phi = \tan^{-1} \left(\frac{x_2}{x_1} \right) \\ x_3 = x_3 \end{cases} \implies \begin{cases} x_1 = \rho \cos \phi \\ x_2 = \rho \sin \phi \\ x_3 = x_3, \end{cases}$$

sendo que as coordenadas curvilíneas estão definidas nos seguintes intervalos: $0 \leq \rho < \infty$, $0 \leq \phi < 2\pi$ e $-\infty < x_3 < \infty$.

O vetor posição pode ser escrito em termos da base canônica; empregando as coordenadas cilíndricas, o vetor posição pode ser escrito como

$$\mathbf{r} = \rho \cos \phi \hat{\mathbf{x}}_1 + \rho \sin \phi \hat{\mathbf{x}}_2 + x_3 \hat{\mathbf{x}}_3.$$

Assim, os vetores da bases $\{e_i\}$, $\{\hat{e}_i\}$ e os fatores de escala $\{h_i\}$ para este sistema ficam escritos

$$\begin{cases} e_\rho = \frac{\partial \mathbf{r}}{\partial \rho} = \cos \phi \hat{x}_1 + \sin \phi \hat{x}_2 \\ e_\phi = \frac{\partial \mathbf{r}}{\partial \phi} = -\rho \sin \phi \hat{x}_1 + \rho \cos \phi \hat{x}_2 \\ e_3 = \hat{x}_3 \end{cases} \implies \begin{cases} \hat{e}_\rho = \cos \phi \hat{x}_1 + \sin \phi \hat{x}_2 \\ \hat{e}_\phi = -\sin \phi \hat{x}_1 + \cos \phi \hat{x}_2 \\ \hat{e}_3 = \hat{x}_3 \end{cases} \implies \begin{cases} h_\rho = 1 \\ h_\phi = \rho \\ h_3 = 1. \end{cases}$$

Invertendo o sistema de equações acima, resulta a transformação inversa

$$\begin{cases} \hat{x}_1 = \cos \phi \hat{e}_\rho - \sin \phi \hat{e}_\phi \\ \hat{x}_2 = \sin \phi \hat{e}_\rho + \cos \phi \hat{e}_\phi \\ \hat{x}_3 = \hat{e}_3. \end{cases}$$

Pode-se ver claramente que o sistema cilíndrico é ortonormal, $\hat{e}_\rho \cdot \hat{e}_\phi = \hat{e}_\rho \cdot \hat{x}_3 = \hat{e}_\phi \cdot \hat{x}_3 = 0$ e $\hat{e}_\rho \times \hat{e}_\phi = \hat{x}_3$, bem como qualquer permutação cíclica dos vetores de base. Assim, o vetor posição pode também ser escrito como

$$\mathbf{r} = \rho \hat{e}_\rho + x_3 \hat{x}_3.$$

O vetor velocidade em coordenadas cilíndricas é diretamente obtido de (1.13), resultando

$$\mathbf{v} \equiv \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt} = h_\rho \dot{\rho} \hat{e}_\rho + h_\phi \dot{\phi} \hat{e}_\phi + \dot{x}_3 h_3 \hat{e}_3,$$

ou seja,

$$\mathbf{v} = \dot{\rho} \hat{e}_\rho + \rho \dot{\phi} \hat{e}_\phi + \dot{x}_3 \hat{e}_3.$$

Este resultado pode ser verificado pela derivação direta de $\mathbf{r}(t)$.

Para se obter a aceleração (a derivada segunda do vetor posição), é útil primeiro obter-se as derivadas temporais dos vetores de base, ou seja,

$$\begin{cases} \dot{\hat{e}}_\rho \doteq \frac{d\hat{e}_\rho}{dt} = -\sin \phi \dot{\phi} \hat{x}_1 + \cos \phi \dot{\phi} \hat{x}_2 \\ \dot{\hat{e}}_\phi \doteq \frac{d\hat{e}_\phi}{dt} = -\cos \phi \dot{\phi} \hat{x}_1 - \sin \phi \dot{\phi} \hat{x}_2 \\ \dot{\hat{e}}_3 \doteq \frac{d\hat{e}_3}{dt} = 0. \end{cases}$$

Então,

$$\begin{aligned} \mathbf{a} \equiv \dot{\mathbf{v}} &= \frac{d\mathbf{v}}{dt} = \ddot{\rho} \hat{e}_\rho + (\dot{\rho} \dot{\phi} + \rho \ddot{\phi}) \hat{e}_\phi + \dot{\rho} \dot{\phi} \hat{e}_\phi + \rho \dot{\phi} \dot{\hat{e}}_\phi + \ddot{x}_3 \hat{e}_3 \\ &= \ddot{\rho} \hat{e}_\rho + (\dot{\rho} \dot{\phi} + \rho \ddot{\phi}) \hat{e}_\phi - (\sin \phi \dot{\rho} \dot{\phi} + \cos \phi \rho \dot{\phi}^2) \hat{x}_1 \\ &\quad + (\cos \phi \dot{\rho} \dot{\phi} - \sin \phi \rho \dot{\phi}^2) \hat{x}_2 + \ddot{x}_3 \hat{e}_3, \end{aligned}$$

resultando em

$$\mathbf{a} = (\ddot{\rho} - \rho \dot{\phi}^2) \hat{e}_\rho + (\rho \ddot{\phi} + 2\dot{\rho} \dot{\phi}) \hat{e}_\phi + \ddot{x}_3 \hat{e}_3.$$

O tensor de métrica é representado pela matriz

$$\mathbf{g} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Os vetores elementos de arco e superfície e o elemento de volume ficam escritos respectivamente,

$$\begin{aligned} d\ell &= d\rho \hat{e}_\rho + \rho d\phi \hat{e}_\phi + dx_3 \hat{x}_3 \\ d\sigma &= \rho d\phi dx_3 \hat{e}_\rho + d\rho dx_3 \hat{e}_\phi + \rho d\rho d\phi \hat{x}_3 \\ d^3r &= \rho d\rho d\phi dx_3. \end{aligned}$$

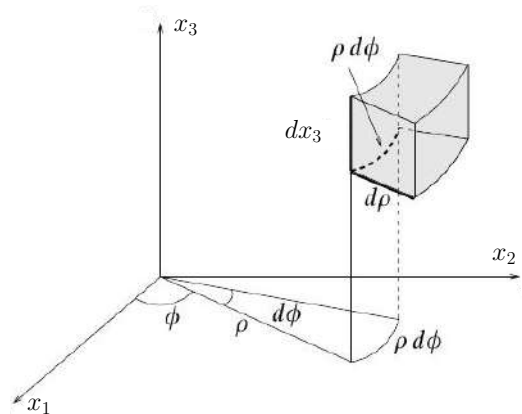


Figura 1.8: Elemento de volume em coordenadas cilíndricas.

Em particular, o elemento de volume pode ser visualizado na figura 1.8.

Finalmente, os operadores diferenciais ficam, para $\psi = \psi(\rho, \phi, x_3)$ e $\mathbf{A} = \mathbf{A}(\rho, \phi, x_3)$,

$$\begin{aligned}\nabla\psi &= \frac{\partial\psi}{\partial\rho}\hat{\mathbf{e}}_\rho + \frac{1}{\rho}\frac{\partial\psi}{\partial\phi}\hat{\mathbf{e}}_\phi + \frac{\partial\psi}{\partial x_3}\hat{\mathbf{x}}_3 \\ \nabla\cdot\mathbf{A} &= \frac{1}{\rho}\frac{\partial}{\partial\rho}(\rho A_\rho) + \frac{1}{\rho}\frac{\partial A_\phi}{\partial\phi} + \frac{\partial A_3}{\partial x_3} \\ \nabla\times\mathbf{A} &= \frac{1}{\rho}\begin{vmatrix} \hat{\mathbf{e}}_\rho & \rho\hat{\mathbf{e}}_\phi & \hat{\mathbf{x}}_3 \\ \partial/\partial\rho & \partial/\partial\phi & \partial/\partial x_3 \\ A_\rho & \rho A_\phi & A_3 \end{vmatrix} \\ \nabla^2\psi &= \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\psi}{\partial\rho}\right) + \frac{1}{\rho^2}\frac{\partial^2\psi}{\partial\phi^2} + \frac{\partial^2\psi}{\partial x_3^2}.\end{aligned}$$

1.5.2 COORDENADAS POLARES ESFÉRICAS

Este sistema de coordenadas é usualmente empregado em problemas com simetria esférica. As coordenadas polares esféricas são identificadas pelo conjunto $\{q_1, q_2, q_3\} \rightarrow \{r, \theta, \phi\}$, sendo ilustradas na figura 1.9a.

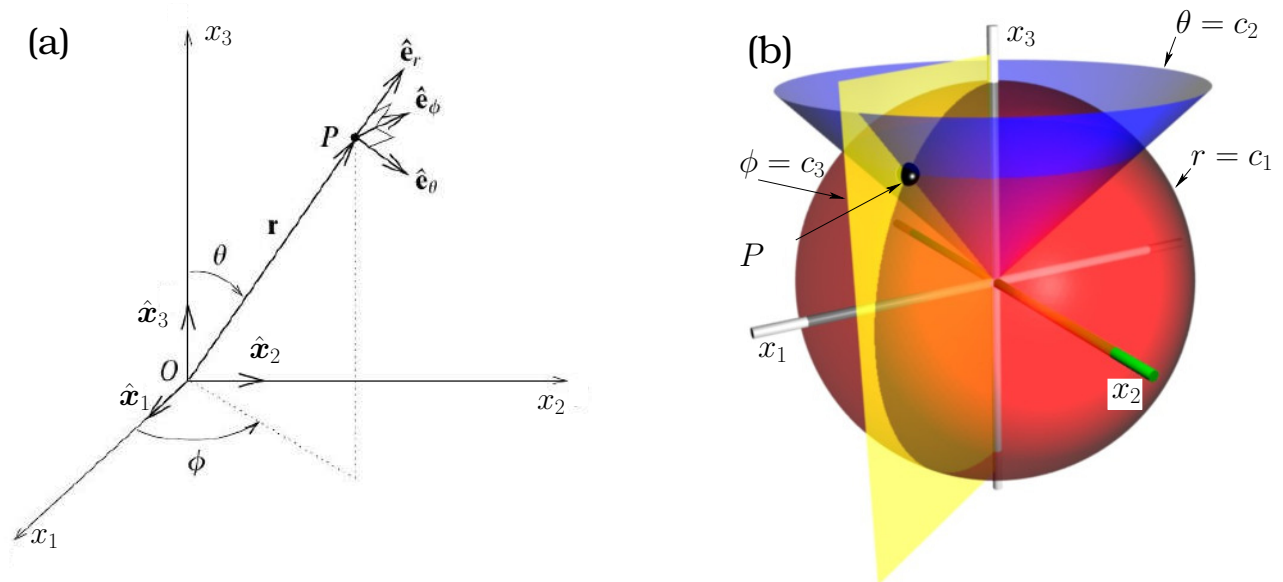


Figura 1.9: (a) Coordenadas esféricas e seus vetores unitários. (b) Superfícies coordenadas. O ponto P está localizado na intersecção das superfícies.

As leis de transformação para este sistema são as seguintes:

$$\begin{cases} r = \sqrt{x_1^2 + x_2^2 + x_3^2} \\ \theta = \cos^{-1}\left(\frac{x_3}{r}\right) \\ \phi = \tan^{-1}\left(\frac{x_2}{x_1}\right) \end{cases} \implies \begin{cases} x_1 = r \sin \theta \cos \phi \\ x_2 = r \sin \theta \sin \phi \\ x_3 = r \cos \theta, \end{cases}$$

sendo que as coordenadas estão definidas nos intervalos: $0 \leq r < \infty$, $0 \leq \theta \leq \pi$ e $0 \leq \phi < 2\pi$.

O vetor posição pode ser escrito agora como

$$\mathbf{r} = r \sin \theta \cos \phi \hat{\mathbf{x}}_1 + r \sin \theta \sin \phi \hat{\mathbf{x}}_2 + r \cos \theta \hat{\mathbf{x}}_3.$$

Assim, as bases $\{e_i\}$, $\{\hat{e}_i\}$ e os fatores de escala $\{h_i\}$ para o sistema de coordenadas esféricas são:

$$\begin{cases} \mathbf{e}_r = \frac{\partial \mathbf{r}}{\partial r} = \sin \theta \cos \phi \hat{\mathbf{x}}_1 + \sin \theta \sin \phi \hat{\mathbf{x}}_2 + \cos \theta \hat{\mathbf{x}}_3 \\ \mathbf{e}_\theta = \frac{\partial \mathbf{r}}{\partial \theta} = r \cos \theta \cos \phi \hat{\mathbf{x}}_1 + r \cos \theta \sin \phi \hat{\mathbf{x}}_2 - r \sin \theta \hat{\mathbf{x}}_3 \\ \mathbf{e}_\phi = \frac{\partial \mathbf{r}}{\partial \phi} = -r \sin \theta \sin \phi \hat{\mathbf{x}}_1 + r \sin \theta \cos \phi \hat{\mathbf{x}}_2 \end{cases}$$

$$\implies \begin{cases} \hat{\mathbf{e}}_r = \sin \theta \cos \phi \hat{\mathbf{x}}_1 + \sin \theta \sin \phi \hat{\mathbf{x}}_2 + \cos \theta \hat{\mathbf{x}}_3 \\ \hat{\mathbf{e}}_\theta = \cos \theta \cos \phi \hat{\mathbf{x}}_1 + \cos \theta \sin \phi \hat{\mathbf{x}}_2 - \sin \theta \hat{\mathbf{x}}_3 \\ \hat{\mathbf{e}}_\phi = -\sin \theta \sin \phi \hat{\mathbf{x}}_1 + \sin \theta \cos \phi \hat{\mathbf{x}}_2 \end{cases} \implies \begin{cases} h_r = 1 \\ h_\theta = r \\ h_\phi = r \sin \theta. \end{cases}$$

Percebe-se claramente que o sistema esférico é ortonormal, com $\hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_\theta = \hat{\mathbf{e}}_r \cdot \hat{\mathbf{e}}_\phi = \hat{\mathbf{e}}_\theta \cdot \hat{\mathbf{e}}_\phi = 0$ e $\hat{\mathbf{e}}_r \times \hat{\mathbf{e}}_\theta = \hat{\mathbf{e}}_\phi$, bem como qualquer permutação cíclica dos vetores de base.

Em termos da base ortonormal esférica, a base canônica pode agora ser escrita

$$\begin{cases} \hat{\mathbf{x}}_1 = \sin \theta \cos \phi \hat{\mathbf{e}}_r + \cos \theta \cos \phi \hat{\mathbf{e}}_\theta - \sin \phi \hat{\mathbf{e}}_\phi \\ \hat{\mathbf{x}}_2 = \sin \theta \sin \phi \hat{\mathbf{e}}_r + \cos \theta \sin \phi \hat{\mathbf{e}}_\theta + \cos \phi \hat{\mathbf{e}}_\phi \\ \hat{\mathbf{x}}_3 = \cos \theta \hat{\mathbf{e}}_r - \sin \theta \hat{\mathbf{e}}_\theta. \end{cases}$$

Assim, o vetor posição pode ser escrito simplesmente como

$$\mathbf{r} = r \hat{\mathbf{e}}_r.$$

O tensor de métrica é representado por

$$\mathbf{g} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}.$$

Os vetores elementos de arco e superfície e o elemento de volume são, respectivamente,

$$\begin{aligned} d\ell &= dr \hat{\mathbf{e}}_r + r d\theta \hat{\mathbf{e}}_\theta + r \sin \theta d\phi \hat{\mathbf{e}}_\phi \\ d\sigma &= r^2 \sin \theta d\theta d\phi \hat{\mathbf{e}}_r + r \sin \theta dr d\phi \hat{\mathbf{e}}_\theta + r dr d\theta \hat{\mathbf{e}}_\phi \\ d^3r &= r^2 \sin \theta dr d\theta d\phi. \end{aligned}$$

O elemento de volume para este sistema pode ser visualizado na figura 1.10.

O vetor velocidade é fornecido por (1.13), resultando em

$$\begin{aligned} \mathbf{v} \equiv \dot{\mathbf{r}} &= \frac{d\mathbf{r}}{dt} = h_r \dot{r} \hat{\mathbf{e}}_r + h_\theta \dot{\theta} \hat{\mathbf{e}}_\theta + h_\phi \dot{\phi} \hat{\mathbf{e}}_\phi \\ &= \dot{r} \hat{\mathbf{e}}_r + r \dot{\theta} \hat{\mathbf{e}}_\theta + r \sin \theta \dot{\phi} \hat{\mathbf{e}}_\phi. \end{aligned}$$

Este resultado pode ser verificado por derivação direta de $\mathbf{r}(t)$.

Para se obter a aceleração, primeiro calcula-se

$$\begin{cases} \dot{\hat{\mathbf{e}}}_r = (\cos \theta \cos \phi \dot{\theta} - \sin \theta \sin \phi \dot{\phi}) \hat{\mathbf{x}}_1 + (\cos \theta \sin \phi \dot{\theta} + \sin \theta \cos \phi \dot{\phi}) \hat{\mathbf{x}}_2 - \sin \theta \dot{\theta} \hat{\mathbf{x}}_3 \\ \quad = \dot{\theta} \hat{\mathbf{e}}_\theta + \sin \theta \dot{\phi} \hat{\mathbf{e}}_\phi \\ \dot{\hat{\mathbf{e}}}_\theta = -(\sin \theta \cos \phi \dot{\theta} + \cos \theta \sin \phi \dot{\phi}) \hat{\mathbf{x}}_1 + (-\sin \theta \sin \phi \dot{\theta} + \cos \theta \cos \phi \dot{\phi}) \hat{\mathbf{x}}_2 - \cos \theta \dot{\theta} \hat{\mathbf{x}}_3 \\ \quad = -\dot{\theta} \hat{\mathbf{e}}_r + \cos \theta \dot{\phi} \hat{\mathbf{e}}_\phi \\ \dot{\hat{\mathbf{e}}}_\phi = -\cos \phi \dot{\theta} \hat{\mathbf{x}}_1 - \sin \phi \dot{\theta} \hat{\mathbf{x}}_2 \\ \quad = -\sin \theta \dot{\phi} \hat{\mathbf{e}}_r - \cos \theta \dot{\phi} \hat{\mathbf{e}}_\theta. \end{cases}$$

Então,

$$\mathbf{a} \equiv \dot{\mathbf{v}} = \ddot{r} \hat{\mathbf{e}}_r + (\dot{r} \dot{\theta} + r \ddot{\theta}) \hat{\mathbf{e}}_\theta + (\dot{r} \sin \theta \dot{\phi} + r \cos \theta \dot{\theta} \dot{\phi} + r \sin \theta \ddot{\phi}) \hat{\mathbf{e}}_\phi + \ddot{r} \hat{\mathbf{e}}_r + r \ddot{\theta} \hat{\mathbf{e}}_\theta + r \sin \theta \ddot{\phi} \hat{\mathbf{e}}_\phi,$$

resultando

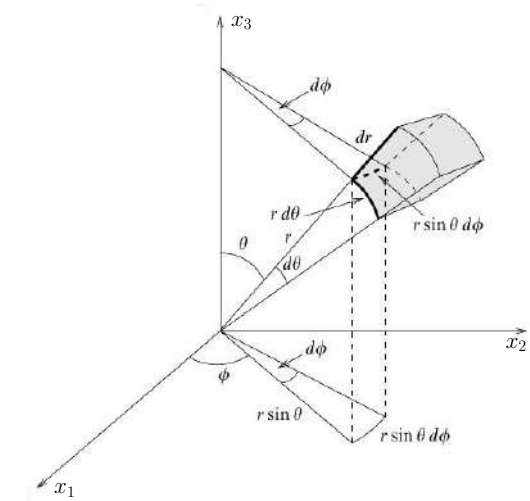


Figura 1.10: Elemento de volume em coordenadas esféricas.

$$\mathbf{a} = \left(\ddot{r} - r\dot{\theta}^2 - r \sin^2 \theta \dot{\phi}^2 \right) \hat{\mathbf{e}}_r + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta} - r \sin \theta \cos \theta \dot{\phi}^2 \right) \hat{\mathbf{e}}_\theta + \left(r \sin \theta \ddot{\phi} + 2 \sin \theta \dot{r}\dot{\phi} + 2r \cos \theta \dot{\theta}\dot{\phi} \right) \hat{\mathbf{e}}_\phi.$$

Por sua vez, os operadores diferenciais ficam, para $\psi = \psi(r, \theta, \phi)$ e $\mathbf{A} = \mathbf{A}(r, \theta, \phi)$,

$$\begin{aligned} \nabla \psi &= \frac{\partial \psi}{\partial r} \hat{\mathbf{e}}_r + \frac{1}{r} \frac{\partial \psi}{\partial \theta} \hat{\mathbf{e}}_\theta + \frac{1}{r \sin \theta} \frac{\partial \psi}{\partial \phi} \hat{\mathbf{e}}_\phi \\ \nabla \cdot \mathbf{A} &= \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 A_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta A_\theta) + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi} \\ \nabla \times \mathbf{A} &= \frac{1}{r^2 \sin \theta} \begin{vmatrix} \hat{\mathbf{e}}_r & r \hat{\mathbf{e}}_\theta & r \sin \theta \hat{\mathbf{e}}_\phi \\ \partial/\partial r & \partial/\partial \theta & \partial/\partial \phi \\ A_r & r A_\theta & r \sin \theta A_\phi \end{vmatrix} \\ \nabla^2 \psi &= \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2}. \end{aligned}$$

1.5.3 COORDENADAS ELÍPTICAS CILÍNDRICAS

Este sistema de coordenadas é empregado com frequência muito menor que os anteriores, mas é adequado para tratar sistemas que ocorrem sobre superfícies elípticas, ao invés de esféricas. Trata-se de um sistema de coordenadas ortogonais resultantes da projeção sobre o eixo x_3 de um sistema de coordenadas elíptico sobre o plano $x_1 - x_2$. Desta forma, as superfícies coordenadas são superfícies elípticas e hiperbólicas confocais, sendo que os focos estão usualmente posicionados nos pontos $x_1 = -a$ e $x_1 = a$ ($a > 0$).

As coordenadas elípticas são definidas por $\{q_1, q_2, q_3\} \rightarrow \{u, v, x_3\}$, sendo

$$\begin{cases} x_1 = a \cosh u \cos v \\ x_2 = a \sinh u \sin v \\ x_3 = x_3, \end{cases}$$

onde $0 \leq u < \infty$ e $0 \leq v < 2\pi$ estão relacionados com as coordenadas Cartesianas por

$$\frac{x_1^2}{a^2 \cosh^2 u} + \frac{x_2^2}{a^2 \sinh^2 u} = 1 \quad \text{e} \quad \frac{x_1^2}{a^2 \cos^2 v} - \frac{x_2^2}{a^2 \sin^2 v} = 1.$$

As coordenadas elípticas cilíndricas e algumas superfícies coordenadas podem ser visualizadas na figura 1.11.

Escrevendo o vetor posição como

$$\mathbf{r} = a \cosh u \cos v \hat{\mathbf{x}}_1 + a \sinh u \sin v \hat{\mathbf{x}}_2 + x_3 \hat{\mathbf{x}}_3,$$

resultam as bases e os fatores de escala

$$\begin{cases} \mathbf{e}_u = \frac{\partial \mathbf{r}}{\partial u} = a \sinh u \cos v \hat{\mathbf{x}}_1 + a \cosh u \sin v \hat{\mathbf{x}}_2 \\ \mathbf{e}_v = \frac{\partial \mathbf{r}}{\partial v} = -a \cosh u \sin v \hat{\mathbf{x}}_1 + a \sinh u \cos v \hat{\mathbf{x}}_2 \\ \mathbf{e}_3 = \frac{\partial \mathbf{r}}{\partial x_3} = \hat{\mathbf{x}}_3 \end{cases} \Rightarrow \begin{cases} \hat{\mathbf{e}}_u = (\sinh^2 u + \sin^2 v)^{-1/2} (\sinh u \cos v \hat{\mathbf{x}}_1 + \cosh u \sin v \hat{\mathbf{x}}_2) \\ \hat{\mathbf{e}}_v = (\sinh^2 u + \sin^2 v)^{-1/2} (-\cosh u \sin v \hat{\mathbf{x}}_1 + \sinh u \cos v \hat{\mathbf{x}}_2) \\ \hat{\mathbf{e}}_3 = \hat{\mathbf{x}}_3 \end{cases} \Rightarrow \begin{cases} h_u = h_v = a \sqrt{\sinh^2 u + \sin^2 v} \\ h_3 = 1. \end{cases}$$

Novamente observa-se que o sistema é ortogonal, $\hat{\mathbf{e}}_u \cdot \hat{\mathbf{e}}_v = \hat{\mathbf{e}}_u \cdot \hat{\mathbf{x}}_3 = \hat{\mathbf{e}}_v \cdot \hat{\mathbf{x}}_3 = 0$ e $\hat{\mathbf{e}}_u \times \hat{\mathbf{e}}_v = \hat{\mathbf{x}}_3$, bem como todas as permutações cíclicas.

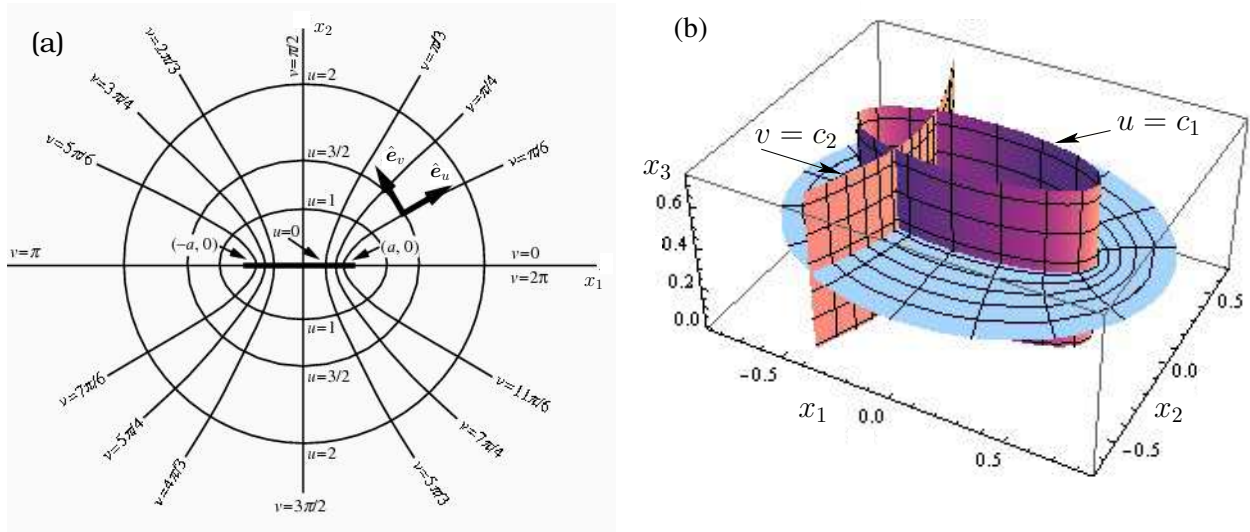


Figura 1.11: (a) Coordenadas elípticas cilíndricas e seus vetores unitários. (b) Algumas superfícies coordenadas.

Agora, os vetores da base canônica podem ser escritos em termos da nova base como

$$\begin{cases} \hat{x}_1 = (\sinh^2 u + \sen^2 v)^{-1/2} (\sinh u \cos v \hat{e}_u - \cosh u \sen v \hat{e}_v) \\ \hat{x}_2 = (\sinh^2 u + \sen^2 v)^{-1/2} (\cosh u \sen v \hat{e}_u + \sinh u \cos v \hat{e}_v) \\ \hat{x}_3 = \hat{e}_3. \end{cases}$$

Assim, o vetor posição pode ser escrito em termos da nova base como

$$\mathbf{r} = \frac{a \sinh u \cosh u}{\sqrt{\sinh^2 u + \sen^2 v}} \hat{e}_u - \frac{a \sen v \cos v}{\sqrt{\sinh^2 u + \sen^2 v}} \hat{e}_v + x_3 \hat{x}_3.$$

O tensor de métrica é dado por

$$\mathbf{g} = \begin{pmatrix} a^2 (\sinh^2 u + \sen^2 v) & 0 & 0 \\ 0 & a^2 (\sinh^2 u + \sen^2 v) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Já os elementos de arco, superfície e volume ficam

$$\begin{aligned} d\ell &= a \sqrt{\sinh^2 u + \sen^2 v} du \hat{e}_u + a \sqrt{\sinh^2 u + \sen^2 v} dv \hat{e}_v + dx_3 \hat{x}_3 \\ d\sigma &= a \sqrt{\sinh^2 u + \sen^2 v} dv dx_3 \hat{e}_u + a \sqrt{\sinh^2 u + \sen^2 v} du dx_3 \hat{e}_v + a^2 (\sinh^2 u + \sen^2 v) du dv \hat{x}_3 \\ d^3r &= a^2 (\sinh^2 u + \sen^2 v) du dv dx_3. \end{aligned}$$

Finalmente, os operadores diferenciais ficam

$$\begin{aligned} \nabla \psi &= \frac{1}{a \sqrt{\sinh^2 u + \sen^2 v}} \left(\frac{\partial \psi}{\partial u} \hat{e}_u + \frac{\partial \psi}{\partial v} \hat{e}_v \right) + \frac{\partial \psi}{\partial x_3} \hat{x}_3 \\ \nabla \cdot \mathbf{A} &= \frac{1}{a (\sinh^2 u + \sen^2 v)} \left[\frac{\partial}{\partial u} \left(\sqrt{\sinh^2 u + \sen^2 v} A_u \right) + \frac{\partial}{\partial v} \left(\sqrt{\sinh^2 u + \sen^2 v} A_v \right) \right] + \frac{\partial A_3}{\partial x_3} \\ \nabla \times \mathbf{A} &= \frac{1}{a^2 (\sinh^2 u + \sen^2 v)} \begin{vmatrix} a \sqrt{\sinh^2 u + \sen^2 v} \hat{e}_u & a \sqrt{\sinh^2 u + \sen^2 v} \hat{e}_v & \hat{x}_3 \\ \frac{\partial}{\partial u} & \frac{\partial}{\partial v} & \frac{\partial}{\partial x_3} \\ a \sqrt{\sinh^2 u + \sen^2 v} A_u & a \sqrt{\sinh^2 u + \sen^2 v} A_v & A_3 \end{vmatrix} \\ \nabla^2 \psi &= \frac{1}{a^2 (\sinh^2 u + \sen^2 v)} \left(\frac{\partial^2 \psi}{\partial u^2} + \frac{\partial^2 \psi}{\partial v^2} \right) + \frac{\partial^2 \psi}{\partial x_3^2}. \end{aligned}$$

1.6 LINHAS DE FORÇA E SUPERFÍCIES EQUIPOTENCIAIS

Em muitas situações, uma melhor compreensão do comportamento de certos campos vetoriais é possibilitada fazendo uso do conceito de *linhas de força* ou de *superfícies equipotenciais*.

1.6.1 LINHAS DE FORÇA DE UM CAMPO VETORIAL

Seja então o campo vetorial $A = A(\mathbf{r})$ no espaço vetorial \mathbb{E}^3 . Suas linhas de força são as curvas no \mathbb{E}^3 cujas tangentes são paralelas a $A(\mathbf{r})$ em todos os pontos. Uma representação das linhas de força para um campo vetorial arbitrário pode ser vista na figura 1.12. Para traçar estas linhas, faz-se necessário obter um conjunto de equações diferenciais cujas soluções corresponderão às mesmas. Sendo então $d\ell$ o elemento de arco na direção e no sentido de A em todos os pontos, conforme representado na figura 1.12, uma condição suficiente para a obtenção das equações diferenciais desejadas é

$$d\ell \times A = 0.$$

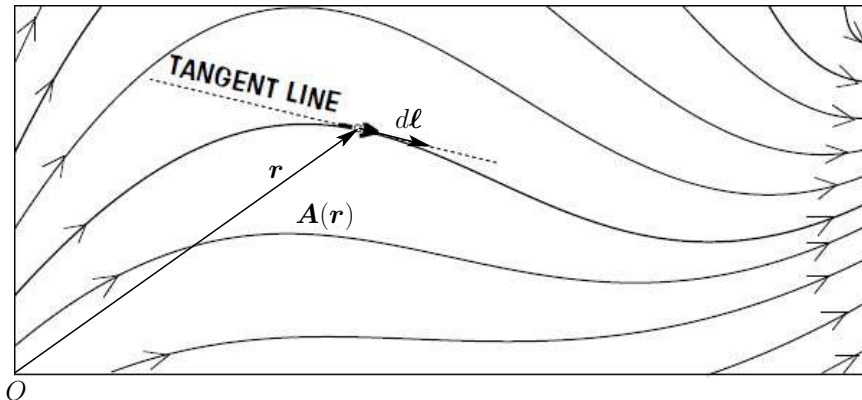


Figura 1.12: Linhas de força do campo $A(\mathbf{r})$. O elemento de arco $d\ell$ é tangencial ao campo no ponto \mathbf{r} .

Considerando agora o sistema de coordenadas curvilíneas ortogonais $\{q_1, q_2, q_3\}$, o elemento de arco $d\ell$ é dado por (1.12) e $A = \sum_{i=1}^3 A_i(q_1, q_2, q_3) \hat{e}_i$. Portanto,

$$d\ell \times A = \begin{vmatrix} \hat{e}_1 & \hat{e}_2 & \hat{e}_3 \\ h_1 dq_1 & h_2 dq_2 & h_3 dq_3 \\ A_1 & A_2 & A_3 \end{vmatrix} = 0,$$

de onde resulta o sistema de equações

$$\begin{aligned} h_1 A_2 dq_1 - h_2 A_1 dq_2 &= 0 \\ h_3 A_1 dq_3 - h_1 A_3 dq_1 &= 0 \\ h_2 A_3 dq_2 - h_3 A_2 dq_3 &= 0, \end{aligned} \tag{1.27a}$$

cuja solução irá fornecer a expressão para as linhas de força.

O mesmo sistema de equações pode ser escrito de uma forma reduzida como

$$h_1 \frac{dq_1}{A_1} = h_2 \frac{dq_2}{A_2} = h_3 \frac{dq_3}{A_3}, \tag{1.27b}$$

a qual facilita a parametrização das linhas de força do campo vetorial.

Exemplo 1.1. Linhas de força de uma carga puntiforme.

O campo elétrico de uma carga elétrica puntiforme situada na origem do sistema de coordenadas pode ser escrita em coordenadas esféricas como

$$\mathbf{E}(\mathbf{r}) = K \frac{q}{r^2} \hat{e}_r,$$

onde K é uma constante que depende do sistema de unidades empregado. As expressões para as coordenadas esféricas e seus fatores de escala são $q_1 = r$, $q_2 = \theta$, $q_3 = \varphi$, $h_1 = 1$, $h_2 = r$, $h_3 = r \sin \theta$. Então o sistema de equações (1.27) se reduz a

$$\begin{aligned} -K \frac{q}{r} d\theta &= 0 \\ \text{sen } \theta K \frac{q}{r} d\varphi &= 0, \end{aligned}$$

ou seja, $\theta = \text{cte}$ e $\varphi = \text{cte}$.

O exemplo a seguir mostra obtém as linhas de força de um dipolo elétrico.

Exemplo 1.2. Linhas de força de um dipolo elétrico.

O campo elétrico de um dipolo de momento \mathbf{p} centrado na origem do sistema de coordenadas é dado por

$$\mathbf{E}(\mathbf{r}) = \frac{K}{r^5} [3(\mathbf{p} \cdot \mathbf{r})\mathbf{r} - r^2\mathbf{p}].$$

Em coordenadas esféricas, este campo fica escrito

$$\mathbf{E}(\mathbf{r}) = \frac{Kp}{r^3} (2 \cos \theta \hat{\mathbf{e}}_r + \text{sen } \theta \hat{\mathbf{e}}_\theta).$$

Então, o sistema de equações (1.27) fica

$$\begin{aligned} \text{sen } \theta dr - 2r \cos \theta d\theta &= 0 \\ d\varphi &= 0. \end{aligned}$$

A segunda equação mostra que $\varphi = \text{cte}$. ao longo de uma linha de força. Já a primeira fica escrita

$$\frac{dr}{2r} = \cotan \theta d\theta \implies \ln r^{1/2} = \ln(\text{sen } \theta) + C \implies r = L \text{sen}^2 \theta,$$

sendo L uma constante que parametriza uma determinada linha de força.

O exemplo a seguir obtém as linhas de força de um quadrupolo elétrico.

Exemplo 1.3. Linhas de força de um quadrupolo elétrico.

Seja um quadrupolo elétrico gerado pela seguinte distribuição de 3 cargas elétricas posicionadas ao longo do eixo z : 1 carga $+q$ em $z = -b$, uma carga $-2q$ em $z = 0$ e uma carga $+q$ em $z = +b$. Para esta distribuição em particular, não existem termos nem de monopolo (carga total nula) nem de dipolo ($\mathbf{p}_{\text{total}} = \mathbf{p} - \mathbf{p} = 0$). Portanto, a primeira contribuição não nula em uma expansão de multipolos provém do termo de quadrupolo elétrico.

O potencial eletrostático desta distribuição, observado em um ponto $r \gg b$ é dado por

$$\Phi(\mathbf{r}) = \frac{Kqb^2}{r^3} (3 \cos^2 \theta - 1).$$

Então o campo elétrico fica, em coordenadas esféricas,

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= -\nabla\Phi(\mathbf{r}) = -\frac{\partial\Phi}{\partial r} \hat{\mathbf{e}}_r - \frac{1}{r} \frac{\partial\Phi}{\partial\theta} \hat{\mathbf{e}}_\theta, \\ \mathbf{E}(\mathbf{r}) &= 3 \frac{Kqb^2}{r^4} [(3 \cos^2 \theta - 1) \hat{\mathbf{e}}_r + \text{sen}(2\theta) \hat{\mathbf{e}}_\theta]. \end{aligned}$$

Assim, o sistema de equações (1.27) fica

$$\begin{aligned} \frac{1}{r} \text{sen}(2\theta) dr - (3 \cos^2 \theta - 1) d\theta &= 0 \\ d\varphi &= 0. \end{aligned}$$

A segunda equação novamente fornece $\varphi = \text{cte}$. Já a primeira pode ser escrita

$$\frac{1}{r} dr = \frac{3 \cos^2 \theta - 1}{\text{sen}(2\theta)} d\theta.$$

Integrando ambos os lados, obtém-se

$$\ln r = \frac{1}{2} \ln(|\cos \theta|) + \ln(\text{sen } \theta) + C,$$

a qual pode ser escrita como

$$r^2 = L \text{sen}^2 \theta |\cos \theta|.$$

REFERÊNCIAS

ARFKEN, G. B.; WEBER, H. J. **Mathematical Methods for Physicists**. Sixth. New York: Elsevier, 2005. 1182 + xii pp. ISBN 0-12-059876-0.

CHOW, T. L. **Mathematical Methods for Physicists: A Concise Introduction**. Cambridge: Cambridge, 2000. 555 pp. ISBN 0521655447.

FUNÇÕES DE UMA VARIÁVEL COMPLEXA



MUITAS VEZES, pergunta-se o porquê da necessidade de se estudar variáveis complexas quando na física estamos interessados apenas nas soluções reais. Poder-se-ia esperar que um estudo de funções reais de variáveis reais seria suficiente para se conhecer as soluções fisicamente relevantes. A resposta é que em muitas situações é desejável estender nosso estudo a valores complexos das variáveis e das soluções por razões de completicidade e conveniência. Por exemplo, o conjunto dos números reais não forma uma base suficiente para a representação das raízes de equações polinomiais ou algébricas. Além disso, o conhecimento do comportamento de uma função complexa $f(z)$, para todos os valores complexos de z , nos fornece uma visão mais completa de suas principais propriedades (mesmo suas propriedades para z real), do que o conhecimento de seu comportamento para somente valores reais de z . A localização, no plano complexo, dos zeros e dos infinitos de f (isto é, a posição das raízes de $f(z) = 0$ e de $1/f(z) = 0$) nos fornece informações sobre o comportamento de f para todos os valores de z . Adicionalmente, uma integral de $f(z)$ ao longo de valores reais de z pode ser modificada em uma integral ao longo de uma trajetória conveniente no plano complexo, de forma a simplificar consideravelmente o seu cálculo.

Integrais no plano complexo possuem uma ampla variedade de aplicações úteis na física e na matemática. Dentre estas, pode-se destacar:

- Cálculo de integrais definidas.
- Inversão de séries de potências.
- Cálculo de produtos infinitos.
- Obtenção de soluções de equações diferenciais para grandes valores da variável (soluções assintóticas).
- Investigação da estabilidade de sistemas potencialmente oscilatórios.
- Inversão de transformadas integrais.

Algumas destas propriedades serão tratadas ao longo deste capítulo.

Em se tratando de soluções de equações da física-matemática, uma solução complexa deve ser tratada como uma função ou número complexos até o momento em que se quer compará-la com um valor medido, físico. Neste momento, devemos associar a parte real e/ou imaginária ou outra quantidade real derivada do número complexo (tal como o módulo) com parâmetros físicos reais. Assim, mencionando somente dois exemplos, o índice de refração real de uma onda eletromagnética propagando-se em um meio ativo torna-se uma quantidade complexa quando a absorção da energia transportada pela onda é incluída. A energia real associada com um nível de energia atômico ou nuclear torna-se complexa quando o tempo de vida finito do nível de energia é considerado.

Mas a mais importante razão para se estudar funções complexas é a compreensão que se pode obter a respeito das propriedades gerais das funções. Por exemplo, as singularidades da função podem estar relacionadas com singularidades físicas, tais como as causadas por fontes, cargas elétricas pontuais, *etc.* É possível, a partir do conhecimento das singularidades de uma função complexa, especificar-se a função completamente.

Estes serão alguns dos tópicos abordados neste capítulo.

2.1 NÚMEROS E VARIÁVEIS COMPLEXOS

O sistema numérico em uso atualmente é o resultado de um desenvolvimento gradual na matemática que se iniciou na Idade Antiga. Os números naturais (inteiros positivos) $\{0, 1, 2, \dots\}$ foram utilizados inicialmente para a contagem. O conjunto dos números naturais é representado pelo símbolo \mathbb{N} e diz-se que um dado número natural n pertence a \mathbb{N} ($n \in \mathbb{N}$).

Os inteiros negativos e o conceito do zero foram então introduzidos para permitir soluções de equações tais como $x+3 = 2$. Cria-se então o conjunto dos números inteiros $\{\dots, -2, -1, 0, 1, 2, \dots\}$, representado pelo símbolo \mathbb{Z} . Observa-se aqui que o conjunto \mathbb{N} é um sub-conjunto de \mathbb{Z} . Diz-se então que \mathbb{N} está contido em \mathbb{Z} ($\mathbb{N} \subset \mathbb{Z}$), ou que \mathbb{Z} contém \mathbb{N} ($\mathbb{Z} \supset \mathbb{N}$).

Para permitir a solução de equações tais como $bx = a$, para todos os inteiros a e b (com $b \neq 0$), os números racionais ($x = a/b$) foram introduzidos. Representa-se o conjunto de todos os números racionais por $\mathbb{Q} = \{x \mid x = p/q, \text{ com } (p, q) \in \mathbb{Z} \text{ e } q \neq 0\}$. Nota-se aqui que \mathbb{Q} contém \mathbb{Z} , consistindo em aqueles $x \in \mathbb{Q} \mid q = 1$.

Posteriormente, os números irracionais foram introduzidos quando descobriu-se que números tais como as soluções da equação

$$x^2 - 2 = 0 \implies x = \pm\sqrt{2} = \pm 1.41423\dots$$

ou a razão entre o perímetro de uma circunferência de raio unitário e o seu diâmetro, igual a $\pi = 3.14159265359\dots$, não podem ser expressos por números racionais. O conjunto dos números irracionais é representado pelo símbolo \mathbb{Q}' . Nota-se aqui que \mathbb{Q} não contém nem está contido em \mathbb{Q}' , sendo ambos conjuntos de números completamente distintos.

A reunião, ou a *união*, dos números racionais com os irracionais formam o conjunto dos números reais, representado pelo símbolo \mathbb{R} ($\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}'$). Disciplinas usuais de cálculo apresentam seus teoremas e resultados considerando somente números pertencentes ao conjunto \mathbb{R} . Contudo, este conjunto ainda está incompleto para aplicações em álgebra e para a análise matemática.

Os números complexos foram descobertos na Idade Média, ao se pesquisar as raízes de certas equações quadráticas, tais como

$$z^2 + 1 = 0 \implies z = \pm\sqrt{-1}.$$

É óbvio, pelo nome dado, que eles foram considerados de maneira suspeita. Leonhard Paul Euler (1707-1783), em 1777, introduziu o símbolo

$$i = \sqrt{-1}.$$

Carl Friedrich Gauss (1777-1855), na sua tese de doutorado em 1799, forneceu aos números complexos a agora familiar expressão algébrica $z = x + iy$, bem como a sua representação geométrica (vetorial) e, com isso, ajudou a desvendar parte de seu mistério. Neste século, a tendência tem sido definir os números complexos como símbolos abstratos sujeitos a certas regras formais de manipulação.

Como o número $\sqrt{-1}$ não possui representação possível dentro do conjunto de números reais, chamou-se este número de imaginário puro e atribuiu-se a ele símbolo $i = \sqrt{-1}$. Além disso, definiu-se um conjunto mais amplo de números, denominado conjunto dos números complexos $\mathbb{C} \supset \mathbb{R}$, o qual contém todos os números complexos, tendo o conjunto dos números reais como um sub-conjunto.

Um número complexo nada mais é que um par ordenado de dois números reais x e y . Assim, o número complexo z pode ser representado de, pelo menos, duas maneiras:

$$z = (x, y) = x + iy,$$

sendo a última representação a preferida neste texto. Deve-se notar que o ordenamento é significativo; assim, $a + ib \neq b + ia$.

Uma propriedade imediata do número i pode ser deduzida observando-se que $i^2 = i \cdot i = -1$, $i^3 = i^2 \cdot i = -i$, $i^4 = i^2 \cdot i^2 = 1$, $i^5 = i \cdot i^4 = i$, \dots . Da mesma forma,

$$\begin{aligned} i^{-1} &= \frac{1}{i} = \frac{i}{i \cdot i} = -i \\ i^{-2} &= \frac{1}{i^2} = -1 \end{aligned}$$

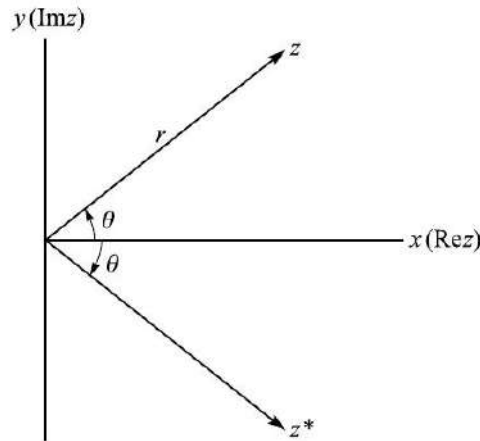


Figura 2.1: Plano complexo ou diagrama de Argand.

$$i^{-3} = -\frac{1}{i} = i$$

$$i^{-4} = 1$$

... ..,

resultando

$$i^{\pm 2n} = (-1)^n \tag{2.1a}$$

$$i^{\pm(2n+1)} = \pm(-1)^n i, \text{ para } n = 0, 1, 2, 3, \dots \tag{2.1b}$$

2.1.1 REPRESENTAÇÕES VETORIAL E POLAR

Em muitas situações, é conveniente empregar-se uma representação gráfica da variável complexa. Traçando-se x — a parte real de z — como a abcissa e y — a parte imaginária de z — como a ordenada, obtém-se o plano complexo, ou *plano de Argand*, ilustrado na figura 2.1.

Em álgebra linear, frequentemente utiliza-se o conceito de um vetor posição r como membro do espaço vetorial \mathcal{R}^2 . Assim, pode-se representar o vetor r fazendo-se uso da base canônica $\{\hat{i}, \hat{j}\}$:

$$r = x\hat{i} + y\hat{j},$$

sendo x a componente de r na direção definida por \hat{i} e y a componente de r ao longo de \hat{j} .

Da mesma forma, pode-se interpretar o número complexo $z = x + iy$, de acordo com a representação gráfica da figura 2.1, como um vetor (ou *fator*), sendo x a componente ao longo do eixo real e y a componente ao longo do eixo imaginário. Assim, tal quantidade satisfaz as regras usuais de adição de vetores e de multiplicação por escalar. Inspeccionando a figura 2.1, relações simples de trigonometria mostram que

$$x = r \cos \theta$$

$$y = r \sen \theta,$$

sendo $r \in \mathbb{R}$ ($0 \leq r < \infty$) denominado *módulo* ou *magnitude* de z e $\theta \in \mathbb{R}$ ($0 \leq \theta < 2\pi$)¹ chamado de *argumento* ou *fase* de z . Portanto,

$$z = r (\cos \theta + i \sen \theta), \tag{2.2a}$$

onde

$$r = \sqrt{x^2 + y^2} \tag{2.2b}$$

$$\theta = \tan^{-1} (y/x). \tag{2.2c}$$

¹Ver uma definição mais geral para o intervalo de variação de θ a seguir.

2.1.2 FÓRMULA DE EULER

Uma representação equivalente à representação algébrica de z dada por (2.2) é a chamada *representação polar*:

$$z = re^{i\theta}. \quad (2.3)$$

Demonstração. A partir das seguintes séries de McLaurin:

$$\begin{aligned} \operatorname{sen} x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ \operatorname{cos} x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \\ e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \end{aligned}$$

e das potências (2.1a,b), obtemos

$$\begin{aligned} e^{i\theta} &= \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!} = \sum_{n=0}^{\infty} \frac{(i\theta)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(i\theta)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n+1}}{(2n+1)!}, \end{aligned}$$

ou seja,

$$e^{i\theta} = \cos \theta + i \operatorname{sen} \theta.$$

Esta é a conhecida *Fórmula de Euler*. □

2.2 ÁLGEBRA DE NÚMEROS COMPLEXOS

Sendo $z = x + iy \in \mathbb{C}$ um número complexo qualquer, as seguintes operações e definições se aplicam:

Parte real de z : a parte real de z é o número $x \in \mathbb{R}$. Esta operação é representada por

$$\operatorname{Re} z = x.$$

Parte imaginária de z : a parte imaginária de z é o número $y \in \mathbb{R}$. Esta operação é representada por

$$\operatorname{Im} z = y.$$

Complexo conjugado de z : o complexo conjugado de z , representado por \bar{z} ou z^* , tal que $z^* \in \mathbb{C}$, é definido por $z^* = x - iy$. Na figura 2.1, pode-se observar a representação vetorial de z^* .

Módulo de z : é o número $|z| \in \mathbb{R}$ tal que

$$|z| = |x + iy| = \sqrt{x^2 + y^2} = \sqrt{z \cdot z^*}.$$

Fase ou argumento de z : número $\theta \in \mathbb{R}$ tal que $\theta_0 \leq \theta < \theta_0 + 2\pi$, dado por

$$\theta \equiv \arg(z) = \tan^{-1} \left(\frac{y}{x} \right).$$

Usualmente, toma-se $\theta_0 = 0$, mas outros textos podem usar, por exemplo, $-\pi \leq \theta < \pi$.

As seguintes operações algébricas estão definidas para dois números $z_1 = a + ib = r_1 e^{i\theta_1}$ e $z_2 = c + id = r_2 e^{i\theta_2}$ quaisquer, tais que $\{z_1, z_2\} \in \mathbb{C}$. Os números $\{r_1, r_2\} \in \mathbb{R}$ são, respectivamente, os módulos de z_1 e z_2 e $\{\theta_1, \theta_2\} \in \mathbb{R}$ são os respectivos argumentos.

Identidade: Se $z_1 = z_2$, então $\operatorname{Re} z_1 = \operatorname{Re} z_2$ e $\operatorname{Im} z_1 = \operatorname{Im} z_2$; ou, de forma equivalente, $r_1 = r_2$ e $\theta_1 = \theta_2 + 2k\pi$.

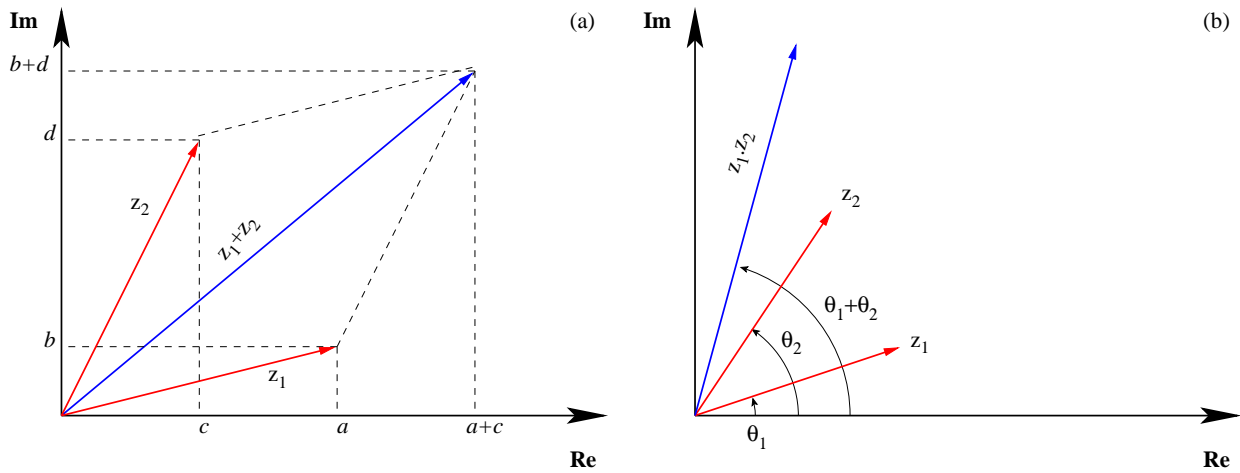


Figura 2.2: (a) Representação da operação $z_1 + z_2$. (b) Representação da operação $z_1.z_2$.

Adição: $z_1 + z_2 = (a + ib) + (c + id) = (a + c) + (b + d)i$. Esta operação está representada na figura 2.2(a).

Subtração: $z_1 - z_2 = (a + ib) - (c + id) = (a - c) + (b - d)i$.

Conjugação complexa da adição: $(z_1 + z_2)^* = z_1^* + z_2^*$.

Multiplicação por real: Dado um $h \in \mathbb{R}$,

$$h.z_1 = h(a + ib) = ha + ihb.$$

Multiplicação de complexos:

$$z_1.z_2 = (a + ib)(c + id) = (ac - bd) + (ad + bc)i,$$

ou, em termos da forma polar,

$$z_1.z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)} = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \text{sen}(\theta_1 + \theta_2)]. \quad (2.4)$$

Esta operação está representada na figura 2.2(b).

Divisão de complexos:

$$\frac{z_1}{z_2} = z_1 \frac{z_2^*}{z_2.z_2^*} = \frac{z_1.z_2^*}{|z_2|^2}, \text{ ou}$$

$$\frac{z_1}{z_2} = \frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{ac + bd}{c^2 + d^2} - i \frac{ad - bc}{c^2 + d^2}.$$

Ou, em termos da forma polar,

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)} = \frac{r_1}{r_2} [\cos(\theta_1 - \theta_2) + i \text{sen}(\theta_1 - \theta_2)].$$

Conjugação complexa do produto: $(z_1.z_2)^* = z_1^*.z_2^*$.

Outras operações algébricas, como potenciação e radiciação, serão vistas nas seções seguintes.

O valor absoluto de z ainda possui as seguintes propriedades. Sendo $\{z_1, z_2, \dots, z_n\}$ números complexos, então

1. $|z_1 z_2 \dots z_n| = |z_1| |z_2| \dots |z_n|$.
2. $\left| \frac{z_1}{z_2} \right| = \frac{|z_1|}{|z_2|}$, desde que $z_2 \neq 0$.
3. $|z_1 + z_2 + \dots + z_n| \leq |z_1| + |z_2| + \dots + |z_n|$.

$$4. |z_1 \pm z_2| \geq ||z_1| - |z_2||.$$

Números complexos $z = re^{i\theta}$ com $r = 1$ possuem $|z| = 1$ e são denominados *unimodulares*. Pode-se imaginar estes números situados ao longo da circunferência de raio unitário sobre o plano complexo. Pontos especiais nesta circunferência são:

$$\theta = 0. \quad z = e^{i0} = 1.$$

$$\theta = \pi/2. \quad z = e^{i\pi/2} = i.$$

$$\theta = \pi. \quad z = e^{-i\pi} = -1.$$

$$\theta = 3\pi/2 \text{ ou } \theta = -\pi/2. \quad z = e^{i3\pi/2} = e^{-i\pi/2} = -i.$$

2.2.1 FÓRMULA DE MOIVRE

Sejam $z_1 = r_1(\cos \theta_1 + i \operatorname{sen} \theta_1) = r_1 e^{i\theta_1}$ e $z_2 = r_2(\cos \theta_2 + i \operatorname{sen} \theta_2) = r_2 e^{i\theta_2}$ dois números complexos. Então o produto de ambos já foi calculado em (2.4):

$$z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)} = r_1 r_2 [\cos(\theta_1 + \theta_2) + i \operatorname{sen}(\theta_1 + \theta_2)].$$

Generalizando agora este resultado para n números complexos, $\{z_1, z_2, \dots, z_n\}$, obtém-se

$$z_1 z_2 \dots z_n = r_1 r_2 \dots r_n e^{i(\theta_1 + \theta_2 + \dots + \theta_n)} = r_1 r_2 \dots r_n [\cos(\theta_1 + \theta_2 + \dots + \theta_n) + i \operatorname{sen}(\theta_1 + \theta_2 + \dots + \theta_n)].$$

Agora, se $z_1 = z_2 = \dots = z_n = z$, onde se escreve $z = r(\cos \theta + i \operatorname{sen} \theta)$. Resulta então,

$$z^n = r^n (\cos \theta + i \operatorname{sen} \theta)^n = r^n [\cos(n\theta) + i \operatorname{sen}(n\theta)], \quad (2.5)$$

a qual é a fórmula para a n -ésima potência ($n > 0$) de z . Cancelando os termos r^n em ambos os lados de (2.5), resulta a Fórmula de Moivre:

$$\boxed{(\cos \theta + i \operatorname{sen} \theta)^n = \cos(n\theta) + i \operatorname{sen}(n\theta)}. \quad (2.6)$$

2.2.2 RAÍZES DE NÚMEROS COMPLEXOS

A Fórmula de Moivre (2.6) permite que se obtenha uma expressão para as raízes de um número complexo. Para tanto, considera-se a seguinte equação polinomial:

$$z^n - w = 0, \quad (2.7)$$

onde $n \in \mathbb{N}$ e $w \in \mathbb{C}$. Buscar a solução de (2.7) é equivalente a buscar a raiz n de w . Esta solução poderia ser escrita simplesmente como $z = \sqrt[n]{w}$, mas esta forma dá a entender que existe somente uma solução de (2.7), o que não é verdade. O número de soluções (ou raízes) de um polinômio de grau n , como a função no lado esquerdo de (2.7), é bem determinado, de acordo com o Teorema Fundamental da Álgebra abaixo.

Teorema 2.1 (Teorema Fundamental da Álgebra). *Toda equação polinomial de grau n , cujos coeficientes são reais ou complexos, possui exatamente n raízes reais ou complexas.*

Portanto, deve-se buscar n soluções para (2.7), o que é equivalente a procurar as n raízes de w , as quais serão denominadas z_0, z_1, \dots, z_{n-1} . Concentrando-nos inicialmente em z_0 , tal que

$$z_0^n = w,$$

pode-se usar para ambos as suas formas polares dadas por (2.2a),

$$z_0 = |z_0|(\cos \theta + i \operatorname{sen} \theta) \text{ e } w = |w|(\cos \alpha + i \operatorname{sen} \alpha).$$

Então, de acordo com (2.6),

$$|z_0|^n (\cos \theta + i \operatorname{sen} \theta)^n = |z_0|^n [\cos(n\theta) + i \operatorname{sen}(n\theta)] = |w|(\cos \alpha + i \operatorname{sen} \alpha), \quad (2.8)$$

possibilitando-nos a identificar

$$|z_0|^n = |w| \implies |z_0| = \sqrt[n]{|w|},$$

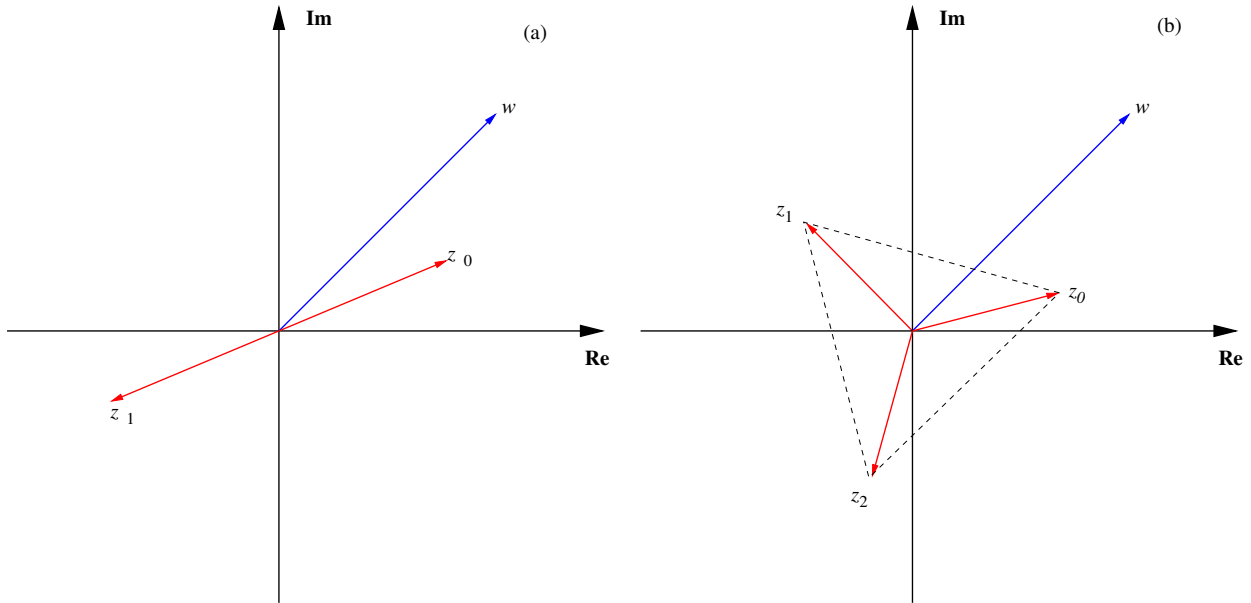


Figura 2.3: (a) Raízes quadradas z_0 e z_1 de $w = 1 + i$. (b) Raízes cúbicas w_0, w_1 e w_2 de $z = 1 + i$.

$$n\theta = \alpha \implies \theta = \frac{\alpha}{n}.$$

Portanto, a raiz principal de (2.7) é dada por

$$z_0 = \sqrt[n]{|w|} \left[\cos\left(\frac{\alpha}{n}\right) + i \operatorname{sen}\left(\frac{\alpha}{n}\right) \right] = \sqrt[n]{|w|} e^{i\alpha/n}. \quad (2.9a)$$

Contudo, como já foi mencionado, existem outras $n - 1$ raízes distintas de w . Estas outras raízes podem ser determinadas levando-se em conta as identidades

$$\cos(\beta \pm 2k\pi) = \cos \beta \text{ e } \operatorname{sen}(\beta \pm 2k\pi) = \operatorname{sen} \beta, \text{ para } k = 0, 1, 2, 3, \dots$$

Assim, retornando-se a (2.8), pode-se escrever a relação entre as fases como

$$n\theta - 2k\pi = \alpha \implies \theta = \frac{\alpha + 2k\pi}{n}.$$

Constata-se facilmente que se α for substituído por $\alpha + 2k\pi$ em (2.9a), haverá sempre um número total de n arcos tais que

$$0 \leq \frac{\alpha + 2k\pi}{n} \leq 2\pi, \text{ para } k = 0, 1, \dots, n - 1,$$

os quais são geometricamente distintos sobre o plano complexo. Se fossem considerados os valores $k = n, n + 1, \dots$, isto simplesmente repetiria os arcos anteriormente encontrados.

Portanto, as n raízes de (2.7), incluindo z_0 , são:

$$z_k = \sqrt[n]{|w|} \left[\cos\left(\frac{\alpha + 2k\pi}{n}\right) + i \operatorname{sen}\left(\frac{\alpha + 2k\pi}{n}\right) \right] = \sqrt[n]{|w|} e^{i(\alpha + 2k\pi)/n}, \quad (k = 0, 1, \dots, n - 1). \quad (2.9b)$$

Exemplo 2.1 (Raízes quadradas). Dado o número $w = 1 + i$, encontre as suas raízes quadradas.

Solução: há exatamente 2 raízes quadradas para w . Inicialmente, escreve-se w na forma polar:

$$w = \sqrt{2} \left(\cos \frac{\pi}{4} + i \operatorname{sen} \frac{\pi}{4} \right) \implies |w| = \sqrt{2} \text{ e } \alpha = \frac{\pi}{4},$$

sendo que $\pi/4 \rightsquigarrow 45^\circ$. De acordo com (2.9b), $n = 2, k = 0, 1$, e as raízes são:

$$z_0 = \sqrt[4]{2} \left(\cos \frac{\pi}{8} + i \operatorname{sen} \frac{\pi}{8} \right),$$

$$z_1 = \sqrt[4]{2} \left(\cos \frac{9\pi}{8} + i \operatorname{sen} \frac{9\pi}{8} \right),$$

sendo que $\pi/8 \rightsquigarrow 22,5^\circ$ e $9\pi/8 \rightsquigarrow 202,5^\circ$, de tal forma que as raízes z_0 e z_1 são antiparalelas no plano complexo. Estas raízes encontram-se representadas no diagrama da figura 2.3(a).

Exemplo 2.2 (Raízes cúbicas). Dado o número $w = 1 + i$, encontre as suas raízes cúbicas.

Solução: há exatamente 3 raízes cúbicas para w . Dado w na forma polar:

$$z = \sqrt{2} \left(\cos \frac{\pi}{4} + i \operatorname{sen} \frac{\pi}{4} \right) \implies |z| = \sqrt{2} \text{ e } \theta = \frac{\pi}{4},$$

sendo que $\pi/4 \rightsquigarrow 45^\circ$. Agora, de acordo com (2.9b), $n = 3$, $k = 0, 1, 2$, e as raízes são:

$$\begin{aligned} z_0 &= \sqrt[3]{2} \left[\cos \left(\frac{\pi}{12} \right) + i \operatorname{sen} \left(\frac{\pi}{12} \right) \right], \\ z_1 &= \sqrt[3]{2} \left[\cos \left(\frac{3\pi}{4} \right) + i \operatorname{sen} \left(\frac{3\pi}{4} \right) \right], \\ z_2 &= \sqrt[3]{2} \left[\cos \left(\frac{17\pi}{12} \right) + i \operatorname{sen} \left(\frac{17\pi}{12} \right) \right], \end{aligned}$$

sendo que $\pi/12 \rightsquigarrow 15^\circ$, $3\pi/4 \rightsquigarrow 135^\circ$ e $17\pi/12 \rightsquigarrow 255^\circ$, de tal forma que z_0 , z_1 e z_2 estão nos vértices de um triângulo equilátero. Estas raízes encontram-se representadas no diagrama da figura 2.3(b).

2.3 FUNÇÕES DE UMA VARIÁVEL COMPLEXA

Seja $\mathbb{D} \subseteq \mathbb{C}$ um conjunto de números complexos $z = x + iy$. Uma *função* $f(z)$, definida em \mathbb{D} é uma operação que atribui a cada $z \in \mathbb{D}$ um outro número complexo $w \in \mathbb{I}$, onde $\mathbb{I} \subseteq \mathbb{C}$. O número w é denominado o *valor* de $f(z)$ em z , isto é,

$$w = f(z).$$

O conjunto \mathbb{D} é denominado o *domínio de definição de $f(z)$* e o conjunto \mathbb{I} é denominado a *imagem de $f(z)$* .

Deve ser enfatizado que tanto o domínio de definição quanto a operação são necessários para que a função seja bem definida. Quando o domínio não é especificado, deve-se supor que o maior conjunto possível é tomado. Assim, se é mencionada simplesmente a função $f(z) = 1/z$, o domínio é subentendido como o conjunto de todos os pontos não nulos no plano complexo.

Existem dois tipos básicos de funções complexas:

Funções unívocas. Uma função é denominada *unívoca* em \mathbb{D} se a cada valor de z corresponde um único valor de w .

Funções plurívocas. Uma função é denominada *plurívoca* em \mathbb{D} se a um determinado valor de z corresponder mais de um valor de w . Uma função plurívoca pode ser considerada como uma coleção de funções unívocas, onde cada membro desta coleção é chamado de *ramo da função plurívoca*. É usual tomar-se um membro em particular da coleção como o *ramo principal da função plurívoca* e o valor da função correspondente a este ramo é denominado *valor principal*.

Como exemplos de funções unívocas ou plurívocas, pode-se tomar:

1. $w = z^2$ – função unívoca ou simplesmente função.
2. $w = \sqrt{z}$ – função plurívoca, pois a cada valor de z correspondem dois valores de w , de acordo com (2.9b). Assim:

$$\text{se } z = re^{i\theta}, \text{ então } \sqrt{z} = w_k = \sqrt{r}e^{i(\theta+2k\pi)/2}, \text{ onde } k = 0, 1,$$

$$\text{Para } k = 0: w_0 = \sqrt{r}e^{i\theta/2} \longrightarrow \text{ramo principal.}$$

$$\text{Para } k = 1; w_1 = \sqrt{r}e^{i\theta/2}e^{i\pi} = -\sqrt{r}e^{i\theta/2} \longrightarrow \text{segundo ramo.}$$

2.3.1 TRANSFORMAÇÕES OU MAPEAMENTOS

Nota-se agora que w , como o valor de $f(z)$, também é uma variável complexa e que, portanto, pode ser escrita na forma

$$w = u + iv,$$

sendo $\{u, v\} \subset \mathbb{R}$, ao passo que $f(z)$ também pode ser separada em partes real e imaginária,

$$f(z) = u(x, y) + iv(x, y),$$

onde $u(x, y)$ e $v(x, y)$ são ambas funções reais. Igualando as partes real e imaginária das expressões acima, obtém-se

$$\operatorname{Re} w = u = u(x, y), \quad \operatorname{Im} w = v = v(x, y).$$

Se $w = f(z)$ é uma função unívoca de z , então pode-se imaginar o *plano complexo de z* e, a cada ponto neste plano, corresponde um ponto no *plano complexo de w* . Se $f(z)$ for plurívoca, então um ponto no plano complexo de z é mapeado em mais de um ponto no plano complexo de w . Pontos no plano z são *mapeados* em pontos no plano w , enquanto que curvas no plano z são *mapeadas* em curvas no plano w . A figura 2.4 ilustra o processo de mapeamento.

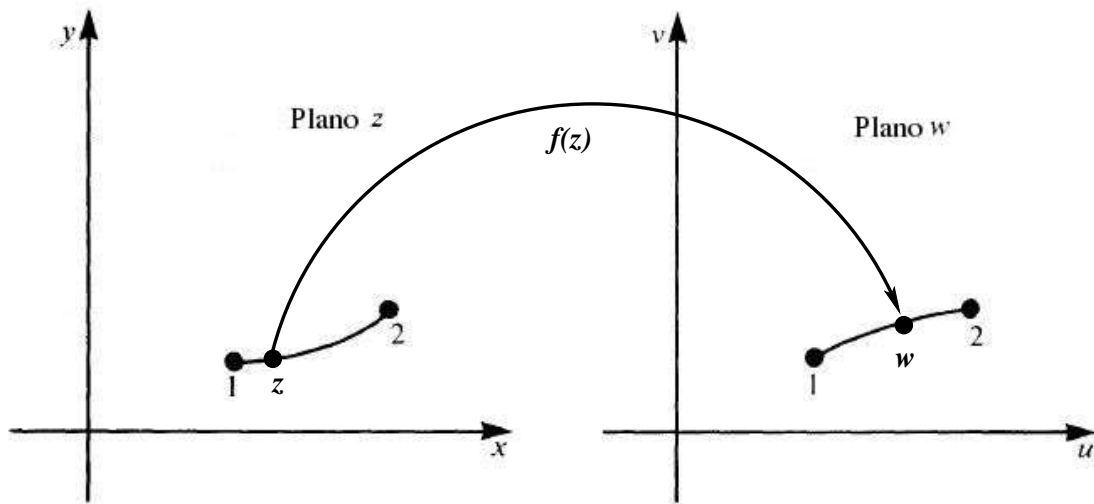


Figura 2.4: A função $w = f(z)$ mapeia pontos no plano z em pontos no plano w .

2.3.2 PONTOS DE RAMIFICAÇÃO, LINHAS DE RAMIFICAÇÃO E SUPERFÍCIES DE RIEMANN

Comparando agora o comportamento de uma função unívoca, $f_1(z) = z^2$, e de uma função plurívoca, $f_2(z) = \sqrt{z}$, percebe-se uma diferença importante entre ambas. Suponha que seja permitido que $z = re^{i\theta}$ execute uma revolução completa em torno do ponto $z = 0$, no sentido anti-horário e mantendo $r = \text{cte.}$, partindo de um determinado ponto no seu plano complexo. Esta operação consiste em substituir

$$\theta \longrightarrow \theta + 2\pi \tag{2.10}$$

na fórmula polar de z . Observando agora o comportamento de $f_1(z)$ e $f_2(z)$ frente a transformação (2.10), observa-se que

$$\begin{aligned} f_1(z) &\rightarrow r^2 e^{i2\theta} e^{i4\pi} \rightarrow f_1(z) \\ f_2(z) &\rightarrow r^{1/2} e^{i\theta/2} e^{i\pi} \rightarrow -f_2(z). \end{aligned}$$

Pode-se constatar que $f_1(z)$ permanece inalterada frente a transformação (2.10), porém $f_2(z)$ muda de sinal. Como o plano complexo possui por definição uma variação total de fase igual a 2π , a transformação (2.10) levou $f_2(z)$ a um valor distinto daquele que apresentava no início. De fato, $f_2(z)$ somente retornará ao valor inicial através de uma nova rotação completa. Ou seja,

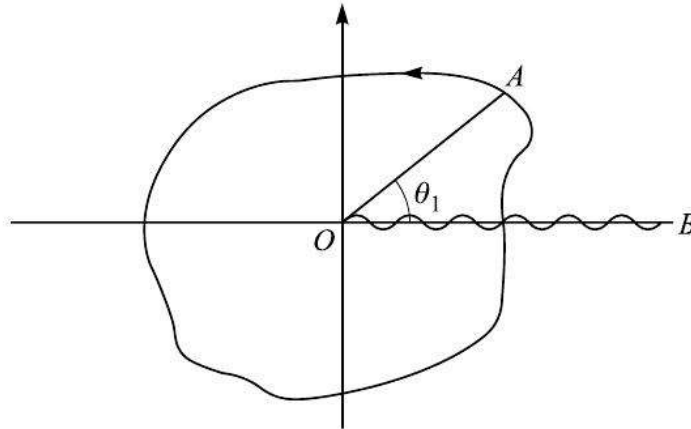


Figura 2.5: Linha de ramificação para a função $w = \sqrt{z}$.

$f_2(z) = \sqrt{z}$ não apresenta simetria frente a uma rotação de 2π radianos, mas sim frente a uma rotação $\theta \rightarrow \theta + 4\pi$, em cuja situação

$$f_2(z) \rightarrow r^{1/2} e^{i\theta/2} e^{i2\pi} \rightarrow f_2(z).$$

Pode-se descrever o que se sucede com a função $f_2(z) = \sqrt{z}$ afirmando-se que quando $0 \leq \theta < 2\pi$, o mapeamento do plano z para o plano w permanece sobre um dos ramos da função plurívoca $f_2(z)$, enquanto que no intervalo $2\pi \leq \theta < 4\pi$, o mapeamento leva ao outro ramo da função. Claramente, sobre cada ramo a função $f_2(z)$ é unívoca e, para assim mantê-la, estabelece-se uma barreira artificial ligando a origem ao infinito ao longo de alguma reta sobre o plano complexo de z . A função permanecerá unívoca desde que esta barreira não seja cruzada.

Para a função \sqrt{z} , esta linha é usualmente traçada ao longo do eixo real positivo e é denominada **linha de ramificação**, enquanto que o ponto O , de onde parte a linha de ramificação, é denominado **ponto de ramificação**. A figura 2.5 mostra esta linha de ramificação como uma linha sinuosa sobre o eixo real positivo. É importante enfatizar aqui que uma volta em torno de um outro ponto qualquer, distinto da origem, de tal forma que esta não esteja dentro da área delimitada pelo caminho fechado, não leva a um outro ramo da função \sqrt{z} . Ou seja, o ponto O é o único ponto de ramificação desta função.

George Friedrich Bernhard Riemann (1826–1866) sugeriu uma outra interpretação para a linha de ramificação definida acima. Imagina-se o plano z composto por duas folhas sobrepostas uma à outra. Corta-se agora as duas folhas ao longo da linha OB vista na figura 2.5 e junta-se a borda inferior da folha de baixo à borda superior da folha de cima. Da mesma forma, juntam-se as outras duas bordas. Partindo-se então do primeiro quadrante da folha de cima, realiza-se uma volta completa sobre o plano z em torno de O . Ao se cruzar a linha de ramificação, passa-se para o primeiro quadrante da folha de baixo; ao se realizar mais um volta completa em torno da origem, retorna-se à folha de cima ao se cruzar pela segunda vez a linha de ramificação. Desta maneira, a função \sqrt{z} permanece unívoca sobre um domínio no qual $0 \leq \theta < 4\pi$.

A coleção de duas folhas para a garantia da unicidade da função \sqrt{z} é denominada de **superfície de Riemann**. Cada **folha de Riemann** corresponde a um ramo da função e, sobre cada folha, a função é unívoca. O conceito de superfície de Riemann possui a vantagem de possibilitar a obtenção dos vários valores de uma função plurívoca de uma maneira contínua. A figura 2.6 ilustra as duas folhas de Riemann da função \sqrt{z} .

2.3.3 EXEMPLOS DE FUNÇÕES UNÍVOCAS OU PLURÍVOCAS

Além das funções $f_1(z) = z^2$ e $f_2(z) = \sqrt{z}$ já abordadas, outras funções de uma variável complexa que com frequência surgem são as seguintes.

Função exponencial. Definida por

$$w = e^z = e^{x+iy} = e^x (\cos y + i \operatorname{sen} y).$$

Funções trigonométricas. Define-se as funções trigonométricas em termos das funções exponenciais.

$$\left. \begin{aligned} \operatorname{sen} z &= \frac{e^{iz} - e^{-iz}}{2i} \\ \operatorname{cos} z &= \frac{e^{iz} + e^{-iz}}{2} \end{aligned} \right\} \cos^2 z + \operatorname{sen}^2 z = 1.$$

Funções trigonométricas hiperbólicas. De maneira análoga, define-se

$$\left. \begin{aligned} \operatorname{senh} z &= \frac{e^z - e^{-z}}{2} \\ \operatorname{cosh} z &= \frac{e^z + e^{-z}}{2} \end{aligned} \right\} \operatorname{cosh}^2 z - \operatorname{senh}^2 z = 1.$$

É possível mostrar as seguintes relações entre as funções trigonométricas circulares e as hiperbólicas:

$$\begin{aligned} \operatorname{sen} iz &= i \operatorname{senh} z & \operatorname{senh} iz &= i \operatorname{sen} z \\ \operatorname{cos} iz &= \operatorname{cosh} z & \operatorname{cosh} iz &= \operatorname{cos} z. \end{aligned}$$

Função logarítmica. Esta é uma outra função plurívoca, definida por

$$w = \ln z = \ln \left[r e^{i(\theta + 2k\pi)} \right] = \ln r + i(\theta + 2k\pi), \quad k = 0, 1, 2, \dots$$

Como se pode notar, esta função possui infinitos ramos, sendo $w = \ln r + i\theta$, para $0 \leq \theta < 2\pi$, o ramo principal. A superfície de Riemann para esta função está representada na figura 2.7.

2.4 O CÁLCULO DIFERENCIAL DE FUNÇÕES DE UMA VARIÁVEL COMPLEXA

Nesta seção serão definidos os conceitos de limites, continuidade e de derivação de uma função de uma variável complexa.

2.4.1 LIMITE DE UMA FUNÇÃO COMPLEXA

Dados os números $\{z, z_0, w_0\} \subset \mathbb{C}$, diz-se que o número w_0 é o limite de $f(z)$ à medida que z se aproxima de z_0 , o que é escrito como

$$\lim_{z \rightarrow z_0} f(z) = w_0,$$

se:

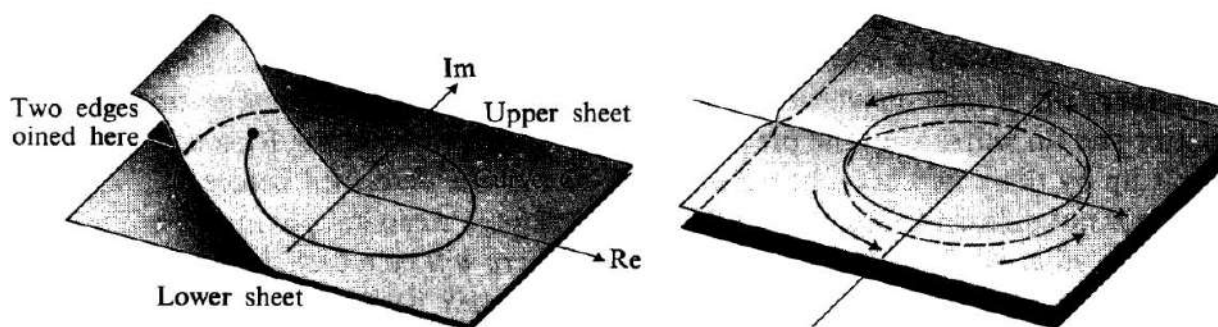


Figura 2.6: Folhas de Riemann da função \sqrt{z} .

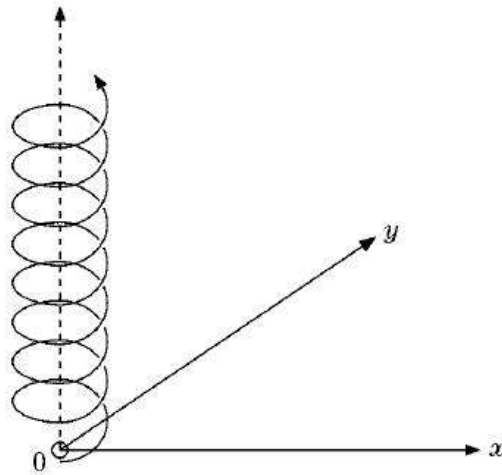


Figura 2.7: Superfície de Riemann da função $\ln z$.

1. A função $f(z)$ está definida e é unívoca em uma vizinhança de $z = z_0$, com a possível exceção do próprio ponto z_0 .
2. Dado um número real positivo qualquer ϵ , arbitrariamente pequeno, existe um outro número real positivo δ tal que

$$|f(z) - w_0| < \epsilon \text{ sempre que } 0 < |z - z_0| < \delta.$$

É importante observar o seguinte:

- O limite w_0 deve ser sempre o mesmo para um dado z_0 , independente da maneira como é realizado o limite $z \rightarrow z_0$.
- Se $f(z)$ é uma função plurívoca, o limite para $z \rightarrow z_0$ depende do particular ramo em que se encontra a vizinhança de z_0 .

A figura 2.8 ilustra as vizinhanças dos pontos $z = x + iy$ e $w_0 = u + iv$ nos respectivos planos complexos.

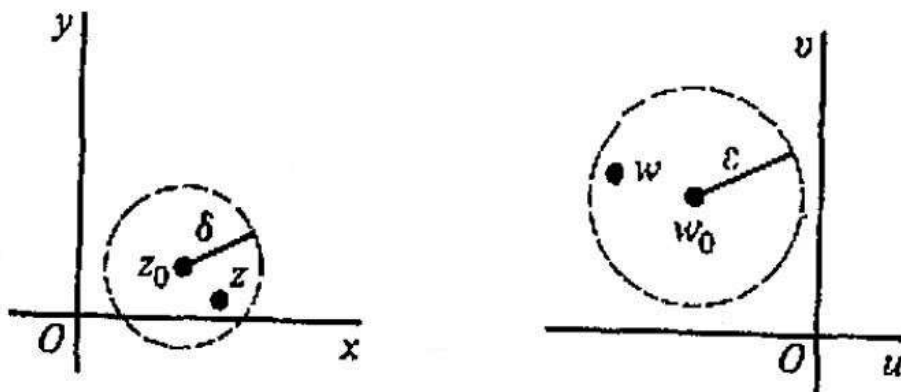


Figura 2.8: Vizinhanças dos pontos z_0 e w_0 nos respectivos planos complexos.

Exemplo 2.3 (Cálculos de limites). Encontre os seguintes limites:

(a) Se $f(z) = z^2$, prove que $\lim_{z \rightarrow z_0} f(z) = z_0^2$.

(b) Encontre $\lim_{z \rightarrow z_0} f(z)$ se

$$f(x) = \begin{cases} z^2, & z \neq z_0 \\ 0, & z = z_0. \end{cases}$$

Solução.

(a) Deve-se mostrar que para qualquer $\epsilon > 0$ é sempre possível encontrar-se um $\delta > 0$ (dependendo, em geral, de ϵ) tal que $|z^2 - z_0^2| < \epsilon$ sempre que $0 < |z - z_0| < \delta$.

Para tanto, considera-se $\delta < 1$. Neste caso, $0 < |z - z_0| < \delta$ implica que

$$\begin{aligned} |z - z_0||z + z_0| &< \delta |z + z_0| = \delta |z - z_0 + 2z_0|, \\ |z^2 - z_0^2| &< \delta (|z - z_0| + 2|z_0|) < \delta (1 + 2|z_0|). \end{aligned}$$

Para um $\epsilon \leq 1$ escolhe-se então $\delta = \epsilon / (1 + 2|z_0|)$, ou seja, $\delta < \epsilon \forall z_0 \in \mathbb{C}$, de tal maneira que

$$|z^2 - z_0^2| < \epsilon,$$

provando-se o limite.

(b) Não há diferença entre este problema e o problema da parte (a), uma vez que em ambos os casos o ponto $z = z_0$ foi excluído. Portanto, $\lim_{z \rightarrow z_0} f(z) = z_0^2$. Nota-se que o valor do limite não necessariamente é igual ao valor de $f(z_0)$.

Teorema 2.2 (Propriedades dos limites). Se $\lim_{z \rightarrow z_0} f(z) = w_1$ e $\lim_{z \rightarrow z_0} g(z) = w_2$, então as seguintes propriedades de limites são válidas:

- $\lim_{z \rightarrow z_0} [f(z) + g(z)] = \lim_{z \rightarrow z_0} f(z) + \lim_{z \rightarrow z_0} g(z) = w_1 + w_2$.
- $\lim_{z \rightarrow z_0} [f(z)g(z)] = \left[\lim_{z \rightarrow z_0} f(z) \right] \left[\lim_{z \rightarrow z_0} g(z) \right] = w_1 w_2$.
- $\lim_{z \rightarrow z_0} \frac{f(z)}{g(z)} = \frac{\lim_{z \rightarrow z_0} f(z)}{\lim_{z \rightarrow z_0} g(z)} = \frac{w_1}{w_2}$, desde que $w_2 \neq 0$.

2.4.2 CONTINUIDADE

Seja $f(z)$ definida e unívoca em uma vizinhança de $z = z_0$, assim como em $z = z_0$. A função $f(z)$ é dita contínua em $z = z_0$ se

$$\lim_{z \rightarrow z_0} f(z) = f(z_0).$$

Observa-se que isso implica em três condições que devem ser satisfeitas:

1. O limite deve existir.
2. $f(z_0)$ deve existir, isto é, $f(z)$ deve ser definida em $z = z_0$.
3. O limite deve ser igual a $f(z_0)$.

Pontos no plano z onde $f(z)$ deixa de ser contínua são denominados **descontinuidades de $f(z)$** .

Se o limite $\lim_{z \rightarrow z_0} f(z)$ existe mas não é igual a $f(z_0)$, então z_0 é denominado uma **descontinuidade removível**, pois é sempre possível redefinir-se $f(z)$ para se obter uma função contínua.

Teorema 2.3 (Teoremas de continuidade). Os seguintes teoremas de continuidade são válidos.

- Se $f(z)$ e $g(z)$ são contínuas em $z = z_0$, então também são contínuas:

$$f(z) + g(z), f(z)g(z) \text{ e } \frac{f(z)}{g(z)}, \text{ desde que } g(z_0) \neq 0.$$

- Se $w = f(z)$ é contínua em $z = z_0$ e $z = g(\xi)$ é contínua em $\xi = \xi_0$ e se $\xi_0 = f(z_0)$, então a função $w = g[f(z)]$ é contínua em $z = z_0$.

Uma função contínua de uma função contínua também é contínua.

- Se $f(z)$ é contínua em uma região fechada do plano complexo, então ela é limitada nessa região; isto é, existe uma constante real positiva M tal que $|f(z)| < M$ para todos os pontos z dentro dessa região.
- Se $f(z)$ é contínua em uma região, então as partes real e imaginária de $f(z)$ também são contínuas nessa região.

2.4.3 DERIVADAS DE FUNÇÕES COMPLEXAS

Dada uma função $f(z)$, contínua e unívoca de uma variável complexa z , em uma dada região $R \subseteq \mathbb{C}$, a derivada

$$f'(z) \equiv \frac{df}{dz}$$

em algum ponto fixo $z_0 \in R$ é definida como

$$f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}, \quad (2.11)$$

desde que este limite exista de forma **independente da maneira como** $\Delta z \rightarrow 0$. Aqui, $\Delta z = z - z_0$, sendo $z \in R$ algum ponto na vizinhança de z_0 .

Teorema 2.4. *Se uma função $f(z)$ possui derivada em $z = z_0$, então ela é necessariamente contínua em $z = z_0$.*

Demonstração. Supondo que $f'(z_0)$ exista, então

$$\lim_{z \rightarrow z_0} [f(z_0 + \Delta z) - f(z_0)] = \lim_{z \rightarrow z_0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} \lim_{z \rightarrow z_0} \Delta z = 0,$$

ou seja,

$$\lim_{z \rightarrow z_0} f(z_0 + \Delta z) = f(z_0).$$

□

Se $f'(z)$ existe em z_0 e em todos os pontos em uma dada vizinhança de z_0 , então $f(z)$ é dita **analítica em** z_0 . A função $f(z)$ é analítica na região R se ela é analítica em todos os pontos $z \in R$. Contudo, nem toda a função contínua é diferenciável em $z = z_0$.

Exemplo 2.4. Dada a a função $f(z) = z^*$, mostre que embora esta seja contínua em qualquer $z_0 \in \mathbb{C}$, sua derivada dz^*/dz não existe em z_0 .

Solução. Pela definição (2.11),

$$\begin{aligned} \frac{dz^*}{dz} &= \lim_{\Delta z \rightarrow 0} \frac{(z + \Delta z)^* - z^*}{\Delta z} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{(x + iy + \Delta x + i\Delta y)^* - (x + iy)^*}{\Delta x + i\Delta y} \\ &= \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{x - iy + \Delta x - i\Delta y - (x - iy)}{\Delta x + i\Delta y} = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\Delta x - i\Delta y}{\Delta x + i\Delta y}. \end{aligned}$$

Se $\Delta y = 0$, o limite resulta em $\lim_{\Delta x \rightarrow 0} \Delta x / \Delta x = 1$. Por outro lado, se $\Delta x = 0$, o limite resulta em $\lim_{\Delta y \rightarrow 0} (-\Delta y) / \Delta y = -1$. Portanto, como o valor do limite depende da maneira como $\Delta z \rightarrow 0$, a derivada de $f(z) = z^*$ não existe em nenhum ponto e, portanto, a função não é analítica em nenhum ponto.

Exemplo 2.5. Dada a função $g(z) = |z|^2$, mostre que esta somente é diferenciável em $z = 0$.

Solução. Pela definição (2.11),

$$\begin{aligned} g'(z) &= \lim_{\Delta z \rightarrow 0} \frac{|z + \Delta z|^2 - |z|^2}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{(z + \Delta z)(z^* + \Delta z^*) - zz^*}{\Delta z} \\ &= \lim_{\Delta z \rightarrow 0} \frac{z^* \Delta z + z \Delta z^* + \Delta z \Delta z^*}{\Delta z} = z^* + z \lim_{\Delta z \rightarrow 0} \frac{\Delta z^*}{\Delta z} + \lim_{\Delta z \rightarrow 0} \Delta z^*. \end{aligned}$$

Pode-se considerar 2 possibilidades:

1. $z = 0$. Neste caso,

$$g'(z)|_{z=0} = 0,$$

e a derivada existe.

2. $z \neq 0$. Neste caso, se $g'(z)$ existe, então a derivada deve existir independente da maneira como se toma o limite. Assim:

- Se $\Delta z = \Delta x$, então $\Delta z^* = \Delta z$ e o limite fica

$$g'(z) = z^* + z.$$

- Se $\Delta z = i\Delta y$, então $\Delta z^* = -\Delta z$ e o limite fica

$$g'(z) = z^* - z.$$

Portanto, a função $g(z)$ não é analítica, pois somente possui derivada em $z = 0$.

Teorema 2.5 (Regras de derivação). *As regras de derivação para somas, produtos e quocientes de funções são, em geral, as mesmas que as regras para funções reais. Isto é, se $f'(z_0)$ e $g'(z_0)$ existem, então*

- $(f + g)'(z_0) = f'(z_0) + g'(z_0)$.
- $(fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0)$.
- $\left(\frac{f}{g}\right)'(z_0) = \frac{f'(z_0)g(z_0) - f(z_0)g'(z_0)}{[g(z_0)]^2}$, se $g(z_0) \neq 0$.

2.4.4 AS CONDIÇÕES DE CAUCHY-RIEMANN

Para testar se uma função $f(z)$ é analítica, Cauchy e Riemann criaram um teste simples mas extremamente importante para testar a analiticidade de $f(z)$. Para deduzir as condições de Cauchy-Riemann, retorna-se à definição (2.11),

$$f'(z_0) = \lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z}. \tag{2.12}$$

Escrevendo $f(z) = u(x, y) + iv(x, y)$, obtém-se

$$f'(z) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{u(x + \Delta x, y + \Delta y) - u(x, y)}{\Delta x + i\Delta y} + i \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{v(x + \Delta x, y + \Delta y) - v(x, y)}{\Delta x + i\Delta y}.$$

Existe um número infinito de maneiras para Δz tender a zero sobre o plano complexo. Consideram-se duas possibilidades (ver figura 2.9): ao longo de x ou ao longo de y . Supondo-se que se tome primeiro a rota ao longo de x , mantendo y constante, isto é, $\Delta y = 0$. Neste caso,

$$f'(z) = \lim_{\Delta x \rightarrow 0} \frac{u(x + \Delta x, y) - u(x, y)}{\Delta x} + i \lim_{\Delta x \rightarrow 0} \frac{v(x + \Delta x, y) - v(x, y)}{\Delta x} = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x}.$$

Agora, toma-se a rota ao longo de y , mantendo x constante ($\Delta x = 0$). Neste caso,

$$f'(z) = \lim_{\Delta y \rightarrow 0} \frac{u(x, y + \Delta y) - u(x, y)}{i\Delta y} + i \lim_{\Delta y \rightarrow 0} \frac{v(x, y + \Delta y) - v(x, y)}{i\Delta y} = -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}.$$

A condição necessária para que $f(z)$ seja analítica é que o limite deve resultar sempre no mesmo valor, independente do caminho adotado sobre o plano complexo. Portanto, uma condição necessária para que $f(z)$ seja analítica é

$$\frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = -i \frac{\partial u}{\partial y} + \frac{\partial v}{\partial y},$$

de onde resultam as **condições de Cauchy-Riemann**

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ e } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{2.13}$$

Estas relações fornecem também duas expressões úteis para a derivada de $f(z)$:

$$f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y}. \tag{2.14}$$

Podemos estabelecer então o seguinte teorema.

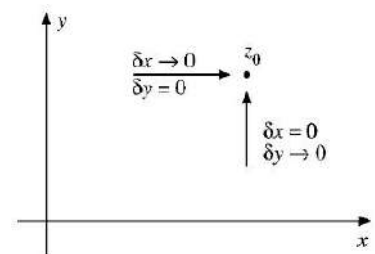


Figura 2.9: Caminhos alternativos para z_0 .

Teorema 2.6 (Condição necessária). Se a derivada $f'(z)$ de um função $f(z) = u(x, y) + iv(x, y)$ existe em um ponto $z = x + iy$, então as derivadas parciais de primeira ordem de $u(x, y)$ e $v(x, y)$ com respeito a x e a y devem existir neste ponto e satisfazer as relações de Cauchy-Riemann (2.13). Além disso, $f'(z)$ pode ser determinada pelas expressões (2.14).

Exemplo 2.6 (Condições de Cauchy-Riemann). Seja a função $f(z) = z^2 = x^2 - y^2 + i2xy$. Neste caso, $u(x, y) = x^2 - y^2$ e $v(x, y) = 2xy$. Para estas funções,

$$\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y} \text{ e } \frac{\partial u}{\partial y} = -2y = -\frac{\partial v}{\partial x}.$$

Portanto, as relações de Cauchy-Riemann são satisfeitas e $f'(z)$ pode ser obtida por (2.14),

$$f'(z) = 2x + i2y = 2z.$$

Exemplo 2.7 (Condições de Cauchy-Riemann). Seja agora a função $f(z) = |z|^2 = x^2 + y^2$. Neste caso, $u(x, y) = x^2 + y^2$ e $v(x, y) = 0$. Portanto, embora as derivadas parciais existam,

$$\frac{\partial u}{\partial x} = 2x, \quad \frac{\partial u}{\partial y} = 2y, \quad \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} = 0,$$

estas não satisfazem as relações (2.13) e, portanto, a função $f(z)$ não possui derivada.

As condições de Cauchy-Riemann fornecem uma condição *necessária* para que a função seja diferenciável em algum ponto $z = z_0$. Contudo, não há garantia até este momento de que estas condições sejam *suficientes* para garantir a existência desta derivada. Um teorema mais geral, apresentado a seguir, estabelece as condições necessária e suficiente para a existência da derivada de $f(z)$.

Teorema 2.7 (Condição necessária e suficiente). Dada a função $f(z) = u(x, y) + iv(x, y)$, se $u(x, y)$ e $v(x, y)$ são contínuas com derivadas parciais de primeira ordem e que satisfazem as condições de Cauchy-Riemann (2.13) em todos os pontos em uma região $R \subseteq \mathbb{C}$, então $f(z)$ é analítica em R .

Demonstração. Para provar este teorema, é necessário empregar o seguinte teorema do cálculo de funções reais de 2 variáveis: se $h(x, y)$, $\partial h/\partial x$ e $\partial h/\partial y$ são contínuas em uma região R em torno do ponto (x_0, y_0) , então existe uma função $H(\Delta x, \Delta y)$ tal que $H(\Delta x, \Delta y) \rightarrow 0$ à medida que $(\Delta x, \Delta y) \rightarrow (0, 0)$ e

$$h(x_0 + \Delta x, y_0 + \Delta y) - h(x_0, y_0) = \frac{\partial h}{\partial x} \Big|_{(x_0, y_0)} \Delta x + \frac{\partial h}{\partial y} \Big|_{(x_0, y_0)} \Delta y + H(\Delta x, \Delta y) \sqrt{(\Delta x)^2 + (\Delta y)^2}.$$

Retornando então à definição de derivada (2.12)

$$\lim_{\Delta z \rightarrow 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z},$$

sendo z_0 qualquer ponto que pertence a R e $\Delta z = \Delta x + i\Delta y$. Pode-se escrever então

$$f(z_0 + \Delta z) - f(z_0) = [u(x_0 + \Delta x, y_0 + \Delta y) - u(x_0, y_0)] + i[v(x_0 + \Delta x, y_0 + \Delta y) - v(x_0, y_0)],$$

$$f(z_0 + \Delta z) - f(z_0) = \frac{\partial u}{\partial x} \Big|_{(x_0, y_0)} \Delta x + \frac{\partial u}{\partial y} \Big|_{(x_0, y_0)} \Delta y + H(\Delta x, \Delta y) \sqrt{(\Delta x)^2 + (\Delta y)^2} + i \left[\frac{\partial v}{\partial x} \Big|_{(x_0, y_0)} \Delta x + \frac{\partial v}{\partial y} \Big|_{(x_0, y_0)} \Delta y + G(\Delta x, \Delta y) \sqrt{(\Delta x)^2 + (\Delta y)^2} \right],$$

onde $H(\Delta x, \Delta y) \rightarrow 0$ e $G(\Delta x, \Delta y) \rightarrow 0$ quando $(\Delta x, \Delta y) \rightarrow (0, 0)$.

Empregando agora as condições de Cauchy-Riemann (2.13), obtém-se

$$f(z_0 + \Delta z) - f(z_0) = \left[\frac{\partial u}{\partial x} \Big|_{(x_0, y_0)} + i \frac{\partial v}{\partial x} \Big|_{(x_0, y_0)} \right] (\Delta x + i\Delta y)$$

$$+ [H(\Delta x, \Delta y) + iG(\Delta x, \Delta y)] \sqrt{(\Delta x)^2 + (\Delta y)^2},$$

portanto,

$$\frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} = \frac{\partial u}{\partial x} \Big|_{(x_0, y_0)} + i \frac{\partial v}{\partial x} \Big|_{(x_0, y_0)} + [H(\Delta x, \Delta y) + iG(\Delta x, \Delta y)] \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\Delta x + i\Delta y}.$$

Assim, no limite $(\Delta x, \Delta y) \rightarrow (0, 0)$,

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\Delta x + i\Delta y} \rightsquigarrow \left| \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\Delta x + i\Delta y} \right| = 1.$$

Ou seja,

$$f'(z_0) = \frac{\partial u}{\partial x} \Big|_{(x_0, y_0)} + i \frac{\partial v}{\partial x} \Big|_{(x_0, y_0)},$$

o que mostra que o limite e, portanto, $f'(z)$ existem em todos os pontos em R . As condições de Cauchy-Riemann são, portanto necessárias e suficientes para garantir a existência de $f'(z)$ em R . \square

2.4.5 FUNÇÕES ANALÍTICAS

Uma função $f(z)$ é analítica em um ponto z_0 se a sua derivada $f'(z)$ existe não somente em z_0 mas em todos os pontos z dentro de uma vizinhança de z_0 . As seguintes definições são feitas, com respeito a funções analíticas:

- Uma função é dita analítica em um domínio $R \subseteq \mathbb{C}$ se ela é analítica em todos os pontos $z \in R$. Uma função analítica também é denominada **regular** ou **holomórfica**.
- Se a função $f(z)$ é analítica sobre todo o plano z complexo, ela é denominada **inteira**.
- Uma função $f(z)$ é denominada **singular em** $z = z_0$ se ela não é diferenciável neste ponto. O ponto z_0 é denominado **ponto singular** ou **singularidade** de $f(z)$.

2.4.6 FUNÇÕES HARMÔNICAS

Se $f(z) = u(x, y) + iv(x, y)$ é analítica em alguma região R do plano complexo, então em todos os pontos desta região as condições de Cauchy-Riemann (2.13) são satisfeitas:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{e} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

e, portanto,

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} \quad \text{e} \quad \frac{\partial^2 u}{\partial y^2} = -\frac{\partial^2 v}{\partial y \partial x},$$

desde que as derivadas segundas existam. Igualando a ambas as expressões acima, obtém-se que $u(x, y)$ e $v(x, y)$ satisfazem a Equação de Laplace:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{2.15a}$$

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \tag{2.15b}$$

sobre toda a região R .

Toda a função que satisfaz as equações de Laplace (2.15) é denominada de **função harmônica**. Como ambas as funções u e v satisfazem a (2.15), estas são denominadas **funções harmônicas conjugadas**.

2.4.7 PONTOS SINGULARES

Um ponto onde $f(z)$ deixa de ser analítica é denominado ponto singular ou singularidade de $f(z)$. As condições de Cauchy-Riemann não são obedecidas em um ponto singular.

Existem várias tipos de pontos singulares distintos:

- 1. Pontos singulares isolados.** O ponto $z = z_0$ é denominado ponto singular isolado de $f(z)$ se for possível encontrar-se uma quantidade $\delta > 0$ tal que o círculo $|z - z_0| = \delta$ circunda apenas o ponto singular z_0 . Se não for possível encontrar-se um $\delta > 0$, o ponto z_0 é denominado ponto singular não isolado.
- 2. Polos.** O ponto singular $z = z_0$ é denominado **polo de ordem n** de $f(z)$ se for possível encontrar-se um número inteiro positivo n tal que

$$\lim_{z \rightarrow z_0} (z - z_0)^n f(z) = A \neq 0. \quad (2.16)$$

Exemplos:

- $f(z) = 1/(z - 2)$ possui um **polo simples** ou de ordem 1 em $z = 2$.
- $f(z) = 1/(z - 2)^2$ possui um **polo duplo** ou de ordem 2 em $z = 2$.
- $f(z) = 1/(z - 2)^3$ possui um polo de ordem 3 em $z = 2$.

- 3. Ponto de ramificação.** Uma função possui um ponto de ramificação em $z = z_0$ se, após circular-se em torno de z_0 , retornando ao ponto de partida sobre o plano z , a função não retorna ao seu valor inicial sobre o plano w . Os pontos de ramificação são singularidades das funções plurívocas. Por exemplo, a função $f(z) = \sqrt{z}$ possui um ponto de ramificação em $z = 0$.
- 4. Singularidades removíveis.** O ponto singular $z = z_0$ é denominado uma singularidade removível se $\lim_{z \rightarrow z_0} f(z)$ existe. Por exemplo, a função $f(z) = \sin z/z$ possui um ponto singular em $z = 0$, mas $\lim_{z \rightarrow 0} \sin z/z = 1$. Neste caso, pode-se redefinir a função $f(z)$ para esta esteja definida em z_0 .
- 5. Singularidades essenciais.** Uma função possui uma singularidade essencial em z_0 se esta não possui polos, em qualquer ordem, que sejam eliminados pela multiplicação por $(z - z_0)^n$, para qualquer valor finito de n . Por exemplo, a função $f(z) = e^{1/(z-2)}$ possui uma singularidade essencial em $z = 2$.
- 6. Singularidades no infinito.** Uma função $f(z)$ possui uma singularidade em $z \rightarrow \infty$ se esta for do mesmo tipo que a singularidade de $f(1/w)$, para $w \rightarrow 0$. Por exemplo, $f(z) = z^2$ possui um polo de ordem 2 no infinito, uma vez que $f(1/w) = 1/w^2$ possui um polo duplo em $w = 0$.

2.5 INTEGRAÇÃO NO PLANO COMPLEXO

Integração complexa é uma ferramenta muito importante na física-matemática. Por exemplo, com frequência surgem integrais de funções reais que não podem ser calculadas pelos métodos usuais de integração, mas que podem ser resolvidas estendendo-se a definição do integrando para o conjunto dos números complexos e realizando-se a integração neste plano. Na análise matemática, o método de integração complexa possibilita demonstrações de algumas propriedades básicas de funções analíticas, as quais seriam muito difíceis de ser realizadas sem o recurso da integração complexa.

O resultado mais importante na teoria da integração complexa é o teorema integral de Cauchy, a partir do qual as fórmulas integrais de Cauchy são derivadas. A outra propriedade de fundamental importância para a física-matemática é o teorema dos resíduos, que possibilita o cálculo de integrais que não poderiam ser realizadas de outra maneira. Estes resultados serão abordados nas seções posteriores. Nesta seção, será introduzido o conceito de integral de caminho sobre o plano complexo e apresentadas as propriedades matemáticas fundamentais das integrais complexas.

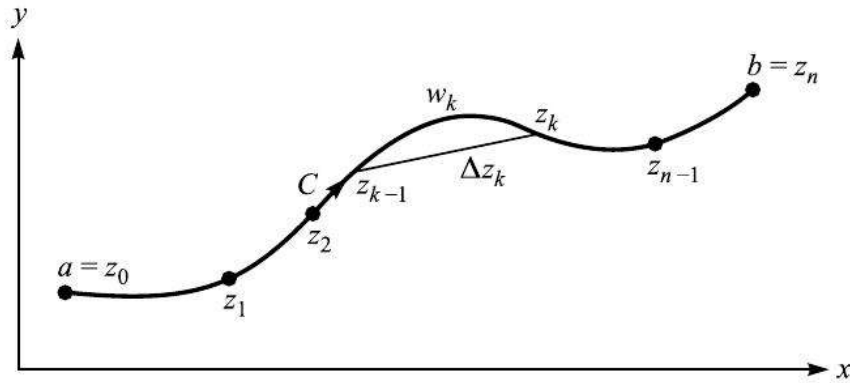


Figura 2.10: Caminho C ao longo do qual a integração complexa é realizada.

2.5.1 INTEGRAIS DE CAMINHO NO PLANO COMPLEXO

Uma integral de caminho, também denominada **integral de linha**, possui uma relação com a derivada no plano complexo exatamente igual à que existe para funções reais. Se a função $F(z)$ é dada pela integral indefinida $F(z) = \int f(z) dz$, então a derivada de $F(z)$ é dada por $F'(z) = f(z)$. Em outras palavras uma integral indefinida no plano complexo é a operação inversa da derivação no mesmo plano.

Por outro lado, o plano complexo é definido a partir de duas variáveis independentes reais. Neste caso, poder-se-ia pensar que uma integral (definida) no plano complexo seria equivalente a uma integral de superfície de uma função real de duas variáveis. Contudo, na análise das funções complexas, a função $f(z)$ é integrada *ao longo de um caminho no plano complexo*. Para tanto, pode-se parametrizar o caminho ao longo do plano z fazendo-se uso de um parâmetro real t :

$$z(t) = x(t) + iy(t) \text{ para } a \leq t \leq b,$$

o qual define um caminho sobre o plano complexo à medida que t varia de a a b . Diz-se que este curva é **suave** se existe um vetor tangente à mesma ao longo de todos os pontos; isto implica que dx/dt e dy/dt existem são contínuas e não são nulas simultaneamente para $a \leq t \leq b$.

Sendo C uma curva suave sobre o plano z complexo, como mostra a figura 2.10, assume-se que a mesma possui um comprimento finito. Dada agora a função $f(z)$, contínua sobre todos os pontos ao longo de C , subdivide-se C em n partes por meio dos pontos $\{z_0, z_1, z_2, \dots, z_n\}$, arbitrariamente escolhidos, mas com $z_0 = a$ e $z_n = b$. Para cada arco de C que conecta os pontos z_{k-1} e z_k ($k = 1, 2, \dots, n$), escolhe-se um ponto w_k ($z_{k-1} \leq w_k \leq z_k$) e forma-se a soma

$$S_n = \sum_{k=1}^n f(w_k) \Delta z_k, \text{ onde } \Delta z_k = z_k - z_{k-1}.$$

Fazendo-se agora com que o número de subdivisões n aumente indefinidamente, de tal forma que o maior dos $|\Delta z_k|$ tenda a zero, a soma S_n aproxima-se de um limite. Se este limite existe e possui o mesmo valor, independente das escolhas dos $\{z_k\}$ e dos $\{w_k\}$ ao longo de C , então este limite é denominado a **integral de caminho (ou de linha) de $f(z)$ ao longo de C** e é denotado por:

$$S = \lim_{n \rightarrow \infty} S_n = \lim_{\substack{n \rightarrow \infty \\ |\Delta z|_{\max} \rightarrow 0}} \sum_{k=1}^n f(w_k) \Delta z_k \equiv \int_C f(z) dz = \int_a^b f(z) dz. \quad (2.17)$$

Quando o caminho é fechado, isto é, quando $b = a$ (ou $z_n = z_0$), a integral de linha é denominada **integral de contorno de $f(z)$** , a qual é denotada por

$$S = \oint_C f(z) dz.$$

Teorema 2.8 (Teorema de existência). *Se o caminho C é suave por partes e $f(z)$ é contínua ao longo de C , então $\int_C f(z) dz$ sempre existe.*

2.5.2 PROPRIEDADES MATEMÁTICAS DAS INTEGRAIS DE LINHA

A integral de linha de $f(z) = u(x, y) + iv(x, y)$ ao longo de um caminho C pode sempre ser expressa em termos de integrais reais de caminho como

$$\int_C f(z) dz = \int_C (u + iv)(dx + idy) = \int_C (u dx - v dy) + i \int_C (v dx + u dy),$$

onde a curva C pode ser aberta ou fechada, mas o sentido de integração deve sempre ser especificado, por exemplo através do uso de um parâmetro t . Invertendo-se o sentido de variação de t , inverte-se o sinal da integral.

Integrais complexas são, portanto, redutíveis a integrais reais de caminho e possuem as seguintes propriedades:

(1) $\int_C [f(z) + g(z)] dz = \int_C f(z) dz + \int_C g(z) dz.$

(2) $\int_C kf(z) dz = k \int_C f(z) dz,$ sendo $k \in \mathbb{C}$ uma constante.

(3) $\int_a^b f(z) dz = - \int_b^a f(z) dz,$ sendo $\{a, b\} \in \mathbb{C}.$

(4) $\int_a^b f(z) dz = \int_a^m f(z) dz + \int_m^b f(z) dz,$ sendo $m \in \mathbb{C}.$

(5) $\left| \int_C f(z) dz \right| \leq ML,$ onde $M = \max |f(z)|$ ao longo de C e L é o comprimento de $C.$

(6) $\left| \int_C f(z) dz \right| \leq \int_C |f(z)| |dz|.$

A propriedade (5), em particular, é bastante útil e será bastante utilizada, porque ao se trabalhar com integrais de linha complexas, com frequência é necessário estabelecer-se limites nos seus valores absolutos.

Demonstração. (Propriedade 5). Retornando à definição (2.17),

$$\int_C f(z) dz = \lim_{\substack{n \rightarrow \infty \\ |\Delta z|_{\max} \rightarrow 0}} \sum_{k=1}^n f(w_k) \Delta z_k.$$

Mas,

$$\left| \sum_{k=1}^n f(w_k) \Delta z_k \right| \leq \sum_{k=1}^n |f(w_k)| |\Delta z_k| \leq M \sum_{k=1}^n |\Delta z_k| \leq ML,$$

onde se fez uso do fato de que $|f(z)| \leq M$ para todos os pontos z ao longo de C e que $\sum |\Delta z_k|$ representa a soma de todas as cordas juntando os pontos z_{k-1} e z_k ao longo de C e que esta soma não pode ser maior que o comprimento L de C . Tomando-se agora o limite para $n \rightarrow \infty$ em ambos os lados, resulta a propriedade (5). A propriedade (6) também segue desta demonstração. \square

Exemplo 2.8. Calcule a integral $\int_C (z^*)^2 dz$, sendo C a linha reta ligando os pontos $z = 0$ e $z = 1 + 2i$.

Solução. Uma vez que

$$(z^*)^2 = (x - iy)^2 = x^2 - y^2 - 2ixy,$$

resulta

$$\int_C (z^*)^2 dz = \int_C [(x^2 - y^2) dx + 2xy dy] + i \int_C [-2xy dx + (x^2 - y^2) dy].$$

Para parametrizar a curva C , pode-se escolher $x(t)$ e $y(t)$ dados por

$$x(t) = t, \quad y(t) = 2t, \quad \text{para } (0 \leq t \leq 1),$$

ou, simplesmente, pode-se escrever $y = 2x$. Portanto,

$$\int_C (z^*)^2 dz = \int_0^1 5x^2 dx + i \int_0^1 (-10x^2) dx = \frac{5}{3} - \frac{10}{3}i.$$

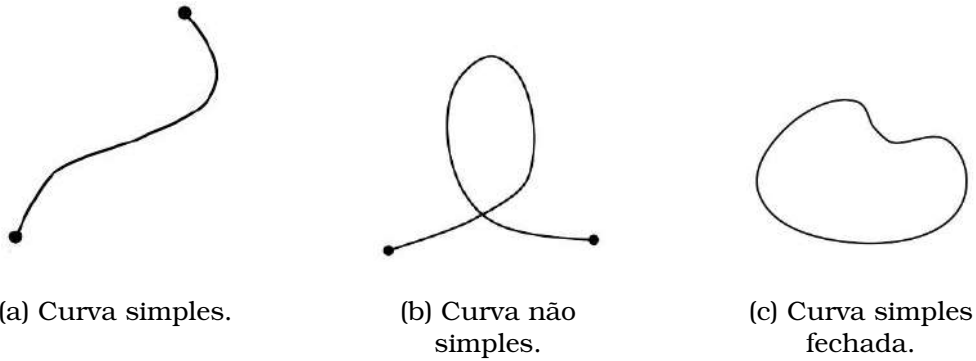


Figura 2.11: Exemplos de curvas simples ou não simples.

Exemplo 2.9. Calcule a integral de caminho

$$\oint \frac{dz}{(z - z_0)^{n+1}},$$

sendo C uma circunferência de raio r centrada em z_0 e n é um número inteiro. Uma ilustração deste contorno pode ser vista na figura 2.8 à esquerda.

Solução. Por conveniência, escolhe-se $z - z_0 = re^{i\theta}$, onde θ é o parâmetro cuja variação ($0 \leq \theta < 2\pi$) determina o contorno C . Então, $dz = ire^{i\theta}d\theta$ e a integral fica:

$$\oint \frac{dz}{(z - z_0)^{n+1}} = \int_0^{2\pi} \frac{ire^{i\theta}d\theta}{r^{n+1}e^{i(n+1)\theta}} = \frac{i}{r^n} \int_0^{2\pi} e^{-in\theta}d\theta = \frac{i}{r^n} \int_0^{2\pi} (\cos n\theta - i \sin n\theta) d\theta = \begin{cases} 2\pi i, & n = 0 \\ 0, & |n| \geq 1. \end{cases}$$

Este é um resultado importante, que será utilizado diversas vezes nas seções posteriores.

2.5.3 TIPOS DE CURVAS E DOMÍNIOS NO PLANO COMPLEXO

Nesta seção serão brevemente definidos os tipos de curvas e domínios no plano complexo que serão considerados nas seções posteriores.

2.5.3.1 TIPOS DE CURVAS NO PLANO COMPLEXO

Uma curva C é dita **simples** (também denominada **arco de Jordan**) se esta não se intersecciona em nenhum ponto, isto é, $z(t_1) \neq z(t_2)$ se $t_1 \neq t_2$, para $a \leq t \leq b$. A exceção $z(b) = z(a)$ é permitida para um contorno fechado, em cuja situação o contorno é dito **contorno simples** ou **curva simples fechada** ou ainda **curva ou contorno de Jordan**. A figura 2.11 mostra exemplos de curvas simples e de curvas não simples.

2.5.3.2 DOMÍNIOS SIMPLEMENTE OU MULTIPLAMENTE CONEXOS

Um **domínio** ou **região simplesmente conexa** D é uma região no plano complexo tal que toda curva simples fechada Γ dentro de D delimita somente pontos que pertencem a D . Uma outra definição: uma região D é dita simplesmente conexa se qualquer curva simples fechada Γ contida dentro de D pode ser reduzida a um ponto de tal maneira que nenhum ponto contido em Γ abandone D .

Uma região que não é simplesmente conexa é dita **multiplamente conexa**. De acordo com as definições, deve então existir pelo menos uma curva simples fechada Γ contida em D que cerca pontos que não pertencem a D . Ou, alternativamente, uma região multiplamente conexa é aquela que não pode ser reduzida a um ponto sem que abandone (mesmo que momentaneamente) a região D . A figura 2.12 apresenta exemplos de regiões simplesmente e multiplamente conexas.

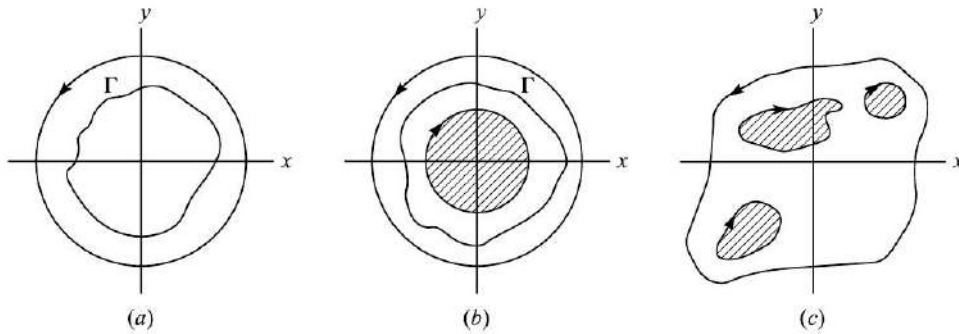


Figura 2.12: Exemplos de regiões: (a) simplesmente conexa e (b) e (c) multiplamente conexas.

2.5.3.3 CONVENÇÃO PARA O PERCURSO DE UM CONTORNO FECHADO

Considera-se uma região D do plano complexo, composta por pontos no interior e ao longo de um contorno simples fechado Γ . O contorno é percorrido no **sentido positivo** se todos os pontos de D se situarem à esquerda de um observador que se desloca ao longo de Γ . Este sentido positivo consiste no percurso *anti-horário* indicado pelas setas nos contornos Γ , representados nas figuras 2.12, e no percurso *horário* nos contornos interiores a Γ .

2.6 O TEOREMA DE CAUCHY-GOURSAT

O teorema de Cauchy-Goursat é um dos mais importantes resultados da análise matemática das funções complexas. Este teorema possui diversas consequências teóricas e práticas a respeito das propriedades analíticas das funções de variável complexa e serve de base para outros teoremas importantes como o teorema dos resíduos.

Para se realizar uma das demonstrações existentes do teorema de Cauchy, faz-se uso do teorema de Green, válido para integrais de linha e de superfície de funções de duas variáveis reais.

2.6.1 O TEOREMA DE GREEN NO PLANO

Teorema 2.9. *Sejam $P(x, y)$ e $Q(x, y)$ duas funções reais tais que suas derivadas são contínuas sobre toda uma região R , delimitada por um contorno fechado simples C . Neste caso, as funções P e Q satisfazem a seguinte identidade,*

$$\oint_R [P(x, y) dx + Q(x, y) dy] = \iint_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy. \quad (2.18)$$

Observação. A demonstração do teorema (2.18) baseia-se no teorema de Stokes e não será apresentada aqui.

2.6.2 O TEOREMA DE CAUCHY-GOURSAT

Serão apresentadas aqui duas versões do mencionado teorema, a primeira para um domínio simplesmente conexo e a segunda para uma região multiplamente conexa. A demonstração apresentada para a primeira versão foi obtida originalmente por Augustin-Louis Cauchy (1789 – 1857) no início do século XIX e supõe que as derivadas da função $f(z)$ são contínuas sobre o domínio D . Posteriormente, Édouard Jean-Baptiste Goursat (1858 – 1936) mostrou que a condição de continuidade de $f'(z)$ não é necessária para a validade do teorema. Por esta razão, o teorema leva o nome de ambos os matemáticos franceses.

Teorema 2.10 (Teorema de Cauchy-Goursat). *Se uma função $f(z) = u(x, y) + iv(x, y)$ é analítica em todos os pontos de um domínio simplesmente conexo D , então para todo contorno simples fechado C no interior de D ,*

$$\oint_C f(z) dz = 0. \quad (2.19)$$

Demonstração. Escreve-se o lado direito de (2.19) da seguinte maneira:

$$\oint_C f(z) dz = \oint_C (u + iv)(dx + idy) = \oint_C (u dx - v dy) + i \oint_C (v dx + u dy).$$

Aplicando-se o teorema de Green (2.18) a cada integral de contorno acima, obtém-se

$$\oint_C f(z) dz = - \iint_D \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) dx dy + i \iint_D \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy.$$

Como a função $f(z)$ é suposta analítica, então as funções $u(x, y)$ e $v(x, y)$ satisfazem as condições de Cauchy-Riemann (2.13). Portanto,

$$\oint_C f(z) dz = - \iint_D \left(\frac{\partial v}{\partial x} - \frac{\partial v}{\partial x} \right) dx dy + i \iint_D \left(\frac{\partial u}{\partial x} - \frac{\partial u}{\partial x} \right) dx dy = 0,$$

o que demonstra o teorema. \square

Quando o domínio D é multiplamente conexo, a seguinte versão do teorema de Cauchy-Goursat se aplica.

Teorema 2.11 (Teorema de Cauchy-Goursat em regiões multiplamente conexas). *Seja C um contorno simples fechado e seja $\{C_j\}$ ($j = 1, \dots, n$) um conjunto finito de contornos fechados simples interiores a C , tais que não existam duas regiões R_j e $R_{j'}$, interiores a C_j e $C_{j'}$, que compartilhem pontos no plano complexo em comum entre si. Seja R a região do plano complexo que consiste em todos os pontos no interior e ao longo de C , exceto por aqueles pontos no interior de cada contorno C_j . Finalmente, seja B o contorno completo que delimita R e que consiste no contorno C mais os contornos C_j , sendo todos estes contornos percorridos no sentido que mantém os pontos de R à direita de B . Se a função $f(z)$ é analítica em R , então*

$$\oint_B f(z) dz = 0. \quad (2.20)$$

Demonstração. A situação descrita no teorema está ilustrada pela figura 2.13a. O contorno mostrado na figura é composto por C , juntamente com os contornos C_1, \dots, C_n e os segmentos de reta $L_1^1, L_1^2, \dots, L_1^n$ e $L_2^1, L_2^2, \dots, L_2^n$. Desta maneira a região R passa de multiplamente conexa a simplesmente conexa. Aproximando-se agora os pares de segmentos de reta L_1^1 e L_2^1, L_1^2 e L_2^2, \dots, L_1^n e L_2^n , de tal forma que a distância entre os mesmos se torne infinitesimalmente pequena, as integrais de caminho de $f(z)$ em cada par de segmentos se anulam mutuamente, isto é,

$$\int_{L_1^1} f(z) dz = - \int_{L_2^1} f(z) dz, \quad \int_{L_1^2} f(z) dz = - \int_{L_2^2} f(z) dz, \quad \dots \quad \int_{L_1^n} f(z) dz = - \int_{L_2^n} f(z) dz,$$

de tal forma que o contorno restante é exatamente o contorno B descrito no teorema. Como a região R é agora simplesmente conexa e a função $f(z)$ é analítica em R , de (2.19) resulta

$$\oint_B f(z) dz = 0. \quad \square$$

O teorema de Cauchy-Goursat (2.19) possui consequências importantes, algumas das quais serão apresentadas nesta e nas seções posteriores.

Teorema 2.12 (Deformação do contorno de integração). *Seja $f(z)$ uma função analítica sobre uma região R delimitada pelo contorno simples fechado C e pelo conjunto de contornos $\{C_j\}$ ($j = 1, \dots, n$), interiores a C e que envolvem n buracos que podem conter singularidades isoladas ou não isoladas. Então,*

$$\oint_C f(z) dz = \sum_{j=1}^n \oint_{C_j} f(z) dz, \quad (2.21)$$

sendo que tanto a integral ao longo de C quanto as integrais nos contornos C_j são realizadas no sentido anti-horário.²

²O que fica evidenciado pelo símbolo \circlearrowleft .

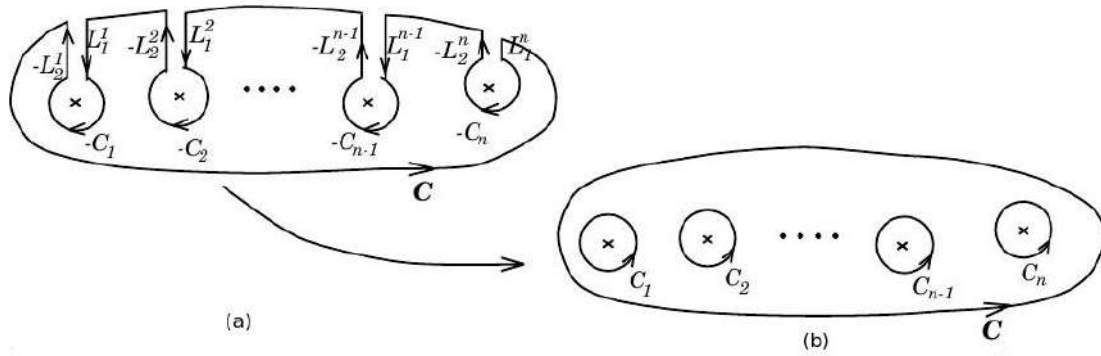


Figura 2.13: Contorno B que transforma uma região multiplamente conexa em uma região simplesmente conexa.

Demonstração. Esta situação também está ilustrada na figura 2.13. Ao se considerar o contorno B na figura 2.13a, o teorema de Cauchy (2.20) afirma que

$$\oint_B f(z) dz = \left[\int_C + \sum_{j=1}^n \left(\int_{-C_j} + \int_{L_1^j} + \int_{-L_2^j} \right) \right] f(z) dz = 0,$$

onde

$$\int_{-L_2^j} f(z) dz = - \int_{L_2^j} f(z) dz \implies \left(\int_{L_1^j} + \int_{-L_2^j} \right) f(z) dz = \left(\int_{L_1^j} - \int_{L_2^j} \right) f(z) dz.$$

Ao se reduzir a distância de cada par de segmentos de reta L_1^j e L_2^j assintoticamente a zero, as integrais de linha percorrem o mesmo caminho sobre o plano complexo, resultando então que

$$\left(\int_{L_1^j} - \int_{L_2^j} \right) f(z) dz \xrightarrow{L_2^j \rightarrow L_1^j} 0.$$

Neste caso, os caminhos C, C_1, \dots, C_n se tornam contornos fechados simples, resultando que

$$\oint_C f(z) dz + \sum_{j=1}^n \oint_{-C_j} f(z) dz = 0,$$

onde se deve notar o símbolo \oint . Como os contornos C_j são percorridos no sentido horário, de acordo com a figura (2.13)a, resulta que

$$\oint_{-C_j} = - \oint_{C_j},$$

onde na segunda integral o contorno é agora percorrido no sentido anti-horário, e a situação se torna semelhante à ilustrada pela figura (2.13)b. Portanto,

$$\oint_C f(z) dz - \sum_{j=1}^n \oint_{C_j} f(z) dz = 0,$$

de onde resulta o teorema. □

Nas situações em que a identidade (2.21) é válida, é comum afirmar-se que o contorno C foi **deformado** nos contornos C_1, C_2, \dots, C_n .

Teorema 2.13 (Independência do caminho). Se a função $f(z)$ é analítica em uma região simplesmente conexa R , então dados dois pontos z_0 e z quaisquer, contidos em R , a integral

$$\int_{z_0}^z f(z) dz$$

independe do caminho ligando os pontos z_0 e z , desde que este caminho esteja totalmente contido em R .

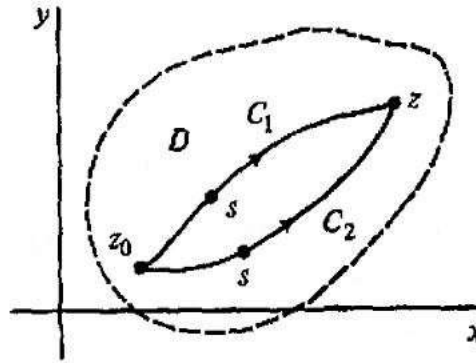


Figura 2.14: Independência nos caminhos C_1 e C_2 .

Demonstração. A situação está ilustrada na figura 2.14. Sendo C_1 e C_2 dois caminhos quaisquer, contidos em R e que ligam os pontos z_0 e z , então, de acordo com o teorema de Cauchy (2.19),

$$\int_{-C_1} f(z) dz + \int_{C_2} f(z) dz = 0, \quad \text{mas} \quad \int_{-C_1} f(z) dz = - \int_{C_1} f(z) dz.$$

Portanto

$$\int_{C_1} f(z) dz = \int_{C_2} f(z) dz,$$

o que demonstra o teorema. □

Teorema 2.14 (Teorema de analiticidade). *Seja $f(z)$ uma função contínua em uma região simplesmente conexa R e sejam z_0 e z dois pontos contidos em R , os quais são conectados por um caminho C , também contido em R . Então, se $\int_{z_0}^z f(s) ds$ independe de C ,*

$$F(z) = \int_{z_0}^z f(s) ds \text{ é analítica em } R \text{ e } F'(z) = f(z).$$

Demonstração. Sendo $z + \Delta z$ qualquer ponto contido em R e distinto de z , mas que esteja em uma vizinhança de z . Então

$$F(z + \Delta z) - F(z) = \int_{z_0}^{z+\Delta z} f(s) ds - \int_{z_0}^z f(s) ds.$$

Pela propriedade (4) da seção 2.5.2, resulta que

$$F(z + \Delta z) - F(z) = \int_z^{z+\Delta z} f(s) ds.$$

Dividindo ambos os lados por Δz , pode-se escrever

$$\frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) = \frac{1}{\Delta z} \int_z^{z+\Delta z} f(s) ds - f(z) = \frac{1}{\Delta z} \int_z^{z+\Delta z} [f(s) - f(z)] ds.$$

Como a função $f(z)$ é contínua em R , então para um número positivo ϵ , deve existir um outro número positivo δ tal que

$$|f(s) - f(z)| < \epsilon$$

sempre que $|s - z| < \delta$. Desta forma, se $z + \Delta z$ é próximo o suficiente de z de tal forma que $|\Delta z| < \delta$, então

$$\left| \int_z^{z+\Delta z} [f(s) - f(z)] ds \right| < \epsilon |\Delta z|$$

e, portanto,

$$\left| \frac{F(z + \Delta z) - F(z)}{\Delta z} - f(z) \right| < \frac{1}{|\Delta z|} \epsilon |\Delta z| = \epsilon.$$

No limite em que $\Delta z \rightarrow 0$, $\epsilon \rightarrow 0$, resultando que

$$\lim_{\Delta z \rightarrow 0} \frac{F(z + \Delta z) - F(z)}{\Delta z} \equiv F'(z) = f(z).$$

Portanto, a derivada de $F(z)$ existe em todos os pontos z pertencentes a R . Como consequência, $F(z)$ é analítica em R e sua derivada é igual a $f(z)$, demonstrando o teorema. \square

Teorema 2.15 (Teorema de Morera). *Se uma função $f(z)$ é contínua em uma região simplesmente conexa R e $\oint_C f(z) dz = 0$ para todo contorno simples fechado C no interior de R , então $f(z)$ é analítica em R .*

Observação. O teorema de Morera é a recíproca do teorema de Cauchy.

Exemplo 2.10 (Cálculo de integrais no plano complexo).

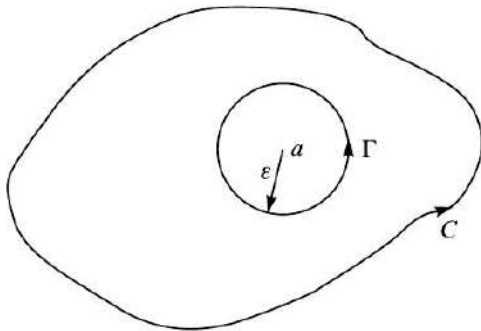


Figura 2.15: Exemplo de cálculo de integrais sobre o plano complexo.

Calcule $\oint_C dz/(z-a)$, onde C é um contorno fechado simples qualquer, quando $z = a$ está (a) fora de C e (b) dentro de C .

Solução.

(a) Se $z = a$ está fora de C , então $f(z) = 1/(z-a)$ é analítica em todos os pontos internos e ao longo de C . Portanto, pelo teorema de Cauchy (2.20),

$$\oint_C \frac{dz}{z-a} = 0.$$

(b) Se $z = a$ está dentro de C e Γ é uma circunferência de raio ϵ centrada em $z = a$, de tal forma que Γ está totalmente contido em C (figura 2.15), então pelo teorema (2.21),

$$\oint_C \frac{dz}{z-a} = \oint_{\Gamma} \frac{dz}{z-a}.$$

Agora, o contorno Γ é dado por todos os pontos z tais $|z-a| = \epsilon$. Pode-se descrever o contorno na figura 2.15 através do parâmetro θ tal que

$$z - a = \epsilon e^{i\theta}, \text{ quando } 0 \leq \theta \leq 2\pi.$$

Então $dz = i\epsilon e^{i\theta} d\theta$ e

$$\oint_{\Gamma} \frac{dz}{z-a} = \int_0^{2\pi} \frac{i\epsilon e^{i\theta} d\theta}{\epsilon e^{i\theta}} = i \int_0^{2\pi} d\theta = 2\pi i.$$

2.7 FÓRMULAS INTEGRAIS DE CAUCHY

Uma das consequências mais importantes do teorema integral de Cauchy são as fórmulas integrais que também levam o seu nome. Pode-se introduzir estas fórmulas através do seguinte teorema.

Teorema 2.16 (Fórmula integral de Cauchy). *Seja $f(z)$ uma função analítica em uma região simplesmente conexa R e z_0 é um ponto qualquer no interior de R , a qual é delimitada pelo contorno simples C , então*

$$f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z-z_0} dz, \quad (2.22)$$

sendo a integração em (2.22) realizada ao longo de C no sentido positivo (anti-horário).

Demonstração. Para provar o teorema (2.22), toma-se uma circunferência Γ , centrada em z_0 e com raio r , como ilustrado na figura 2.16. Então, de acordo com o teorema (2.21),

$$\oint_C \frac{f(z)}{z-z_0} dz = \oint_{\Gamma} \frac{f(z)}{z-z_0} dz.$$

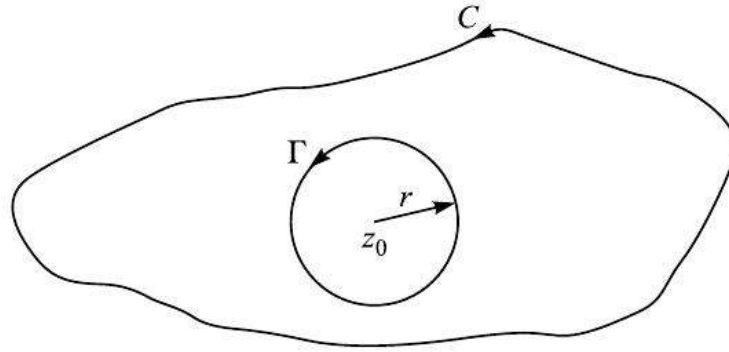


Figura 2.16: Fórmula integral de Cauchy.

Agora, a circunferência Γ é descrita por $|z - z_0| = r$, ou seja, usando o parâmetro θ , $z - z_0 = re^{i\theta}$, sendo $(0 \leq \theta \leq 2\pi)$ e $dz = ire^{i\theta}d\theta$. Então

$$\oint_{\Gamma} \frac{f(z)}{z - z_0} dz = \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{re^{i\theta}} ire^{i\theta} d\theta = i \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta.$$

Tomando-se agora o limite $r \rightarrow 0$ em ambos os lados e fazendo uso da continuidade de $f(z)$, resulta

$$\oint_C \frac{f(z)}{z - z_0} dz = \lim_{r \rightarrow 0} i \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta = i \int_0^{2\pi} \lim_{r \rightarrow 0} f(z_0 + re^{i\theta}) d\theta = i \int_0^{2\pi} f(z_0) d\theta = 2\pi i f(z_0).$$

Ou seja,

$$f(z_0) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{z - z_0} dz.$$

□

Uma forma conveniente para a fórmula integral (2.22) é escrevê-la como

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(z')}{z' - z} dz', \tag{2.23}$$

para enfatizar o fato que z pode ser um ponto qualquer em C .

Exemplo 2.11. Calcule a integral $\oint_C e^z dz / (z^2 + 1)$, sendo C a circunferência de raio unitário e com centro em: (a) $z = i$ e (b) $z = -i$.

Solução.

(a) Escrevendo-se a integral na forma

$$\oint_C \left(\frac{e^z}{z + i} \right) \frac{dz}{z - i},$$

percebe-se que $f(z) = e^z / (z + i)$ é analítica dentro e sobre a circunferência de raio unitário centrada em $z_0 = i$. Portanto, pela fórmula integral de Cauchy (2.22), temos

$$\oint_C \left(\frac{e^z}{z + i} \right) \frac{dz}{z - i} = 2\pi i f(i) = 2\pi i \frac{e^i}{2i} = \pi e^i = \pi (\cos 1 + i \operatorname{sen} 1).$$

(b) Para a circunferência centrada em $z_0 = -i$, define-se $f(z) = e^z / (z - i)$, a qual é novamente analítica dentro e sobre C . Então, usando novamente (2.22) resulta

$$\oint_C \left(\frac{e^z}{z - i} \right) \frac{dz}{z + i} = 2\pi i f(-i) = 2\pi i \frac{e^{-i}}{-2i} = -\pi e^{-i} = -\pi (\cos 1 - i \operatorname{sen} 1).$$

A fórmula integral de Cauchy pode ser generalizada para derivadas de ordem mais alta de $f(z)$. Isto é descrito pelo teorema a seguir.

Teorema 2.17 (Fórmula integral de Cauchy para derivadas de ordens mais altas). Seja $f(z)$ uma função analítica em uma região simplesmente conexa R e z_0 é um ponto qualquer no interior de R , a qual é delimitada pelo contorno simples C , então

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_C \frac{f(z) dz}{(z - z_0)^{n+1}} \quad (2.24)$$

e as derivadas de $f(z)$ em z_0 existem em todas as ordens.

Demonstração. Uma prova simples, porém incompleta, do teorema (2.24) pode ser feita por intermédio da indução matemática. Derivando-se ambos os lados de (2.23) em relação a z , obtém-se

$$f'(z) = \frac{1!}{2\pi i} \oint_C \frac{f(s)}{(s - z)^2} ds.$$

Derivando-se novamente,

$$f''(z) = \frac{2!}{2\pi i} \oint_C \frac{f(s)}{(s - z)^3} ds.$$

E assim sucessivamente, resultando, para a derivada de ordem n , na fórmula (2.24). A demonstração completa desta fórmula integral pode ser obtida na bibliografia citada. \square

A fórmula (2.24), obtida para uma região simplesmente conexa, pode ser estendida para o caso onde o contorno simples C é substituído pelo contorno B da figura 2.13, composto por um contorno exterior C e por um conjunto $\{C_i\}$ de contornos interiores. Para tanto, basta assumir que em (2.24) o ponto z_0 pertence ao domínio definido por B e que $f(z)$ é analítica neste domínio. Desta forma, a fórmula integral de Cauchy pode ser estendida a regiões multiplamente conexas.

Exemplo 2.12. Calcule

$$\oint_C \frac{e^{2z}}{(z + 1)^4} dz,$$

sendo C um contorno simples que não passa por $z = -1$. Considere 2 casos: (a) C não envolve $z = -1$ e (b) C envolve $z = -1$.

Solução.

(a) Neste caso, a função $f(z) = e^{2z}/(z + 1)^4$ é analítica dentro e sobre C . Portanto, pelo teorema de Cauchy,

$$\oint_C \frac{e^{2z}}{(z + 1)^4} dz = 0.$$

(b) Chamando agora $f(z) = e^{2z}$, esta função é analítica dentro e sobre C . Portanto, de acordo com o teorema (2.24),

$$\oint_C \frac{e^{2z} dz}{(z + 1)^4} = \frac{2\pi i}{3!} f^{(3)}(-1). \text{ Como } f^{(3)}(-1) = 8e^{-2}, \text{ resulta } \oint_C \frac{e^{2z} dz}{(z + 1)^4} = \frac{8\pi}{3} e^{-2} i.$$

2.8 REPRESENTAÇÃO EM SÉRIES DE FUNÇÕES ANALÍTICAS

Será apresentado agora um ponto muito importante: a representação em séries de funções analíticas. Inicialmente será discutida a noção de convergência de uma série complexa. Grande parte das definições e teoremas válidos para séries de termos reais podem ser aplicadas às séries complexas com uma pequena ou nenhuma modificação.

2.8.1 SÉRIES COMPLEXAS

Nesta seção, serão consideradas séries complexas em geral, cujos termos são funções complexas,

$$f_1(z) + f_2(z) + f_3(z) + \dots + f_n(z) + \dots \quad (2.25)$$

A soma dos n primeiros termos desta série,

$$S_n(z) = f_1(z) + f_2(z) + f_3(z) + \cdots + f_n(z) = \sum_{i=1}^n f_i(z),$$

é denominada a n -ésima soma parcial da série (2.25). A soma dos termos restantes, partindo do $(n+1)$ -ésimo termo, é denominada de **resto da série**.

2.8.1.1 CONVERGÊNCIA DA SÉRIE

Associa-se agora à série (2.25) a sequência de suas somas parciais S_1, S_2, \dots, S_n . Se esta sequência de somas parciais é convergente, então a série converge; se a sequência é divergente, então a série também diverge. De uma maneira mais formal, a série (2.25) é dita convergente à soma $S(z)$ em uma região R do plano complexo se para qualquer $\epsilon > 0$ existe um inteiro N , dependente de ϵ e do valor de z sob consideração, tal que

$$|S_n(z) - S(z)| < \epsilon \text{ para todo } n > N.$$

Neste caso, escreve-se

$$S(z) = \lim_{n \rightarrow \infty} S_n(z) \equiv \sum_{n=1}^{\infty} f_n(z). \quad (2.26)$$

A diferença $S_n(z) - S(z)$ é o resto $R_n(z)$. Assim, a definição de convergência da série (2.25) demanda que

$$\lim_{n \rightarrow \infty} |R_n(z)| = 0.$$

Teorema 2.18 (Teoremas de convergência). *Os seguintes teoremas de convergência são válidos:*

1. Uma condição necessária, mas não suficiente, para que $S_n(z)$ convirja no limite $n \rightarrow \infty$ é

$$\lim_{n \rightarrow \infty} f_n(z) = 0.$$

2. A multiplicação de cada termo de uma série por uma constante não nula não afeta a convergência, assim como a remoção ou adição de um número finito de termos.
3. Uma condição necessária e suficiente para que a série de termos complexos

$$\sum_{n=1}^{\infty} f_n(z) = f_1(z) + f_2(z) + f_3(z) + \cdots + f_n(z) + \cdots$$

seja convergente é que as séries das respectivas partes reais e imaginárias dos termos $f_n(z)$ sejam convergentes. Além disso, se

$$\sum_{n=1}^{\infty} \operatorname{Re} f_n \quad \text{e} \quad \sum_{n=1}^{\infty} \operatorname{Im} f_n$$

convergem às respectivas funções $R(z)$ e $I(z)$, então a série complexa converge para $S(z) = R(z) + iI(z)$.

2.8.1.2 CONVERGÊNCIA ABSOLUTA

Se os valores absolutos dos termos em (2.25),

$$\sum_{n=1}^{\infty} |f_n(z)| = |f_1(z)| + |f_2(z)| + |f_3(z)| + \cdots + |f_n(z)| + \cdots,$$

formam uma série convergente, então a série (2.25) é dita **absolutamente convergente**.

Se a série (2.25) converge, mas não é absolutamente convergente, então esta é dita **condicionalmente convergente**. A partir da definição de convergência, pode-se demonstrar os teoremas a seguir.

Teorema 2.19 (Teoremas de convergência absoluta). *Os seguintes teoremas são válidos.*

1. Se $\sum_{n=1}^{\infty} |f_n(z)|$ converge, então $\sum_{n=1}^{\infty} f_n(z)$ também converge (*condição suficiente*).
2. A soma, diferença ou o produto de séries absolutamente convergentes é convergente.

2.8.1.3 CONVERGÊNCIA UNIFORME

Seja

$$R_n(z) = f_{n+1}(z) + f_{n+2}(z) + \dots = S(z) - S_n(z)$$

o **resto da série** $S(z)$, dada por (2.26), sobre uma região R . A série $S(z)$ é dita **uniformemente convergente** em R se, dado um número real positivo ϵ , é possível encontrar um número inteiro positivo N , tal que para todo $z \in R$,

$$|R_n(z)| = |S(z) - S_n(z)| < \epsilon, \text{ para todo } n > N.$$

2.8.2 TESTES DE CONVERGÊNCIA

Os testes de convergência determinam uma condição necessária e suficiente para a convergência de uma determinada série. Há vários testes, dentre os quais os principais serão citados abaixo sem as demonstrações, as quais podem ser obtidas na bibliografia recomendada.

2.8.2.1 TESTES DE CONVERGÊNCIA ABSOLUTA

Para testar a convergência absoluta de uma série, os seguintes testes existem:

Teste da comparação. Se $\sum_{n=1}^{\infty} g_n(z)$ converge absolutamente em R e $|f_n(z)| \leq |g_n(z)|, \forall z \in R$, então $\sum_{n=1}^{\infty} f_n(z)$ também converge absolutamente.

Teste da razão. Dentre todos os testes de convergência, o mais útil é o teste da razão, o qual se aplica a séries complexas, além de séries reais. Dada a série $\sum_{n=1}^{\infty} f_n(z)$, esta converge absolutamente na região R se

$$0 < |r(z)| = \lim_{n \rightarrow \infty} \left| \frac{f_{n+1}(z)}{f_n(z)} \right| < 1 \quad (2.27)$$

e diverge se $|r(z)| > 1$. Quando $|r(z)| = 1$, este teste não fornece informação conclusiva a respeito da convergência da série.

Teste da raiz. Dada a série $\sum_{n=1}^{\infty} f_n(z)$, esta converge absolutamente na região R se

$$0 < |r(z)| = \lim_{n \rightarrow \infty} \left| \sqrt[n]{|f_n(z)|} \right| < 1$$

e diverge se $|r(z)| > 1$. Quando $|r(z)| = 1$, este teste não fornece informação conclusiva a respeito da convergência da série.

Exemplo 2.13 (Teste da razão). Mostre que a série complexa

$$S(z) = \sum_{n=0}^{\infty} (2^{-n} + ie^{-n})$$

converge.

Solução. Pode-se aplicar o teste da razão separadamente para as partes real e imaginária:

$$\lim_{n \rightarrow \infty} \left| \frac{2^{-n-1}}{2^{-n}} \right| = \frac{1}{2} < 1 \quad \text{e} \quad \lim_{n \rightarrow \infty} \left| \frac{e^{-n-1}}{e^{-n}} \right| = \frac{1}{e} < 1.$$

Portanto, a série converge absolutamente.

2.8.2.2 TESTE DE CONVERGÊNCIA UNIFORME

O seguinte teste verifica a convergência uniforme de uma série.

Teste de Weierstrass. Se $|f_n(z)| \leq M_n$, sendo M_n independente de z em uma região R e $\sum_{n=1}^{\infty} M_n$ converge, então $\sum_{n=1}^{\infty} f_n(z)$ converge uniformemente em R .

2.8.3 SÉRIES DE POTÊNCIAS E SÉRIES DE TAYLOR

Séries de potências constituem-se em uma das mais importantes ferramentas da análise matemática em geral e da teoria de funções complexas em particular. Isto porque séries de potências com raios de convergência não nulos podem sempre representar funções analíticas. Como exemplo, a série

$$S_1(z) = \sum_{n=0}^{\infty} a_n z^n \quad (2.28)$$

claramente define uma função analítica, desde que a série convirja.

Nesta seção, o interesse estará restrito às séries que apresentam convergência absoluta. Neste caso, o teste da razão (2.27) demanda que

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} z^{n+1}}{a_n z^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| |z| < 1 \implies \frac{1}{|z|} > \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|.$$

O resultado acima mostra que a série irá convergir absolutamente para todos os pontos z cujos módulos satisfazem a condição $|z| < R$, sendo R o **raio de convergência da série**, definido por

$$\frac{1}{R} = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|. \quad (2.29)$$

A série (2.28) está centrada na origem do plano complexo; assim, o raio de convergência R em (2.29) define uma circunferência centrada na origem. De forma similar, a série

$$S_2(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (2.30)$$

converge para todos os pontos z dentro da circunferência de raio R centrada em z_0 .

A série de potências mais importante na análise matemática é a série de Taylor. Na análise complexa é possível realizar-se uma expansão de Taylor para toda e qualquer função analítica. Esta propriedade é devida ao homônimo *teorema de Taylor*,³ exposto a seguir.

Teorema 2.20 (Teorema de Taylor). *Seja $f(z)$ uma função analítica sobre a região R , delimitada pela circunferência C centrada em a e de raio R_0 . Se z é um ponto interior a C , então $f(z)$ pode ser desenvolvida em uma série de potências centrada em $z = a$,*

$$f(z) = f(a) + f'(a)(z - a) + \frac{f''(a)}{2!}(z - a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(z - a)^n + \dots = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(z - a)^n, \quad (2.31)$$

a qual converge para $f(z)$ quando $|z - z_0| < R_0$. Esta série é denominada **Série de Taylor**.

Demonstração. Seja z qualquer ponto interior à circunferência C_0 . Portanto, se $|z - a| = r$, então $r < R_0$. Seja agora uma outra circunferência C_1 , centrada em a e de raio r_1 , tal que $r < r_1 < R_0$. Uma vez que z está dentro de C_1 e $f(z)$ é analítica no interior e sobre C_1 , a fórmula integral de Cauchy (2.23) é válida, a qual é escrita da seguinte maneira:

$$f(z) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(w) dw}{w - z} = \frac{1}{2\pi i} \oint_{C_1} \frac{f(w)}{(w - a)} \left[\frac{1}{1 - (z - a)/(w - a)} \right] dw.$$

Nota-se que, na integral acima, como w está sempre ao longo do contorno C_1 e z é um ponto interior a C_1 , então

$$\left| \frac{z - a}{w - a} \right| < 1, \quad \forall w.$$

Agora, a partir da fórmula de progressão geométrica

$$1 + q + q^2 + \dots + q^n = \frac{1 - q^{n+1}}{1 - q} = \frac{1}{1 - q} - \frac{q^{n+1}}{1 - q},$$

válida para $q \neq 1$, obtém-se

$$\frac{1}{1 - q} = 1 + q + q^2 + \dots + q^n + \frac{q^{n+1}}{1 - q}.$$

³Devido ao matemático inglês Brook Taylor (1685 - 1731).

Identificando-se então $q = (z - a) / (w - a)$, e inserindo a progressão geométrica resultante na fórmula integral acima, resulta

$$f(z) = \frac{1}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)} + \frac{(z-a)}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^2} + \frac{(z-a)^2}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^3} \\ + \dots + \frac{(z-a)^n}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^{n+1}} + \frac{(z-a)^{n+1}}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^{n+1}(w-z)}.$$

Contudo, nesta situação as fórmulas integrais de Cauchy (2.24) garantem que

$$\frac{1}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^{n+1}} = \frac{f^{(n)}(a)}{n!};$$

portanto, pode-se escrever a expressão para $f(z)$ como

$$f(z) = f(a) + f'(a)(z-a) + \frac{f''(a)}{2!}(z-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(z-a)^n + R_n(z),$$

sendo

$$R_n(z) = \frac{(z-a)^{n+1}}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^{n+1}(w-z)}$$

o resto do desenvolvimento de $f(z)$ na forma de um polinômio de grau n . Porém, uma vez que $|z-a| = r$ e $|w-a| = r_1$ em C_1 , nota-se que

$$|w-z| = |(w-a) - (z-a)| \geq |w-a| - |z-a| = r_1 - r > 0.$$

Em consequência, se $M > 0$ é o módulo do maior valor de $f(z)$ ao longo de C_1 , pode-se estabelecer um limite superior para $|R_n(z)|$, dado por

$$|R_n(z)| = \left| \frac{(z-a)^{n+1}}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w-a)^{n+1}(w-z)} \right| \leq \frac{r^{n+1}M}{r_1^n(r_1-r)} = \frac{Mr}{r_1-r} \left(\frac{r}{r_1}\right)^n.$$

Portanto, como $r/r_1 < 1$, se $n \rightarrow \infty$, o resto da série resultante tenderá a zero,

$$\lim_{n \rightarrow \infty} R_n(z) = 0,$$

e a série converge uniformemente para $f(z)$, o que demonstra o teorema de Taylor. \square

No caso particular onde $a = 0$, a série de Taylor (2.31) para $f(z)$ é denominada **série de Maclaurin**.⁴

Com as fórmulas integrais de Cauchy (2.24) e a série de Taylor (2.31), ficam estabelecidas duas propriedades fundamentais das funções analíticas:

1. Elas possuem derivadas de todas as ordens.
2. Elas sempre podem ser representadas por uma série de Taylor.

O mesmo não pode ser dito sobre as funções reais; existem funções reais que possuem derivadas de todas as ordens, mas que não podem ser representadas por uma série de Taylor.

As principais propriedades de séries de potências podem ser resumidas nos seguintes teoremas.

Teorema 2.21. *Os seguintes teoremas sobre séries de potências são válidos:*

1. Uma série de potências converge uniformemente e absolutamente em qualquer região interior ao seu raio de convergência.
2. Uma série de potências pode ser diferenciada termo a termo em qualquer ponto interno ao seu raio de convergência.
3. Uma série de potências pode ser integrada termo a termo ao longo de qualquer curva C , desde que C esteja contida dentro do seu círculo de convergência.
4. Uma série de potências representa uma função analítica em cada ponto de seu círculo de convergência.

⁴Colin Maclaurin (1698 – 1746), matemático escocês.

2.8.4 SÉRIES DE TAYLOR DE FUNÇÕES ELEMENTARES

Séries de Taylor de funções analíticas são similares às series de funções reais. Em geral, basta trocar a variável independente real pela variável complexa z para que as séries reais sejam *continuadas*⁵ para o plano complexo. A lista a seguir mostra as séries de Taylor de algumas funções elementares; no caso de funções plurívocas, a série apresentada representa o ramo principal.

$$\frac{1}{1+z} = \sum_{n=0}^{\infty} (-1)^n z^n = 1 - z + z^2 - \dots, \quad |z| < 1 \quad (2.32a)$$

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \dots, \quad |z| < \infty \quad (2.32b)$$

$$\operatorname{sen} z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots, \quad |z| < \infty \quad (2.32c)$$

$$\operatorname{cos} z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots, \quad |z| < \infty \quad (2.32d)$$

$$\operatorname{senh} z = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots, \quad |z| < \infty \quad (2.32e)$$

$$\operatorname{cosh} z = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots, \quad |z| < \infty \quad (2.32f)$$

$$\ln(1+z) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1} z^n}{n} = z - \frac{z^2}{2} + \frac{z^3}{3} - \dots, \quad |z| < 1. \quad (2.32g)$$

2.8.5 SÉRIES DE LAURENT

Em muitas aplicações, é comum deparar-se com funções que não são analíticas sobre todo o plano complexo.⁶ Tipicamente, estas funções não são analíticas em um ou vários pontos ou até em uma região do plano. Por consequência, séries de Taylor não podem ser empregadas nas vizinhanças destes pontos. Contudo, em muitos casos, uma representação em série contendo potências tanto positivas quanto negativas pode ainda ser obtida, a qual é válida nesta vizinhança. Esta série é denominada **série de Laurent** e é válida para aquelas funções que são analíticas dentro e sobre um anel caracterizado pelos raios R_1 e R_2 ($R_1 < R_2$); isto é, a função é analítica nos pontos

$$R_1 \leq |z - z_0| \leq R_2,$$

sendo $z = a$ um dos pontos singulares da função. Esta situação está representada na figura 2.17a. A série de Laurent foi obtida pela primeira vez pelo matemático francês Pierre Alphonse Laurent (1813 - 1854).

Teorema 2.22 (Série de Laurent). *Seja $f(z)$ uma função analítica ao longo dos contornos circulares concêntricos C_1 e C_2 , de raios R_1 e R_2 ($R_1 < R_2$), respectivamente, ambos centrados em $z = z_0$, bem como na região anelar delimitada por C_1 e C_2 . Então em cada ponto z nesta região, a função $f(z)$ pode ser representada pela série*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}, \quad (2.33a)$$

onde

$$a_n = \frac{1}{2\pi i} \oint_{C_2} \frac{f(z) dz}{(z - z_0)^{n+1}} \quad (2.33b)$$

$$b_n = \frac{1}{2\pi i} \oint_{C_1} \frac{f(z)}{(z - z_0)^{-n+1}} dz, \quad (2.33c)$$

⁵Ver seção 2.10.

⁶Isto é, não são inteiras.

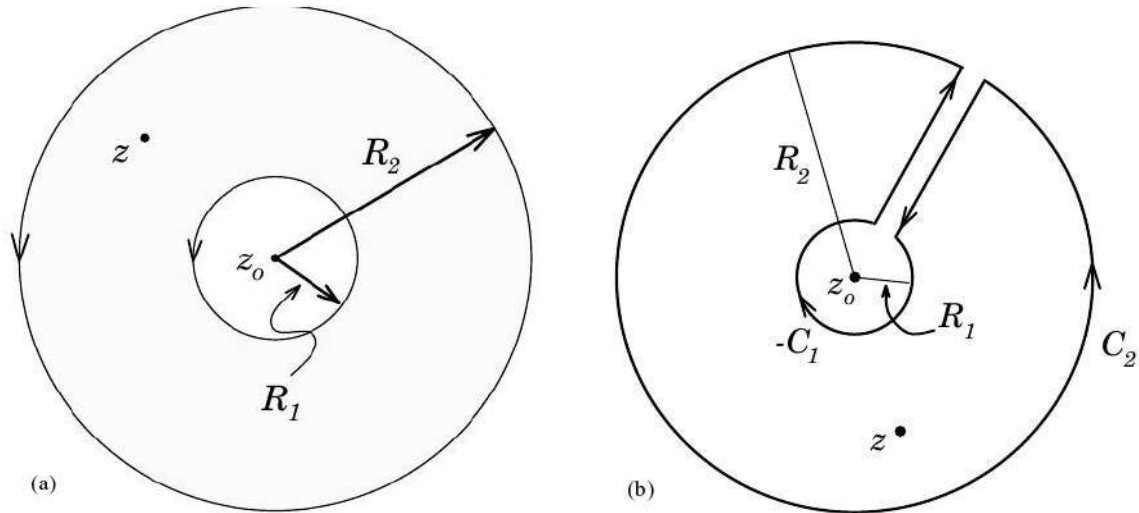


Figura 2.17: (a) Anel $R_1 \leq |z - z_0| \leq R_2$ que representa a região de analiticidade da função. (b) Contornos de integração interior (C_1) e exterior (C_2), utilizados para a derivação da série de Laurent.

sendo cada contorno de integração realizado no sentido anti-horário.

Demonstração. Realiza-se a integração ao longo do contorno mostrado na figura 2.17b. Como $f(z)$ é analítica ao longo e no interior do contorno e z é um ponto que pertence a esta região, a fórmula integral de Cauchy (2.22) pode ser utilizada, resultando em

$$f(z) = \frac{1}{2\pi i} \oint_{C_2} \frac{f(w) dw}{w - z} - \frac{1}{2\pi i} \oint_{C_1} \frac{f(w) dw}{w - z}.$$

Na primeira integral, escreve-se

$$\frac{1}{w - z} = \frac{1}{w - z_0} \frac{1}{1 - \frac{(z - z_0)}{(w - z_0)}} = \frac{1}{w - z_0} \sum_{j=0}^{\infty} \left(\frac{z - z_0}{w - z_0} \right)^j,$$

sendo que a última identidade é válida porque $|z - z_0| < |w - z_0| = R_2$, para todo w ao longo de C_2 . Já na segunda integral, escreve-se

$$-\frac{1}{w - z} = \frac{1}{z - z_0 - (w - z_0)} = \frac{1}{z - z_0} \frac{1}{1 - \frac{(w - z_0)}{(z - z_0)}} = \frac{1}{z - z_0} \sum_{j=0}^{\infty} \left(\frac{w - z_0}{z - z_0} \right)^j,$$

sendo que agora a última identidade é válida porque $|w - z_0| = R_1 < |z - z_0|$. Então, pode-se escrever $f(z)$ como

$$\begin{aligned} f(z) &= \sum_{j=0}^{\infty} \left[\frac{1}{2\pi i} \oint_{C_2} \frac{f(w) dw}{(w - z_0)^{j+1}} \right] (z - z_0)^j + \sum_{j=0}^{\infty} \left[\frac{1}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w - z_0)^{-j}} \right] \frac{1}{(z - z_0)^{j+1}} \\ &= \sum_{j=0}^{\infty} \left[\frac{1}{2\pi i} \oint_{C_2} \frac{f(w) dw}{(w - z_0)^{j+1}} \right] (z - z_0)^j + \sum_{j=1}^{\infty} \left[\frac{1}{2\pi i} \oint_{C_1} \frac{f(w) dw}{(w - z_0)^{-j+1}} \right] \frac{1}{(z - z_0)^j}, \end{aligned}$$

de onde resulta (2.33). □

Em (2.33a,b), a série com os coeficientes $\{a_n\}$ é denominada a **parte analítica**, ao passo que em (2.33a,c) a série com os coeficientes $\{b_n\}$ é denominada a **parte principal** da série de Laurent. Se a parte principal for nula, a série de Laurent se reduz à série de Taylor (2.31).

Uma vez que as funções $f(z)/(z - z_0)^{n+1}$ e $f(z)/(z - z_0)^{-n+1}$ são analíticas sobre toda a região R na figura 2.17b, qualquer contorno simples fechado C contido dentro deste anel pode ser usado como caminho de integração, desde que percorrido no sentido positivo, no lugar dos contornos circulares C_1 e C_2 . Assim, a série de Laurent (2.33) pode ser generalizada como

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n, \quad (R_1 < |z - z_0| < R_2), \quad (2.34a)$$

sendo

$$c_n = \frac{1}{2\pi i} \oint_C \frac{f(z) dz}{(z - z_0)^{n+1}}, \quad (n = 0, \pm 1, \pm 2, \dots). \quad (2.34b)$$

2.8.6 TEOREMAS DE EXISTÊNCIA E UNICIDADE

A seguir serão apresentados, sem demonstração, dois teoremas de existência e unicidade das séries de Laurent. Demonstrações destes teoremas podem ser encontradas na bibliografia citada.

Teorema 2.23 (Teorema de existência). A série de Laurent (2.34a,b) de uma função $f(z)$, analítica na região anelar $R_1 \leq |z - z_0| \leq R_2$, converge uniformemente para $f(z)$ para $\rho_1 \leq |z - z_0| \leq \rho_2$, sendo $R_1 < \rho_1$ e $R_2 > \rho_2$.

Teorema 2.24 (Teorema de unicidade). Dada uma função analítica $f(z)$, se esta pode ser representada pela série uniformemente convergente

$$f(z) = \sum_{n \rightarrow -\infty}^{\infty} b_n (z - z_0)^n$$

na região anelar $R_1 \leq |z - z_0| \leq R_2$, então $b_n = c_n$, para $n = 0, \pm 1, \pm 2, \dots$, sendo o coeficiente c_n dado por (2.34b).

2.8.7 ALGUMAS TÉCNICAS DE CONSTRUÇÃO DE SÉRIES DE TAYLOR E LAURENT

Nos exemplos a seguir, serão ilustradas algumas das técnicas mais comuns para a construção das séries de Taylor e Laurent. Em muitas situações, não ocorre a necessidade de se calcular explicitamente os coeficientes $\{c_n\}$ da série, pois o teorema de unicidade acima garante que qualquer série que represente uma função $f(z)$ é, de fato, a única série de Laurent para a mesma. Usualmente, para se obter a série de Laurent de uma dada função $f(z)$, basta fazer uso de expansões de Taylor de funções conhecidas e realizar substituições de forma apropriada.

Exemplo 2.14 (Uso de séries geométricas). Seja

$$f(z) = \frac{1}{z - a}.$$

Sabendo-se que

$$1 + z + z^2 + z^3 + \dots = \sum_{n=0}^{\infty} z^n = \frac{1}{1 - z}, \quad (|z| < 1),$$

pode-se escrever, para $|z| < |a|$:

$$f(z) = \frac{1}{z - a} = -\frac{1}{a} \frac{1}{1 - z/a} = -\frac{1}{a} \sum_{n=0}^{\infty} \left(\frac{z}{a}\right)^n, \quad (|z| < |a|).$$

Esta é a série de Taylor de $f(z)$ em torno de $z = 0$. Seu raio de convergência é $R = |a|$, porque a uma distância R da origem existe o ponto $z = a$, onde $f(z)$ não é analítica. Este é o único ponto onde $f(z)$ não é analítica.

Portanto, $f(z)$ deve possuir uma série de Laurent em torno de $z = 0$ válida para $|z| > |a|$. Escrevendo-se

$$f(z) = \frac{1}{z - a} = \frac{1}{z} \frac{1}{1 - a/z},$$

se $|z| > |a|$, $|a/z| < 1$ e é possível desenvolver:

$$\frac{1}{1 - a/z} = \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n.$$

Portanto,

$$f(z) = \frac{1}{z-a} = \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{a}{z}\right)^n = \sum_{n=0}^{\infty} \frac{a^n}{z^{n+1}}, \quad (|z| > |a|).$$

Esta é a série de Laurent desejada.

A função $f(z)$ pode ser expandida por este método em torno de qualquer ponto $z = b$:

$$f(z) = \frac{1}{z-a} = \frac{1}{(z-b) - (a-b)} \xrightarrow{w=z-b} \frac{1}{w - (a-b)}, \quad (b \neq a).$$

Então,

$$f(z) = -\frac{1}{a-b} \sum_{n=0}^{\infty} \frac{w^n}{(a-b)^n} = -\frac{1}{a-b} \sum_{n=0}^{\infty} \frac{(z-b)^n}{(a-b)^n}, \quad (|z-b| < |a-b|)$$

ou

$$f(z) = \sum_{n=0}^{\infty} \frac{(a-b)^n}{(z-b)^{n+1}}, \quad (|z-b| > |a-b|).$$

Exemplo 2.15 (Decomposição em frações racionais). Seja

$$f(z) = \frac{1}{z^2 - (2+i)z + 2i}.$$

Esta função não é analítica nos pontos $z = i$ e $z = 2$; portanto, ela deve possuir uma série de Taylor em torno de $z = 0$, válida para $|z| < 1$ e duas séries de Laurent em torno de $z = 0$, válidas para $1 < |z| < 2$ e $|z| > 2$, respectivamente. Para se obter estas três séries, usa-se a identidade:

$$f(z) = \frac{1}{z^2 - (2+i)z + 2i} = \frac{1}{(z-i)(z-2)} = \frac{1}{2-i} \left(\frac{1}{z-2} - \frac{1}{z-i} \right).$$

Para $|z| < 1$. Neste caso, pode-se usar diretamente a série geométrica:

$$\frac{1}{z-2} = -\frac{1}{2} \frac{1}{1-z/2} = -\frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n, \quad (|z| < 2)$$

$$\frac{1}{z-i} = i \frac{1}{1+iz} = i \sum_{n=0}^{\infty} (-iz)^n, \quad (|z| < 1).$$

Subtraindo as séries, pode-se colocar em evidência fatores proporcionais à mesma potência de z ,

$$\frac{1}{z-2} - \frac{1}{z-i} = -\frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n - i \sum_{n=0}^{\infty} (-iz)^n = -\sum_{n=0}^{\infty} \left[\frac{1}{2^{n+1}} + (-1)^n i^{n+1} \right] z^n,$$

de onde se obtém a série de Taylor de $f(z)$, válida para $|z| < 1$.

Para $1 < |z| < 2$. Neste caso, escreve-se:

$$\frac{1}{z-2} = -\frac{1}{2} \frac{1}{1-z/2} = -\frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z}{2}\right)^n, \quad (|z| < 2)$$

$$\frac{1}{z-i} = \frac{1}{z} \frac{1}{1-i/z} = \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{i}{z}\right)^n, \quad (|z| > 1).$$

Subtraindo-se ambas as séries, obtém-se a série de Laurent para $1 < |z| < 2$.

Para $|z| > 2$. Neste caso, escreve-se:

$$\frac{1}{z-2} = \frac{1}{z} \frac{1}{1-2/z} = \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{2}{z}\right)^n, \quad (|z| > 2)$$

$$\frac{1}{z-i} = \frac{1}{z} \frac{1}{1-i/z} = \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{i}{z}\right)^n, \quad (|z| > 1).$$

Subtraindo-se ambas as séries, obtém-se a série de Laurent para $|z| > 2$, a qual é composta somente pela parte principal.

Exemplo 2.16 (Uso de séries de Taylor conhecidas). Fazendo-se uso das séries de Taylor para as funções e^z e $\operatorname{sen} z$, expressões (2.32b) e (2.32c), respectivamente, as seguintes séries de Laurent podem ser obtidas:

$$\begin{aligned} \frac{\operatorname{sen}(z^2)}{z^4} &= \frac{1}{z^2} + \sum_{n=1}^{\infty} \frac{(-1)^n z^{4n-2}}{(2n+1)!} = \frac{1}{z^2} - \frac{z^2}{3!} + \frac{z^6}{5!} - \frac{z^{10}}{7!} + \dots, \quad |z| > 0 \\ \frac{e^z}{z^2} &= \frac{1}{z^2} + \frac{1}{z} + \sum_{n=2}^{\infty} \frac{z^{n-2}}{n!} = \frac{1}{z^2} + \frac{1}{z} + \frac{1}{2!} + \frac{z}{3!} + \frac{z^2}{4!} + \dots, \quad |z| > 0 \\ e^{1/z} &= 1 + \sum_{n=1}^{\infty} \frac{1}{n!z^n} = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \frac{1}{3!z^3} + \dots, \quad |z| > 0. \end{aligned}$$

Exemplo 2.17 (Obtenção da série de Laurent por diferenciação). Seja, por exemplo,

$$f(z) = \frac{1}{(z-1)^2}.$$

Para esta função, não se pode aplicar diretamente a expressão para a série geométrica. Contudo, sabendo-se que para $z \neq 0$

$$\frac{1}{(z-1)^2} = \frac{d}{dz} \left(\frac{1}{1-z} \right),$$

agora pode-se usar a série geométrica, resultando

$$f(z) = \frac{1}{(z-1)^2} = \frac{d}{dz} \left(\frac{1}{1-z} \right) = \frac{d}{dz} \sum_{n=0}^{\infty} z^n.$$

Esta série pode ser diferenciada termo a termo dentro de seu círculo de convergência ($|z| < 1$), de onde se obtém:

$$f(z) = \frac{1}{(z-1)^2} = \sum_{n=0}^{\infty} (n+1) z^n = 1 + 2z + 3z^2 + 4z^3 + \dots.$$

Exemplo 2.18 (Obtenção da série de Laurent por integração). Seja, por exemplo,

$$f(z) = \ln(1+z) = \ln|1+z| + i \arg(1+z),$$

onde se assume que o plano z fica restrito ao ramo principal da função logarítmica.

Sabendo-se que

$$\frac{d}{dz} \ln(1+z) = \frac{1}{1+z}, \text{ então } \ln(1+z) = \int_0^z \frac{dw}{1+w},$$

pode-se desenvolver

$$\frac{1}{1+z} = 1 - z + z^2 - z^3 + z^4 + \dots = \sum_{n=0}^{\infty} (-1)^n z^n, \quad (|z| < 1),$$

e integrar termo a termo:

$$\ln(1+z) = \int_0^z \frac{dw}{1+w} = \sum_{n=0}^{\infty} (-1)^n \int_0^z w^n dw = \sum_{n=0}^{\infty} \frac{(-1)^n z^{n+1}}{n+1} = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \dots.$$

2.8.8 SÉRIES DE LAURENT DE FUNÇÕES ELEMENTARES

Sem demonstração, apresenta-se a seguir algumas séries de Laurent de funções elementares desenvolvidas em torno de $z_0 = 0$:

$$\operatorname{cotan} z = \frac{1}{z} - \frac{z}{3} - \frac{z^3}{45} - \frac{2z^5}{945} - \dots - \frac{(-1)^{n-1} 2^{2n} B_{2n} z^{2n-1}}{(2n)!} - \dots \quad (|z| < \pi) \quad (2.35a)$$

$$\operatorname{cotanh} z = \frac{1}{z} + \frac{z}{3} - \frac{z^3}{45} + \frac{2z^5}{945} - \dots + \frac{2^{2n} B_{2n} z^{2n-1}}{(2n)!} + \dots \quad (|z| < \pi) \quad (2.35b)$$

$$\operatorname{cosec} z = \frac{1}{z} + \frac{z}{6} + \frac{7z^3}{360} + \frac{31z^5}{15120} + \cdots + \frac{(-1)^{n-1} 2(2^{2n-1} - 1) B_{2n}}{(2n)!} z^{2n-1} + \cdots \quad (|z| < \pi) \quad (2.35c)$$

$$\operatorname{cosech} z = \frac{1}{z} - \frac{z}{6} + \frac{7z^3}{360} - \frac{31z^5}{15120} + \cdots - \frac{2(2^{2n-1} - 1) B_{2n}}{(2n)!} z^{2n-1} + \cdots \quad (|z| < \pi), \quad (2.35d)$$

sendo $\{B_n\}$ os números de Bernoulli, dados por

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}, B_6 = \frac{1}{42}, B_{2n-1} = 0, B_n = -n! \sum_{k=0}^{n-1} \frac{B_k}{k!(n+1-k)!} \quad (n \geq 2).$$

2.8.9 CLASSIFICAÇÃO DE SINGULARIDADES

Na seção 2.8.7 observou-se que sempre que a função $f(z)$ possui alguma singularidade em um dado ponto z_0 , a sua série de Laurent desenvolvida em torno deste ponto possui uma parte principal não nula, a qual pode conter um número finito ou infinito de termos. O número de termos contidos na parte principal serve como um critério adicional, alternativo aos critérios definidos na seção 2.4.7, de classificação do tipo de singularidade que $f(z)$ possui em z_0 .

2.8.9.1 POLOS

Se a série de Laurent da função $f(z)$ possuir um número finito de termos na sua parte principal, então esta singularidade é um polo, cuja ordem é dada pela potência mais alta na parte principal.

Exemplo 2.19 (Polos). A função

$$f(z) = \frac{e^{2z}}{(z-1)^3} = \frac{e^2}{(z-1)^3} + \frac{2e^2}{(z-1)^2} + \frac{2e^2}{z-1} + \frac{4}{3}e^2 + \frac{2}{3}e^2(z-1) + \cdots \quad (2.36)$$

possui um polo de ordem 3 em $z = 1$.

2.8.9.2 SINGULARIDADES ESSENCIAIS

Se a série de Laurent da função $f(z)$ possuir um número infinito de termos na sua parte principal, então a função possui uma singularidade essencial.

Exemplo 2.20 (Singularidades essenciais). A função

$$f(z) = e^{1/z} = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \frac{1}{3!z^3} + \cdots$$

possui uma singularidade essencial em $z = 0$.

2.8.9.3 SINGULARIDADES REMOVÍVEIS

Neste caso, a série de Laurent de $f(z)$ possui apenas parte analítica.

Exemplo 2.21 (Singularidades removíveis). A função

$$f(z) = \frac{\operatorname{sen} z}{z} = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \frac{z^6}{7!} + \cdots$$

possui uma singularidade removível em $z = 0$.

2.9 INTEGRAÇÃO NO PLANO COMPLEXO PELO MÉTODO DOS RESÍDUOS

Nesta seção, o teorema de Cauchy será estendido a casos onde o integrando não é analítico; por exemplo, se o integrando possui pontos singulares isolados. Cada singularidade isolada

contribui com um termo ao resultado da integral, sendo este termo proporcional ao **resíduo da singularidade**. Esta propriedade, discutida pelo *teorema dos resíduos*, é muito útil para o cálculo de integrais definidas, não somente no plano complexo, mas também puramente reais. Em muitas situações, o teorema dos resíduos consiste no único método conhecido de solução da integral. O mesmo teorema também é útil na solução de certas equações diferenciais ordinárias ou parciais.

2.9.1 RESÍDUOS

Seja $f(z)$ unívoca e analítica no interior e sobre um contorno fechado simples C , exceto em um ponto $z = z_0$, o qual por hipótese é interno a C . Se o ponto z_0 é uma singularidade isolada de $f(z)$, então existe, de acordo com o teorema 2.22, um número real $R_1 > 0$ tal que para $0 < |z - z_0| < R_1$ a função $f(z)$ pode ser desenvolvida em termos de uma série de Laurent (2.34),

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n + \frac{c_{-1}}{z - z_0} + \frac{c_{-2}}{(z - z_0)^2} + \dots, \quad (2.37)$$

onde

$$c_n = \frac{1}{2\pi i} \oint_C \frac{f(z) dz}{(z - z_0)^{n+1}}.$$

Em particular, para $n = -1$ obtém-se que

$$c_{-1} = \frac{1}{2\pi i} \oint_C f(z) dz. \quad (2.38)$$

O número complexo c_{-1} , o qual é o coeficiente de $1/(z - z_0)$ na expansão (2.37), é denominado o **resíduo de $f(z)$ no ponto singular isolado z_0** .

Este resultado também pode ser obtido de uma maneira mais formal integrando-se (2.37) em ambos os lados ao longo de C :

$$\oint_C dz f(z) = \sum_{n=0}^{\infty} c_n \oint_C dz (z - z_0)^n + \oint_C dz \frac{c_{-1}}{z - z_0} + \oint_C dz \frac{c_{-2}}{(z - z_0)^2} + \dots.$$

De acordo com o teorema de Cauchy (teorema (2.10)) ou com o exemplo (2.9),

$$\oint_C \frac{dz}{(z - z_0)^{n+1}} = 2\pi i \delta_{n0},$$

resultando em

$$\oint_C dz f(z) = 2\pi i c_{-1},$$

de onde resulta o resíduo (2.38).

É comum usar-se também a notação

$$\text{Res } f(z_0) \equiv \frac{1}{2\pi i} \oint_C f(z) dz = c_{-1}.$$

A fórmula (2.38) consiste em um método poderoso para calcular certas integrais ao longo de contornos simples fechados. Para tanto, basta conhecer o valor do coeficiente c_{-1} da série de Laurent associada à função que está sendo integrada.

Exemplo 2.22. Calcule a integral

$$\oint_C \frac{e^{-z} dz}{(z - 1)^2}, \text{ sendo } C \text{ definido por } |z| = 2.$$

Solução. O único ponto singular do integrando é $z = 1$, um polo simples interior à circunferência $|z| = 2$. Desenvolvendo e^{-z} em uma série de Taylor em torno do ponto $z = 1$, resulta a série de Laurent

$$\frac{e^{-z}}{(z - 1)^2} = \frac{e^{-1}}{(z - 1)^2} - \frac{e^{-1}}{z - 1} + e^{-1} \sum_{n=2}^{\infty} \frac{(-1)^n}{n!} (z - 1)^{n-2}, \quad (|z - 1| > 0),$$

cujo resíduo em $z = 1$ é $c_{-1} = -e^{-1}$. Portanto,

$$\oint_C \frac{e^{-z} dz}{(z - 1)^2} = -\frac{2\pi i}{e}.$$

Exemplo 2.23. Calcule a integral

$$\oint_C \exp\left(\frac{1}{z^2}\right) dz$$

ao longo do mesmo contorno do exemplo anterior.

Solução. O ponto singular do integrando agora é $z = 0$, a qual é uma singularidade essencial. Empregando a série de MacLaurin para e^z , pode-se escrever

$$\exp\left(\frac{1}{z^2}\right) = \sum_{n=0}^{\infty} \frac{1}{n!z^{2n}} = 1 + \frac{1}{z^2} + \frac{1}{2!z^4} + \dots, \quad (|z| > 0).$$

Portanto, o resíduo em $z = 0$ é $c_{-1} = 0$ e, assim,

$$\oint_C \exp\left(\frac{1}{z^2}\right) dz = 0.$$

Observação. Se $f(z)$ for uma função analítica em $z = z_0$, o resíduo $\text{Res } f(z) = c_{-1}$ é, obviamente zero. Contudo, se z_0 for um ponto singular isolado, o resíduo neste ponto pode ou não ser nulo.

2.9.2 TEOREMA DOS RESÍDUOS

Se uma função $f(z)$ possui um número finito de pontos singulares no interior de algum contorno simples fechado C , a integral de $f(z)$ ao longo deste contorno será dada pela soma dos respectivos resíduos da função. Este resultado é garantido pelo *teorema dos resíduos de Cauchy*, descrito a seguir.

Teorema 2.25 (Teorema dos resíduos). *Seja $f(z)$ uma função analítica no interior e ao longo de um contorno simples fechado C , exceto em um número finito de pontos singulares isolados z_1, \dots, z_n localizados no interior de C . Se $b_j = \text{Res } f(z_j)$ ($j = 1, \dots, n$) são os respectivos resíduos de $f(z)$ nestes pontos singulares, então*

$$\oint_C f(z) dz = 2\pi i \sum_{j=1}^n b_j = 2\pi i (b_1 + b_2 + \dots + b_n). \quad (2.39)$$

Demonstração. Considera-se o contorno exterior C representado na figura 2.18. Deformando-se este contorno com os segmentos de reta e as circunferências $\{C_j\}$ ilustradas na figura, o contorno $B = C + \sum_j (-C_j)$ passa a ser simplesmente conexo. Então, de acordo com o teorema de deformação do contorno (teorema 2.12), resulta

$$\oint_C f(z) dz = \sum_{j=1}^n \oint_{C_j} f(z) dz.$$

Como $f(z)$ possui uma expansão de Laurent (2.37) em torno de cada ponto singular z_j , resulta a expressão (2.39). \square

2.9.3 CÁLCULO DE RESÍDUOS

Alguns métodos básicos de obtenção dos resíduos de uma função $f(z)$ serão agora discutidos.

2.9.3.1 PRIMEIRO MÉTODO: DIRETO DA DEFINIÇÃO

Calcula-se o resíduo de $f(z)$ no ponto z_0 direto da definição (2.38),

$$\text{Res } f(z_0) = \frac{1}{2\pi i} \oint_C f(z) dz,$$

sendo C o contorno que envolve somente o ponto singular z_0 . Este método é pouco utilizado, mas pode ser útil se $f(z)$ tem a primitiva ($F'(z) = f(z)$) conhecida e possui um ponto de ramificação em $z = z_0$.

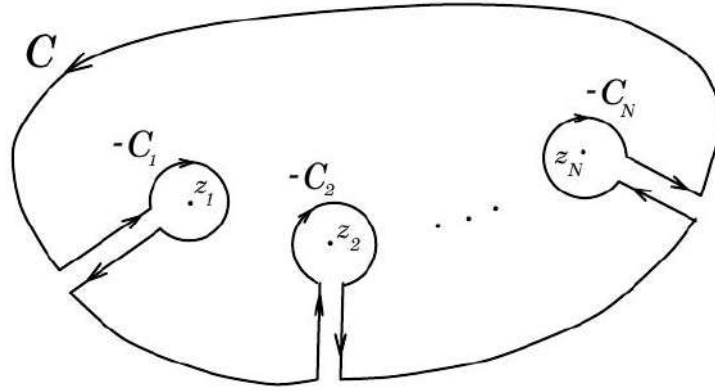


Figura 2.18: Contorno C utilizado na demonstração do teorema dos resíduos. Os pontos $\{z_j\}$ ($j = 1, \dots, n$) são pontos singulares do integrando.

Exemplo 2.24. Calcule resíduo de $f(z) = 1/z$ em $z = 0$.

Solução. Esta função possui a primitiva conhecida, $F(z) = \ln z$, sendo que o ponto $z = 0$ é um ponto de ramificação, com a linha de ramificação ao longo do eixo real positivo. No ramo principal, $F(z) = \ln r + i\theta$, onde $z = re^{i\theta}$. Para evitar a linha de ramificação, o contorno C deve ser inicialmente desconexo (aberto), sendo o mesmo fechado por um processo de limite, conforme mostra a figura 2.19. Assim,

$$\text{Res } f(z_0 = 0) = \frac{1}{2\pi i} \lim_{B \rightarrow A} \int_A^B \frac{dz}{z} = \frac{1}{2\pi i} \lim_{B \rightarrow A} (\ln B - \ln A) = \frac{1}{2\pi i} \lim_{\epsilon \rightarrow 0} i(2\pi - \epsilon) = 1.$$

2.9.3.2 SEGUNDO MÉTODO: POLOS DE ORDEM m EM $z = z_0$

Se $f(z)$ é analítica no interior e ao longo de um contorno fechado simples C , exceto por um polo de ordem m em $z = z_0$, então

$$\text{Res } f(z_0) = \frac{1}{(m-1)!} \lim_{z \rightarrow z_0} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)].$$

Demonstração. Se $f(z)$ possui um polo de ordem m em $z = z_0$, então da definição de polo (2.16) e pela propriedade das séries de Laurent (2.36), segue ser possível escrever $f(z) = g(z)/(z - z_0)^m$, sendo $g(z)$ uma função analítica em R . Então

$$\text{Res } f(z_0) = \frac{1}{2\pi i} \oint_C f(z) dz = \frac{1}{2\pi i} \oint_C \frac{g(z)}{(z - z_0)^m} dz.$$

Pela fórmula integral de Cauchy (2.24), resulta

$$\text{Res } f(z_0) = \frac{1}{(m-1)!} g^{(m-1)}(z_0) = \frac{1}{(m-1)!} \lim_{z \rightarrow z_0} \frac{d^{m-1}}{dz^{m-1}} [(z - z_0)^m f(z)].$$

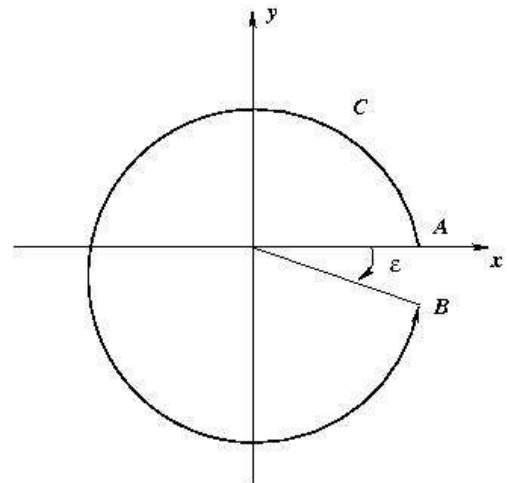


Figura 2.19: Contorno de integração para o exemplo 2.24.

Exemplo 2.25. Calcule os resíduos de:

(a) $f(z) = \frac{z}{(z-1)(z+1)^2}$.

Solução. Os polos são: $z = 1$ (polo simples) e $z = -1$ (polo duplo). Então,

$$\text{Res } f(1) = \lim_{z \rightarrow 1} (z-1) \frac{z}{(z-1)(z+1)^2} = \frac{1}{4},$$

$$\text{Res } f(-1) = \lim_{z \rightarrow -1} \frac{d}{dz} \left[(z+1)^2 \frac{z}{(z-1)(z+1)^2} \right] = -\frac{1}{4}.$$

$$(b) f(z) = \frac{\tan z}{z^2}.$$

Solução. Há somente um polo simples em $z = 0$, pois

$$\lim_{z \rightarrow 0} z \frac{\tan z}{z^2} = \lim_{z \rightarrow 0} \frac{\tan z}{z} = \lim_{z \rightarrow 0} \frac{1}{\cos z} = 1.$$

Portanto, $\text{Res } f(0) = 1$.

$$(c) f(z) = \cot z.$$

Solução. Os polos são $z = n\pi$, os quais são de primeira ordem. Então,

$$\text{Res } f(n\pi) = \lim_{z \rightarrow n\pi} (z - n\pi) \cot z = \lim_{z \rightarrow n\pi} \cos z \lim_{z \rightarrow n\pi} \frac{z - n\pi}{\sin z} = (-1)^n (-1)^n = 1.$$

2.9.3.3 TERCEIRO MÉTODO: RESÍDUO DE UMA FUNÇÃO RACIONAL

Este método se aplica quando a função $f(z)$ possui um polo simples em z_0 e pode ser escrita na forma racional

$$f(z) = \frac{p(z)}{q(z)},$$

sendo $p(z)$ e $q(z)$ funções analíticas, com $q(z_0) = 0$ e $p(z_0) \neq 0$. Neste caso,

$$\text{Res } f(z_0) = \frac{p(z_0)}{q'(z_0)}, \text{ desde que } q'(z_0) \neq 0.$$

Demonstração. Como z_0 por hipótese é um polo simples, pode-se escrever

$$\begin{aligned} (z - z_0) f(z) &= (z - z_0) \frac{p(z_0) + p'(z_0)(z - z_0) + p''(z_0)(z - z_0)^2/2! + \dots}{\underbrace{q(z_0)}_{=0} + q'(z_0)(z - z_0) + q''(z_0)(z - z_0)^2/2! + \dots} \\ &= \frac{p(z_0) + p'(z_0)(z - z_0) + p''(z_0)(z - z_0)^2/2! + \dots}{q'(z_0) + q''(z_0)(z - z_0)/2! + \dots}. \end{aligned}$$

Então,

$$\text{Res } f(z_0) = \lim_{z \rightarrow z_0} (z - z_0) f(z) = \frac{p(z_0)}{q'(z_0)}.$$

□

Exemplo 2.26. Calcule o resíduo em $z = 0$ para

$$f(z) = \frac{e^z}{\sin z}.$$

Solução. O resíduo é dado por:

$$\text{Res } f(0) = \left. \frac{e^z}{\cos z} \right|_{z=0} = 1.$$

2.9.3.4 QUARTO MÉTODO: PELO DESENVOLVIMENTO EM SÉRIE DE LAURENT

Quando $z = z_0$ é uma singularidade essencial, este é o único método disponível. Deve-se então primeiramente construir a série de Laurent de $f(z)$ a partir dos métodos discutidos na seção 2.8.7, ou a partir de outro método. Então, o resíduo será diretamente fornecido pelo coeficiente c_{-1} , de acordo com a definição (2.38).

Exemplo 2.27. Calcule o resíduo de:

$$(a) f(z) = e^{1/z} \text{ em } z = 0.$$

Solução. De acordo com o exemplo 2.16,

$$e^{1/z} = 1 + \frac{1}{z} + \frac{1}{2!z^2} + \frac{1}{3!z^3} + \dots, \quad (c_{-1} = 1),$$

então $\text{Res } f(0) = 1$.

$$(b) f(z) = e^{-1/z^2} \text{ em } z = 0.$$

Solução. Aqui pode-se usar novamente o exemplo 2.16, resultando

$$e^{-1/z^2} = 1 - \frac{1}{z^2} + \frac{1}{2!z^4} - \frac{1}{3!z^6} + \dots, \quad (c_{-1} = 0),$$

então $\text{Res } f(0) = 0$.

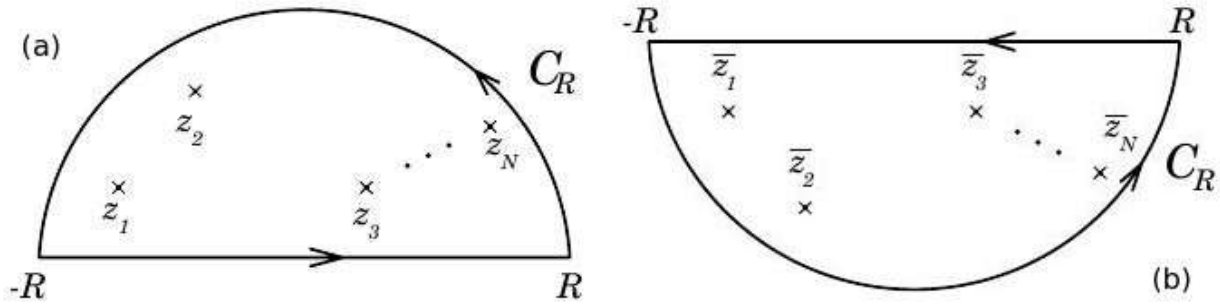


Figura 2.20: (a) Contorno C_R no semi-plano superior. (b) Contorno C_R no semi-plano inferior.

2.9.4 CÁLCULO DE INTEGRAIS DEFINIDAS OU IMPRÓPRIAS

O teorema dos resíduos pode ser usado para calcular uma grande variedade de integrais, tanto definidas quanto impróprias, reais ou complexas. É necessário, contudo, escolher-se o contorno integração de uma forma adequada. Alguns dos procedimentos mais frequentes são ilustrados nas seções seguintes. As integrais ilustradas podem aparecer em problema físicos, particularmente na solução de problemas de contorno em equações diferenciais parciais.

2.9.4.1 INTEGRAIS DO TIPO I: FUNÇÕES RACIONAIS

Aqui serão consideradas integrais do tipo

$$\int_{-\infty}^{\infty} \frac{N(x)}{D(x)} dx,$$

onde $N(x)$ e $D(x)$ são dois polinômios reais que satisfazem as seguintes condições:

1. $D(x) \neq 0, \forall x \in \mathbb{R}$, isto é, $D(x)$ não possui raízes no eixo real. Isto implica que o grau do polinômio é necessariamente par.
2. O grau de $D(x)$ é maior que o grau de $N(x)$ por um fator de 2, no mínimo.

A última hipótese é necessária para garantir a convergência da integral.

Considera-se então a seguinte integral de contorno

$$\oint_C \frac{N(z)}{D(z)} dz = \int_{-R}^R \frac{N(x)}{D(x)} dx + \int_{C_R} \frac{N(z)}{D(z)} dz,$$

onde $R > 0$. As funções $N(z)$ e $D(z)$ são as *continuações analíticas*⁷ dos polinômios reais para o plano complexo, obtidas pela substituição $x \rightarrow z$. A curva simples C_R consiste em um semi-círculo de raio R localizado ou no semi-plano complexo superior ou no inferior e que fecha o contorno C com a reta $[-R, R]$ no eixo real, conforme é ilustrado na figura 2.20. Desta forma, o contorno fechado C pode envolver parte das ou todas as N raízes de $D(z)$ no semi-plano superior ($\{z_j\}$, onde $j = 1, \dots, N$) quando C_R está nesta região (figura 2.20a) ou os seus complexos conjugados ($\{\bar{z}_j\}$, $j = 1, \dots, N$) quando C_R está no semi-plano inferior (figura 2.20b). Ao se fazer $R \rightarrow \infty$, o contorno C engloba todas as N raízes de $D(z)$ em um dos semi-planos.

Portanto, pelo teorema dos resíduos (2.39),

$$\lim_{R \rightarrow \infty} \oint_C \frac{N(z)}{D(z)} dz = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{N(x)}{D(x)} dx + \lim_{R \rightarrow \infty} \int_{C_R} \frac{N(z)}{D(z)} dz = 2\pi i \sum_{j=1}^N \text{Res} \frac{N(z)}{D(z)} \Big|_{z=z_j},$$

sendo $\{z_j\}$ as raízes de $D(z)$ englobadas por C . Portanto,

$$\int_{-\infty}^{\infty} \frac{N(x)}{D(x)} dx = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{N(x)}{D(x)} dx = 2\pi i \sum_{j=1}^N \text{Res} \frac{N(z)}{D(z)} \Big|_{z=z_j} - \lim_{R \rightarrow \infty} \int_{C_R} \frac{N(z)}{D(z)} dz.$$

Mostra-se agora, por intermédio de um teorema, que a integral ao longo de C_R se anula quando $R \rightarrow \infty$, bastando para isso que $|N(z)/D(z)| \rightarrow 1/|z^2|$ quando $|z| \rightarrow \infty$.

⁷Ver seção 2.10.

Teorema 2.26. Seja $F(z)$ uma função analítica ao longo do semi-círculo C_R , de raio R , tal que $|F(z)| \leq M/R^k$, sendo $z = Re^{i\theta}$ e onde $k > 1$ e M são constantes, então

$$\lim_{R \rightarrow \infty} \int_{C_R} F(z) dz = 0.$$

Demonstração. Pela propriedade (5) das integrais de linha (seção 2.5.2), sendo $A = \max(|F(z)|)$ ao longo de C_R e L a extensão de C_R , então

$$\left| \int_{C_R} F(z) dz \right| \leq \frac{M}{R^k} \pi R = \frac{\pi M}{R^{k-1}}.$$

Assim,

$$\lim_{R \rightarrow \infty} \int_{C_R} F(z) dz = \lim_{R \rightarrow \infty} \left| \int_{C_R} F(z) dz \right| = 0.$$

□

O corolário a seguir particulariza o teorema 2.26 para polinômios.

Corolário. Seja $F(z)$ uma função racional,

$$F(z) = \frac{N(z)}{D(z)} = \frac{a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0}{b_m z^m + b_{m-1} z^{m-1} + \dots + b_1 z + b_0},$$

com $m \geq n + 2$, então

$$\lim_{R \rightarrow \infty} \int_{C_R} \frac{N(z)}{D(z)} dz = 0.$$

Demonstração. De acordo com o teorema 2.26, ao longo de C_R ,

$$\left| \int_{C_R} \frac{N(z)}{D(z)} dz \right| \leq \frac{|a_n| R^n + |a_{n-1}| R^{n-1} + \dots + |a_0|}{||b_m| R^m - |b_{m-1}| R^{m-1} - \dots - |b_0||} \pi R.$$

Então

$$\lim_{R \rightarrow \infty} \left| \int_{C_R} \frac{N(z)}{D(z)} dz \right| \leq \lim_{R \rightarrow \infty} \frac{\pi |a_n| R^{n+1}}{|b_m| R^m} = \lim_{R \rightarrow \infty} \frac{\pi |a_n|}{|b_m|} \frac{1}{R^{m-n+1}} = 0.$$

□

Assim, de acordo com o teorema 2.26, resulta

$$\int_{-\infty}^{\infty} \frac{N(x)}{D(x)} dx = 2\pi i \sum_{j=1}^N \operatorname{Res} \frac{N(z)}{D(z)} \Big|_{z=z_j}. \quad (2.40)$$

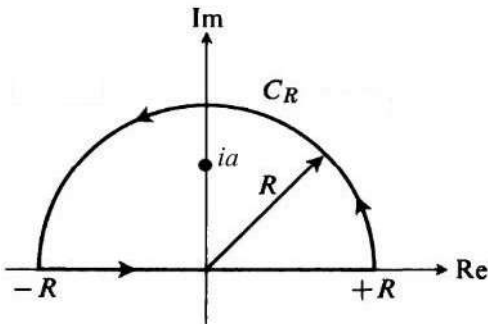


Figura 2.21: Contorno de integração para integrais do tipo I.

Exemplo 2.28. Calcule a integral

$$I = \int_{-\infty}^{\infty} \frac{dx}{x^2 + a^2}, \quad (\text{sendo } a > 0).$$

Solução. Como $F(z) = 1/(z^2 + a^2)$ satisfaz a condição estipulada no teorema 2.26, então

$$\lim_{R \rightarrow \infty} \int_{C_R} \frac{dz}{z^2 + a^2} = 0.$$

Por outro lado, as raízes de $D(z)$ são $z = \pm ia$. Pode-se escolher C_R dentro do semi-plano superior ou inferior. Escolhendo-se C_R conforme a figura ao lado, temos de (2.40),

$$I = 2\pi i \operatorname{Res} \left(\frac{1}{z^2 + a^2} \right) \Big|_{z=ia}.$$

Pelo método de cálculo de resíduos para funções que possuem polos no plano complexo, apresentado na seção 2.9.3.2,

$$I = 2\pi i \lim_{z \rightarrow ia} \frac{z - ia}{z^2 + a^2} = 2\pi i \frac{1}{2ia} = \frac{\pi}{a}.$$

Se o contorno escolhido fosse no semi-plano inferior, o resultado seria o mesmo.

Exemplo 2.29. Calcule a integral

$$I = \int_{-\infty}^{\infty} \frac{x^2}{x^4 + 1} dx.$$

Solução. A função $F(z) = z^2 / (z^4 + 1)$ novamente satisfaz a condição estipulada no teorema 2.26. Agora as 4 raízes de $D(z)$ são determinadas pela equação $z^4 = -1 = e^{i\pi}$, resultando em (ver eq. 2.9) $z_1 = e^{i\pi/4}$, $z_2 = e^{i3\pi/4}$, $z_3 = e^{i5\pi/4}$ e $z_4 = e^{i7\pi/4}$. Traçando novamente C_R no semi-plano superior, somente será necessário considerar as raízes z_1 e z_2 . Então,

$$I = 2\pi i \left[\operatorname{Res} \frac{z^2}{z^4 + 1} \Big|_{z=z_1} + \operatorname{Res} \frac{z^2}{z^4 + 1} \Big|_{z=z_2} \right].$$

Pelo método apresentado na seção 2.9.3.3, resulta

$$I = 2\pi i \left(\frac{z_1^2}{4z_1^3} + \frac{z_2^2}{4z_2^3} \right) = 2\pi i \left(\frac{e^{-i\pi/4}}{4} + \frac{e^{-i3\pi/4}}{4} \right) = \pi \frac{e^{i\pi/4} + e^{-i\pi/4}}{2} = \pi \cos \left(\frac{\pi}{4} \right) = \frac{\pi}{\sqrt{2}}.$$

2.9.4.2 INTEGRAIS DO TIPO II: FUNÇÕES RACIONAIS DE FUNÇÕES TRIGONOMÉTRICAS

Agora serão consideradas integrais do tipo

$$\int_0^{2\pi} F(\operatorname{sen} \theta, \operatorname{cos} \theta) d\theta,$$

onde $F(\operatorname{sen} \theta, \operatorname{cos} \theta)$ é uma função racional de $\operatorname{sen} \theta$ e $\operatorname{cos} \theta$.

Realizando a substituição

$$z = e^{i\theta}, \quad dz = ie^{i\theta} d\theta$$

e usando $\operatorname{cos} \theta = (e^{i\theta} + e^{-i\theta}) / 2 = (z + 1/z) / 2$ e $\operatorname{sen} \theta = (e^{i\theta} - e^{-i\theta}) / 2i = (z - 1/z) / 2i$, resulta

$$\int_0^{2\pi} F(\operatorname{sen} \theta, \operatorname{cos} \theta) d\theta = \oint_{C_0} \frac{dz}{iz} F \left(\frac{z - 1/z}{2i}, \frac{z + 1/z}{2} \right),$$

sendo C_0 a circunferência de raio unitário centrada na origem. Como $F(x, y)$ é uma função racional, a integral complexa acima pode ser obtida a partir do teorema dos resíduos, os quais serão determinados novamente pelas raízes de um polinômio. Portanto, se o polinômio resultante no denominador possui N raízes dentro do círculo de raio unitário, determinadas pelo conjunto $\{z_j\}$ ($j = 1, \dots, N$),

$$\int_0^{2\pi} F(\operatorname{sen} \theta, \operatorname{cos} \theta) d\theta = 2\pi i \sum_{j=1}^N \operatorname{Res} \left[\frac{1}{iz} F \left(\frac{z - 1/z}{2i}, \frac{z + 1/z}{2} \right) \right] \Big|_{z=z_j}. \quad (2.41)$$

Exemplo 2.30. Calcule a integral

$$I = \int_0^{2\pi} \frac{\cos 3\theta}{5 - 4 \cos \theta} d\theta.$$

Solução. Realizando a transformação $z = e^{i\theta}$, resulta

$$\cos 3\theta = \frac{e^{i3\theta} + e^{-i3\theta}}{2} = \frac{z^3 + z^{-3}}{2}$$

e

$$I = \frac{1}{2} \oint_{C_0} \frac{dz}{iz} \frac{z^6 + 1}{z^2 (5z - 2z^2 - 2)} = -\frac{1}{2i} \oint_{C_0} dz \frac{z^6 + 1}{z^3 (2z - 1)(z - 2)}.$$

O integrando possui os seguintes polos:

- $z = 0$: polo de ordem 3.
- $z = 1/2$: polo de ordem 1.
- $z = 2$: polo de ordem 1 (fora do círculo $|z| = 1$).

Os resíduos são os seguintes:

$$\begin{aligned} \operatorname{Res} F(0) &= \frac{1}{2!} \lim_{z \rightarrow 0} \frac{d^2}{dz^2} \left[z^3 \frac{z^6 + 1}{z^3 (2z - 1)(z - 2)} \right] = \frac{21}{8} \\ \operatorname{Res} F\left(\frac{1}{2}\right) &= \lim_{z \rightarrow 1/2} \left[\left(z - \frac{1}{2}\right) \frac{z^6 + 1}{z^3 (2z - 1)(z - 2)} \right] = -\frac{65}{24}. \end{aligned}$$

Portanto,

$$I = -\frac{1}{2i} 2\pi i \left(\frac{21}{8} - \frac{65}{24} \right) = \frac{\pi}{12}.$$

2.9.4.3 INTEGRAIS DO TIPO III: INTEGRAIS DE FOURIER

Tratam-se de integrais do tipo

$$I_{\pm} = \int_{-\infty}^{\infty} F(x) e^{\pm ikx} dx \quad (k > 0),$$

onde $F(x)$ é uma função racional que satisfaz as condições do teorema 2.26. As partes real e imaginária do integrando determinam as integrais

$$I_c = \int_{-\infty}^{\infty} F(x) \cos kx dx \quad \text{e} \quad I_s = \int_{-\infty}^{\infty} F(x) \sin kx dx.$$

Novamente, é feita a hipótese de ausência de singularidades de $F(x)$ no eixo real. Quando é necessário calcular integrais dos tipos I_c ou I_s , inicialmente substitui-se a integral em questão por I_+ (ou I_- , dependendo das singularidades de $F(x)$) e novamente calcula-se a integral ao longo do semi-círculo ilustrado na figura 2.20a. Isto é, calcula-se

$$\oint_C F(z) e^{ikz} dz = \lim_{R \rightarrow \infty} \left[\int_{-R}^R F(x) e^{ikx} dx + \int_{C_R} F(z) e^{ikz} dz \right].$$

As condições que $F(z)$ deve satisfazer para que a integração ao longo de C_R se anule para $R \rightarrow \infty$ são dadas pelo Lema de Jordan.

Lema 2.1 (Lema de Jordan). *Seja $F(z)$ uma função analítica ao longo do semi-círculo C_R , de raio R , tal que $|F(z)| \leq M/R^\alpha$, sendo $z = Re^{i\theta}$ e onde $\alpha > 0$ e M são constantes, então*

$$\lim_{R \rightarrow \infty} \int_{C_R} F(z) e^{ikz} dz = 0.$$

Demonstração. Sendo $z = Re^{i\theta}$, então

$$\int_{C_R} F(z) e^{ikz} dz = \int_0^\pi F(Re^{i\theta}) \exp(ikRe^{i\theta}) iRe^{i\theta} d\theta.$$

Pelas propriedades (5) e (6) da seção 2.5.2,

$$\left| \int_0^\pi F(Re^{i\theta}) \exp(ikRe^{i\theta}) iRe^{i\theta} d\theta \right| \leq \int_0^\pi |F(Re^{i\theta}) \exp(ikRe^{i\theta}) iRe^{i\theta}| d\theta,$$

mas

$$\exp(ikRe^{i\theta}) = e^{ikR(\cos\theta + i\operatorname{sen}\theta)} \leq e^{-kR\operatorname{sen}\theta},$$

portanto,

$$\left| \int_0^\pi F(Re^{i\theta}) \exp(ikRe^{i\theta}) iRe^{i\theta} d\theta \right| \leq \int_0^\pi |F(Re^{i\theta})| e^{-kR\operatorname{sen}\theta} R d\theta \leq \frac{M}{R^{\alpha-1}} \int_0^\pi e^{-kR\operatorname{sen}\theta} d\theta.$$

Como $\operatorname{sen}(\pi - \theta) = \operatorname{sen}\theta$, pode-se alterar o intervalo de integração para $[0, \pi/2]$. Além disso, como mostra a figura 2.22, $\operatorname{sen}\theta \geq 2\theta/\pi$ neste intervalo. Assim,

$$\left| \int_0^\pi F(Re^{i\theta}) \exp(ikRe^{i\theta}) iRe^{i\theta} d\theta \right| \leq \frac{2M}{R^{\alpha-1}} \int_0^{\pi/2} e^{-2kR\theta/\pi} d\theta = \frac{\pi M}{kR^\alpha} (1 - e^{-kR}).$$

Portanto,

$$\lim_{R \rightarrow \infty} \int_{C_R} F(z) e^{ikz} dz = 0.$$

□

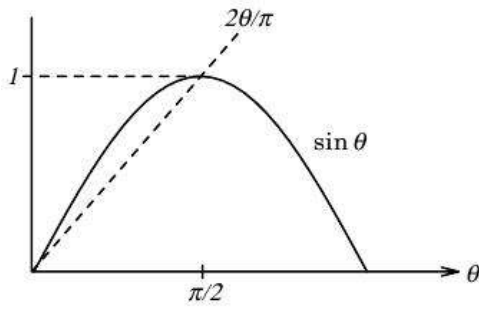


Figura 2.22: Demonstração gráfica da desigualdade $\sin \theta \geq 2\theta/\pi$ em $0 \leq \theta \leq \pi/2$.

Devido ao Lema de Jordan, o teorema dos resíduos garante que

$$I_+ = I_c + iI_s = \int_{-\infty}^{\infty} F(x)e^{ikx} dx = 2\pi i \sum_j \text{Res} [F(z_j) e^{ikz_j}], \tag{2.42a}$$

sendo $\{z_j\}$ os pontos singulares de $F(z)$ na região delimitada pelo contorno C ilustrado na figura 2.26.

Caso os pontos singulares de $F(z)$ sejam tais que se faz necessário considerar-se C_R percorrendo o semi-círculo no semi-plano inferior ilustrado na figura 2.20b, deve-se calcular então I_- . As conclusões são similares às de I_+ , salvo que os resíduos são computados no semi-plano inferior e

$$I_- = I_c - iI_s = -2\pi i \sum_j \text{Res} [F(z_j) e^{-ikz_j}]. \tag{2.42b}$$

Exemplo 2.31. Calcule a integral

$$I = \int_0^{\infty} \frac{\cos \alpha x}{1+x^2} dx.$$

Solução. Pode-se escrever

$$I = \frac{1}{2} \text{Re} \int_{-\infty}^{\infty} \frac{e^{i\alpha x}}{1+x^2} dx.$$

Identifica-se $F(z) = 1/(1+z^2)$, e esta satisfaz as condições do Lema de Jordan, pois ao longo de C_R ,

$$|F(z)| = \frac{1}{|1+z^2|} \leq \frac{1}{R^2-1} \xrightarrow{R \rightarrow \infty} \frac{1}{R^2}.$$

Os resíduos de $F(z)$ estão em $z = \pm i$, sendo estes polos simples. Assim, de (2.42a),

$$\int_{-\infty}^{\infty} \frac{e^{i\alpha x}}{1+x^2} dx = 2\pi i \text{Res} \left(\frac{e^{i\alpha x}}{1+z^2} \right) \Big|_{z=i} = \pi e^{-\alpha}.$$

Portanto,

$$I = \frac{\pi}{2} e^{-\alpha}.$$

2.9.4.4 INTEGRAIS DO TIPO IV: INTEGRANDO COM POLOS NO EIXO REAL

Seja $F(z)$ uma função **meromórfica**, ou seja, uma função que possui apenas polos em um domínio finito no plano complexo. Supõe-se que $F(z)$ possua, no mínimo, um polo ao longo do eixo real. Supõe-se também que $|F(z)| \rightarrow 0$ para $|z| \rightarrow \infty$. Deseja-se calcular agora integrais do tipo

$$\int_{-\infty}^{\infty} F(x) dx = \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} dx.$$

Devido a presença do polo no eixo real, para que o teorema dos resíduos permaneça válido, o contorno de integração não pode passar pela referida singularidade; torna-se necessário, portanto, que o contorno seja deformado nas vizinhanças do polo real. A partir desta situação surge a definição da **parte principal de Cauchy** da integral.

PARTE PRINCIPAL DE CAUCHY

Seja $F(z) = f(z)/(z-x_0)$ uma função meromórfica que possui, no mínimo, um polo simples no eixo real em $z = x_0$. Ao se considerar a integral $\int_a^b F(x) dx$, com $a < x_0 < b$ tais que $z = x_0$ é o único polo de $F(z)$ neste intervalo, a convergência desta integral depende da existência do seguinte limite:

$$I_{PP} = \lim_{\epsilon \rightarrow 0^+} \left(\int_a^{x_0-\epsilon} \frac{f(x)}{x-x_0} dx + \int_{x_0+\epsilon}^b \frac{f(x)}{x-x_0} dx \right).$$

Diz-se que $\int_a^b F(x)dx$ é convergente se e somente se I_{PP} existe e é finito, caso contrário a integral é divergente. O limite I_{PP} pode existir, em algumas situações, mesmo se $\lim_{x \rightarrow x_0} f(x) \rightarrow \pm\infty$. Isto ocorre, por exemplo, se $f(x)$ é simétrica em uma vizinhança de x_0 .

Caso I_{PP} exista e seja finita, esta define a parte principal de Cauchy da integral $\int_a^b F(x)dx$, a qual é representada pelos símbolos f_a^b ou \mathcal{P}_a^b :

$$\int_a^b \frac{f(x)}{x-x_0} dx \equiv \mathcal{P}_a^b \int_a^b \frac{f(x)}{x-x_0} dx = \lim_{\epsilon \rightarrow 0^+} \left(\int_a^{x_0-\epsilon} \frac{f(x)}{x-x_0} dx + \int_{x_0+\epsilon}^b \frac{f(x)}{x-x_0} dx \right). \quad (2.43a)$$

A parte principal de Cauchy no infinito existe se

$$\int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} dx \equiv \mathcal{P}_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} dx = \lim_{\epsilon \rightarrow 0^+} \left(\int_{-\infty}^{x_0-\epsilon} \frac{f(x)}{x-x_0} dx + \int_{x_0+\epsilon}^{\infty} \frac{f(x)}{x-x_0} dx \right) \quad (2.43b)$$

existe e é finito.

CÁLCULO DE INTEGRAIS DO TIPO IV

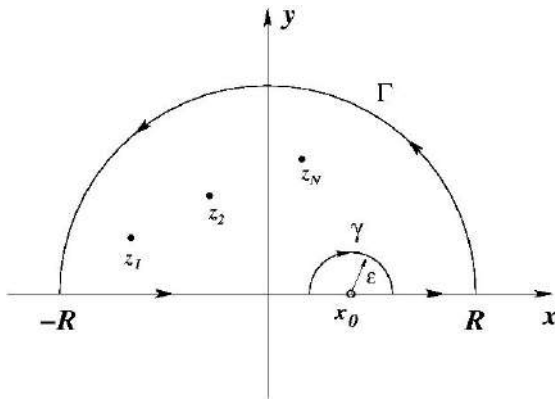


Figura 2.23: Contorno de integração para integrais do tipo IV.

Para calcular integrais do tipo IV pelo teorema dos resíduos, considera-se o contorno

$$C = \Gamma + \gamma + (-R, x_0 - \epsilon) + (x_0 + \epsilon, R)$$

apresentado na figura 2.23, o qual é composto por um semi-círculo exterior Γ , dois segmentos de reta ao longo do eixo real nos intervalos $(-R, x_0 - \epsilon)$ e $(x_0 + \epsilon, R)$, e do semi-círculo interior γ . Na figura 2.23, a curva Γ foi traçada no semi-plano superior para englobar os polos $\{z_j\}$ ($j = 1, \dots, N$) de $f(z)$. Caso estes polos se encontrem no semi-plano inferior, a curva Γ pode ser traçada nesta região. Supõe-se também que $\epsilon (> 0)$ seja pequeno o suficiente para que nenhum dos polos $\{z_j\}$ se encontre na região interior da curva γ .

Assim, pelo teorema dos resíduos,

$$\oint_C \frac{f(z)}{z-x_0} dz = 2\pi i \sum_{j=1}^N \text{Res} \left(\frac{f(z)}{z-x_0} \right) \Big|_{z=z_j}.$$

Dividindo o contorno C nas suas curvas constituintes,

$$\int_{\Gamma} \frac{f(z)}{z-x_0} dz + \int_{-R}^R \frac{f(x)}{x-x_0} dx + \int_{\gamma} \frac{f(z)}{z-x_0} dz = 2\pi i \sum_{j=1}^N \text{Res} \left(\frac{f(z)}{z-x_0} \right) \Big|_{z=z_j},$$

Toma-se agora o limite $R \rightarrow \infty$, supondo que $F(z)$ ao longo de Γ satisfaça as condições prescritas pelo teorema 2.26. Neste caso, $\int_{\Gamma} F(z)dz = 0$. A integral ao longo de γ pode ser calculada, tomando-se para ela o limite $\epsilon \rightarrow 0$:

$$\int_{\gamma} \frac{f(z)}{z-x_0} dz = \lim_{\epsilon \rightarrow 0} \int_{\pi}^0 \frac{f(x_0 + \epsilon e^{i\theta})}{\epsilon e^{i\theta}} i\epsilon e^{i\theta} d\theta = -i\pi f(x_0),$$

sendo neste caso suposto que $f(x_0)$ exista e seja finita. Este valor da integral em γ é muitas vezes denominado de **semi-resíduo de $f(x)$ em x_0** .

Portanto, obtém-se o seguinte resultado para a parte principal,

$$\int_{-\infty}^{\infty} \frac{f(x)}{x-x_0} dx = \pi i f(x_0) + 2\pi i \sum_{j=1}^N \text{Res} \left(\frac{f(z)}{z-x_0} \right) \Big|_{z=z_j}. \quad (2.44a)$$

Caso a função $F(x)$ possua mais de um polo no eixo real, o resultado (2.44a) pode ser facilmente generalizado. Sendo novamente $\{z_j\}$ ($j = 1, \dots, N$) o conjunto de polos de $F(z)$ fora do eixo real e $\{x_\ell\}$ ($\ell = 1, \dots, M$) o conjunto de polos ao longo do eixo real, a forma generalizada de (2.44a) é

$$\int_{-\infty}^{\infty} F(x) dx = \pi i \sum_{\ell=1}^M \operatorname{Res} F(x_\ell) + 2\pi i \sum_{j=1}^N \operatorname{Res} F(z_j). \quad (2.44b)$$

Exemplo 2.32. Calcule a integral

$$\int_{-\infty}^{\infty} \frac{\operatorname{sen} x}{x} dx.$$

Solução. É conveniente calcular

$$\int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx.$$

A função e^{iz}/z satisfaz as condições estabelecidas pelo Lema de Jordan, ou seja, $|z^{-1}| \leq M/R^\alpha$ ($\alpha > 0$) ao longo de Γ . Portanto, $\int_{\Gamma} (e^{iz}/z) dz = 0$ e a equação (2.44a) pode ser aplicada, de onde resulta

$$\int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx = i\pi.$$

Uma vez que $e^{ix} = \cos x + i \operatorname{sen} x$, obtém-se

$$\int_{-\infty}^{\infty} \frac{\operatorname{sen} x}{x} dx = \pi.$$

Observação. Nota-se que, neste caso, $f = \int$, pois $x = 0$ é uma singularidade removível. É interessante também que o mesmo resultado fornece

$$\int_{-\infty}^{\infty} \frac{\cos x}{x} dx = 0,$$

o que é esperado, uma vez que o integrando neste caso é ímpar.

Exemplo 2.33. Calcule a integral

$$\int_{-\infty}^{\infty} \frac{\operatorname{sen} x}{x(a^2 - x^2)} dx, \quad (a > 0).$$

Solução. Novamente toma-se

$$F(z) = \frac{e^{iz}}{z(a^2 - z^2)},$$

sendo que $F(z)$ satisfaz as condições do Lema de Jordan e os seus polos são $z = 0$ e $z = \pm a$, todos ao longo do eixo real. Neste caso, de acordo com (2.44b),

$$\int_{-\infty}^{\infty} \frac{e^{ix} dx}{x(a^2 - x^2)} = i\pi [\operatorname{Res} F(0) + \operatorname{Res} F(a) + \operatorname{Res} F(-a)].$$

Calculando os resíduos,

$$\operatorname{Res} F(0) = z \frac{e^{iz}}{z(a^2 - z^2)} \Big|_{z=0} = \frac{1}{a^2},$$

$$\operatorname{Res} F(a) = (z - a) \frac{e^{iz}}{z(a^2 - z^2)} \Big|_{z=a} = -\frac{e^{ia}}{2a^2}$$

$$\operatorname{Res} F(-a) = (z + a) \frac{e^{iz}}{z(a^2 - z^2)} \Big|_{z=-a} = -\frac{e^{-ia}}{2a^2}.$$

Portanto,

$$\int_{-\infty}^{\infty} \frac{e^{ix} dx}{x(a^2 - x^2)} = i \frac{\pi}{a^2} (1 - \cos a)$$

e o resultado desejado fica

$$\int_{-\infty}^{\infty} \frac{\operatorname{sen} x dx}{x(a^2 - x^2)} = \frac{\pi}{a^2} (1 - \cos a).$$

Observação. Novamente observa-se também que

$$\int_{-\infty}^{\infty} \frac{\cos x dx}{x(a^2 - x^2)} = 0.$$

2.9.4.5 INTEGRAIS DO TIPO V: INTEGRAÇÃO AO LONGO DE LINHAS DE RAMIFICAÇÃO

Neste caso, o contorno de integração deve ser construído de forma a evitar tanto os pontos de singularidades essenciais e polos no eixo real quanto a linha de ramificação.

Para ilustrar este tipo de integração, serão consideradas integrais do tipo

$$\int_0^{\infty} x^{\lambda-1} G(x) dx, \quad (\text{sendo } 0 < \lambda < 1)$$

e a função $G(z)$ é racional, analítica em $z = 0$ e não possui polos ao longo do eixo real positivo. Supõe-se ainda que

$$\lim_{\substack{|z| \rightarrow 0 \\ |z| \rightarrow \infty}} |z^{\lambda-1} G(z)| = 0.$$

A função $f(z) = z^{\lambda-1}$ é plurívoca, com um ponto de ramificação em $z = 0$, o que pode ser comprovado por uma rotação do fador z em torno da origem,

$$f(z) = z^{\lambda-1} \xrightarrow{\theta \rightarrow \theta + 2\pi} e^{i2\pi(\lambda-1)} z^{\lambda-1}.$$

Como $\lambda - 1$ não é inteiro, o valor de $f(z)$ em $\theta + 2\pi$ difere do valor em θ . A outra singularidade de $f(z)$ está no infinito, significando que a linha de ramificação deve unir estes dois pontos singulares. Usualmente, esta linha é traçada ao longo do eixo real positivo. O número de folhas de Riemann para $f(z)$ depende da natureza do número λ . Se este número é racional, então há um número finito de folhas de Riemann; contudo, se λ é irracional, então existem infinitas folhas de Riemann. Devido a este fato, o contorno de integração C a ser adotado deve evitar o cruzamento da linha de ramificação para evitar a necessidade de se considerar os outros ramos da função $z^{\lambda-1}$.

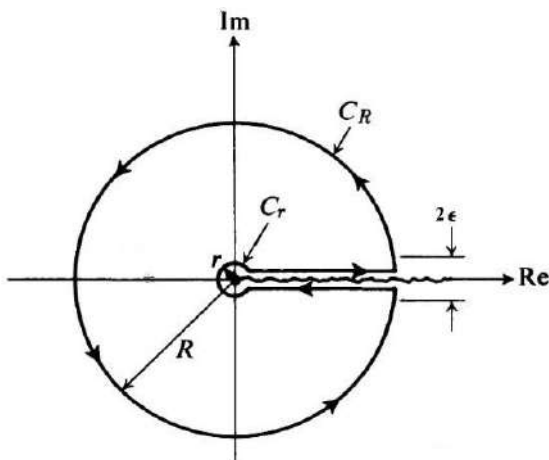


Figura 2.24: Contorno de integração para integrais do tipo V.

O contorno de integração usualmente empregado para resolver uma integral do tipo V é apresentado na figura 2.24, a qual mostra também a linha de ramificação no intervalo $[0, \infty)$. Este tipo de contorno é denominado *buraco de fechadura*⁸ e consiste em um arco de circunferência exterior C_R , de raio R e centrado na origem, com uma variação angular $\delta_R \leq \theta \leq 2\pi - \delta_R$. O número $0 < \delta_R \ll 1$ pode ser tão pequeno quanto se queira, mas nunca é nulo, justamente para evitar que C_R toque a linha de ramificação. O contorno C contém também o arco de circunferência interior C_r , de raio r e centrado na origem, com uma variação angular $\delta_r \leq \theta \leq 2\pi - \delta_r$, sendo $0 < \delta_r \ll 1$ também tão pequeno quanto se queira. Finalmente, os arcos C_r e C_R são ligados pelos segmentos de reta $[r, R]$ e $[R, r]$,⁹ os quais distam da linha de ramificação por uma distância $0 < \epsilon \ll 1$. Obviamente, $\epsilon = r \sin \delta_r$.

Como a função $z^{\lambda-1}$ é descontínua ao longo da linha de ramificação, o valor das integrais ao longo dos segmentos de reta não se cancelam. Na reta superior, $\arg(z) \simeq 0$ e, portanto,

$$f(z) = z^{\lambda-1} = (x + i\epsilon)^{\lambda-1}.$$

Na reta inferior, $\arg(z) \simeq 2\pi$ e, portanto,

$$f(z) = e^{i2\pi(\lambda-1)} z^{\lambda-1} = e^{i2\pi(\lambda-1)} (x - i\epsilon)^{\lambda-1}.$$

Supondo agora que a função $R(z)$ possua N singulares isoladas (polos e/ou singularidades essenciais) nos pontos $\{z_j\}$ ($j = 1, \dots, N$), o teorema dos resíduos (2.39) aplicado ao contorno simples fechado $C = C_R + C_r + [r, R] + [R, r]$ garante que

⁸Do inglês *keyhole*.

⁹Rigorosamente, os segmentos de reta percorrem o intervalo $[r \cos \delta_r, R \cos \delta_R]$, em ambos os sentidos.

$$\oint_C z^{\lambda-1} G(z) dz = \int_{C_R} z^{\lambda-1} G(z) dz + \int_{C_r} z^{\lambda-1} G(z) dz + \int_r^R (x+i\epsilon)^{\lambda-1} G(x+i\epsilon) dx + \int_R^r e^{2\pi i(\lambda-1)} (x-i\epsilon)^{\lambda-1} G(x-i\epsilon) dx = 2\pi i \sum_{j=1}^N \text{Res} [z^{\lambda-1} G(z)]|_{z_j}.$$

A integral em C_R se anula no limite $R \rightarrow \infty$, desde que as condições do seguinte teorema sejam satisfeitas.

Teorema 2.27. *Seja uma função $F(z)$ analítica ao longo do arco de circunferência C_R , de raio R centrado na origem, tal que $|F(z)| \leq M/R^\alpha$ ao longo de C_R , sendo $\alpha > 1$ e $M > 0$ constantes, então*

$$\lim_{R \rightarrow \infty} \int_{C_R} F(z) dz = 0.$$

Demonstração. Sendo a curva C_R parametrizada pelo ângulo θ , o qual varia no intervalo $[\theta_1, \theta_2]$, então

$$\left| \int_{C_R} F(z) dz \right| \leq \int_{\theta_1}^{\theta_2} |F(z)| R d\theta \leq \frac{M}{R^{\alpha-1}} \int_{\theta_1}^{\theta_2} d\theta = \frac{M(\theta_2 - \theta_1)}{R^{\alpha-1}}.$$

Portanto,

$$\lim_{R \rightarrow \infty} \left| \int_{C_R} F(z) dz \right| = 0,$$

de onde segue o teorema. □

Por outro lado, a integral ao longo de C_r pode ser calculada parametrizando-se $z = re^{i\theta}$ e tomando-se o limite $r \rightarrow 0$:

$$\lim_{r \rightarrow 0} \int_{C_r} z^{\lambda-1} G(z) dz = \lim_{r \rightarrow 0} \left[ir^\lambda \int_{\delta_r}^{2\pi-\delta_r} e^{i\lambda\theta} G(re^{i\theta}) d\theta \right] = \left[\frac{G(0)}{\lambda} (e^{i2\pi\lambda} e^{-i\lambda\delta_r} - e^{i\lambda\delta_r}) \right] \lim_{r \rightarrow 0} r^\lambda.$$

Como $G(z)$ por hipótese é analítica em $z = 0$ e $\lambda > 0$, resulta que

$$\lim_{r \rightarrow 0} \int_{C_r} z^{\lambda-1} G(z) dz = 0.$$

Lembrando também que $r \rightarrow 0$ implica em $\epsilon \rightarrow 0$, restam as seguintes integrais,

$$\int_0^\infty x^{\lambda-1} G(x) dx + e^{2\pi i(\lambda-1)} \int_\infty^0 x^{\lambda-1} G(x) dx = 2\pi i \sum_{j=1}^N \text{Res} [z^{\lambda-1} G(z)]|_{z_j},$$

as quais podem ser escritas

$$\begin{aligned} \int_0^\infty x^{\lambda-1} G(x) dx + e^{2\pi i(\lambda-1)} \int_\infty^0 x^{\lambda-1} G(x) dx &= (1 - e^{2\pi i(\lambda-1)}) \int_0^\infty x^{\lambda-1} G(x) dx \\ &= -e^{i\pi\lambda} (e^{i\pi\lambda} - e^{-i\pi\lambda}) \int_0^\infty x^{\lambda-1} G(x) dx = 2i(-1)^{\lambda+1} \text{sen } \pi\lambda \int_0^\infty x^{\lambda-1} G(x) dx. \end{aligned}$$

Portanto, resulta

$$\int_0^\infty x^{\lambda-1} G(x) dx = (-1)^{-(\lambda+1)} \frac{\pi}{\text{sen } \pi\lambda} \sum_{j=1}^N \text{Res} [z^{\lambda-1} G(z)]|_{z_j}. \tag{2.45}$$

Exemplo 2.34. Calcule a integral

$$I = \int_0^\infty \frac{x^{\lambda-1}}{1+x} dx, \quad (0 < \lambda < 1).$$

Solução. Como $0 < \lambda < 1$, o integrando ao longo da curva C_R , para $R \gg 1$, pode ser escrito

$$\left| \frac{z^{\lambda-1}}{1+z} \right| \simeq \frac{1}{R^{2-\lambda}}.$$

Então, de acordo com o teorema 2.27,

$$\lim_{R \rightarrow \infty} \int_{C_R} \frac{z^{\lambda-1}}{1+z} dz = 0$$

e o resultado (2.45) é válido neste caso. O único polo de $G(z) = (1+z)^{-1}$ está em $z = -1$ e o resíduo do integrando neste ponto é

$$\text{Res} \left(\frac{z^{\lambda-1}}{1+z} \right) \Big|_{z=-1} = (-1)^{\lambda-1}.$$

Portanto,

$$I = \int_0^{\infty} \frac{x^{\lambda-1}}{1+x} dx = (-1)^{-(\lambda+1)} \frac{\pi}{\text{sen } \pi \lambda} (-1)^{\lambda-1} = \frac{\pi}{\text{sen } \pi \lambda}.$$

2.9.4.6 OUTROS TIPOS DE INTEGRAIS

Existem diversos outros exemplos de integrais que podem ser calculadas usando o teorema dos resíduos por intermédio de uma escolha adequada do contorno de integração. Nesta seção serão apresentados alguns exemplos relevantes para a física.

INTEGRAIS DE FRESNEL. DIFRAÇÃO DE ONDAS ELETROMAGNÉTICAS.

As integrais

$$C = \int_0^{\infty} \cos(tx^2) dx \text{ e } S = \int_0^{\infty} \text{sen}(tx^2) dx$$

correspondem aos valores assintóticos das funções de Fresnel, as quais descrevem a difração de ondas eletromagnéticas em um único obstáculo pontiagudo (difração de canto).

Para a obtenção dos valores de C e S , considere-se a seguinte integral

$$I = \int_0^{\infty} e^{itx^2} dx,$$

cujas partes real e imaginária fornecem, respectivamente, C e S . Por conveniência, assume-se que $t > 0$; para este caso, o contorno de integração adequado pode ser visto na figura 2.25. Como a função e^{itz^2} é analítica dentro e ao longo do contorno

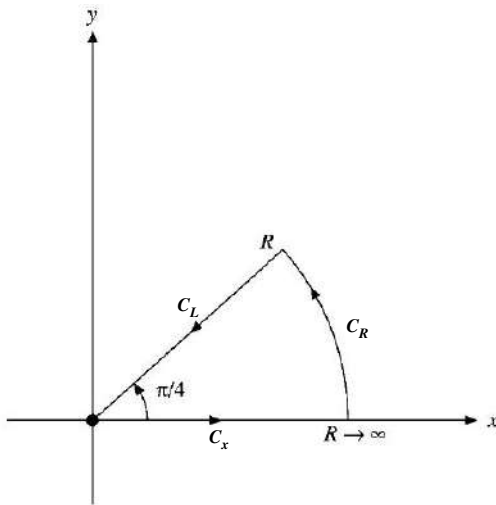


Figura 2.25: Contorno de integração para as integrais de Fresnel.

$C = C_x + C_R + C_L$, pode-se escrever

$$\oint_C e^{itz^2} dz = \left(\int_{C_x} + \int_{C_R} + \int_{C_L} \right) e^{itz^2} dz = 0.$$

A integral em C_R pode ser estimada usando argumento semelhante ao exposto pelo Lema de Jordan. Ao longo deste contorno, $z^2 = R^2 (\cos 2\theta + i \text{sen } 2\theta)$. Como demonstrado graficamente na figura 2.22, $\text{sen } 2\theta \geq 4\theta/\pi$ em $0 \leq \theta \leq \pi/4$, portanto

$$\begin{aligned} \left| \int_{C_R} e^{itz^2} dz \right| &= \left| \int_0^{\pi/4} e^{itR^2(\cos 2\theta + i \text{sen } 2\theta)} i R e^{i\theta} d\theta \right| \\ &\leq \int_0^{\pi/4} R e^{-tR^2 \text{sen } 2\theta} d\theta \end{aligned}$$

$$\leq \int_0^{\pi/4} R e^{-tR^2 4\theta/\pi} d\theta = \frac{\pi}{4tR} (1 - e^{-tR^2}).$$

Ou seja, $\left| \int_{C_R} e^{itz^2} dz \right| \rightarrow 0$ para $R \rightarrow 0$.

Por outro lado, ao longo de C_x , $z = x$ e ao longo de C_L , $z = re^{i\pi/4}$ ($0 \leq r \leq R$). Assim,

$$\int_{C_x} e^{itz^2} dz = \int_0^R e^{itx^2} dx \text{ e } \int_{C_L} e^{itz^2} dz = \int_R^0 e^{-tr^2} e^{i\pi/4} dr.$$

Portanto, no limite $R \rightarrow \infty$,

$$\int_0^\infty e^{itx^2} dx = e^{i\pi/4} \int_0^\infty e^{-tr^2} dr.$$

A integral $J = \int_0^\infty e^{-tr^2} dr$ pode ser calculada da seguinte maneira:

$$J^2 = \left(\int_0^\infty e^{-tx^2} dx \right) \left(\int_0^\infty e^{-ty^2} dy \right) = \int_0^\infty \int_0^\infty e^{-t(x^2+y^2)} dx dy.$$

Usando coordenadas polares, resulta

$$J^2 = \int_0^{\pi/2} d\theta \int_0^\infty d\rho \rho e^{-t\rho^2} = \frac{\pi}{4t}.$$

Finalmente, obtém-se

$$\int_0^\infty e^{itx^2} dx = \frac{e^{i\pi/4}}{2} \sqrt{\frac{\pi}{t}} = \left(\cos \frac{\pi}{4} + i \operatorname{sen} \frac{\pi}{4} \right) \frac{1}{2} \sqrt{\frac{\pi}{t}}.$$

Portanto, as integrais de Fresnel valem

$$S = C = \frac{1}{2} \sqrt{\frac{\pi}{2t}}.$$

INTEGRAL DE PLANCK E LEI DE STEFAN-BOLTZMANN. A Lei de Planck determina a intensidade espectral da radiação eletromagnética emitida por um corpo negro em equilíbrio termodinâmico a uma temperatura absoluta T . Esta intensidade é dada por

$$I(\nu, T) = \frac{2h}{c^2} \frac{\nu^3}{e^{h\nu/k_B T} - 1},$$

sendo $I(\nu, T)$ a potência emitida por unidade de área da superfície emissora por unidade de ângulo sólido por frequência ν . As quantidades h , k_B e c são, respectivamente, as constantes de Planck e Boltzmann e a velocidade da luz no vácuo. A Lei de Stefan-Boltzmann fornece a potência total emitida por unidade de área do corpo negro, ou seja,

$$j = \pi \int_0^\infty d\nu I(\nu, T) = \frac{2\pi h}{c^2} \left(\frac{k_B T}{h} \right)^4 I_P, \text{ sendo } I_P = \int_0^\infty \frac{u^3 du}{e^u - 1}$$

a integral que se deseja calcular, a qual é obtida por uma simples troca de variáveis de integração.

Para se obter I_P , considera-se a integral

$$f(k) = \int_0^\infty \frac{\operatorname{sen}(kx)}{e^x - 1} dx.$$

Obviamente,

$$I_P = - \left. \frac{d^3 f}{dk^3} \right|_{k=0}.$$

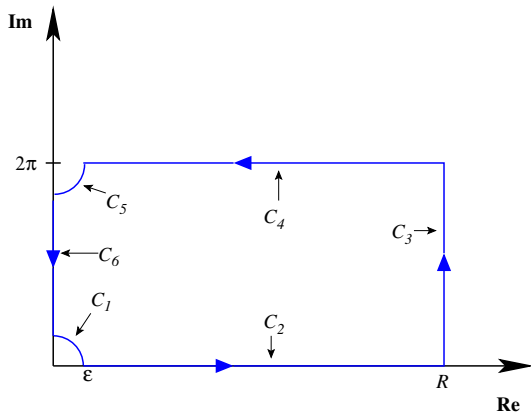


Figura 2.26: Contorno de integração empregado para o cálculo da integral de Planck. Observa-se que $\epsilon \ll 1$ e $R \gg 1$.

Portanto, pode-se obter I_P conhecendo-se $f(k)$. A integral que define esta função também pode ser escrita como

$$f(k) = \text{Im} \int_0^{\infty} \frac{e^{ikx}}{e^x - 1} dx.$$

Este cálculo será realizado utilizando-se o contorno $C = C_1 + \dots + C_6$ ilustrado na figura 2.26 e a integral de contorno

$$\oint_C \frac{e^{ikz}}{e^z - 1} dz = 0,$$

a qual é nula pelo teorema dos resíduos (2.39) porque o integrando possui apenas polos simples nos pontos $z = 2n\pi i$, ($n = 0, \pm 1, \pm 2, \dots$). As integrais nos contornos C_1, \dots, C_6 devem ser realizadas separadamente. Nos contornos C_2 e C_4 obtém-se

$$\left(\int_{C_2} + \int_{C_4} \right) \frac{e^{ikz}}{e^z - 1} dz = (1 - e^{-2\pi k}) \int_{\epsilon}^R \frac{e^{ikx}}{e^x - 1} dx.$$

Já ao longo dos contornos C_1 e C_5 observa-se que $|z| = \epsilon \ll 1$, o que permite o uso da série de Laurent (2.35d). Para a integral I_1 pode-se escrever $z = \epsilon e^{i\theta}$, resultando

$$\begin{aligned} I_1(\epsilon) &\equiv \int_{C_1} \frac{e^{ikz}}{e^z - 1} dz = \frac{1}{2} \int_{C_1} e^{z/2} e^{ikz} \text{cosech} \left(\frac{z}{2} \right) dz \\ &= \frac{1}{2} \int_{C_1} e^{z/2} e^{ikz} \left(\frac{2}{z} - \frac{z}{12} + \frac{7z^3}{2880} - \dots \right) dz = -i \int_0^{\pi/2} \exp \left[\left(\frac{1}{2} + ik \right) \epsilon e^{i\theta} \right] d\theta - \mathcal{O}(\epsilon^2). \end{aligned}$$

Portanto,

$$\lim_{\epsilon \rightarrow 0} I_1 = -i \frac{\pi}{2}.$$

Já para o contorno C_5 pode-se escrever $z = 2\pi i + \epsilon e^{i\theta}$ e

$$I_5(\epsilon) \equiv \int_{C_5} \frac{e^{ikz}}{e^z - 1} dz = \frac{1}{2} \int_{C_5} e^{z/2} e^{ikz} \text{cosech} \left(\frac{z}{2} \right) dz.$$

Neste caso é necessária a série de Laurent de $\text{cosech } z$ em torno de $z_0 = 2\pi i$, a qual é simplesmente

$$\text{cosech } z = -\frac{1}{z - 2\pi i} + \frac{z - 2\pi i}{6} - \frac{7(z - 2\pi i)^3}{360} + \dots,$$

resultando,

$$I_5(\epsilon) = \frac{1}{2} \int_{C_5} e^{z/2} e^{ikz} \left[-\frac{2}{z - 2\pi i} + \frac{z - 2\pi i}{12} + \dots \right] dz = ie^{-2\pi k} \int_0^{-\pi/2} \exp \left[\left(\frac{1}{2} + ik \right) \epsilon e^{i\theta} \right] d\theta + \mathcal{O}(\epsilon^2).$$

Portanto,

$$\lim_{\epsilon \rightarrow 0} I_5 = -i \frac{\pi}{2} e^{-2\pi k}.$$

Já a integral I_3 pode ser escrita

$$I_3(R) \equiv \int_{C_3} \frac{e^{ikz}}{e^z - 1} dz = ie^{-(1-ik)R} \int_0^{2\pi} \frac{e^{-ky}}{e^{iy} - e^{-R}} dy.$$

Observa-se agora que no limite $R \rightarrow \infty$ a integral é finita, pois o denominador do integrando pode ser substituído por e^{iy} , o que torna a integração trivial. Portanto, $\lim_{R \rightarrow \infty} I_3 = 0$. Finalmente, a integral I_6 pode ser escrita

$$I_6 \equiv \int_{C_6} \frac{e^{ikz}}{e^z - 1} dz = -i \int_{\epsilon}^{2\pi - \epsilon} \frac{e^{-ky}}{e^{iy} - 1} dy = -i \int_{\epsilon}^{2\pi - \epsilon} \frac{e^{-ky} e^{-iy/2}}{e^{iy/2} - e^{-iy/2}} dy = -\frac{1}{2} \int_{\epsilon}^{2\pi - \epsilon} \frac{e^{-ky} e^{-iy/2}}{\text{sen}(y/2)} dy.$$

Assim, pelo teorema dos resíduos,

$$(1 - e^{-2\pi k}) \int_{\epsilon}^R \frac{e^{ikx}}{e^x - 1} dx = \frac{1}{2} \int_{\epsilon}^{2\pi - \epsilon} \frac{e^{-ky} e^{-iy/2}}{\operatorname{sen}(y/2)} dy - I_1 - I_5 - I_3(R)$$

e, portanto, $f(k)$ é dada por

$$\begin{aligned} f(k) &= \lim_{\substack{\epsilon \rightarrow 0 \\ R \rightarrow \infty}} \operatorname{Im} \int_{\epsilon}^R \frac{e^{ikx}}{e^x - 1} dx = -\frac{1}{2(1 - e^{-2\pi k})} \int_0^{2\pi} e^{-ky} dy + \frac{\pi}{2} \frac{1 + e^{-2\pi k}}{1 - e^{-2\pi k}} \\ &= -\frac{1}{2k} + \frac{\pi}{2} \frac{1 + e^{-2\pi k}}{1 - e^{-2\pi k}} = -\frac{1}{2k} + \frac{\pi}{2} \operatorname{coth}(\pi k) \\ &\stackrel{(2.35b)}{\rightarrow} \frac{\pi^2 k}{6} - \frac{\pi^4 k^3}{90} + \frac{\pi^6 k^5}{945} - \dots + \frac{\pi B_{2n}}{(2n)!} (2\pi k)^{2n-1} + \dots \quad (2.46) \end{aligned}$$

Com este resultado, é possível finalmente obter-se I_P :

$$I_P = \frac{\pi^4}{15},$$

resultando na seguinte expressão para a constante de Stefan-Boltzmann,

$$j = \sigma T^4, \text{ sendo } \sigma = \frac{2\pi^5 k_B^4}{15h^3 c^2}.$$

Como um bônus, o resultado (2.46) pode ser usado para fornecer o valor das integrais

$$\int_0^{\infty} \frac{x^{2n+1} dx}{e^x - 1} = (-1)^n \left. \frac{d^{2n+1} f}{dk^{2n+1}} \right|_{k=0} = (-1)^n 2^{2n} \pi^{2(n+1)} \frac{B_{2(n+1)}}{n+1}, \quad (n = 0, 1, 2, \dots).$$

2.10 CONTINUAÇÃO ANALÍTICA

Com frequência, a representação conhecida de uma dada função é válida somente em uma região limitada do plano complexo, como uma série de Laurent com raio de convergência finito, por exemplo. A região de convergência (ou domínio) desta série se estende até a singularidade mais próxima, distinta do ponto em torno do qual se realiza a expansão, que pode ou não conter uma singularidade. Com frequência também é desejável obter-se uma outra forma matemática que represente a mesma função em uma outra região do plano complexo, ou que seja válida em uma porção maior do mesmo. O processo de estender o intervalo de validade de uma representação ou, de forma mais geral, estender a região de domínio de uma função analítica é denominado **continuação analítica**.

Definição 2.1. Um elemento de função analítica $(f(z), D)$ é uma função analítica $f(z)$ no interior de seu domínio de definição D . Um elemento de função (f_2, D_2) é uma continuação analítica de outro elemento (f_1, D_1) se $D_1 \cap D_2 \neq \emptyset$ e se $f_1(z) = f_2(z)$ em $D_1 \cap D_2$.

Em outras palavras, seja $f_1(z)$ uma função analítica da variável z em um domínio D_1 do plano complexo. Supõe-se agora ser possível encontrar uma outra função $f_2(z)$, a qual é analítica em um outro domínio D_2 . Se ocorrer uma intersecção $D_1 \cap D_2$ não nula de ambos os domínios, conforme representado na figura 2.27, e se $f_1(z) = f_2(z)$ em $D_1 \cap D_2$, então se diz que $f_2(z)$ é a continuação analítica de $f_1(z)$ em D_2 , e vice-versa. Da mesma forma, pode existir uma terceira função $f_3(z)$, analítica em D_3 , e se $f_3(z) = f_1(z)$ em $D_3 \cap D_1$ e $f_3(z) = f_2(z)$ em $D_3 \cap D_2$, então $f_3(z)$ pode ser a continuação analítica das anteriores em D_3 . Esta situação também está representada na figura 2.27. Portanto, existe uma função analítica $f(z)$ com o domínio $D = D_1 \cup D_2 \cup D_3$, tal que $f(z) = f_1(z)$ em D_1 , $f(z) = f_2(z)$ em D_2 e $f(z) = f_3(z)$ em D_3 . Nota-se que basta a intersecção entre os domínios ser composta simplesmente por um arco que a continuação analítica existe e é única, desde que as três regiões não englobem um ponto singular ou um ponto de ramificação de uma função plurívoca.

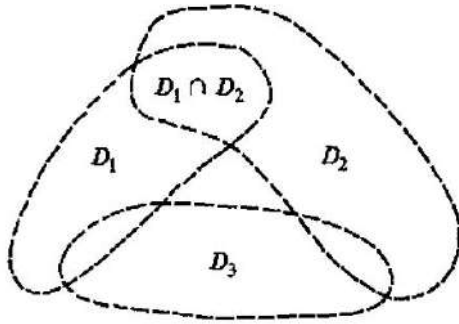


Figura 2.27: Representação de três possíveis continuações analíticas da mesma função $f(z)$.

Contudo, a situação descrita acima nem sempre ocorre. Somente é garantida a existência da função analítica $f(z)$ no domínio $D = D_1 \cup D_2 \cup D_3$ se $D_1 \cap D_2 \cap D_3 \neq \emptyset$. Caso a intersecção ocorra somente entre pares de domínios e se a região entre os domínios D_1, D_2 e D_3 contiver pontos singulares, como por exemplo, pontos de ramificação, então $f_3(z) \neq f_1(z)$, porque a região $D_1 \cap D_3$ será parte do domínio de diferentes folhas de Riemann para cada função. O exemplo a seguir mostra uma situação onde isto ocorre.

Exemplo 2.35. Considera-se o seguinte ramo da função $f(z) = \sqrt{z}$:

$$f_1(z) = \sqrt{r}e^{i\theta/2}, \quad D_1 : (r > 0, 0 < \theta < \pi).$$

Uma continuação analítica de $f_1(z)$ através do eixo real negativo e para o semi-plano inferior é:

$$f_2(z) = \sqrt{r}e^{i\theta/2}, \quad D_2 : \left(r > 0, \frac{\pi}{2} < \theta < 2\pi\right).$$

Claramente, em $D_1 \cap D_2 : (r > 0, \pi/2 < \theta < \pi)$ (segundo e terceiro quadrantes), $f_1(z) = f_2(z)$. Uma continuação analítica de f_2 através do eixo real positivo e para o semi-plano superior pode ser definida então como

$$f_3(z) = \sqrt{r}e^{i\theta/2}, \quad D_3 : \left(r > 0, \pi < \theta < \frac{5\pi}{2}\right).$$

Claramente, agora, em $D_2 \cap D_3 : (r > 0, \pi < \theta < 2\pi)$ (terceiro e quarto quadrantes), $f_3(z) = f_2(z)$, mas em $D_1 \cap D_3 : (r > 0, 0 < \theta < \pi/2)$ (primeiro quadrante), $f_3(z) \neq f_1(z)$; de fato, $f_3(z) = -f_1(z)$. Isto ocorre porque os três domínios circundam o ponto de ramificação na origem.

TEOREMAS DE EXISTÊNCIA E UNICIDADE

Os seguintes teoremas, apresentados sem demonstração, mostram quais são as condições necessárias e suficientes para a existência e unicidade das continuações analíticas.

Teorema 2.28. Se uma função $f(z)$ é analítica em todo o domínio D e $f(z) = 0$ em todos os pontos de uma região $R \subset D$ ou de um arco C , interior a D , então $f(z) = 0$ em todos os pontos de D .

Observação. Este teorema é muito importante porque, em primeiro lugar, ele garante que todas as raízes de $f(z)$ são isoladas. Contudo, a sua importância aqui está no fato de que ele garante a unicidade das continuações analíticas. Sejam as funções $f_1(z)$ e $f_2(z)$ mencionadas na definição 2.1. Definindo-se agora a função $g(z) = f_1(z) - f_2(z)$ em $R = D_1 \cap D_2$, obviamente $g(z) = 0$ em R ; de onde se conclui que $g(z) = 0$ em todo o domínio $D_1 \cup D_2$. Segue então o seguinte teorema de unicidade.

Teorema 2.29. Uma função $f(z)$ que é analítica em todo o domínio D é determinada de forma única sobre D pelos seus valores sobre uma região, ou ao longo de um arco, contidos no interior de D .

Exemplo. A função inteira $f(z) = e^z$ é a única que pode assumir os valores de $f(x) = e^x$, ao longo do eixo real. Além disso, uma vez que e^{-z} também é inteira e $e^x e^{-x} = 1$ ($\forall x \in \mathbb{R}$), a função $h(x) = e^x e^{-x} - 1$ é nula sobre todo o eixo real e, portanto, a única função que representa a continuação analítica de $h(x)$ fora do eixo real é $h(z) = e^z e^{-z} - 1 = 0$. Segue então que a identidade $e^{-z} = 1/e^z$ é válida sobre todo o plano complexo.

Corolário. Como corolário a este teorema, qualquer forma polinomial de funções $f_k(x)$ que satisfaça a identidade

$$P[f_1(x), f_2(x), \dots, f_n(x)] = 0$$

tem a sua forma mantida,

$$P[f_1(z), f_2(z), \dots, f_n(z)] = 0,$$

ao longo de todo o domínio D .

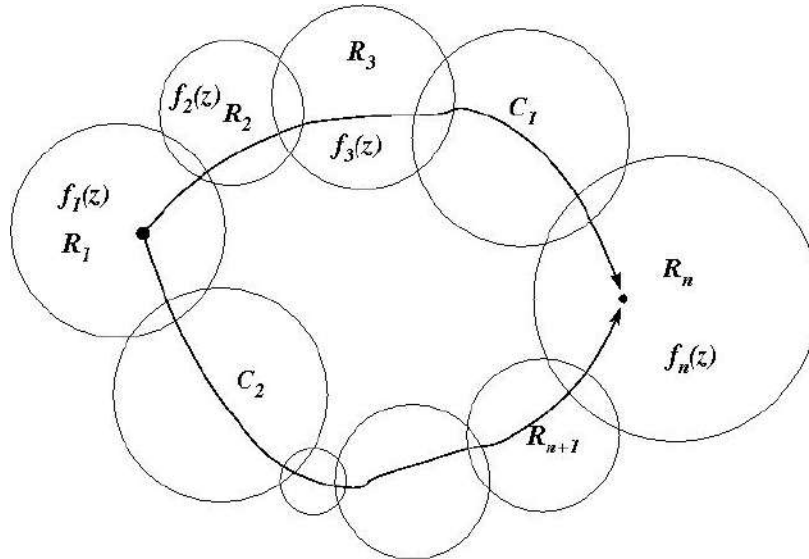


Figura 2.28: Continuação analítica da função $f_1(z)$ para a região R_n por dois caminhos distintos.

Exemplo. Dadas as funções trigonométricas $\sen x$ e $\cos x$, estas satisfazem a forma polinomial $\sen^2 x + \cos^2 x - 1 = 0$. Portanto, a identidade $\sen^2 z + \cos^2 z = 1$ é válida sobre todo o plano complexo.

Teorema 2.30 (Teorema da monodromia). Se uma função $f_1(z)$, definida no domínio R_1 , é continuada analiticamente a uma região R_n ao longo de dois caminhos diferentes, então as duas continuações analíticas serão idênticas se não houver singularidades contidas entre os dois caminhos.

A propriedade descrita no teorema acima é ilustrada na figura 2.28, na qual um ponto na região R_1 é ligado a um outro ponto na região R_n por dois caminhos simples (C_1 e C_2) distintos. Os domínios $R_1, R_2, \dots, R_n, \dots$ podem ser definidos pelos raios de convergência das séries de Taylor que representam a mesma função $f(z)$ em diferentes regiões do plano complexo. Se não houver pontos singulares na região interna aos círculos na figura 2.28, então a continuação analítica de $f(z)$ da região R_1 a R_n pelo caminho C_1 será equivalente à continuação analítica ao longo de C_2 . Portanto,

$$f(z) = \begin{cases} f_1(z), & \text{em } R_1 \\ f_2(z), & \text{em } R_2 \\ \vdots & \vdots \\ f_n(z), & \text{em } R_n \\ \vdots & \vdots \end{cases}$$

As funções $f_1(z), f_2(z), \dots, f_n(z), \dots$ acima são denominadas os **elementos da função analítica completa** $f(z)$.

Exemplo 2.36. Considera-se a função $f_1(z)$ definida por

$$f_1(z) = \sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \dots$$

Sabe-se que esta série de Taylor converge na região $R_1 : |z| < 1$ para a função $f(z) = 1/(1 - z)$, cujo domínio é todo o plano complexo exceto o ponto $z = 1$. Como o domínio de $f(z)$ se intersecciona

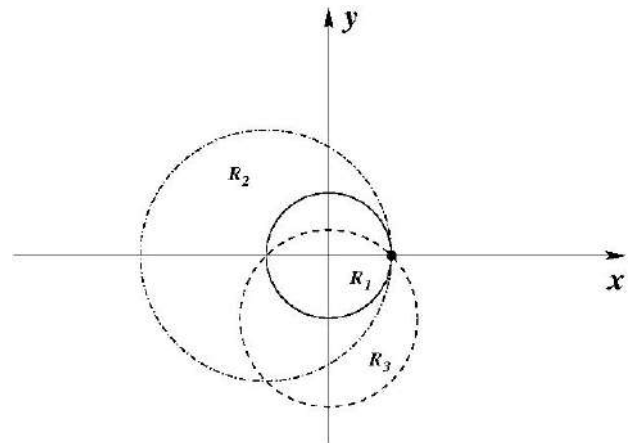


Figura 2.29: Domínios de $f(z) = (1 - z)^{-1}$.

com o domínio de $f_1(z)$, esta última é a única continuação analítica de $f_1(z)$ possível para $|z| \geq 1$ (exceto $z = 1$). Outras possíveis continuações analíticas de $f(z)$ para a região à esquerda de $z = 1$ (ao longo do eixo real) são:

$$f_2(z) = \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{z+1}{2} \right)^n \quad (\text{Região } R_2 : |z+1| < 2), \quad f_3(z) = \frac{1}{1+i} \sum_{n=0}^{\infty} \left(\frac{z+i}{1+i} \right)^n \quad (\text{Região } R_3 : |z+i| < 2).$$

A figura 2.29 ilustra estas três regiões.

REFERÊNCIAS

- ABLOWITZ, Mark J.; FOKAS, Athanassios S. **Complex Variables. Introduction and Applications**. Second. New York: Cambridge, 2003. (Cambridge Texts in Applied Mathematics). 647 pp. ISBN 978-0-521-53429-1. Disponível em: <<http://www.cambridge.org/9780521534291>>.
- CHURCHILL, R. V.; BROWN, J. W.; VERHEY, R. F. **Complex Variables and Applications**. Third. New York: McGraw-Hill, 1976. 332 pp. ISBN 0070108552.
- EISBERG, R. M. **Introdução à Física Moderna**. [S.l.]: Guanabara Dois, 1979. 643 pp.

TEORIA DE GRUPOS ABSTRATOS



ESTE CAPÍTULO tem o objetivo básico de colocar o leitor em contato com um dos ramos mais ativos da física e da matemática dos dias de hoje: a Teoria de Grupos e suas aplicações ao estudo dos fenômenos físicos.

As origens históricas da teoria de grupos, como uma disciplina ou área da matemática, remontam a três áreas distintas: à teoria de números, à teoria de equações algébricas e à geometria. A teoria de números obteve contribuições importantes do matemático e físico suíço Leonhard Euler (1707 – 1783), juntamente com desenvolvimentos no campo da aritmética modular oferecidos pelo matemático e físico alemão Carl Friederich Gauss (1777 – 1855), o qual também identificou os grupos aditivos e multiplicativos de campos quadráticos. Resultados iniciais acerca dos grupos de permutação foram obtidos pelos filósofos naturais italianos Joseph-Louis Lagrange (1736 – 1813) e Paolo Ruffini (1765 – 1822) e pelo matemático norueguês Niels Henrik Abel (1802 – 1829).

Contudo o termo *grupo* foi proposto pela primeira vez pelo matemático francês Évariste Galois (1811 – 1832), o qual estabeleceu a conexão, conhecida como teoria de Galois, entre a teoria de grupos e a teoria de corpos abstratos. Na geometria, grupos são importantes na geometria projetiva e para geometrias não-Euclidianas. Finalmente, os grupos de permutação foram definitivamente estabelecidos pelo matemático britânico Arthur Cayley (1821 – 1895) e pelo francês Augustin-Louis Cauchy (1789 – 1857).

As diferentes contribuições para a teoria de grupos foram unificadas em meados de 1880. Desde então, a mesma fomentou o surgimento de outros campos na matemática tais como a álgebra abstrata e a teoria da representação (tratada no capítulo 5), entre outros.

A importância do estudo da Teoria de Grupos em física surgiu, basicamente, com o livro do matemático, físico e filósofo alemão Hermann Klaus Hugo Weyl (1885 – 1955) intitulado *Gruppentheorie und Quantenmechanik*,¹ publicado em 1928, no qual o autor mostra que existe uma íntima relação entre as leis gerais da Teoria Quântica e a Teoria de Grupos, ao observar que todos os números quânticos, com exceção do número quântico principal n , são índices que caracterizam as representações de grupo.

Uma das grandes aplicações práticas da Teoria de Grupos em física é vista no livro do físico húngaro-norte-americano Eugene Paul Wigner (1902 – 1995) intitulado *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*.² Neste livro, publicado em 1944, evidencia-se que todas as regras da espectroscopia atômica podem ser bem entendidas fazendo-se o estudo das simetrias observadas nos resultados espectroscópicos. Neste estudo, Wigner empregou a teoria originalmente criada por Évariste Galois em 1832.

O grande momento da aplicação em Física da Teoria de Grupos em Partículas Elementares ocorreu em 1961, com a publicação de dois artigos independentes dos físicos: o norte-americano Murray Gell-Mann (1929 –) e o israelense Yuval Ne'eman (1925 – 2006). Nestes trabalhos, admitindo que a Hamiltoniana de Interações Fortes fosse invariante pelo grupo SU(3) os autores obtiveram, entre outros resultados, uma classificação coerente dos hádrons (usando as representações de octetos desse grupo) e a previsão da existência de novas partículas elementares, dentre as quais a partícula Ω^- . Esta partícula foi detectada em 1964, em uma experiência sobre o espalhamento de káons por prótons ($K^- + p \rightarrow \Omega^- + K^+ + K^0$).

Deve se observar que anteriormente, em 1956, o físico japonês Shoichi Sakata (1911 – 1970) havia sem sucesso usado o grupo SU(3) para classificar as Partículas Elementares. Observe-se ainda que em 1964 Gell-Mann e, independentemente, o físico russo-norte-americano George

¹Teoria de Grupos e Mecânica Quântica.

²Teoria de Grupos e sua Aplicação à Mecânica Quântica dos Espectros Atômicos.

Zweig (1937 –) usaram uma outra representação do SU(3) (na forma de tripletos) para prever a existência dos quarks. Estes até o momento não foram observados isoladamente. Um outro grande momento da aplicação em Física da Teoria de Grupos ocorreu no começo da década de 1970 quando o físico norte-americano Nobel Kenneth Geddes Wilson (1936 – 2013) e o físico britânico Michael Ellis Fisher (1931 –) aplicaram o Grupo de Renormalização aos fenômenos críticos (transições de fases), retomando o que havia sido considerado por Gell-Mann e pelo físico norte-americano Francis Eugene Low (1921 – 2007), em 1954.

De um modo geral, a aplicação da Teoria de Grupos a problemas físicos é dividida em duas grandes áreas: considerações sobre simetria e considerações sobre problemas de autovalores. Como exemplo do primeiro tipo, pode-se mencionar o estudo da simetria de um cristal, de fundamental importância na Física da Matéria Condensada (Espectroscopia, Cristalografia, *etc.*). No segundo tipo, um exemplo relevante é o estudo de invariâncias das equações de autovalores resultantes de transformações de coordenadas (translações e rotações).

3.1 DEFINIÇÕES E CLASSIFICAÇÕES INICIAIS

A estrutura matemática denominada *grupo* faz parte de um conjunto amplo de entidades abstratas denominadas *estruturas algébricas*, as quais consistem em determinados conjuntos de objetos, para os quais estão definidas uma ou mais operações finitas, e no estudo das consequentes propriedades matemáticas dessas operações sobre o conjunto. Além dos objetos denominados *grupos* (os quais serão abordados aqui), também fazem parte deste conjunto outros conceitos tais como *semigrupo*, *monóide*, *corpo*, *espaço vetorial*, *álgebra*, *anel*, *módulo*, *etc.* Neste capítulo somente serão definidos e discutidos em detalhes os *grupos*. Algumas das outras estruturas algébricas existentes serão somente exemplificadas na seção 3.7.

Serão feitas inicialmente uma série de definições pertinentes à teoria de grupos, a partir das quais as classificações básicas bem como alguns dos principais teoremas e consequências de interesse para a física serão então apresentados e discutidos.

Definição 3.1 (Grupo). Um conjunto G , composto pelos elementos a, b, c, \dots , é denominado um *grupo* frente a uma dada operação, denominada *produto* e representada pelo símbolo $*$, se seus elementos satisfizerem as seguintes propriedades:

(PG 1) Condição de clausura: $\forall a, b \in G, a * b \in G$.

(PG 2) Condição de associatividade: $\forall a, b, c \in G, (a * b) * c = a * (b * c)$.

(PG 3) Elemento neutro ou unitário ou identidade: $\exists I \in G$ tal que $\forall a \in G, a * I = I * a = a$.

(PG 4) Elemento inverso: $\forall a \in G, \exists a^{-1} \in G$ tal que $a * a^{-1} = a^{-1} * a = I$.

O conjunto G pode ser representado por $G = \{a, b, c, \dots\}$, ao passo que o grupo formado pelo mesmo frente a operação $*$ é representado por

$$\mathcal{G} \doteq \{a, b, c, \dots; *\} = \{G; *\}.$$

Definição 3.2 (Grupo comutativo ou Abelian). Dado um grupo $\mathcal{G} = \{G; *\}$, com $G = \{a, b, c, \dots\}$, se para todos $a, b \in G$ ocorre

$$a * b = b * a,$$

diz-se que o grupo é *comutativo* ou *Abeliano*.³

A partir das definições acima, os seguintes exercícios podem ser propostos.

Exercício 3.1. Na definição 3.1, partindo dos lados esquerdos das propriedades (PG 3) e (PG 4), demonstre os respectivos lados direitos.

Demonstração. Considera-se o produto $(a^{-1} * a) * a^{-1}$. Manipulando-se o mesmo, obtém-se

$$\begin{aligned} (a^{-1} * a) * a^{-1} &\stackrel{(2)}{=} a^{-1} * (a * a^{-1}) \stackrel{(4)}{=} a^{-1} * I \\ &\stackrel{(3)}{=} a^{-1}. \end{aligned} \tag{3.1}$$

³Em homenagem ao matemático norueguês Niels Henrik Abel (1802 – 1829).

Mas $a^{-1} \in \mathcal{G}$, pela propriedade (1). Assim, a partir da propriedade (4), existe um elemento $c \in \mathcal{G}$ o qual é o inverso de a^{-1} , isto é, tal que

$$a^{-1} * c = I.$$

Multiplicando-se agora a primeira e a última expressões em (3.1) à direita com o elemento c , obtém-se

$$[(a^{-1} * a) * a^{-1}] * c = a^{-1} * c = I.$$

Simultaneamente, o lado esquerdo deste resultado resulta em

$$[(a^{-1} * a) * a^{-1}] * c \stackrel{(2)}{=} (a^{-1} * a) * (a^{-1} * c) = (a^{-1} * a) * I \stackrel{(3)}{=} a^{-1} * a.$$

Comparando-se os dois últimos resultados, conclui-se que

$$a^{-1} * a = I,$$

o qual é o lado direito da propriedade (4), como se queria obter.

Dado agora o produto $I * a$,

$$I * a \stackrel{(4)}{=} (a * a^{-1}) * a \stackrel{(2)}{=} a * (a^{-1} * a).$$

Empregando o resultado acima,

$$I * a = a * I \stackrel{(3)}{=} a,$$

o qual é o lado direito da propriedade (3), como se queria obter. \square

Exercício 3.2. Dado o grupo $\mathcal{G} = \{a, b, \dots, I, \dots; *\}$, mostre que o elemento identidade é único.

Demonstração. A resolução é obtida via *reductio ad absurdum*. Suponha, ao contrário, que exista um outro elemento $I' \in \mathcal{G}$ ($I' \neq I$) tal que para todo $a \in \mathcal{G}$ e pela propriedade (PG 3),

$$I' * a = a * I' = a.$$

Neste caso, se $a = I$, então

$$I' * I = I.$$

Por outro lado, também da propriedade (PG 3),

$$a * I = a.$$

Se agora for tomado $a = I'$, então

$$I' * I = I'.$$

Comparando-se ambos os resultados, necessariamente conclui-se que $I = I'$, contrariando a premissa de que $I' \neq I$. Portanto, o elemento I é único. \square

Outro exercício possível consiste na demonstração da unicidade do elemento inverso, isto é, para cada elemento $a \in \mathcal{G}$ existe um único elemento $a^{-1} \in \mathcal{G}$ tal que a propriedade (PG 4) é satisfeita. A partir das definições e exercícios apresentados acima, algumas propriedades adicionais de grupos podem ser demonstradas.

Propriedade 3.1 (Inverso do produto). Dado o grupo $\mathcal{G} = \{G; *\}$, uma vez que para todo $a \in \mathcal{G}$ o seu elemento inverso a^{-1} é único, segue então que, para $a, b, \dots, y, z \in \mathcal{G}$,

$$\begin{aligned} & (a * b * \dots * y * z) * (z^{-1} * y^{-1} * \dots * b^{-1} * a^{-1}) \\ & \stackrel{(2)}{=} (a * b * \dots * y) * (z * z^{-1}) * (y^{-1} * \dots * b^{-1} * a^{-1}) \\ & \stackrel{(4,3)}{=} (a * b * \dots * y) * (y^{-1} * \dots * b^{-1} * a^{-1}) \\ & = (a * b * \dots) * (y * y^{-1}) * (\dots * b^{-1} * a^{-1}) \\ & \vdots \\ & = I. \end{aligned}$$

Por conseguinte, pela propriedade (PG 1), como ambos os produtos iniciais pertencem a \mathcal{G} , segue de (PG 4) que a inversa de um produto é o produto dos inversos na ordem reversa, isto é,

$$(a * b * \dots * y * z)^{-1} = z^{-1} * y^{-1} * \dots * b^{-1} * a^{-1}. \quad (3.2)$$

Propriedade 3.2 (Axioma da divisão). Dado o grupo $\mathcal{G} = \{G; *\}$, para cada par de elementos $a, b \in \mathcal{G}$, existem outros elementos únicos $c, d \in \mathcal{G}$ tais que

$$a * c = b \text{ e } d * a = b.$$

Ou seja, $c = a^{-1} * b$ e $d = b * a^{-1}$. Em geral, $c \neq d$, exceto se o grupo for Abeliano.

Propriedade 3.3 (Lei do cancelamento). Dado o grupo $\mathcal{G} = \{G; *\}$, se existem os elementos $a, b, c \in \mathcal{G}$ tais que

$$a * b = a * c \text{ ou } b * a = c * a,$$

então, necessariamente,

$$b = c.$$

3.1.1 CLASSIFICAÇÃO DOS GRUPOS

As definições apresentadas a seguir permitem estabelecer a classificação básica dos diferentes tipos de grupos.

Definição 3.3 (Ordem do grupo). O número total de elementos de um grupo é denominada a *ordem* ou a *cardinalidade* do grupo. Dado o grupo $\mathcal{G} = \{G; *\}$, a seguinte operação que resulta na ordem do mesmo pode ser definida:

$$g = \text{ord}(\mathcal{G}) = |\mathcal{G}|. \quad (3.3)$$

Desta forma, os grupos podem ser classificados inicialmente em:

Finitos: com um número finito de elementos. Um grupo finito pode ter seus elementos organizados em uma tabela de multiplicações de grupo, discutida na seção 3.2.1.

Infinitos: com um número infinito de elementos. Um grupo infinito pode ser classificado como *discreto* ou *contínuo*.

Definição 3.4 (Grupos infinitos discretos). Um grupo infinito \mathcal{G} é *discreto* se os seus elementos forem *enumeráveis*, isto é, se for possível estabelecer uma relação unívoca entre cada elemento de \mathcal{G} com um elemento do conjunto \mathbb{N}^* . Um exemplo é próprio conjunto dos números inteiros frente a operação de soma algébrica.

Definição 3.5 (Grupos contínuos). Um grupo infinito \mathcal{G} cujos elementos são inumeráveis, mas que podem ser identificados por um conjunto de parâmetros contínuos é denominado *grupo contínuo*.

Pode-se afirmar que dois elementos do grupo estão “arbitrariamente próximos” entre si, de tal forma que eles podem ser distinguidos através da variação de um conjunto de parâmetros $\{\epsilon_i\}$ ($i = 1, \dots$). Ou seja, estes podem ser identificados por funções do tipo $f(\epsilon_1, \epsilon_2, \dots)$ e as operações de produto de grupo e inversão operam sobre essas funções. Uma classificação mais moderna para este tipo de grupo é *grupo topológico*.

As seguintes classificações podem ser aplicadas a um grupo contínuo:

Dimensão: o número de parâmetros necessários para a identificação dos elementos do grupo. Se \mathcal{G} é um grupo contínuo, mas tal que cada elemento do mesmo pode ser identificado por um conjunto finito de n ($= 1, 2, \dots$) parâmetros reais, então esse grupo é denominado *grupo contínuo de n parâmetros* ou *grupo contínuo de dimensão n* . Neste último caso, a a operação $n = \dim(\mathcal{G})$ fornece a dimensão do mesmo.

Compacidade: um grupo contínuo é *compacto* se o espaço de seus parâmetros é compacto, ou seja, se o espaço é fechado (contém todos os seus pontos limites) e limitado (a maior distância entre qualquer ponto do espaço e todos os demais é finita). Uma outra maneira de verificar se o espaço de parâmetros é compacto é se todas as sequências de Cauchy,⁴ inclusive o limite, estiverem contidas no espaço.

A partir destas definições e classificações básicas, alguns exemplos de grupos podem ser agora apresentados.

⁴Ver definição 4.36.

3.1.2 EXEMPLOS DE GRUPOS

Serão apresentados agora alguns exemplos importantes de grupos, bem como as notações mais comuns adotadas para as suas identificações.

3.1.2.1 GRUPOS INFINITOS DISCRETOS

$\{\mathbb{Z}; +\}$: grupo formado pelo conjunto de todos os números inteiros

$$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

frente a operação de soma algébrica. Trata-se de um grupo Abeliano infinito, com $I = 0$ e, para todo $n \in \mathbb{Z}$, seu elemento inverso é simplesmente $-n$.

$\{\mathbb{Q}; +\}$: o conjunto de todos os números racionais

$$\mathbb{Q} = \{n/m, \text{ tais que } n, m \in \mathbb{Z} \text{ e } m \neq 0\},$$

frente a operação de soma algébrica. Trata-se de um grupo Abeliano infinito.

$\{\mathbb{Q}^*; \times\}$: o conjunto de todos os racionais exceto 0, frente a operação de produto algébrico. Grupo Abeliano infinito.

3.1.2.2 GRUPOS CONTÍNUOS COMPACTOS

Grupo ortogonal $O(n)$: grupo formado pelas matrizes $n \times n$ reais ortogonais frente a multiplicação matricial. Se $A \in O(n)$ é uma matriz real quadrada $n \times n$ e \tilde{A} a sua transposta,⁵ então $A\tilde{A} = \tilde{A}A = I_n$, onde I_n é a matriz identidade.

Um exemplo importante do grupo $O(n)$ é formado pelo conjunto das matrizes de rotações ou reflexões de vetores no espaço Euclidiano \mathbb{R}^n .

Dimensão: O número total de elementos em uma matriz $n \times n$ é n^2 . A condição de ortogonalidade $A\tilde{A} = \tilde{A}A = I_n$ gera n^2 relações entre os elementos da matriz; contudo, como $A\tilde{A}$ é uma matriz simétrica, somente metade das relações envolvendo os elementos fora da diagonal principal são distintos. Portanto, o número de relações entre os elementos da matriz é $n^2 - (n^2 - n)/2 = n(n+1)/2$ e, assim, o número de parâmetros independentes em uma matriz do grupo $O(n)$ (i. e., a dimensão do grupo), é igual a

$$\dim [O(n)] = n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1).$$

Compacidade: Se $A \in O(n)$, então, como $\tilde{A}A = I_n$, resulta que $(\tilde{A}A)_{jj} = \sum_{i=1}^n A_{ij}^2 = 1$ ($j = 1, \dots, n$). Portanto, $|A_{ij}| \leq 1$, de onde segue que o espaço de parâmetros é necessariamente fechado e limitado.

Grupo ortogonal especial $SO(n)$: as matrizes do grupo $O(n)$ possuem sempre determinante igual a ± 1 . O grupo *ortogonal especial* $SO(n)$ é composto somente por matrizes ortogonais com determinante igual a $+1$. A dimensão do grupo $SO(n)$ também é $n(n-1)/2$. A demonstração da compacidade do $SO(n)$ segue a linha de demonstração da compacidade do $O(n)$.

O caso particular do grupo de rotações $SO(2)$ é mencionado a seguir.

Grupo de rotações $SO(2)$: o conjunto das matrizes de rotações próprias⁶ de um vetor no \mathbb{R}^3 em torno do eixo z por um certo ângulo θ ⁷ forma um grupo contínuo Abeliano compacto de dimensão 1, denotado por $SO(2)$. Os elementos deste grupo são as matrizes 2×2 $R = R(\theta)$, cujos componentes são

$$R_{11} = \cos \theta \quad R_{12} = \sin \theta \quad R_{21} = -\sin \theta \quad R_{22} = \cos \theta.$$

O parâmetro contínuo do grupo é o ângulo $0 \leq \theta < 2\pi$, o qual identifica cada elemento do grupo e o produto de grupo é a multiplicação matricial. Dados então $R_1 = R(\theta_1), R_2 = R(\theta_2), R_3 = R(\theta_3) \in SO(2)$, de acordo com a definição 3.1,

⁵Os símbolos \tilde{A} ou A^T serão empregados alternativamente para representar a transposta da matriz A .

⁶Ver seção 6.2.1.

⁷Ver figura 6.1a.

1. $R(\theta_1)R(\theta_2) \doteq \begin{cases} R(\theta_1 + \theta_2), & \theta_1 + \theta_2 < 2\pi \\ R(\theta_1 + \theta_2 - 2\pi), & \theta_1 + \theta_2 \geq 2\pi \end{cases} \in SO(2).$
2. $R_3(R_2R_1) = (R_3R_2)R_1.$
3. Existe $I = R(0) = I_2$ tal que $I_2R(\theta) = R(\theta)I_2.$
4. Para todo $R(\theta)$, existe $R(2\pi - \theta)$ tal que $R(\theta)R(2\pi - \theta) = R(2\pi - \theta)R(\theta) = I_2.$

Grupo unitário $U(n)$: grupo formado pelas matrizes $n \times n$ complexas unitárias frente a multiplicação matricial. Se $U \in U(n)$, e U^\dagger é a sua Hermitiana conjugada, então $UU^\dagger = I_n.$

Dimensão: Dada a matriz $U \in U(n)$, se $u_{ij} = \operatorname{Re} u_{ij} + i \operatorname{Im} u_{ij}$ ($i, j = 1, \dots, n$) denotar genericamente os seus elementos, então o número total de parâmetros envolvidos na matriz é $2n^2$. Agora, a condição de matriz unitária leva às relações

$$UU^\dagger = I_n \implies \sum_{k=1}^n u_{ik}u_{jk}^* = \delta_{ij} \implies \begin{cases} \sum_{k=1}^n |u_{ik}|^2 = 1 \\ \sum_{k=1}^n u_{ik}u_{jk}^* = 0, (i \neq j). \end{cases}$$

Ou seja, há n relações para os casos $i = j$ e $2(n^2 - n)/2$ relações distintas para $i \neq j$. Portanto, a dimensão do $U(n)$ é

$$\dim[U(n)] = 2n^2 - [n + (n^2 - n)] = n^2.$$

Compacidade: a demonstração da compacidade do $U(n)$ também segue da compacidade do $O(n)$.

Grupo unitário especial $SU(n)$: as matrizes do grupo $U(n)$ possuem determinante com módulo unitário. O grupo *unitário especial* $SU(n)$ é composto por aquelas matrizes unitárias cujo determinante é igual a $+1$.

3.1.2.3 GRUPOS CONTÍNUOS NÃO COMPACTOS

$\{\mathbb{R}; +\}$: grupo formado pelo conjunto de todos os números reais frente a operação de soma algébrica.

Grupo Abelian, com $I = 0$ e, para todo $x \in \mathbb{R}^*$, a sua inversa é simplesmente $-x$.

$\{\mathbb{R}^*; \times\}$: grupo formado pelo conjunto de todos os números reais exceto 0, frente a operação de produto algébrico.

Grupo Abelian infinito, com $I = 1$ e para todo $x \in \mathbb{R}^*$, $x^{-1} = 1/x$.

$\{\mathbb{C}; +\}$: grupo Abelian formado pelo conjunto de todos os números complexos frente a operação de adição de números complexos.

$\{\mathbb{C}^*; \times\}$: grupo Abelian formado pelo conjunto de todos os números complexos exceto 0, frente a operação de produto de números complexos.

$\mathcal{R}^3 \doteq \{\mathbb{R}^3; +\}$: grupo formado pelo conjunto de todas as ternas ordenadas

$$\mathbb{R}^3 = \{(x_1, x_2, x_3), \text{ onde } x_i \in \mathbb{R} (i = 1, 2, 3)\},$$

frente a operação aditiva “+” tal que, dados $\mathbf{a} = (a_1, a_2, a_3)$ e $\mathbf{b} = (b_1, b_2, b_3)$ pertencentes a \mathbb{R}^3 ,

$$\mathbf{a} + \mathbf{b} \doteq (a_1 + b_1, a_2 + b_2, a_3 + b_3) \in \mathbb{R}^3.$$

Grupo Abelian não compacto de dimensão 3, onde $I = (0, 0, 0) \in \mathcal{R}^3$ e, para cada $\mathbf{a} \in \mathcal{R}^3$, o seu elemento inverso é $-\mathbf{a} \doteq (-a_1, -a_2, -a_3) \in \mathcal{R}^3$.

$\mathcal{C}^3 \doteq \{\mathbb{C}^3; +\}$: grupo formado pelo conjunto de todas as ternas ordenadas

$$\mathbb{C}^3 = \{(z_1, z_2, z_3), \text{ onde } z_i \in \mathbb{C} (i = 1, 2, 3)\},$$

frente a operação aditiva “+” tal que, dados $\mathbf{a} = (a_1, a_2, a_3)$ e $\mathbf{b} = (b_1, b_2, b_3)$ pertencentes a \mathbb{C}^3 ,

$$\mathbf{a} + \mathbf{b} \doteq (a_1 + b_1, a_2 + b_2, a_3 + b_3) \in \mathbb{C}^3.$$

Grupo Abelian não compacto de dimensão 3, onde $I = (0, 0, 0) \in \mathbb{C}^3$ e, para cada $\mathbf{a} \in \mathbb{C}^3$, o seu elemento inverso é $-\mathbf{a} \doteq (-a_1, -a_2, -a_3) \in \mathbb{C}^3$.

Grupo geral linear $GL(n, \mathbb{K})$: grupo formado pelas matrizes $n \times n$ com componentes pertencentes ao corpo⁸ \mathbb{K} (usualmente, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C}) inversíveis, i. e., com determinante não nulo, frente a multiplicação matricial. $GL(n, \mathbb{K})$ é não-Abeliano se $n > 1$ e tem dimensão n^2 .

Grupo de Lorentz \mathcal{L} : trata-se do grupo formado por todas as operações de isometria do espaço de Minkowski que mantêm a origem do referencial fixa. O produto de grupo é novamente a multiplicação matricial. Uma designação mais correta deste grupo é *grupo de Lorentz homogêneo*.

O grupo de Lorentz é formado por todas as matrizes 4×4 L que realizam a transformação de Lorentz⁹ entre dois referenciais inerciais O e O' , deslocando-se com velocidade relativa v e também por todas as rotações espaciais próprias em torno da origem do referencial. A transformação de Lorentz mantém a norma do espaço tempo de Minkowski invariante, ou seja, as matrizes L mantêm invariante a forma quadrática

$$t^2 - x^2 - y^2 - z^2 \xrightarrow{L} t'^2 - x'^2 - y'^2 - z'^2 \quad (c = 1).$$

O grupo de Lorentz é não-Abeliano e de dimensão 6. O grupo é não-compacto porque o espaço topológico de variação do parâmetro v não é compacto (fechado e delimitado), uma vez que $-c < v < c$.

Exemplificando-se com 2 referenciais inerciais O e O' que se deslocam ao longo dos eixos x e x' com velocidade relativa v , as transformações de Lorentz que mantêm invariante a forma quadrática

$$t^2 - x^2 \xrightarrow{L} t'^2 - x'^2$$

são obtidas a partir dos sistemas de equações

$$\begin{cases} x' = \gamma(x - vt) \\ t' = \gamma(t - \beta x/c) \end{cases} \iff \begin{cases} x = \gamma(x' + vt') \\ t = \gamma(t' + \beta x'/c) \end{cases},$$

sendo c a velocidade da luz no vácuo, $\gamma = (1 - \beta^2)^{-1/2}$ e $\beta = v/c$, então a matriz de transformação de Lorentz é

$$\begin{pmatrix} x' \\ t' \end{pmatrix} = L(v) \begin{pmatrix} x \\ t \end{pmatrix}, \text{ sendo } L(v) = \begin{pmatrix} \gamma & -\gamma v \\ -\gamma \beta/c & \gamma \end{pmatrix}.$$

As matrizes $L(v)$ formam um grupo (de dimensão 1) pois, da definição 3.1, constata-se que:

1. $L_1 L_2 = L(v_1) L(v_2) \in \mathcal{L}$.

Demonstração: realizando-se o produto,

$$\begin{aligned} L_1 L_2 &= \begin{pmatrix} \gamma_1 & -\gamma_1 v_1 \\ -\gamma_1 \beta_1/c & \gamma_1 \end{pmatrix} \begin{pmatrix} \gamma_2 & -\gamma_2 v_2 \\ -\gamma_2 \beta_2/c & \gamma_2 \end{pmatrix} \\ &= [\gamma_1 \gamma_2 (1 + \beta_1 \beta_2)] \begin{pmatrix} 1 & -\frac{\beta_1 + \beta_2}{1 + \beta_1 \beta_2} c \\ -\frac{\beta_1 + \beta_2}{1 + \beta_1 \beta_2} \frac{1}{c} & 1 \end{pmatrix}. \end{aligned}$$

A composição de velocidades na relatividade restrita é

$$v_3 = \frac{v_1 + v_2}{1 + v_1 v_2 / c^2} = \frac{\beta_1 + \beta_2}{1 + \beta_1 \beta_2} c.$$

Portanto,

$$\gamma_1 \gamma_2 (1 + \beta_1 \beta_2) = \gamma_3,$$

ou seja,

$$L_1 L_2 = \begin{pmatrix} \gamma_3 & -\gamma_3 v_3 \\ -\gamma_3 \beta_3/c & \gamma_3 \end{pmatrix} = L_3 \in \mathcal{L}.$$

⁸Ver seção 3.7.1.2.

⁹Nos textos em inglês, esta transformação é usualmente denominada *boost*.

2. $L_1(L_2L_3) = (L_1L_2)L_3$.
3. Existe $I = L(0)$.
4. Para todo $L(v)$ existe $L(-v)$ tal que $L(v)L(-v) = L(-v)L(v) = L(0)$.

Finalmente, uma maneira de se verificar a não-compacidade de \mathcal{L} é obtida considerando-se a rapidez $\eta = \tanh^{-1}(v/c)$. Neste caso, a matriz de Lorentz fica

$$L(\eta) = \begin{pmatrix} \cosh \eta & -\sinh \eta \\ -\sinh \eta & \cosh \eta \end{pmatrix}$$

e a variação do parâmetro fica mapeada em

$$v \mapsto \eta \implies -c < v < c \mapsto \mathbb{R}.$$

3.1.2.4 GRUPOS FINITOS

Z_n : grupo formado pelas n raízes da unidade frente a multiplicação algébrica, ou seja, $Z_n = \{z_p (p = 0, \dots, n-1), \times\}$, sendo que os $\{z_p\}$ satisfazem

$$z^n = 1 \implies z_p = e^{i2p\pi/n}, (p = 0, \dots, n-1).$$

Grupo cíclico C_n : grupo formado a partir de um único elemento $a (\neq I)$ a partir de multiplicações sucessivas $a = a^1, b = a^2, c = a^3$, etc.

Grupo de permutações: grupo formado pelas permutações de um conjunto de N objetos.

Grupo cristalográfico: grupo formado pelas operações de simetria de um cristal ou molécula em uma rede cristalina.

Dentre os grupos finitos, os três últimos citados acima são particularmente importantes para a física e serão abordados em maiores detalhes nas próximas seções.

3.2 GRUPOS FINITOS

Grupos finitos são importantes na física porque certas propriedades de sistemas físicos podem ser agrupadas nestes. Um exemplo é o grupo formado pelas operações de simetria de um cristal, o qual consiste na célula básica de uma rede cristalina em um sólido. Os grupos resultantes, denominados *grupos cristalográficos*, serão estudados em mais detalhe na seção 3.4.

Outro grupo importante, tanto para a física quanto para a matemática, é o grupo formado por todas as operações de permutação (ou de parte destas) possíveis em um conjunto com N objetos. Este grupo é denominado *grupo simétrico*, e o mesmo será estudado em mais detalhes na seção 3.2.3. A importância deste grupo é enfatizada pelo teorema de Cayley, o qual mostra que o mesmo pode ser empregado para representar qualquer grupo finito.

Antes de se estudar estes grupos em particular, discute-se a construção das *tabelas de multiplicação de grupo*, as quais são empregadas para representar o grupo finito de forma diagramática.

3.2.1 TABELA DE MULTIPLICAÇÃO DE GRUPO

Trata-se de um diagrama que representa tanto os elementos de um grupo quanto os resultados de todas operações binárias entre os elementos do mesmo.

Definição 3.6 (Tabela de multiplicação de grupo). Dado um grupo $\mathcal{G} = \{G; *\}$ finito, de ordem g , formado a partir do conjunto

$$G = \{I = a_1, a_2, a_3, \dots, a_g\}.$$

Elabora-se uma tabela quadrada contendo, nas primeiras linha e coluna, os elementos de \mathcal{G} na mesma ordem e, no corpo da tabela, os resultados das operações de multiplicação de grupo realizadas de forma matricial, ou seja, de tal forma que se $a_i \in \mathcal{G}$ está na i -ésima linha da tabela

e $a_j \in \mathcal{G}$ está na j -ésima coluna ($i, j = 1, \dots, g$), então o elemento na posição ij será dado por $a_i * a_j$.

Com esta definição, a tabela de multiplicação do grupo \mathcal{G} pode ser visualizada como

*	I	a_2	a_3	\dots	a_g
I	I	a_2	a_3	\dots	a_g
a_2	a_2	$a_2 * a_2$	$a_2 * a_3$	\dots	$a_2 * a_g$
a_3	a_3	$a_3 * a_2$	$a_3 * a_3$	\dots	$a_3 * a_g$
\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
a_g	a_g	$a_g * a_2$	$a_g * a_3$	\dots	$a_g * a_g$

O conhecimento de parte de uma tabela de multiplicação pode também ser utilizado para a dedução do restante da mesma, graças ao *teorema do rearranjo*, mencionado a seguir.

Teorema 3.1 (Teorema do rearranjo). *Seja $\mathcal{G} = \{G; *\}$ um grupo finito de ordem g obtido a partir do conjunto $G = \{I, a_2, a_3, \dots, a_g\}$. Se $a_k \in \mathcal{G}$ ($k = 1, \dots, g$), então cada elemento de \mathcal{G} ocorre uma e somente uma vez na sequência $I * a_k = a_k, a_2 * a_k, a_3 * a_k, \dots, a_k * a_k, \dots, a_g * a_k$.*

Demonstração. Primeiro, demonstra-se que todos os elementos de \mathcal{G} aparecem pelo menos uma vez na sequência acima. Para tanto, se $a_j \in \mathcal{G}$, então devem existir $a_k, a_r \in \mathcal{G}$ ($k, r = 1, \dots, g$) tais que $a_r = a_j * a_k^{-1}$. Logo, $a_j = a_r * a_k$ deve fazer parte da sequência.

A unicidade de a_j na sequência segue da lei do cancelamento (propriedade 3.3). □

A sequência resultante da multiplicação de a_k por todos os elementos de \mathcal{G} , gerando novamente todos os elementos do grupo, porém em outro arranjo de ordenamento, é muitas vezes representada por

$$\mathcal{G} * a_k = \mathcal{G}, \text{ ou, simplesmente, } \mathcal{G}a_k = \mathcal{G}.$$

O teorema do rearranjo pode ser empregado para completar uma tabela de multiplicação com vacâncias, pois, de acordo com o mesmo, nenhum elemento do grupo pode aparecer mais de uma vez ao longo de uma linha ou de uma coluna.

Exemplo 3.1. O teorema do rearranjo é usado para completar a tabela de multiplicação do grupo $\mathcal{G} = \{I, A, B; *\}$, apresentada abaixo:

*	I	A	B
I	I	A	B
A	A		
B	B		

Resolução: De acordo com o teorema do rearranjo, o elemento na posição (2,3): $A * B$ não pode ser nem A nem B , pois ambos já estão presentes ao longo da linha e da coluna, respectivamente. Logo, $A * B = I$. Com isto, o preenchimento do restante da tabela é trivial. A tabela completa, portanto, é:

*	I	A	B
I	I	A	B
A	A	B	I
B	B	I	A

Observa-se também como a multiplicação de cada elemento de \mathcal{G} pelo grupo simplesmente reorganiza o ordenamento de seus elementos: $I * \mathcal{G} = \{I, A, B\}$, $A * \mathcal{G} = \{A, B, I\}$ e $B * \mathcal{G} = \{B, I, A\}$.

Definição 3.7 (Ordem do elemento do grupo). Dado o grupo $\mathcal{G} = \{G; *\}$ finito de ordem g , qualquer elemento $a \in \mathcal{G}$ quando multiplicado por si mesmo, resultando nos elementos

$$a^2 = a * a, a^3 = a * a^2, \dots,$$

irá necessariamente resultar na identidade $I \in \mathcal{G}$, após um máximo de $g - 1$ multiplicações.

Representando estas operações de multiplicação por

$$a^n = \underbrace{a * \dots * a}_{n \text{ vezes}} \quad (n \geq 1), \tag{3.4a}$$

então deve existir um natural $1 \leq m \leq g$, o qual é o menor número tal que $a^m = I$; ou, em outras palavras, a identidade deve resultar após $m - 1$ multiplicações de a por si mesmo. Este número é denominado a *ordem do elemento do grupo*. A obtenção deste número pode ser representada pela operação

$$m = \text{ord}(a) = |a|,$$

para todo $a \in \mathcal{G}$.

Prosseguindo com a convenção de “potências” do elemento $a \in \mathcal{G}$, se “potências positivas” a^n ($n > 0$) são definidas por (3.4a), então, por extensão,

$$a^0 = I \quad (3.4b)$$

$$a^{-n} = (a^{-1})^n \quad (3.4c)$$

$$a^n * a^m = a^{n+m} \quad (3.4d)$$

$$(a^n)^m = a^{nm}, \quad (n, m \in \mathbb{Z}), \quad (3.4e)$$

onde $a^{-1} \in \mathcal{G}$ é o inverso de a .

Para qualquer grupo $\mathcal{G} = \{G; *\}$, a identidade $I \in \mathcal{G}$ possui sempre a ordem 1, $\text{ord}(I) = 1$, pois $I^1 = I$. Por outro lado, se existe $b \in \mathcal{G}$ tal que $b * b = I$, isto é, b é o seu próprio inverso, então $\text{ord}(b) = 2$ sempre. Para todos os outros elementos $a \in \mathcal{G}$ ($a \neq I$ e $a \neq b$), resulta $\text{ord}(a) > 2$.

Definição 3.8 (Período do elemento do grupo). Dado o grupo $\mathcal{G} = \{G; *\}$ finito de ordem g , seja $x \in \mathcal{G}$ tal que $\text{ord}(x) = n$ ($n \geq 1$). A sequência $x, x^2, x^3, \dots, x^n = I$ é denominada o *período de x* .

3.2.2 GRUPO CÍCLICO \mathcal{C}_n

Dado o grupo $\mathcal{G} = \{G; *\}$ finito e de ordem $g = \text{ord}(\mathcal{G})$. Dado um elemento $a \in \mathcal{G}$ ($a \neq I$), se todos os seus demais elementos podem ser gerados a partir de a pela aplicação sucessiva das multiplicações (3.4), então o conjunto de elementos G pode ser escrito como

$$G = \{a, a^2, a^3, \dots, a^{g-1}, I = a^g\}$$

e o grupo é denominado *cíclico*. O elemento a é denominado o **gerador** de \mathcal{G} e este fato é muitas vezes representado por $\mathcal{G} = \langle a \rangle = \{a^n \mid n = 1, 2, \dots, g\}$.

Um grupo cíclico é sempre Abelianiano.

As notações usualmente empregadas para representar um grupo cíclico são: C_g , $\langle a \rangle$ ou $[a]$, sendo g a ordem do grupo e a o seu gerador.

Definição 3.9 (Grau e base do grupo). Dado um grupo $\mathcal{G} = \{G; *\}$ e os elementos $\{a_1, \dots, a_m\} \in \mathcal{G}$. O número mínimo de geradores necessários para reproduzir a tabela de multiplicações de \mathcal{G} é denominado o *grau do grupo*. Por sua vez, o menor subconjunto de G também capaz de reproduzir a tabela de multiplicações de \mathcal{G} é denominada uma *base do grupo*. Observa-se que um grupo pode conter mais de uma base.

3.2.3 O GRUPO SIMÉTRICO S_n

Considera-se um conjunto χ composto por n objetos quaisquer. Cada elemento desse conjunto pode ser identificado por χ_i ($i = 1, \dots, n$), por exemplo, ou simplesmente pelo índice posicional i . Neste caso, pode-se escrever

$$\chi = \{\chi_1, \chi_2, \dots, \chi_n\} \text{ ou, simplesmente, } \chi = \{1, 2, \dots, n\}.$$

O *grupo simétrico de grau n* sobre o conjunto $\chi = \{1, 2, \dots, n\}$ corresponde ao grupo cujos elementos são iguais a todas as operações de permutação possíveis ao ordenamento original de χ .

O grupo simétrico sobre um conjunto χ qualquer é denotado de diferentes maneiras na literatura: S_χ , Σ_χ , \mathfrak{S}_χ ou $\text{Sym}(\chi)$. Se o conjunto é finito, $\chi = \{1, 2, \dots, n\}$, então o grupo simétrico de grau n sobre χ é denotado por: S_n , Σ_n , \mathfrak{S}_n , ou $\text{Sym}(n)$. Neste caso, a ordem do grupo simétrico finito é simplesmente o número total de permutações possíveis sobre as posições dos elementos de χ , ou seja,

$$\text{ord}(S_n) = |S_n| = n!.$$

Dado o ordenamento original dos elementos de χ , denotado por $\chi = \{1, 2, \dots, n\}$, uma operação de permutação sobre χ irá rearranjar os objetos contidos no mesmo de uma determinada maneira. O ordenamento final obtido por esta operação será identificado por uma sequência numérica contendo novamente os índices $i = 1, \dots, n$, porém num ordenamento que indica a posição final de um dado objeto em relação à inicial.

Estas permutações podem ser representadas com um total de $|S_n|$ operadores através da notação de duas linhas de Cauchy,

$$\pi_p = \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix}, \tag{3.5}$$

onde $p_j = i$, com $i, j = 1, \dots, n$. A primeira linha de π_p indica o ordenamento inicial dos elementos de χ (antes da permutação), enquanto que a segunda linha indica o ordenamento em relação às posições iniciais.

Há ao todo n índices p_j os quais assumem valores entre 1 e n de forma excludente, ou seja, não é possível ocorrer $p_1 = p_2$, por exemplo. Se $p_1 = i$, isto indica que o objeto que estava originalmente na i -ésima posição passou a ocupar a primeira posição no novo ordenamento dos elementos de χ . Por sua vez, $p_2 = \ell$ ($\ell = 1, \dots, n$, mas $\ell \neq i$) indica que o objeto que estava originalmente na ℓ -ésima posição passou a ocupar a segunda posição, e assim consecutivamente.

A operação de permutação π_p sobre χ pode ser interpretado como uma *bijecção* do tipo

$$\chi \xrightarrow{\pi_p} \chi,$$

ou seja, um mapeamento de χ sobre si mesmo. A aplicação de π_p sobre χ resulta no mesmo conjunto de objetos, porém rearranjados. Este “novo” conjunto pode ser identificado por ψ , por exemplo, para ser distinguido de χ . Este procedimento pode ser representado por

$$\psi = \pi_p \chi.$$

Exemplo 3.2 (O grupo S_3). Considere um conjunto de 3 objetos, $\chi = \{1, 2, 3\}$. O número total de permutações possíveis sobre χ é igual a $3! = 6$. Esses operadores podem ser representados por

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

Nota-se que π_1 mantém o ordenamento original inalterado. A figura 3.1 ilustra todas essas permutações na ordem de operadores apresentada acima.

O conjunto de operadores de permutação de 3 objetos é definido então como

$$S_3 \doteq \{\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6\}.$$

Posteriormente, no exercício 3.4, será demonstrado que este conjunto forma um grupo.

Após se realizar duas permutações consecutivas em χ , o ordenamento final pode ser descrito na forma de um operador do tipo (3.5) através de uma *composição de permutações*. Se π_a e π_b são dois operadores do tipo (3.5) e estes são aplicados consecutivamente sobre χ , o ordenamento final pode ser descrito na forma de um terceiro operador π_c , obtido a partir da operação

$$\pi_b(\pi_a \chi) \equiv (\pi_b \circ \pi_a) \chi \equiv \pi_b \pi_a \chi = \pi_c \chi. \tag{3.6}$$

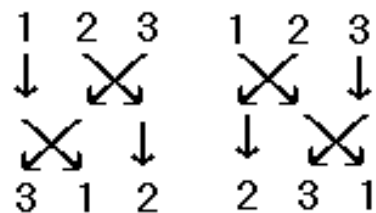
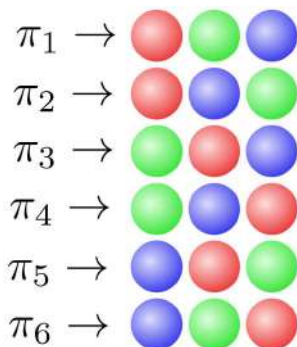


Figura 3.1: As 6 permutações possíveis sobre um conjunto de 3 objetos.

Figura 3.2: Os ordenamentos finais de χ após as composições $\pi_3 \circ \pi_2 = \pi_5$ e $\pi_2 \circ \pi_3 = \pi_4$.

Desde já chama-se a atenção que, em geral, $\pi_b \circ \pi_a \neq \pi_a \circ \pi_b$.

A composição de permutações é um caso particular de uma *composição de funções bijetoras*, ou seja, a operação resultante da aplicação de uma função bijetora à imagem de uma outra função bijetora. A composição de permutações sobre o conjunto χ pode ser representada por

$$\chi \xrightarrow{\pi_a} \chi \xrightarrow{\pi_b} \chi, \text{ ou seja, } \chi \rightsquigarrow \pi_a \chi \rightsquigarrow \pi_b(\pi_a \chi).$$

Exercício 3.3. Considere o mesmo conjunto de 3 objetos, $\chi = \{1, 2, 3\}$ e os operadores de permutação empregados no exemplo 3.2. Realize as composições $\pi_3 \circ \pi_2$ e $\pi_2 \circ \pi_3$ e identifique os operadores resultantes.

Resolução. Realizando-se as permutações, resulta

$$\begin{aligned} (\pi_3 \circ \pi_2)\chi &= \pi_3(\pi_2\chi) = \pi_3 \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \{1, 2, 3\} \right] \\ &= \pi_3 \{1, 3, 2\} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \{1, 3, 2\} = \{3, 1, 2\} = \pi_5\chi. \end{aligned}$$

Ou seja, $\pi_3 \circ \pi_2 = \pi_5$. Por sua vez,

$$\begin{aligned} (\pi_2 \circ \pi_3)\chi &= \pi_2 \left[\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \{1, 2, 3\} \right] \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \{2, 1, 3\} = \{2, 3, 1\} = \pi_4\chi. \end{aligned}$$

Ou seja, $\pi_2 \circ \pi_3 = \pi_4$. A figura 3.2 ilustra estas composições.

3.2.3.1 VERIFICAÇÃO DOS AXIOMAS DE GRUPO

O grupo simétrico S_n sobre o conjunto $\chi = \{1, 2, \dots, n\}$ será formado, portanto, não pelos n objetos que compõe χ , mas sim pelos $|S_n| = n!$ operadores de permutação π_p definidos em (3.5), frente a composição de permutações (3.6), a qual é a operação de multiplicação de grupo. Ou seja,

$$S_n = \{\pi_1, \pi_2, \dots, \pi_{n!}; \circ\}.$$

Para verificar que S_n é de fato um grupo, este deve satisfazer os axiomas apresentados na definição 3.1. Ou seja:

1. **Clausura.** Dados $\pi_a, \pi_b \in S_n$, a composição de permutações $\pi_a \circ \pi_b$ irá simplesmente gerar um rearranjo dos elementos de χ . Portanto, $\pi_a \circ \pi_b \in S_n$.
2. **Associatividade.** A condição de associatividade é satisfeita, porque a composição de funções bijetoras é uma operação associativa.
3. **Identidade.** A operação trivial de permutação

$$I = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \text{ tal que } I\chi = \chi$$

é o elemento identidade de S_n , pois para todo $\pi \in S_n$, $I \circ \pi = \pi \circ I$.

4. **Elemento inverso.** Toda bijeção possui uma função inversa que desfaz a operação inicial. Em consequência, para toda permutação $\pi \in S_n$ existe um elemento $\pi^{-1} \in S_n$ tal que $\pi^{-1} \circ \pi = \pi \circ \pi^{-1} = I$.

Exercício 3.4. Demonstre que o conjunto S_3 forma um grupo, obtenha a ordem de cada elemento e os períodos do grupo e construa a sua tabela de multiplicação.

Resolução. Os elementos de S_3 já foram identificados no exemplo 3.2, com $\pi_1 = I$. Realizando-se todas as composições de permutações possíveis, é fácil verificar que

$$\begin{array}{cccccc} II = I & \pi_2 I = \pi_2 & \pi_3 I = \pi_3 & \pi_4 I = \pi_4 & \pi_5 I = \pi_5 & \pi_6 I = \pi_6 \\ I\pi_2 = \pi_2 & \pi_2\pi_2 = I & \pi_3\pi_2 = \pi_5 & \pi_4\pi_2 = \pi_6 & \pi_5\pi_2 = \pi_3 & \pi_6\pi_2 = \pi_4 \\ I\pi_3 = \pi_3 & \pi_2\pi_3 = \pi_4 & \pi_3\pi_3 = I & \pi_4\pi_3 = \pi_2 & \pi_5\pi_3 = \pi_6 & \pi_6\pi_3 = \pi_5 \end{array}$$

$$\begin{array}{llllll}
 I\pi_4 = \pi_4 & \pi_2\pi_4 = \pi_3 & \pi_3\pi_4 = \pi_6 & \pi_4\pi_4 = \pi_5 & \pi_5\pi_4 = I & \pi_6\pi_4 = \pi_2 \\
 I\pi_5 = \pi_5 & \pi_2\pi_5 = \pi_6 & \pi_3\pi_5 = \pi_2 & \pi_4\pi_5 = I & \pi_5\pi_5 = \pi_4 & \pi_6\pi_5 = \pi_3 \\
 I\pi_6 = \pi_6 & \pi_2\pi_6 = \pi_5 & \pi_3\pi_6 = \pi_4 & \pi_4\pi_6 = \pi_3 & \pi_5\pi_6 = \pi_2 & \pi_6\pi_6 = I.
 \end{array}$$

Portanto, as condições de clausura, associatividade e existência da identidade são satisfeitas. Os elementos inversos são os seguintes:

$$I^{-1} = I \quad \pi_2^{-1} = \pi_2 \quad \pi_3^{-1} = \pi_3 \quad \pi_4^{-1} = \pi_5 \quad \pi_5^{-1} = \pi_4 \quad \pi_6^{-1} = \pi_6.$$

A condição de existência dos elementos inversos foi verificada. As ordens dos elementos de S_3 são:

$$|I| = 1 \quad |\pi_2| = 2 \quad |\pi_3| = 2 \quad |\pi_4| = 3 \quad |\pi_5| = 3 \quad |\pi_6| = 2.$$

Assim, os períodos de S_3 são:

$$I, \quad \{\pi_2, I\}, \quad \{\pi_3, I\}, \quad \{\pi_4, \pi_5, I\}, \quad \{\pi_6, I\}.$$

Finalmente, a tabela de multiplicação de grupo é:

\circ	I	π_2	π_3	π_4	π_5	π_6
I	I	π_2	π_3	π_4	π_5	π_6
π_2	π_2	I	π_4	π_3	π_6	π_5
π_3	π_3	π_5	I	π_6	π_2	π_4
π_4	π_4	π_6	π_2	π_5	I	π_3
π_5	π_5	π_3	π_6	I	π_4	π_2
π_6	π_6	π_4	π_5	π_2	π_3	I

Observe como o teorema do rearranjo é obedecido pela tabela acima.

3.2.3.2 NOTAÇÃO DE CICLOS

Como o número total de operações de permutação possíveis em um conjunto de n objetos é igual a $n!$, a ordem do grupo S_n cresce muito rapidamente com n . A *notação de ciclos* dos operadores $\pi \in S_n$ simplifica a identificação das operações possíveis e também facilita a construção de seus subgrupos.

TRANSPOSIÇÕES. Uma *transposição* é uma permutação que atua somente sobre dois objetos do conjunto χ , com a qual estes objetos têm suas posições trocadas no ordenamento original de χ , mantendo os demais objetos fixos.

Dado o conjunto χ de n objetos, a transposição (mk) ($m, k \leq n$) troca a posição do m -ésimo objeto pela posição do k -ésimo objeto, e vice-versa. Esta notação simplifica a representação do operador de permutação quando este realiza somente uma transposição. Ou seja,

$$\text{se } \pi = \begin{pmatrix} 1 & 2 & \dots & m & \dots & k & \dots & n \\ 1 & 2 & \dots & k & \dots & m & \dots & n \end{pmatrix}, \text{ pode-se escrever } \pi = (mk).$$

A composição de transposições entre (mk) e $(r\ell)$ (com $m, k, r, \ell \leq n$) pode ser representada por $(mk) \circ (r\ell)$ ou, simplesmente, $(mk)(r\ell)$, sempre mantendo a convenção direita \rightarrow esquerda na ordem das permutações. É possível verificar que o grupo S_n pode ser completamente gerado através de composições das $n - 1$ transposições $(12), (13), \dots, (1n)$.

Se uma permutação consistir em um número par de transposições, ela é denominada uma *permutação par*. Se consistir em um número ímpar, é chamada de *permutação ímpar*. A composição de duas permutações pares ou ímpares resulta em uma permutação par, ao passo que a composição de uma permutação par com uma permutação ímpar resulta em uma permutação ímpar.

Exemplo 3.3. Os elementos de S_3 definidos no exemplo 3.2 podem ser construídos pelas transposições (12) e (13) da seguinte maneira:

$$\begin{array}{lll}
 \pi_1 = (12)(12) & \pi_2 = (12)(13)(12) & \pi_3 = (12) \\
 \pi_4 = (12)(13) & \pi_5 = (13)(12) & \pi_6 = (13).
 \end{array}$$

Portanto, $\pi_1 = I$, π_4 e π_5 são permutações pares, enquanto que π_2 , π_3 e π_6 são permutações ímpares.

CICLOS. Dados o conjunto de n objetos $\chi = \{\chi_1, \dots, \chi_n\}$ e o grupo simétrico S_n , composto por todas as permutações de χ . Uma particular permutação aplicada a χ é denominada um *ciclo* se esta atua sobre um subconjunto $\sigma \subseteq \chi$, permutando as posições dos elementos de σ de uma maneira cíclica, mantendo os demais elementos de $\chi \setminus \sigma$ fixos. O conjunto σ é denominado a **órbita** do ciclo.

Dado o grupo S_n , de ordem $n!$, um determinado elemento $\pi \in S_n$ pode ser representado genericamente, na notação de Cauchy, por

$$\pi = \begin{pmatrix} 1 & 2 & \dots & i & \dots & n \\ p_1 & p_2 & \dots & p_i & \dots & p_n \end{pmatrix}.$$

Define-se agora a operação $\pi(i) = p_i$, a qual indica que a i -ésima posição no ordenamento de χ passa a ser ocupada pelo objeto que estava na posição p_i . Ou seja, $\pi(1) = p_1, \pi(2) = p_2, \dots, \pi(n) = p_n$.

Um **ciclo de extensão k** (ou **ciclo- k**) é uma permutação $\pi \in S_n$ para a qual existe um elemento $x \in \chi$ tal que os únicos elementos movidos pela permutação são $x, \pi(x), \pi^2(x), \dots, \pi^k(x) = x$.

Como exemplo, observa-se que a permutação $\psi \in S_5$ dada por

$$\psi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 3 & 5 \end{pmatrix},$$

contém um ciclo de extensão 3, uma vez que, dado $\chi = \{1, 2, 3, 4, 5\}$,

$$\chi \xrightarrow{\psi\chi} \{4, 2, 1, 3, 5\} \xrightarrow{\psi^2\chi} \{3, 2, 4, 1, 5\} \xrightarrow{\psi^3\chi} \{1, 2, 3, 4, 5\}.$$

Ou seja,

$$\begin{array}{lll} \psi(1) = 4, & \psi^2(1) = \psi(4) = 3, & \psi^3(1) = \psi(3) = 1 \\ \psi(3) = 1, & \psi^2(3) = \psi(1) = 4, & \psi^3(3) = \psi(4) = 3 \\ \psi(4) = 3, & \psi^2(4) = \psi(3) = 1, & \psi^3(4) = \psi(1) = 4. \end{array}$$

Observa-se que somente os objetos 1, 3 e 4 são permutados por ψ e de uma maneira cíclica, sendo os objetos 2 e 5 mantidos fixos.

Para escrever um operador π na notação de ciclos, o procedimento é o seguinte:

1. Selecione um objeto qualquer $x \in \chi$, abra parênteses e escreva: $(x$
2. Trace então a órbita de x , i. e., escreva os valores das aplicações sucessivas de π : $(x \pi(x) \pi^2(x) \dots$
3. Repita o procedimento até obter novamente x ; feche então o parênteses sem repetir o valor. Para um ciclo- k resulta então: $(x \pi(x) \dots \pi^{k-1}(x))$.
4. Reinicie o procedimento com um elemento $y \in \chi$ que não esteja na órbita de x ; ou seja: $(x \pi(x) \dots \pi^{k-1}(x)) (y \dots$
5. Repita até que todos os elementos de χ resultem escritos em ciclos.

Empregando este procedimento, o operador ψ acima é denotado, na *notação de ciclos*, por $\psi = (143) = (431) = (314)$.

O ordenamento dos termos no operador pode ser alterado de forma cíclica.

Na notação de ciclos, a ação $\pi\chi$ corresponde à troca cíclica nas posições dos objetos em χ de acordo com a ordem indicada por π , no sentido direita \mapsto esquerda; ou seja,

$$\pi = \begin{pmatrix} p_1 & p_2 & \dots & p_n \end{pmatrix}.$$

Por exemplo, a ação de $\psi = (143)$ resulta em

$$\psi\chi = \begin{pmatrix} 1 & 4 & 3 \end{pmatrix} \{1, 2, 3, 4, 5\} = \{4, 2, 1, 3, 5\}.$$

Um ciclo deve ter uma extensão $k \geq 2$, pois $k = 1$ significa que o objeto não é trocado de lugar. Entretanto, para a permutação ψ acima, pode-se acrescentar os símbolos (2) e (5), $\psi =$

(143) (2) (5), os quais indicam que estes objetos permanecem fixos. Um ciclo de extensão 2 é uma transposição, como definida anteriormente.

Dois ciclos são **disjuntos** se estes movem subconjuntos disjuntos de elementos de χ . Como dois ciclos disjuntos comutam, todos os elementos de S_n podem ser escritos como composições de ciclos disjuntos.

Por exemplo, para o ciclo $\psi \in S_5$ definido acima,

$$\psi = (143) (2) (5) = (2) (5) (143) = (2) (143) (5) = (143),$$

sendo que a última notação omite os objetos que permanecem fixos. O elemento identidade de S_n pode ser representado por $I = (1) (2) \cdots (n)$ ou simplesmente por $I = ()$, mas a extensão da identidade é $k = 1$.

Como qualquer permutação arbitrária, um ciclo de extensão maior que 2 sempre pode ser decomposto em uma composição de ciclos menores de diversas formas. Sendo $\zeta = (i_1 i_2 i_3 \dots i_{k-1} i_k) \in S_n$ um ciclo de extensão k ($2 < k \leq n$, $i_j = 1, \dots, n$, $j = 1, \dots, k$), as seguintes propriedades são válidas:

- A ordem do ciclo é a sua extensão; isto é, $\zeta^k = I$.
- O ordenamento das posições em ζ pode ser alterada de forma cíclica $k - 1$ vezes:

$$\zeta = \underbrace{(i_1 i_2 i_3 \dots i_{k-1} i_k)}_{k \text{ termos}} = (i_2 i_3 \dots i_{k-1} i_k i_1) = \cdots = (i_k i_1 i_2 i_3 \dots i_{k-1}).$$

- O ciclo ζ é igual à composição de dois ciclos, os quais possuem em comum a posição i_j . Um ciclo contém as posições 1 até j , enquanto que o outro ciclo parte de j e contém as posições restantes. A composição de ciclos menores é então realizada na ordem inversa:

$$\zeta = (i_1 \dots i_j \dots i_k) = (i_j \dots i_k) (i_1 \dots i_j).$$

- Um ciclo de extensão $k > 2$ sempre pode ser expresso como a composição de transposições que possuem como posição em comum i_1 ou i_k . Ou seja, o ciclo ζ pode ser decomposto:

$$\zeta = (i_1 i_2 i_3 \dots i_k) = \begin{cases} (i_1 i_2) (i_1 i_3) \cdots (i_1 i_k) \\ \text{ou} \\ (i_1 i_k) (i_2 i_k) \cdots (i_{k-1} i_k). \end{cases}$$

- Dada a permutação ζ , a permutação inversa ζ^{-1} , tal que $\zeta \circ \zeta^{-1} = \zeta^{-1} \circ \zeta = I$, é

$$\zeta^{-1} = (i_k i_{k-1} \dots i_2 i_1).$$

- A paridade do ciclo é obtida diretamente de sua extensão. Se P_ζ é a paridade da permutação ζ , então

$$P_\zeta = (-1)^{k-1} = \begin{cases} +1, & \text{permutação par} \\ -1, & \text{permutação ímpar.} \end{cases}$$

O elemento identidade é sempre uma permutação par.

- A ordem de qualquer permutação é igual à sua extensão, pois se $x \in \chi$ é um elemento movido pela permutação π de extensão k , então a sequência $x, \pi(x), \pi^2(x), \dots, \pi^k(x) = x$ implica que $\pi^k = I$.

Exemplo 3.4. Os elementos de S_3 dados no exemplo 3.2 podem ser representados na notação de ciclos por:

$$\begin{array}{lll} \pi_1 = (1) (2) (3) = () & \pi_2 = (1) (23) = (23) & \pi_3 = (3) (12) = (12) \\ \pi_4 = (123) = (12) (13) & \pi_5 = (132) = (13) (12) & \pi_6 = (2) (13) = (13). \end{array}$$

Exercício 3.5. Considere as seguintes permutações contidas em S_6 :

$$\theta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 5 & 6 & 1 & 4 & 3 \end{pmatrix} \qquad \phi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 3 & 6 & 2 & 1 \end{pmatrix}.$$

Expresse estes elementos na notação por ciclos, bem como as composições $\theta \circ \phi$ e $\phi \circ \theta$.

3.3 SUBGRUPOS, CLASSES LATERAIS E DE CONJUGAÇÃO

Nesta seção serão discutidos alguns teoremas de subdivisões de um grupo finito genérico e suas consequências. Contudo, as definições apresentadas e suas consequências podem também ser aplicadas a grupos infinitos, o que será feito em seções posteriores.

3.3.1 SUBGRUPOS

A principal subdivisão possível de um grupo é um *subgrupo* do mesmo.

Definição 3.10 (Subgrupo). Dado um grupo $\mathcal{G} = \{G; *\}$, um subgrupo \mathcal{H} de \mathcal{G} é formado a partir de um subconjunto $H \subseteq G$ que forma um grupo sob o mesmo produto de grupo $*$. Isto é, o conjunto H deve satisfazer os axiomas de grupo (definição 3.1). Nestas condições o subgrupo $\mathcal{H} = \{H; *\}$ é denotado por $\mathcal{H} \subseteq \mathcal{G}$. O grupo \mathcal{G} é denominado o *sobreggrupo* de \mathcal{H} .

Um subgrupo \mathcal{H} é denominado **próprio** se $\mathcal{H} \subset \mathcal{G}$. Por outro lado, qualquer grupo \mathcal{G} possui dois subgrupos **triviais** ou **impróprios**: $\mathcal{H} = \{I; *\}$ ou $\mathcal{H} = \mathcal{G}$.

Dado o grupo $\mathcal{G} = \{G; *\}$ e um subgrupo $\mathcal{H} \subseteq \mathcal{G}$, algumas propriedades simples de subgrupos podem ser destacadas:

- Se $I \in \mathcal{G}$ é o elemento identidade, então, necessariamente, $I \in \mathcal{H}$.
- Se \mathcal{G} é Abeliano, então \mathcal{H} também o é. A recíproca não é verdadeira.
- Dado qualquer elemento $a \in \mathcal{G}$ de ordem $n = |a|$, o período de a (definição 3.8) forma o subgrupo cíclico

$$\langle a \rangle \subseteq \mathcal{G} : \langle a \rangle = \{a, a^2, \dots, a^n = I; *\}.$$

- O subconjunto de elementos de \mathcal{G} que pertencem a todos os subgrupos de \mathcal{G} forma um subgrupo por si mesmo, embora este conjunto possa ser composto somente pela identidade.

Com relação aos subgrupos cíclicos próprios formados a partir dos períodos de determinados elementos do grupo, uma nova definição é importante, a qual explicita o número mínimo de geradores capazes de reproduzir o sobreggrupo.

Dado um grupo $\mathcal{G} = \{G; *\}$ e o elemento $a_1 \in \mathcal{G}$ que gera o subgrupo cíclico próprio $\langle a_1 \rangle \subseteq \mathcal{G}$. Se $\langle a_1 \rangle \subset \mathcal{G}$, então deve existir outro elemento $a_2 \in \mathcal{G}$ ($a_2 \neq I$), mas $a_2 \notin \langle a_1 \rangle$, tal que $\langle a_2 \rangle \subset \mathcal{G}$. Se $\langle a_1 \rangle \cup \langle a_2 \rangle \neq \mathcal{G}$, então deve existir $a_3 \in \mathcal{G}$ ($a_3 \neq I$) tal que $a_3 \notin \langle a_1 \rangle \cup \langle a_2 \rangle$, o qual gera o subgrupo próprio $\langle a_3 \rangle \subset \mathcal{G}$. E assim por diante, até que existam m subgrupos próprios $\langle a_i \rangle$ ($j = 1, \dots, m$) tais que $\langle a_1 \rangle \cup \langle a_2 \rangle \cup \dots \cup \langle a_m \rangle = \mathcal{G}$.

EXEMPLOS DE SUBGRUPOS

Serão apresentados alguns exemplos de subgrupos obtidos a partir dos exemplos de grupos apresentados até o momento.

Grupo de permutações P_{nm} . Dado o grupo simétrico S_n um subgrupo próprio $P_{nm} \subset S_n$ é denominado um *grupo de permutações de m objetos* se $m < n$ ou seja, a ordem de P_{nm} deve ser menor que a ordem de S_n .

Grupo alternante A_n . Dado o grupo simétrico S_n , para $n > 2$, sempre existirá um subgrupo próprio denominado *alternante*, o que é composto pelas permutações pares de S_n . A ordem de A_n é $|A_n| = n!/2$. Como $I \in \mathcal{G}$ é sempre uma permutação par, o conjunto das permutações ímpares não forma um subgrupo de S_n .

Exemplo 3.5 (Subgrupos de S_3). Dados os elementos de S_3 identificados no exemplo 3.2, suas paridades foram obtidas no exemplo 3.3. Portanto,

$$A_3 = \{\pi_4, \pi_5 = \pi_4^2, I = \pi_4^3\}.$$

No exercício 3.4 foi construída a tabela de multiplicação de S_3 . A partir desta, pode-se obter os grupos de permutações $\langle \pi_2 \rangle = \{\pi_2, I\}$, $\langle \pi_3 \rangle = \{\pi_3, I\}$ e $\langle \pi_6 \rangle = \{\pi_6, I\}$.

Portanto, o grupo S_3 é de grau 4 e sua base é o conjunto $\{\pi_2, \pi_3, \pi_4, \pi_6\}$.

Exercício 3.6. Mostre que o conjunto:

$$e = (), \quad a = (12), \quad b = (34), \quad c = a \circ b = (12)(34)$$

forma um subgrupo de permutação de S_4 .

3.3.2 CLASSES LATERAIS E O TEOREMA DE LAGRANGE

Uma *classe lateral*¹⁰ é um outro subconjunto importante de um grupo. Porém, de forma distinta a um subgrupo, os elementos de uma classe lateral não necessariamente satisfazem os axiomas de grupo. Contudo, a identificação das classes laterais de um grupo fornece informações importantes quanto aos subgrupos do mesmo.

Grupos de ordens 1 – 3 não possuem subgrupos próprios. Isto somente ocorre para grupos de ordem maior ou igual a 4. Dado um grupo de ordem g , a existência ou não de subgrupos próprios é determinada pelo teorema de Lagrange, o qual será exposto a seguir, após a definição das classes laterais.

Dado um grupo $\mathcal{G} = \{G; *\}$ de ordem g e com elemento identidade I , que contém pelo menos um subgrupo próprio $\mathcal{H} = \{H; *\}$ ($\mathcal{H} \subset \mathcal{G}$), de ordem h ($h < g$). Denota-se o conjunto H por

$$H = \{H_1, H_2, \dots, H_h\}, \text{ sendo que } H_j = I \text{ para algum } j = 1, \dots, h.$$

Seja o elemento a tal que $a \in \mathcal{G}$, mas $a \notin \mathcal{H}$. Formam-se então dois novos conjuntos através das multiplicações de a por todos os elementos de H . Estas multiplicações podem ser realizadas pela esquerda ou pela direita; desta forma, os novos conjuntos serão respectivamente representados por

$$\begin{aligned} a\mathcal{H} &\doteq \{a * H_1, a * H_2, \dots, a * H_h\} \\ \mathcal{H}a &\doteq \{H_1 * a, H_2 * a, \dots, H_h * a\}. \end{aligned}$$

Deve ser enfatizado aqui que, em geral, $a * H_j \neq H_j * a$. Definem-se então as seguintes classes laterais.

Definição 3.11 (Classes laterais). Dado um grupo $\mathcal{G} = \{G; *\}$ de ordem g , que contém pelo menos um subgrupo próprio $\mathcal{H} = \{H; *\}$ ($\mathcal{H} \subset \mathcal{G}$), de ordem h ($h < g$). Seja o elemento a tal que $a \in \mathcal{G}$, mas $a \notin \mathcal{H}$. Então,

$$\begin{aligned} a\mathcal{H} = \{a * H_j : j = 1, \dots, h\} &\text{ é uma classe lateral à esquerda de } \mathcal{H} \text{ em } \mathcal{G} \\ \mathcal{H}a = \{H_j * a : j = 1, \dots, h\} &\text{ é uma classe lateral à direita de } \mathcal{H} \text{ em } \mathcal{G}. \end{aligned}$$

A importância na formação das classes laterais reside no seguinte teorema.

Teorema 3.2 (Teorema de Lagrange). Seja $\mathcal{G} = \{G; *\}$ um grupo finito de ordem g . Seja $\mathcal{H} = \{H; *\}$ um subgrupo de \mathcal{G} de ordem h . Então, a ordem de \mathcal{H} é um divisor da ordem de \mathcal{G} , ou seja,

$$\frac{g}{h} = m, \text{ onde } m \in \mathbb{N} \setminus \{0\},$$

onde m é denominado o índice do subgrupo \mathcal{H} sob o grupo \mathcal{G} .

Demonstração. Dados o grupo $\mathcal{G} = \{G; *\}$ de ordem g e com elemento identidade I , e um subgrupo $\mathcal{H} = \{H; *\}$ de ordem h . Se \mathcal{H} é um subgrupo impróprio, então $h = 1$ ou $h = g$ e o teorema é trivialmente satisfeito.

Se \mathcal{H} é um subgrupo próprio de \mathcal{G} , então existe pelo menos um elemento $a_1 \in \mathcal{G}$ tal que $a_1 \notin \mathcal{H}$. Forma-se então a classe lateral $a_1\mathcal{H}$, por exemplo. Se $\mathcal{H} \cup a_1\mathcal{H} \neq \mathcal{G}$, ou seja, se \mathcal{H} e o coset $a_1\mathcal{H}$ não contêm todos os elementos de \mathcal{G} , então existe pelo menos um outro elemento $a_2 \in \mathcal{G}$ mas tal que $a_2 \notin \mathcal{H}$ e $a_2 \notin a_1\mathcal{H}$. Forma-se então um segundo coset $a_2\mathcal{H} = \{a_2H_1, \dots, a_2H_h\}$, cujos elementos não estão contidos nem em \mathcal{H} nem em $a_1\mathcal{H}$. Verifica-se então se $\mathcal{H} \cup a_1\mathcal{H} \cup a_2\mathcal{H} = \mathcal{G}$. Se isto não for satisfeito, o processo é continuado até que se forme um conjunto de classes laterais $\{a_j\mathcal{H}\}$ ($j = 1, \dots, k$, $k \geq 1$) tais que

$$\mathcal{H} \cup a_1\mathcal{H} \cup a_2\mathcal{H} \cup \dots \cup a_k\mathcal{H} = \mathcal{G}.$$

¹⁰Em inglês: *coset*. Este termo será usado como designação compacta de *classe lateral*.

A operação acima em teoria de grupos é escrita como

$$\mathcal{H} + a_1\mathcal{H} + a_2\mathcal{H} + \cdots + a_k\mathcal{H} = \mathcal{G},$$

onde, neste contexto, a operação “+” corresponde à união de conjuntos. Da mesma forma,

$$\mathcal{H} + \mathcal{H}a_1 + \mathcal{H}a_2 + \cdots + \mathcal{H}a_k = \mathcal{G}.$$

Como \mathcal{H} e todas as classes laterais $\{a_j\mathcal{H}\}$ contêm um número de elementos igual a h , então

$$(k+1)h = g \implies \frac{g}{h} = m \in \mathbb{N} \setminus \{0\}.$$

□

A partir do teorema de Lagrange, o seguinte corolário pode ser obtido.

Corolário 3.1 (Lagrange). *Seja $\mathcal{G} = \{G; *\}$ um grupo finito de ordem g . Seja $a \in \mathcal{G}$ um elemento de ordem n . Então,*

$$\frac{g}{n} = m \in \mathbb{N} \setminus \{0\}, \text{ para todo } a \in \mathcal{G}.$$

Exercício 3.7. Verifique o teorema de Lagrange para o grupo S_3 , usando o subgrupo A_3 , e mostre como S_3 é formado pela união de A_3 com suas classes laterais.

Resolução. Como $|S_3| = 3! = 6$, os únicos subgrupos possíveis têm ordens 1, 2, 3 e 6. Os elementos de S_3 foram identificados no exemplo 3.2, enquanto que os elementos de A_3 foram obtidos no exemplo 3.5. Os elementos que não pertencem a A_3 são π_2, π_3 e π_6 . A tabela de multiplicações de S_3 foi deduzida no exercício 3.4. Formam-se então os cosets:

$$\pi_2 A_3 = \{\pi_2, \pi_3, \pi_6\}, \quad \pi_3 A_3 = \{\pi_3, \pi_6, \pi_2\}, \quad \pi_6 A_3 = \{\pi_6, \pi_2, \pi_3\}.$$

Ou seja,

$$S_3 = A_3 + \pi_2 A_3 = A_3 + \pi_3 A_3 = A_3 + \pi_6 A_3.$$

Nota-se que os três cosets são o mesmo, pois $A_3 \cup \pi_2 A_3 = S_3$. Portanto, os subgrupos de S_3 são:

$$\{I\}, \quad \langle \pi_2 \rangle, \quad \langle \pi_3 \rangle, \quad \langle \pi_6 \rangle, \quad A_3, \quad S_3.$$

O exercício a seguir mostra como o teorema de Lagrange pode ser usado para se determinar a tabela de multiplicação de um grupo finito.

Exercício 3.8. Obtenha as tabelas de multiplicação dos possíveis grupos de ordem 6 a partir do teorema de Lagrange.

Resolução. Como $g = 6$, segue do corolário 3.1 que a ordem de qualquer um de seus elementos é um divisor de g , ou seja, 1, 2, 3 ou 6. A primeira estrutura possível consiste em um grupo cíclico cujo gerador é o elemento a , ou seja, $\mathcal{G} = \langle a \rangle = \{a, a^2, a^3, \dots, a^5, a^6 = I\}$.

Para verificarmos outras estruturas, assume-se que não há nenhum elemento de ordem 6. Suponha então que o elemento a seja agora de ordem 3. Neste caso, \mathcal{G} possui o subgrupo $\langle a \rangle = \{a, a^2, I\}$. Necessariamente então \mathcal{G} deve conter um outro elemento $b \notin \langle a \rangle$, com o qual pode-se construir a classe lateral $b\langle a \rangle = \{b, ba, ba^2\}$. Ou seja, $\mathcal{G} = \{I, a, a^2, b, ba, ba^2\}$. Agora, a ordem de b pode ser 2 ou 3. Se $|b| = 3$, o elemento b^2 deve ser um dos listados anteriormente, exceto I . Além disso, as possibilidades $b^2 = b, ba$ ou ba^2 implicam em $b = I, a$ ou a^2 , ou que contradiz a hipótese de que $b \notin \langle a \rangle$. Por outro lado, a possibilidade $b^2 = a$ implica $ba = I$ e $b^2 = a^2$ implica $ba^2 = I$, o que também contradiz a hipótese. Portanto, necessariamente, $|b| \neq 3$, implicando que $|b| = 2$. Estas conclusões já permitem a obtenção dos termos em azul na tabela de multiplicação abaixo. Agora, pelo teorema do rearranjo, $ab = ba$ ou ba^2 . A hipótese $ab = ba$ leva a: $(ab)^2 = (ab)(ab) = (ab)(ba) = a^2$, $(ab)^3 = a^2(ab) = b$, $(ab)^4 = a$, $(ab)^5 = ba^2$ e $(ab)^6 = I$, o que contraria a hipótese de que \mathcal{G} não é cíclico. Portanto, $ab = ba^2$, o que leva aos demais termos (em vermelho) da tabela de multiplicação.

Verifica-se facilmente que a estrutura deste grupo é idêntica à estrutura do S_3 ; ou seja, este grupo é isomórfico¹¹ ao S_3 .

I	a	a^2	b	ba	ba^2
a	a^2	I	ba^2	b	ba
a^2	I	a	ba	ba^2	b
b	ba	ba^2	I	a	a^2
ba	ba^2	b	a^2	I	a
ba^2	b	ba	a	a^2	I

¹¹Ver seção 3.6.

O teorema de Lagrange também pode ser empregado para a decomposição de grupos infinitos, como mostra o exemplo a seguir para um grupo discreto.

Exemplo 3.6. Seja $\mathcal{G} = \{\mathbb{Z}; +\}$ e $\mathcal{H} \subset \mathcal{G}$ tal que

$$\mathcal{H} = \{4n : n \in \mathbb{Z}; +\} = \{\dots, -8, -4, 0, 4, 8, \dots; +\}.$$

Então, pode-se construir os cosets

$$\begin{aligned}(1 + \mathcal{H}) &= \{\dots, -7, -3, 1, 5, 9, \dots\} = \{4n + 1 : n \in \mathbb{Z}\}, \\(2 + \mathcal{H}) &= \{\dots, -6, -2, 2, 6, 10, \dots\} = \{4n + 2 : n \in \mathbb{Z}\}, \\(3 + \mathcal{H}) &= \{\dots, -5, -1, 3, 7, 11, \dots\} = \{4n + 3 : n \in \mathbb{Z}\},\end{aligned}$$

de tal forma que

$$\mathcal{G} = \mathcal{H} + (1 + \mathcal{H}) + (2 + \mathcal{H}) + (3 + \mathcal{H}),$$

pois $(4 + \mathcal{H}) = \{4n + 4 : n \in \mathbb{Z}\} = \mathcal{H}$. Portanto, \mathcal{H} possui índice 4 em \mathcal{G} .

3.3.3 CLASSES DE CONJUGAÇÃO

Nas seções anteriores verificou-se como um grupo pode ser subdividido em subgrupos ou em classes laterais. Verificou-se também que há mais de uma maneira de se dividir um grupo, de forma consistente com o teorema de Lagrange. Como resultado, um determinado elemento do grupo pode resultar em diferentes conjuntos, dependendo de qual subgrupo é usado para se construir as classes laterais.

Deve existir então uma maneira de se particionar um grupo qualquer de tal forma que os subconjuntos resultantes sejam compostos por elementos que possuam algum tipo de “afinidade” entre si. Uma maneira de se realizar este particionamento ordenado do grupo é através das *classes de conjugação*.

Definição 3.12 (Relação de equivalência). Uma dada relação binária “ \sim ” sobre um conjunto S é denominada uma *relação de equivalência sobre S* se e somente se, para quaisquer elementos $a, b, c \in S$, as seguintes propriedades são satisfeitas:

1. **Reflexividade:** $a \sim a$.
2. **Simetria:** Se $a \sim b$ então $b \sim a$.
3. **Transitividade:** Se $a \sim b$ e $b \sim c$, então $a \sim c$.

Uma particular relação de equivalência pode ser empregada para particionar um grupo em *classes*, tais que um determinado elemento do grupo pertence a somente uma classe e todos os membros dessa classe são *equivalentes* entre si.

Existem várias notações distintas para uma relação de equivalência. Dado o conjunto S e o elementos $a, b \in S$, se estes elementos são equivalentes entre si com respeito a uma dada relação de equivalência R , isto pode ser representado por $a \sim b$ ou $a \equiv b$, se não houver ambiguidade quanto à relação de equivalência, ou, em caso contrário, como $a \sim_R b$, $a \equiv_R b$ ou ainda aRb .

Alguns exemplos de relações de equivalência:

- \sim : “É igual a” sobre o conjunto \mathbb{R} .
- \sim : “Têm o mesmo aniversário” sobre um conjunto de pessoas.
- \sim : “É congruente a” sobre o conjunto de todos os triângulos.
- \sim : “Têm a mesma imagem da função” sobre o conjunto dos elementos do domínio da função.
- \sim : “Têm o mesmo cosseno” sobre o conjunto de todos os ângulos.

Percebe-se então que dado um conjunto S qualquer, nem sempre é possível determinar uma relação de equivalência entre todos os seus elementos. Da mesma forma, se dois elemento de S são equivalentes, os demais elementos não necessariamente são equivalentes aos dois primeiros. Portanto, faz-se necessária a definição de uma *classe de equivalência*, a seguir.

Definição 3.13 (Classe de equivalência). Dado um conjunto S e o elemento $a \in S$, a classe de equivalência de a frente a relação de equivalência R é definida como o conjunto $[a]_R$ tal que

1. $[a]_R = \{b \in S \mid a \sim_R b\}$.
2. Todo elemento $a \in S$ pertence a exatamente uma classe.

Quando não há ambiguidade, a classe de equivalência do elemento a pode ser denotada simplesmente por $[a]$ ou C_a .

A propriedade (2) na definição de classe de equivalência é, na verdade, uma consequência da definição da mesma. Para demonstrar isso, seja $a \in S$ e $[a]$ a sua classe de equivalência. Claramente, $a \in [a]$ por reflexividade. Supõe-se agora que exista um outro elemento $b \in S$, formando a classe $[b]$, mas tal que $a \sim b$. Se $c \in S$ é tal que $c \in [b]$, então $b \sim c$. Porém, por transitividade, $a \sim c$, o que significa que $c \in [a]$. De forma inversa, pela lei de reflexividade $b \sim a$; então, se $c \in [a]$, isto implica que $c \in [b]$. Ou seja, necessariamente $[a] = [b]$.

A partir da definição de classe de equivalência, o seguinte corolário pode ser demonstrado:

Corolário 3.2. Dados o conjunto S e os elementos $a, b \in S$, com as classes de equivalência $[a]_R$ e $[b]_R$ frente a relação de equivalência R . Então ou $[a]_R = [b]_R$ ou $[a]_R \cap [b]_R = \emptyset$ e a coleção de todas as classes C_i é uma partição de S , ou seja, todo elemento de S pertence a uma e somente uma classe de equivalência.

Demonstração. Supõe-se inicialmente que $[a]_R$ e $[b]_R$ são subconjuntos distintos, mas com pelo menos um elemento c em comum. Então, $a \sim c$ e $b \sim c$ e pelas leis de simetria e transitividade $a \sim b$. Contudo, isto implica que $[a]_R = [b]_R$, o que contradiz a hipótese de que os conjuntos são distintos. Portanto, $[a]_R \cap [b]_R \neq \emptyset$.

Por outro lado, dadas as classes $[a]_R$ e $[b]_R$, qualquer outro elemento $c \in S$ ou está em uma das classes já definidas ou constitui uma nova classe. Este processo pode ser seguido até que as coleção de todas as classes de equivalência exaurem S , i. e., todo elemento de S está em alguma classe. \square

A partição mencionada no corolário 3.2 possui um nome próprio, de acordo com a definição abaixo.

Definição 3.14 (Espaço fator). Dado o conjunto S , a sua partição dada pela união de todas as suas classes de equivalência frente a relação R , denotada por S/R ,

$$S/R \doteq \{[a]_R \mid a \in S\},$$

é denominada o *espaço fator*, ou o *espaço quociente* ou ainda o *conjunto quociente* de S .

Até este momento, o tipo de relação de equivalência é geral. Para a teoria de grupo em particular, uma relação de equivalência importante é aquela que define uma *classe de conjugação*, através de relações de conjugação entre seus elementos.

Definição 3.15 (Conjugação). Dado um grupo $\mathcal{G} = \{G; *\}$, dois elementos $a, b \in \mathcal{G}$ são ditos *conjugados* se existe um elemento $c \in \mathcal{G}$ tal que¹²

$$c^{-1} * a * c = b.$$

Diz-se também que b é o *equivalente* de a , ou que b é o *transformado* de a por c .

A importância dos elementos conjugados em um grupo está no fato de que estes pertencem a uma classe de equivalência, de acordo com as definições 3.12 e 3.13. Ou seja, se $a, b \in \mathcal{G}$ são conjugados entre si, pode-se dizer que $a \sim b$.

Para verificar que a condição de conjugação realmente consiste em uma relação de equivalência, é necessário que satisfaça os axiomas da definição 3.12, a saber:

1. *Reflexividade:* $a \sim a$, pois dado o elemento identidade $I \in \mathcal{G}$, $I^{-1}aI = IaI = a$.
2. *Simetria:* se $a \sim b$, então existe $c \in \mathcal{G}$ tal que $b = c^{-1}ac$. Portanto, existe $c^{-1} \in \mathcal{G}$ tal que

$$(c^{-1})^{-1} b (c^{-1}) = \left((c^{-1})^{-1} c^{-1} \right) a (cc^{-1}) = a.$$

Ou seja, $a \sim b$ implica em $b \sim a$.

¹²Alguns textos definem a operação de conjugação como $b = c * a * c^{-1}$, a qual é equivalente.

3. *Transitividade*: se $a \sim b$ e $b \sim c$, então existem $d, e \in \mathcal{G}$ tais que $b = d^{-1}ad$ e $c = e^{-1}be$. Portanto,

$$c = e^{-1}(d^{-1}ad)e = (e^{-1}d^{-1})a(de) = (de)^{-1}a(de),$$

de acordo com a propriedade (3.2) do inverso do produto. Como necessariamente $de, (de)^{-1} \in \mathcal{G}$, isto implica em $c \sim a$.

Estes resultados estabelecem que a conjugação entre dois elementos do grupo é uma relação de equivalência. Em consequência, estes elementos pertencem a mesma classe de equivalência, agora denominada *classe de conjugação*.

Definição 3.16 (Classe de conjugação). Dado o grupo $\mathcal{G} = \{G; *\}$ e o elemento $a \in \mathcal{G}$, o conjunto

$$[a] \equiv C_a = \{b^{-1} * a * b, \forall b \in \mathcal{G}\}$$

forma uma classe de equivalência, denominada *classe de conjugação de a*.

Algumas consequências importantes desta definição devem ser mencionadas:

- O elemento $a \in \mathcal{G}$ pertence à sua própria classe de conjugação, pois escolhendo $b = a$, $a^{-1} * a * a = a$; ou seja, $a \in [a]$.
- Se $a, b \in \mathcal{G}$ são tais que $a \sim b$, então $[a] = [b]$. Isto é uma consequência direta do corolário 3.2.
- Se $a, b \in \mathcal{G}$ são tais que $a \not\sim b$, então $[a] \neq [b]$. Também uma consequência do corolário 3.2.
- Se o elemento a é tal que $a^m = I$, então todos os elementos de C_a têm a mesma ordem.
- Se $[I]$ é uma classe de conjugação contendo o elemento identidade $I \in \mathcal{G}$, então, para todo $a \in \mathcal{G}$,

$$a^{-1} * I * a = I.$$

Ou seja, $[I] = \{I\}$; a classe de conjugação de I contém somente este elemento.

- Se $a \in \mathcal{G}$ está em uma **classe por si próprio**, então a conjugação $c = b^{-1} * a * b$ para todo $b \in \mathcal{G}$ deve necessariamente implicar que $c = a$. Mas, como

$$a = (b * b^{-1}) * a * (b * b^{-1}) = b * (b^{-1} * a * b) * b^{-1} = b * c * b^{-1} = b * a * b^{-1},$$

conclui-se que $a * b = b * a$ para todo $b \in \mathcal{G}$. Portanto, a comutação de a com todos os elementos de \mathcal{G} é uma condição necessária e suficiente para que $[a] = \{a\}$. Isto ocorre com $I \in \mathcal{G}$, mas ocorre também em um grupo Abelianiano; ou seja, cada elemento deste grupo forma uma classe por si próprio.

Teorema 3.3 (Centro do grupo). Seja o grupo $\mathcal{G} = \{G; *\}$. Dado o conjunto $S \subseteq \mathcal{G}$, composto pelos elementos de \mathcal{G} que formam classes por si próprios, este conjunto forma um subgrupo Abelianiano de \mathcal{G} , denominado o centro do grupo.

Demonstração. Para todo grupo \mathcal{G} , existe pelo menos um elemento em S que forma uma classe por si próprio: o elemento identidade. Portanto, $I \in S$. Se este é o único elemento em S , então o subgrupo é Abelianiano.

Se houver pelo menos um outro elemento de \mathcal{G} que forma uma classe em si próprio, então S possui pelo menos dois elementos. Assim, se $s_1, s_2 \in S$, então $s_1b = bs_1$ e $s_2b = bs_2$, para todo $b \in \mathcal{G}$. Portanto,

1. *Clausura*: $(s_1s_2)b = s_1(bs_2) = b(s_1s_2)$; ou seja, $s_1s_2 \in S$.
2. *Elemento inverso*: Se $s_1b = bs_1$, então $s_1^{-1}(s_1b)s_1^{-1} = s_1^{-1}(bs_1)s_1^{-1}$ implica em $bs_1^{-1} = s_1^{-1}b$; ou seja, $s_1^{-1} \in S$.
3. *Grupo Abelianiano*: $s_1s_2 = s_1bs_2 = bs_1s_2 = bs_2s_1 = s_2s_1$.

□

3.3.4 SUBGRUPOS INVARIANTES E GRUPO FATOR

A partir das definições e propriedades das classes de conjugações, é possível construir-se uma classe importantes de subgrupos. As definições a seguir pertencem a essa classe.

Definição 3.17 (Subgrupo conjugado). Dados um grupo $\mathcal{G} = \{G; *\}$ e um subgrupo $\mathcal{H} \subseteq \mathcal{G}$. A partir de qualquer elemento $a \in \mathcal{G}$, pode-se formar o conjunto

$$a^{-1}\mathcal{H}a \doteq \{a^{-1} * h * a, \forall h \in \mathcal{H}\},$$

o qual é também um subgrupo de \mathcal{G} , denominado *subgrupo conjugado de \mathcal{H} em \mathcal{G}* .

Tomando diversos elementos distintos de \mathcal{G} , é possível formar-se vários subgrupos conjugados distintos. Dentre os possíveis subgrupos conjugados, pode existir um tipo especial para o qual $a^{-1}\mathcal{H}a = \mathcal{H}$, para todo $a \in \mathcal{G}$. Este é denominado um *subgrupo invariante*.

Definição 3.18 (Subgrupo invariante). Dados um grupo $\mathcal{G} = \{G; *\}$, um subgrupo $\mathcal{H} \subseteq \mathcal{G}$ e o conjunto $\{h_1, h_2\} \subseteq \mathcal{H}$. Se, para todo $a \in \mathcal{G}$, ocorrer que $a^{-1} * h_1 * a = h_2$, ou, de forma equivalente, se $h_1 * a = a * h_2$, então o subgrupo \mathcal{H} é denominado um *subgrupo invariante* (ou *auto-conjugado*, ou ainda *divisor normal*) em \mathcal{G} .

Pode-se representar um subgrupo invariante na forma de classes laterais como

$$a\mathcal{H} = \mathcal{H}a, (\forall a \in \mathcal{G}).$$

Esta notação permite uma definição equivalente para um subgrupo invariante como aquele para o qual as classes laterais à esquerda e à direita são as mesmas, para qualquer $a \in \mathcal{G}$.

Consequências e propriedades importantes desta definição:

- Qualquer grupo \mathcal{G} contém dois subgrupos invariantes triviais ou impróprios: $\mathcal{H} = \mathcal{G}$ e $\mathcal{H} = \{I\}$.
- Se $h \in \mathcal{H}$, então todos os elementos do conjunto $\{a^{-1} * h * a, \forall a \in \mathcal{G}\}$ também estão em \mathcal{H} . Ou seja, um subgrupo $\mathcal{H} \subseteq \mathcal{G}$ é invariante se e somente se ele contiver os elementos de \mathcal{G} em **classes completas**; isto é, \mathcal{H} deve conter todos ou nenhum elemento de uma ou mais classes de \mathcal{G} .
- Dado o grupo $\mathcal{G} = \{G; *\}$ e um subgrupo invariante $\mathcal{H} \subseteq \mathcal{G}$, o produto de duas classes laterais de \mathcal{H} também é uma classe lateral, pois, dados $\{a, b\} \in \mathcal{G}$,

$$(a\mathcal{H})(b\mathcal{H}) = a(\mathcal{H}b)\mathcal{H} = a(b\mathcal{H})\mathcal{H} = (ab)(\mathcal{H}\mathcal{H}) = (ab)\mathcal{H}.$$

Foi introduzida aqui a operação de **multiplicação de cosets**, definida de tal forma que

$$(a\mathcal{H})(b\mathcal{H}) \doteq \{(a * h_1) * (b * h_2), \forall h_1, h_2 \in \mathcal{H}\}.$$

- Igualmente, o produto de \mathcal{H} por uma classe lateral resulta na própria classe lateral,

$$\mathcal{H}(a\mathcal{H}) = (a\mathcal{H})\mathcal{H} = a(\mathcal{H}\mathcal{H}) = a\mathcal{H}.$$

- Da mesma forma, o produto de uma classe lateral pela sua “classe inversa” resulta no subgrupo invariante; isto é, dado $a \in \mathcal{G}$ e o coset $a\mathcal{H}$, deve existir o elemento inverso $a^{-1} \in \mathcal{G}$, cujo coset é $a^{-1}\mathcal{H}$ (o coset inverso). Então,

$$(a^{-1}\mathcal{H})(a\mathcal{H}) = a^{-1}(\mathcal{H}a)\mathcal{H} = a^{-1}(a\mathcal{H})\mathcal{H} = (a^{-1}a)(\mathcal{H}\mathcal{H}) = \mathcal{H}.$$

Nota-se também que o subgrupo normal \mathcal{H} pode ser considerado o elemento identidade frente a operação de multiplicação de cosets.

Outras definições importantes são apresentadas a seguir.

Definição 3.19 (Grupo simples ou semisimples). Um dado grupo \mathcal{G} é denominado:

Simples: caso não possua nenhum subgrupo invariante próprio.

Semisimples: caso nenhum de seus subgrupos invariantes próprios seja Abelian.

Apresenta-se finalmente o importante teorema a seguir.

Teorema 3.4 (Grupo fator). *Dados o grupo $\mathcal{G} = \{G; *\}$ e um subgrupo invariante $\mathcal{H} \subseteq \mathcal{G}$, o conjunto formado por \mathcal{H} e por todas as suas classes laterais forma um grupo, denominado **grupo fator** ou **grupo quociente**, frente a operação de multiplicação de cosets. O grupo fator é representado por \mathcal{G}/\mathcal{H} e sua ordem, se $|\mathcal{G}|$ for finita, é a razão $|\mathcal{G}|/|\mathcal{H}|$.*

Exercício 3.9. Dado o grupo S_3 , obtenha suas classes de conjugação, seus grupos invariantes e seus grupos fatores.

Resolução. O grupo S_3 foi definido no exemplo 3.2 e sua tabela de multiplicação foi obtida no exercício 3.4.

Classes: a primeira classe trivial é $[\pi_1] \equiv [I] = \{I\}$. Para $[\pi_2]$:

$$\begin{array}{lll} I\pi_2I = \pi_2 & \pi_2^{-1}\pi_2\pi_2 = \pi_2 & \pi_3^{-1}\pi_2\pi_3 = \pi_6 \\ \pi_4^{-1}\pi_2\pi_4 = \pi_6 & \pi_5^{-1}\pi_2\pi_5 = \pi_3 & \pi_6^{-1}\pi_2\pi_6 = \pi_3. \end{array}$$

Portanto, $[\pi_2] = [\pi_3] = [\pi_6] = \{\pi_2, \pi_3, \pi_6\}$. Para $[\pi_4]$:

$$\begin{array}{lll} I\pi_4I = \pi_4 & \pi_2^{-1}\pi_4\pi_2 = \pi_5 & \pi_3^{-1}\pi_4\pi_3 = \pi_4 \\ \pi_4^{-1}\pi_4\pi_4 = \pi_4 & \pi_5^{-1}\pi_4\pi_5 = \pi_4 & \pi_6^{-1}\pi_4\pi_6 = \pi_5. \end{array}$$

Portanto, $[\pi_4] = [\pi_5] = \{\pi_4, \pi_5\}$. O grupo S_3 é formado por 3 classes de conjugação distintas.

Subgrupos invariantes: no exemplo 3.5 foi mostrado que os subgrupos próprios de S_3 são: A_3 , $\langle \pi_2 \rangle$, $\langle \pi_3 \rangle$ e $\langle \pi_6 \rangle$. Cada subgrupo será testado agora quanto a sua invariância. Para A_3 :

$$\begin{array}{lll} IA_3I = A_3 & \pi_2^{-1}A_3\pi_2 = \pi_2^{-1}\left\{\begin{array}{l} I \\ \pi_4 \\ \pi_5 \end{array}\right\}\pi_2 = \left\{\begin{array}{l} I \\ \pi_5 \\ \pi_4 \end{array}\right\} = A_3 & \pi_3^{-1}A_3\pi_3 = \left\{\begin{array}{l} I \\ \pi_5 \\ \pi_4 \end{array}\right\} \\ \pi_4^{-1}A_3\pi_4 = \left\{\begin{array}{l} I \\ \pi_4 \\ \pi_5 \end{array}\right\} & \pi_5^{-1}A_3\pi_5 = \left\{\begin{array}{l} I \\ \pi_4 \\ \pi_5 \end{array}\right\} & \pi_6^{-1}A_3\pi_6 = \left\{\begin{array}{l} I \\ \pi_5 \\ \pi_4 \end{array}\right\}. \end{array}$$

Portanto, o subgrupo A_3 é invariante. Para $\langle \pi_2 \rangle$:

$$I\langle \pi_2 \rangle I = \langle \pi_2 \rangle \quad \pi_2^{-1}\langle \pi_2 \rangle \pi_2 = \pi_2^{-1}\left\{\begin{array}{l} I \\ \pi_2 \end{array}\right\}\pi_2 = \langle \pi_2 \rangle \quad \pi_3^{-1}\langle \pi_2 \rangle \pi_3 = \left\{\begin{array}{l} I \\ \pi_6 \end{array}\right\} \neq \langle \pi_2 \rangle.$$

Portanto, $\langle \pi_2 \rangle$ não é invariante. O mesmo pode ser verificado para $\langle \pi_3 \rangle$ e $\langle \pi_6 \rangle$. O único grupo normal é A_3 .

Grupo fator: no exercício 3.7 foi demonstrado que há somente uma única classe lateral de A_3 : $\pi_2 A_3$. Portanto, o grupo fator é

$$S_3/A_3 \doteq \{A_3, \pi_2 A_3\},$$

sendo que, neste grupo, $I = A_3$. Observa-se também que $|S_3/A_3| = |S_3|/|A_3| = 2$.

3.4 GRUPOS DE SIMETRIA

Simetria, tanto para as ciências naturais (física, química, biologia, etc) quanto para a matemática possui significado semelhante, embora o conceito seja aplicado a diferentes objetos ou estruturas.

Para a matemática, o termo *simetria* está relacionado à invariância das estruturas matemáticas abstratas. Por esta razão, simetria não ocorre somente em geometria ou topologia, mas ocorre também na álgebra e na análise, como no estudo de equações diferenciais, por exemplo.

Para a física, o termo *simetria* está relacionado às características de um sistema físico que são preservadas, de acordo com algum tipo de observação, após a aplicação de algum tipo de transformação sobre o mesmo. Em outras palavras, a *simetria de um sistema físico* é uma característica física (estrutural) e/ou matemática do mesmo (observada ou intrínseca) que é mantida invariante frente à transformação aplicada (passiva ou ativa) sobre o mesmo. Algumas das transformações que podem ser aplicadas sobre um sistema físico são também discutidas nas seções 6.2 e 6.6, no contexto da álgebra e análise tensoriais.

Observa-se que as transformações mencionadas são tais que a estrutura ou propriedades do objeto (físico ou matemático) permanecem invariantes frente a aplicação destas transformações. Assim, como cada transformação deste tipo essencialmente leva o sistema para si mesmo (processo denominado também de *automorfismo*), duas transformações distintas e consecutivas terão a mesma consequência.

Assim, o conjunto de todas as transformações aplicadas ao objeto que satisfazem a condição de automorfismo formam um grupo, denominado *grupo de simetria* ou, em outros textos, de *grupo de transformações*. A validade da última designação baseia-se no fato de que, como cada transformação sobre o objeto que não seja isomórfica irá alterar o mesmo, dificilmente o conjunto de tais transformações anisomórficas irá constituir um grupo.

Como o conceito de simetria em física aplica-se tanto à estrutura material do mesmo quanto às suas propriedades matemáticas, um grupo de simetria em particular pode ser tanto finito quanto infinito (discreto ou contínuo). Exemplos de grupos de simetria contínuos já foram apresentados na seção 3.1.2, dentre os quais pode-se mencionar os grupos $O(n)$, $SO(n)$, $U(n)$, $SU(n)$ e o grupo de Lorentz.

O grupo de Lorentz em particular; na verdade, o *grupo de Lorentz homogêneo*, é composto por todas as transformações de Lorentz¹³ descritas por matrizes 4×4 , as quais são determinadas por 3 parâmetros não compactos. O grupo também é composto por todas as rotações no \mathbb{R}^3 em torno da origem, as quais formam o grupo $SO(3)$, compacto de dimensão 3. Estas transformações mantêm invariante a métrica do espaço-tempo de Minkowski e mantêm também a origem fixa. Portanto, o grupo de Lorentz é um grupo de simetria contínuo, não compacto e não Abeliano, de dimensão 6. A *simetria* neste caso (ou *isometria*) consiste no conjunto de transformações que mantêm invariante uma propriedade matemática do sistema: a métrica do espaço-tempo.

O grupo homogêneo de Lorentz é um subgrupo do *grupo de Poincaré* (ou *grupo de Lorentz inhomogêneo*), o qual contém também o conjunto de todas as translações espaciais e deslocamentos temporais que mantêm a métrica do espaço-tempo invariante. O grupo de Poincaré é um grupo de simetria contínuo, não compacto e não Abeliano, com $06 + 04 = 10$ dimensões.

Os exemplos de grupos de simetria até aqui apresentados destacam grupos contínuos e grande parte das transformações isométricas que mantêm invariantes propriedades matemáticas do sistema físico formam grupos contínuos.

Considerando transformações isométricas na estrutura material do sistema físico, os grupos de simetria que surgem são de interesse principalmente nas áreas de física molecular e do estado sólido. Já para a química, o interesse primordial está nas simetrias de moléculas. Estas propriedades fundamentais são importantes para a dinâmica das reações químicas, para a espectroscopia e estrutura dos níveis de energia das moléculas.

Já para a física do estado sólido, as propriedades de simetria das redes cristalinas têm influência na estrutura eletrônica dos cristais, nos espectros de difração, nas suas propriedades elétricas e magnéticas, entre outras. Um *cristal* consiste no ordenamento regular de átomos em três dimensões. Uma rede cristalina é composta por uma *rede de Bravais* e uma *base*. A rede de Bravais é simplesmente um arranjo regular de pontos discretos em 3 dimensões de tal forma que a rede é vista exatamente da mesma forma, independente do ponto de observação. Já a base da rede cristalina consiste em um conjunto de 3 *vetores primitivos* que geram e varrem a rede de Bravais; ou seja, que descrevem a posição de cada átomo da rede cristalina com o uso de até 3 índices inteiros. As figuras 3.3 e 3.4 mostram dois exemplos de redes cristalinas cúbicas (simples e de corpo centrado) e os vetores primitivos usualmente empregados na localização de todos os pontos da rede.

Assumindo-se que os efeitos de superfície são negligenciáveis em um cristal, as transformações que preservam as posições relativas entre os pontos discretos do cristal podem ser de 3 tipos:

1. Rotações por um dado ângulo em torno de algum eixo.
2. Reflexão especular em um plano e inversões.
3. Translações espaciais.

As duas primeiras mantêm pelo menos um ponto do cristal fixo, ao passo que a terceira implica em deslocamentos ao longo da rede cristalina.

A descrição algébrica das transformações de simetria aplicadas sobre objetos físicos e/ou matemáticos é implementada pela noção de ação de grupo. Os elementos que caracterizam o objeto são identificados e descritos por um conjunto e as simetrias desse objeto são descritas pelo *grupo de simetria* do conjunto. Este grupo é um *grupo de permutações* se o conjunto for finito e não constituir um espaço vetorial ou um *grupo de transformações*, se o conjunto forma um espaço vetorial, em cuja situação a ação de grupo atua como *transformações lineares*

¹³Em inglês: *boost*.

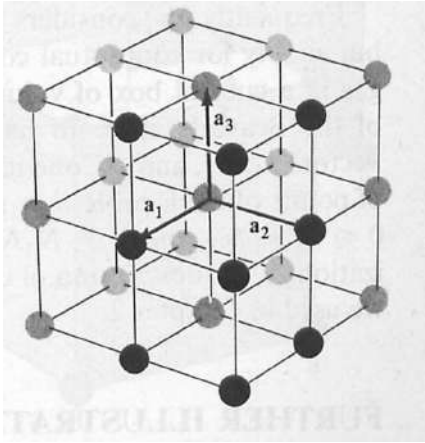


Figura 3.3: Uma rede de Bravais cúbica simples. Os três vetores primitivos \mathbf{a}_1 , \mathbf{a}_2 e \mathbf{a}_3 são mutuamente perpendiculares e têm a mesma magnitude.

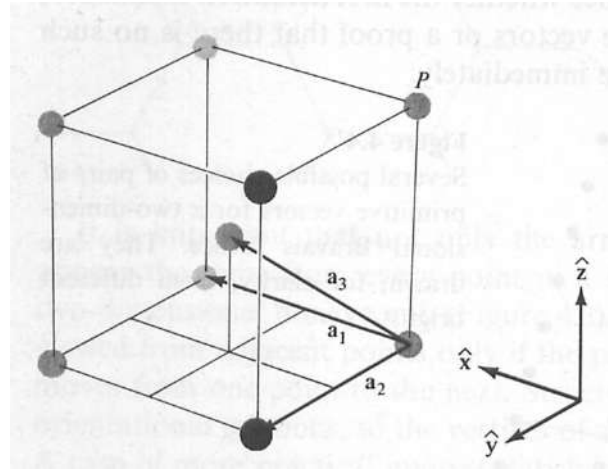


Figura 3.4: Célula unitária de uma rede de Bravais cúbica de corpo centrado (bcc). Os vetores unitários são dados por $\mathbf{a}_1 = a\hat{x}$, $\mathbf{a}_2 = a\hat{y}$ e $\mathbf{a}_3 = a(\hat{x} + \hat{y} + \hat{z})/2$.

no conjunto.¹⁴ Um exemplo de ações de grupo ocorre quando se considera as operações de isometria de um triângulo equilátero, sendo o triângulo descrito por um conjunto de pontos que identificam seus vértices. O grupo de simetria resultante é obtido no exercício 3.10.

Definição 3.20 (Ação de grupo). Sejam $\mathcal{G} = \{G; *\}$ um grupo e C um conjunto. Uma ação (pela esquerda) de grupo φ de G sobre C é o mapeamento

$$\varphi : \mathcal{G} \times C \mapsto C, \quad (g, c) \mapsto \varphi(g, c),$$

para todo $g \in \mathcal{G}$ e todo $c \in C$, que satisfaz os axiomas

Identidade. Sendo I o elemento identidade de \mathcal{G} , $\varphi(I, c) = c$, para todo $c \in C$.

Compatibilidade. Para todo $g, h \in \mathcal{G}$ e $c \in C$, $\varphi(g * h, c) = \varphi(g, \varphi(h, c))$.

São definidos também os seguintes tipos de ações de grupo.

Transitiva. Se C não for o conjunto vazio ($C \neq \emptyset$) e se para todo $c, d \in C$ existir um elemento $g \in \mathcal{G}$ tal que $\varphi(g, c) = d$.

Fiel ou efetiva. Se para todos $g, h \in \mathcal{G}$ distintos ($g \neq h$) existe um elemento $c \in C$ tal que $\varphi(g, c) \neq \varphi(h, c)$. Alternativamente, se para todo $g \in \mathcal{G}$ tal que $g \neq I$ existe um elemento $c \in C$ tal que $\varphi(g, c) \neq c$.

Intuitivamente, para uma ação de grupo fiel, diferentes elementos de \mathcal{G} induzem diferentes permutações dos elementos de C .

Livre ou semiregular. Se, dados $g, h \in \mathcal{G}$, a existência de um elemento $c \in C$ onde $\varphi(g, c) = \varphi(h, c)$ implica $g = h$. De forma equivalente, se $g \in \mathcal{G}$ e existir um elemento $c \in C$ para o qual $\varphi(g, c) = c$, então, necessariamente, $g = I$.

Regular ou transitivo simples. Se φ for tanto transitivo quanto livre. Isto equivale a dizer que para quaisquer $c, d \in C$ existe um e somente um $g \in \mathcal{G}$ tal que $\varphi(g, c) = d$.

n -Transitiva. Se C possuir pelo menos n elementos e existirem subconjuntos $\{c_1, \dots, c_n\}$ e $\{d_1, \dots, d_n\}$ para os quais existe um elemento $g \in \mathcal{G}$ tal que $\varphi(g, c_k) = d_k$ ($k = 1, \dots, n$).

Primitiva. Se as ações de todos os elementos de \mathcal{G} sobre C forem transitivas e a subtração de todos os pares $c, d \in C$ resultantes das ações de grupo resultar no conjunto vazio, i. e., se a ação de grupo não preservar nenhuma partição de C em subconjuntos não vazios.

¹⁴Espaços vetoriais e transformações lineares operando no mesmo são apresentados na seção 4.1.

Logicamente livre. Se \mathcal{G} for um grupo topológico e não existir nenhuma vizinhança U de I tal que a restrição da ação em U seja livre. Isto é, se $\varphi(g, c) = c$ para algum $c \in C$ e algum $g \in U$, então $g = I$.

Irredutível. Se C for um módulo sobre um anel R ,¹⁵ a ação de \mathcal{G} é R -linear e não existir nenhum submódulo próprio invariante.

Após esta longa introdução, discute-se agora alguns dos grupos de simetria mais importantes para a física. Globalmente, os grupos de simetria podem ser divididos em dois tipos, descritos a seguir.

Grupo de simetria pontual. Um *grupo pontual* é um grupo de simetrias geométricas (isometrias) que mantêm pelo menos um ponto fixo. As transformações isométricas consistem usualmente em rotações em torno de um eixo, reflexões especulares e inversões, tais que todos os eixos de rotação e todos os planos de reflexão devem se interceptar em pelo menos um ponto.

Os grupos $O(n)$ e $U(n)$ são grupos de simetria pontuais. Da mesma forma, o grupo homogêneo de Lorentz também, uma vez que as transformações consideradas neste são os *boosts* e rotações no espaço. Grupos pontuais finitos de interesse para a cristalografia são considerados a seguir.

Grupo de simetria espacial. Grupo de simetria formado a partir de translações espaciais. Para que um sistema físico possua simetria espacial, este deve ter a mesma estrutura, para qualquer deslocamento realizado por um observador. Para que isso seja possível, é necessário que o sistema tenha uma extensão infinita ou que as propriedades físicas observadas da rede não sejam afetadas pela sua superfície.

O grupo inhomogêneo de Lorentz é um grupo de simetria espacial, uma vez que a métrica do espaço-tempo (o qual é infinito) não é afetado por translações espaciais e deslocamentos temporais. Redes de Bravais também irão apresentar simetria espacial se estas puderem ser consideradas infinitas.

Os grupos finitos de simetria que são de interesse para a física do estado sólido e para a química serão consideradas em mais detalhes a seguir.

3.4.1 GRUPOS CRISTALOGRAFÍCOS PONTUAIS

Como já mencionado, este é o grupo formado pelas transformações isométricas que mantêm um ponto central fixo e movem as pontos e faces restantes do cristal, de tal forma que este assume, após a transformação, uma configuração indistinguível da inicial, de forma consistente com a simetria translacional da rede cristalina.

Existem ao todo 32 grupos pontuais que preservam a simetria translacional de redes cristalinas em um espaço de 3 dimensões. Este número em particular é fornecido pelo **teorema da restrição cristalográfica**, o qual afirma que as simetrias rotacionais de um cristal somente podem ser de segunda, terceira, quarta e sexta ordens. Redes com outras ordens de simetria (como de 5ª ordem, por exemplo) são denominados *quasecristais*.

Se a origem do sistema de coordenadas for posicionada sobre um dos pontos fixos das transformações, então o grupo pontual é composto pelas seguintes operações:

1. Rotações por um dado ângulo em torno de um eixo que passa pela origem.
2. Reflexões especulares em planos que passam pela origem.
3. Inversão espacial.

Cabe aqui mencionar que estas operações não são totalmente independentes entre si. Qualquer uma das operações acima pode ser expressa por uma combinação adequada das duas restantes. Assim, uma inversão espacial pode ser considerada como uma rotação por π radiano, seguida por uma reflexão no plano perpendicular ao eixo de rotação.

Uma rotação seguida por uma reflexão ou inversão é denominada uma *rotação imprópria*, a qual muda a quiralidade do cristal. Assim, as rotações acima mencionadas, as quais não

¹⁵Ver seção 3.7.2.

mudam a quiralidade do sistema, são também denominadas *rotações próprias*. Pode-se verificar facilmente que o produto de duas rotações próprias ou duas impróprias sempre resulta em uma rotação própria, ao passo que o produto de uma rotação própria por uma imprópria, em qualquer ordem, é sempre uma rotação imprópria. As implicações ao sistema físico oriundas de rotações próprias ou impróprias (ativas ou passivas) são discutidas em maior detalhe na seção 6.6, no contexto de campos tensoriais.

É também conveniente observar aqui que os seguintes pares de operações consecutivas comutam entre si: (a) uma inversão com qualquer outra operação, (b) duas rotações em torno do mesmo eixo, (c) duas rotações por π radianos em torno de eixos perpendiculares, (d) uma rotação e uma reflexão em um plano normal ao eixo de rotação, (e) uma rotação por π radianos e uma reflexão em um plano que passa pelo eixo de rotação, e (f) duas reflexões em planos perpendiculares.

Alguns exemplos de grupos pontuais serão apresentados a seguir. Será feita também uma breve descrição dos 32 grupos pontuais, juntamente com a nomenclatura e notações adotadas. Com respeito às notações, atualmente há dois sistemas empregados: a notação **Schoenflies**, comumente empregada em espectroscopia, e a **notação internacional**, ou **Hermann-Mauguin**, empregada em cristalografia. Os grupos discutidos serão sempre identificados pela notação de Schoenflies, mas a notação internacional equivalente será também mencionada.

Empregando a notação de Schoenflies, os grupos cristalográficos pontuais são os seguintes:

C_n : grupo cíclico composto pelas rotações próprias de um polígono regular de n lados. Este polígono será levado em coincidência consigo mesmo por meio de uma rotação em um ângulo $\psi = 2\pi/n$ em torno de um eixo normal ao plano do polígono e que passa através de seu ponto central. Este eixo de rotação é denominado **eixo de rotação de ordem n** . A operação de rotação do polígono através deste eixo pelo ângulo ψ será denotada por C_n . Aplicações sucessivas deste operador ao polígono são denotadas por $C_n^2, C_n^3, \text{ etc.}$, isto é, rotações através de $2\psi = 4\pi/n, 3\psi, \text{ etc.}$ Claramente, a n -ésima aplicação do operador C_n irá levar o polígono à sua configuração original e isto corresponde então à operação identidade, denotada aqui por $C_n^n = E$.¹⁶ O conjunto destas rotações forma o grupo cíclico C_n (notação de Schoenflies) ou n (notação internacional):

$$C_n = \{E, C_n, C_n^2, \dots, C_n^{n-1}\}.$$

Este grupo é ampliado por produtos diretos de rotações com reflexões, criando-se os grupos:

C_{nh} : criado pela adição de uma reflexão especular em um plano perpendicular (horizontal) ao eixo de rotação. A operação de reflexão neste plano será identificada por σ_h . Claramente, $\sigma_h^2 = E$, constituindo assim o grupo cíclico

$$C_{1h} = \{E, \sigma_h\}. \quad (3.7)$$

C_{nv} : criado pela adição de reflexões especulares em n planos paralelos (verticais) ao eixo de rotação, separados por um ângulo ψ entre si. Esta operação será identificada por σ_v . Claramente, $\sigma_v^2 = E$; assim, cada operador σ_v constitui um grupo cíclico $\{E, \sigma_v\}$.

S_{2n} : (de *Spiegel*, espelho em alemão) denota o grupo que contém somente um eixo rotação imprópria (rotação-reflexão) de ordem $2n$.

D_n : (de *diedral*, ou frente-verso) indica que o grupo possui um eixo de rotação de ordem n mais n eixos de rotação de ordem 2 perpendiculares ao primeiro. Por exemplo, o grupo D_2 possui um eixo de rotação principal de ordem 2 mais dois outros eixos de ordem 2, ortogonais ao primeiro; assim, o grupo D_2 possui três eixos de rotação de ordem 2 mutuamente ortogonais. Este grupo é ampliado por:

D_{nh} : com a adição de uma reflexão especular em um plano perpendicular ao eixo de ordem n .

D_{nd} : com a adição de reflexões especulares em planos paralelos ao eixo de ordem n .

¹⁶ E : do alemão *Einheit* (identidade).

\mathcal{T} : (de tetraedro) indica o grupo que possui a simetria de um tetraedro. Este poliedro possui ao todo 12 eixos de rotação que geram os seus elementos de simetria. Este grupo pode ser ampliado por:

\mathcal{T}_d : inclui 6 rotações impróprias e 6 planos de reflexão, de forma que o grupo possui ao todo 24 elementos.

\mathcal{T}_h : inclui uma inversão, o que gera ao todo 24 elementos também.

\mathcal{O} : (de octaedro) indica o grupo que possui a simetria de um octaedro (ou cubo). Possui ao todo 24 elementos. Ampliado por:

\mathcal{O}_d : inclui rotações impróprias, o que eleva o número total de elementos para 48.

Como já mencionado, de acordo com o teorema da restrição cristalográfica, somente são possíveis eixos de ordens $n = 1, 2, 3, 4$ ou 6 para redes cristalinas no \mathbb{R}^3 , os quais são os valores possíveis para o índice n dos grupos $\mathcal{C}_n, \mathcal{C}_{nh}, \mathcal{C}_{nv}, \mathcal{D}_n, \mathcal{D}_{nh}, \mathcal{D}_{nd}$ e \mathcal{S}_{2n} . É importante mencionar também que $\mathcal{C}_{1v} = \mathcal{C}_{1h}, \mathcal{D}_1 = \mathcal{C}_2, \mathcal{D}_{1h} = \mathcal{C}_{2v}, \mathcal{D}_{1d} = \mathcal{C}_{2h}$ e que os grupos $\mathcal{D}_{4d}, \mathcal{D}_{6d}, \mathcal{S}_8$ e \mathcal{S}_{12} são proibidos.

Os 27 grupos de rotações, juntamente com $\mathcal{T}, \mathcal{T}_d, \mathcal{T}_h, \mathcal{O}$ e \mathcal{O}_h constituem os 32 grupos cristalográficos pontuais. Os exercícios a seguir ilustram a formação de alguns destes grupos.

Exercício 3.10. Construa o grupo pontual de simetrias de um triângulo equilátero.

Resolução: a figura 3.5 ilustra um triângulo equilátero ($n = 3$), com seus vértices e linhas medianas identificadas. O eixo de rotação de ordem 3 atravessa o baricentro em O e é perpendicular ao plano do triângulo. As linhas medianas correspondem às intersecções dos planos verticais (paralelos ao eixo de rotação) com o triângulo. Os números 1 – 6 permanecem sempre fixos, enquanto que os pontos $a, b,$ e c identificam os vértices do triângulo, os quais se movem com o mesmo a cada operação de simetria. As ações dos operadores são denotadas pelas coincidências entre as letras (móveis) e os números (imóveis). Definem-se os seguintes operadores de transformações de simetria:

Rotações próprias: O operador de rotação própria C_3 corresponde a uma rotação por $\psi = 2\pi/3$ radianos ou 120° graus no sentido horário, por convenção. Ou seja, se $\{abc\}$ corresponde à ordem inicial dos vértices, no sentido horário,

$$C_3 \{abc\} \rightarrow \{cab\} \quad C_3^2 \{abc\} \rightarrow \{bca\} \quad E \equiv C_3^3 \{abc\} \rightarrow \{abc\}.$$

Ou seja,

$$C_3 = \{E, C_3, C_3^2\}.$$

A ação de C_3 sobre o triângulo está ilustrada na figura 3.5b.

Reflexões nas medianas: Este polígono não possui reflexões no plano perpendicular ao eixo de rotação, mas existem 3 operações de reflexão σ_v que podem ser definidas, uma para a reflexão em torno de cada mediana, ou seja, σ_{v16} (reflexão em torno da linha 1 – 6), σ_{v24} (em torno da linha 2 – 4) e σ_{v35} (em torno de 3 – 5). Neste caso,

$$\sigma_{v16} \{abc\} \rightarrow \{acb\} \quad \sigma_{v24} \{abc\} \rightarrow \{cba\} \quad \sigma_{v35} \{abc\} \rightarrow \{bac\},$$

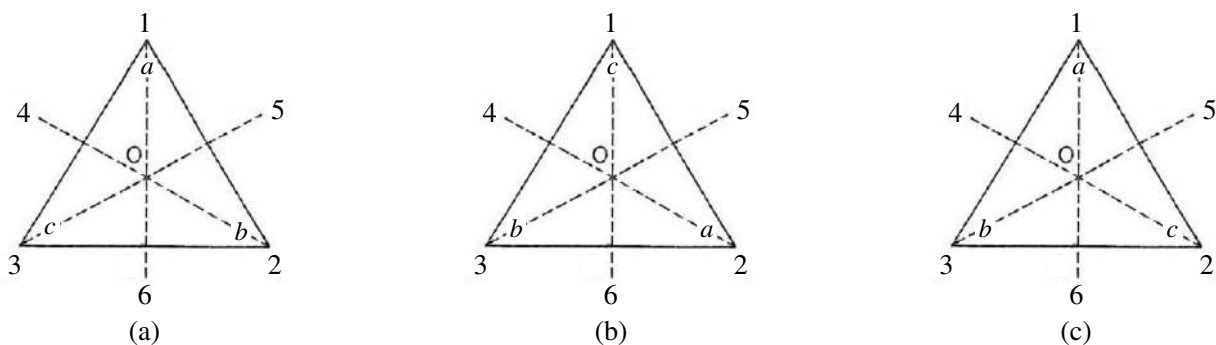


Figura 3.5: Triângulo equilátero: vértices e medianas para a determinação do grupo de simetrias. (a) Configuração original. (b) Ação do operador C_3 . (c) Ação do operador σ_{v16} .

sendo que estes operadores têm ordem igual a 2. A ação de σ_{v16} está ilustrada na figura 3.5c.

Rotações nas medianas: Ao invés de reflexões, podem ser definidas também as 3 operações de rotação por um ângulo de π radianos (ordem 2) em torno de eixos que passam pelas medianas:

$$C_{2,16} \{abc\} \rightarrow \{acb\} \quad C_{2,24} \{abc\} \rightarrow \{cba\} \quad C_{2,35} \{abc\} \rightarrow \{bac\}.$$

Desta forma, o grupo pontual de simetrias de um triângulo equilátero é dado por:

Grupo C_{3v} : (notação internacional: grupo $3m$) composto por

$$C_{3v} = C_3 + (\text{reflexões}) = \{E, C_3, C_3^2, \sigma_{v16}, \sigma_{v24}, \sigma_{v35}\}.$$

Grupo D_3 : (notação internacional: grupo 32) composto por

$$D_3 = \{E, C_3, C_3^2, C_{2,16}, C_{2,24}, C_{2,35}\}.$$

Nota-se que o grupo cíclico C_3 é sempre um subgrupo invariante de C_{3v} ou D_3 .

Exercício 3.11. Construa as tabelas de multiplicações dos grupos pontuais de um triângulo e destaque o isomorfismo existente entre o grupo C_{3v} e os grupos D_3 e S_3 (ver também exercício 3.14).

Resolução: realizando-se todas as multiplicações entre os membros do grupo C_{3v} obtém-se a tabela 3.1. Este grupo é formado por um subgrupo próprio de ordem 3: C_3 e pelas rotações impróprias $\{\sigma_{v16}, \sigma_{v24}, \sigma_{v35}\}$. Pode-se realizar então os mapeamentos entre os elementos de C_{3v} , D_3 e S_3 ilustrados na tabela 3.2. Nota-se também o isomorfismo entre os grupos invariantes: $C_3 \mapsto A_3$.

Tabela 3.1: Tabela de multiplicações do grupo C_{3v} .

E	C_3	C_3^2	σ_{v16}	σ_{v24}	σ_{v35}
C_3	C_3^2	E	σ_{v35}	σ_{v16}	σ_{v24}
C_3^2	E	C_3	σ_{v24}	σ_{v35}	σ_{v16}
σ_{v16}	σ_{v24}	σ_{v35}	E	C_3	C_3^2
σ_{v24}	σ_{v35}	σ_{v16}	C_3^2	E	C_3
σ_{v35}	σ_{v16}	σ_{v24}	C_3	C_3^2	E

Tabela 3.2: Tabela de isomorfismos entre os grupos C_{3v} , D_3 e S_3 .

C_{3v}	D_3	S_3
E	E	I
C_3	C_3	π_5
C_3^2	C_3^2	π_4
σ_{v16}	$C_{2,16}$	π_2
σ_{v24}	$C_{2,24}$	π_6
σ_{v35}	$C_{2,35}$	π_3

Exercício 3.12. Construa o grupo pontual de simetrias de um quadrado e obtenha sua tabela de multiplicações.

Resolução: a figura 3.6 ilustra o quadrado ($n = 4$), identificando seus vértices e os planos de reflexão. O eixo de rotação de ordem 4 atravessa o ponto O e é perpendicular ao plano do quadrado. Os números 1, 2, ..., 8 permanecem sempre fixos, enquanto que os pontos do quadrado

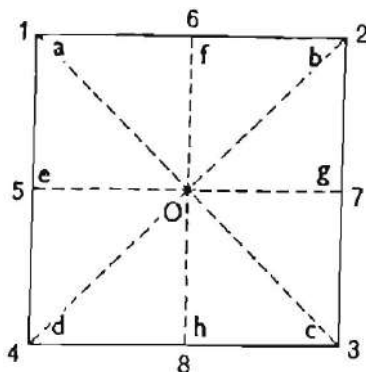


Figura 3.6: Os eixos e planos de simetria de um quadrado.

marcados por a, b, ..., h movem-se com o mesmo frente as transformações de simetria. Assim, os pontos a, b, c e d marcam os vértices do quadrado, enquanto que e, f, g e h marcam os pontos médios de suas arestas. As ações das transformações serão denotadas pelas coincidências entre as letras (móveis) e os números (fixos). Os operadores de simetria são:

Rotações próprias: O operador C_4 consiste em uma rotação horária em torno de O por um ângulo de $\psi = 2\pi/4 = \pi/2$ radianos (90° graus). Identificando cada rotação própria pela variação dos vértices do quadrado, o grupo de rotações C_4 , de ordem 4, é formado pelos operadores:

$$\begin{array}{ll} E\{abcd\} \rightarrow \{abcd\} & C_4\{abcd\} \rightarrow \{dabc\} \\ C_4^2\{abcd\} \rightarrow \{cdab\} & C_4^3\{abcd\} \rightarrow \{bcad\}. \end{array}$$

Ou seja,

$$C_4 = \{E, C_4, C_4^2, C_4^3\}.$$

Reflexões nas medianas e diagonais: As reflexões nas linhas medianas são executadas pelos operadores σ_{v57} , definido como a reflexão especular em torno da linha 5 – 7, e σ_{v68} , gera a reflexão em torno da linha 6 – 8. Já as reflexões nas diagonais são realizadas pelos operadores σ_{v13} e σ_{v24} . Desta forma,

$$\begin{array}{ll} \sigma_{v57}\{abcd\} \rightarrow \{dcba\} & \sigma_{v68}\{abcd\} \rightarrow \{badc\} \\ \sigma_{v13}\{abcd\} \rightarrow \{adcb\} & \sigma_{v24}\{abcd\} \rightarrow \{cbad\}. \end{array}$$

Todos estes operadores têm ordem 2.

Portanto, o grupo de todas as operações de simetria de um quadrado é

$$C_{4v} = C_4 + (\text{reflexões}) = \{E, C_4, C_4^2, C_4^3, \sigma_{v57}, \sigma_{v68}, \sigma_{v13}, \sigma_{v24}\},$$

o qual é de ordem 8.

A tabela de multiplicações do grupo C_{4v} é facilmente obtida:

E	C_4	C_4^2	C_4^3	σ_{v57}	σ_{v68}	σ_{v13}	σ_{v24}
C_4	C_4^2	C_4^3	E	σ_{v13}	σ_{v24}	σ_{v68}	σ_{v57}
C_4^2	C_4	E	C_4	σ_{v68}	σ_{v57}	σ_{v24}	σ_{v13}
C_4^3	E	C_4	C_4^2	σ_{v24}	σ_{v13}	σ_{v57}	σ_{v68}
σ_{v57}	σ_{v24}	σ_{v68}	σ_{v13}	E	C_4^2	C_4^3	C_4
σ_{v68}	σ_{v13}	σ_{v57}	σ_{v24}	C_4^2	E	C_4	C_4^3
σ_{v13}	σ_{v57}	σ_{v24}	σ_{v68}	C_4	C_4^3	E	C_4^2
σ_{v24}	σ_{v68}	σ_{v13}	σ_{v57}	C_4^3	C_4	C_4^2	E

3.4.2 PROJEÇÕES ESTEREOGRÁFICAS

A conexão entre as posições dos átomos em uma rede cristalina no \mathbb{R}^3 com os polígonos regulares permitidos pelo teorema da restrição cristalográfica é realizada através de projeções desses átomos em certos planos, projeções essas que formarão os polígonos estudados. Projeções em planos também são empregados para as faces do cristal.

Uma das projeções mais empregadas em cristalografia é a *projeção estereográfica*. Como esta é obtida para uma estrutura no \mathbb{R}^3 está exemplificado na figura 3.7. Nesta, três átomos em particular de uma molécula ou um cristal encontram-se sobre a superfície de uma esfera, nos pontos P , P'' e P''' , todos situados no hemisfério superior da mesma. No painel (a) observa-se que a partir de cada átomo é traçada uma reta que parte da superfície da esfera até o polo da esfera situado no hemisfério oposto. A intersecção da linha que parte de P até o polo S com o plano equatorial da esfera, representado pela reta \overline{AB} , será a projeção estereográfica deste átomo. Para o átomo em P , a sua projeção, portanto, estará no ponto P' do plano equatorial.

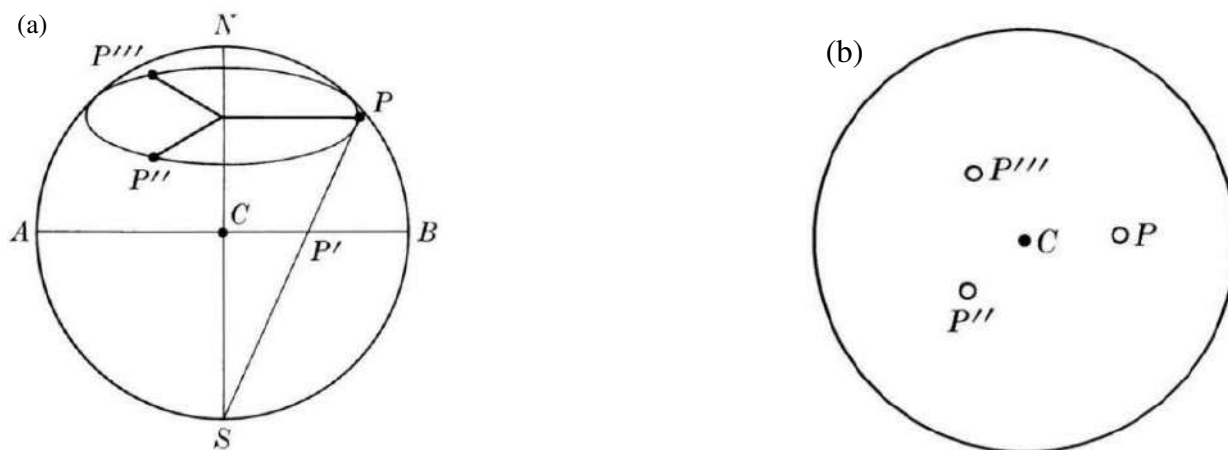


Figura 3.7: (a) Projeção estereográfica dos átomos em P , P'' e P''' na superfície da esfera sobre o plano equatorial. (b) Projeções sobre o plano equatorial dos átomos.

Realizando-se a mesma projeção para os demais átomos em P'' e P''' , a projeção estereográfica desta parte da molécula ou do cristal sobre o plano equatorial é visualizada na figura 3.7b. Observa-se que as projeções, representadas por círculos abertos, compõem os vértices de um triângulo equilátero.

Projeções estereográficas existem para os 32 grupos cristalográficos pontuais. A figura 3.8 mostra estas projeções.

3.4.3 GRUPOS CRISTALOGRÁFICOS ESPACIAIS

Considera-se inicialmente uma rede linear composta por N pontos de rede com uma constante de rede a . Definindo a coordenada medida ao longo da rede linear por x , esta é invariante sob translações por múltiplos de a ao longo da coordenada x .

Idealmente, $N \rightarrow \infty$, mas para se tratar redes finitas (embora com N grande), empregam-se condições de contorno periódicas, ou seja, $x + Na = x$. Definindo T_1 como o operador que realiza a translação ao longo de x pela distância a , a operação de translação T_n denota então a translação pela distância na . Ou seja, se

$$T_1 x = x + a, \text{ então } T_n x = (T_1)^n x = x + na.$$

Devido às condições de contorno periódicas,

$$T_N x = (T_1)^N x = x + Na = x.$$

Portanto, o conjunto

$$\mathcal{T} = \{T_1, T_2, \dots, T_n, \dots, T_{N-1}, T_N = E\}$$

claramente define um grupo cíclico de ordem N . Este grupo será denominado *grupo de translação*.

Os operadores isométricos de translação de uma rede de Bravais no \mathbb{R}^3 também constituem um grupo. Dada a rede gerada pelos vetores primitivos \mathbf{a}_1 , \mathbf{a}_2 e \mathbf{a}_3 e sujeita às condições de contorno periódicas

$$\mathbf{r} + N_1 \mathbf{a}_1 = \mathbf{r}, \quad \mathbf{r} + N_2 \mathbf{a}_2 = \mathbf{r}, \quad \mathbf{r} + N_3 \mathbf{a}_3 = \mathbf{r},$$

onde \mathbf{r} é a posição em algum ponto da rede cristalina e $\{N_1, N_2, N_3\}$ é o número de células em cada direção.

Define-se o operador de translação $T(n_1, n_2, n_3)$ dentro da rede cristalina como

$$T(n_1, n_2, n_3) \mathbf{r} = \mathbf{r} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \equiv \mathbf{r} + \mathbf{t}(n_1, n_2, n_3),$$

onde $1 \leq n_i \leq N_i$. Observa-se que a aplicação sucessiva de dois operadores de translação é equivalente a uma outra translação, *i. e.*,

$$T(m_1, m_2, m_3) T(n_1, n_2, n_3) \mathbf{r} = T(m_1, m_2, m_3) [\mathbf{r} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3]$$

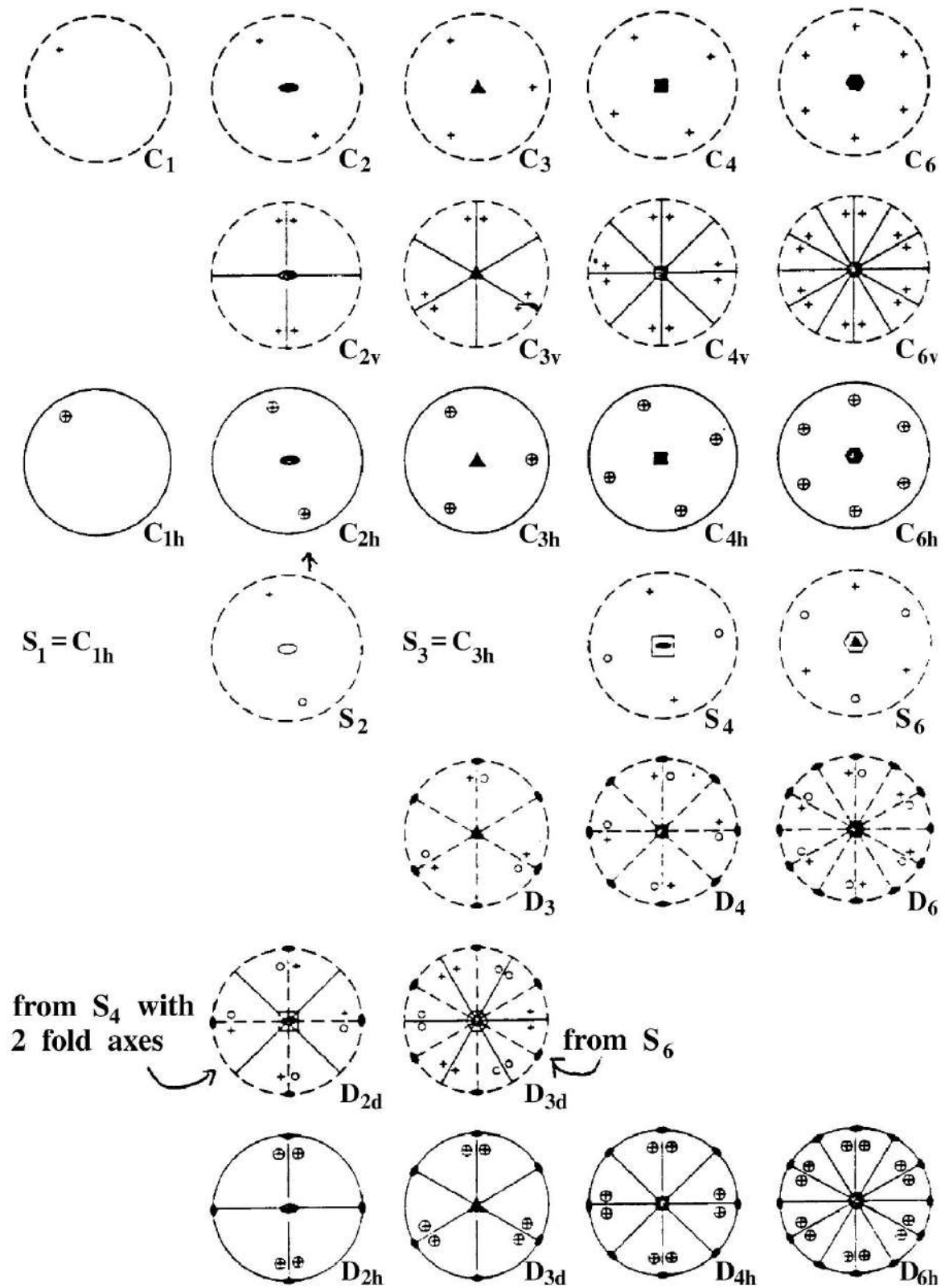


Figura 3.8: Projeções estereográficas para os grupos pontuais. O símbolo "+" significa acima do plano, "O" abaixo do plano e "⊕" no plano.

$$= \mathbf{r} + (n_1 + m_1) \mathbf{a}_1 + (n_2 + m_2) \mathbf{a}_2 + (n_3 + m_3) \mathbf{a}_3 = T(m_1 + n_1, n_2 + m_2, n_3 + m_3) \mathbf{r}.$$

Dessa forma, observa-se que o conjunto de operadores de translação forma um grupo de ordem $N_1 N_2 N_3$ no qual o elemento identidade é $E = T(0, 0, 0) \equiv T(N_1, N_2, N_3)$ e onde o elemento inverso de $T(n_1, n_2, n_3)$ é $T(N_1 - n_1, N_2 - n_2, N_3 - n_3)$. Observa-se também que as operações de translação comutam e, portanto, o grupo é Abelianiano, porém não é cíclico.

O tipo de simetria translacional recém descrita forma um grupo cristalográfico espacial e muitas redes cristalinas reais seguem o mesmo. De fato, cristais simples são invariantes tanto frente às operações de simetria pontual quanto frente a translações. Contudo, cristais mais complexos demandam outros tipos de simetrias. Pode ocorrer que o cristal seja invariante frente a uma operação pontual seguida de uma translação, mas não é isométrico frente a apenas uma das operações. A complexidade apresentada pelas redes cristalinas reais tem como consequência que existem ao todo 230 grupos cristalográficos conhecidos.

3.5 PRODUTO DIRETO DE GRUPOS

Trata-se de um artifício que permite descrever um grupo em termos de seus subgrupos próprios invariantes, mas serve também para combinar grupos distintos e então formar um grupo superior, denominado de *grupo do produto direto*. Neste segundo caso, um grupo do produto direto resulta, por exemplo, quando o conjunto completo de operações de simetria do sistema em consideração (físico ou matemático) pode ser dividido em dois ou mais subconjuntos (formando grupos) tais que os operadores de um subconjunto comutam com os operadores do outro.

Um exemplo já foi apresentado na seção 3.4.1. Em um grupo cristalográfico pontual, um operador de inversão espacial comuta com qualquer operador de rotação própria, por exemplo. Por isso, um determinado cristal que possui um eixo de rotação de ordem n tem o seu grupo de isometrias C_n ampliado pelo produto direto com o grupo $\{E, J\}$, sendo J o operador de inversão espacial. Um outro exemplo consiste no grupo formado pelos operadores que comutam com o hamiltoniano de um determinado sistema eletrônico. Se a interação spin-órbita for desconsiderada em primeira aproximação, então os operadores que atuam sobre a parte orbital da função de onda são independentes dos operadores que atuam sobre os spins (dos elétrons ou dos núcleos). Neste caso, o grupo dos operadores de momentum angular é composto pelo produto dos operadores orbitais pelos de spin.

O produto direto de subgrupos será definido inicialmente.

Definição 3.21. Um grupo $\mathcal{G} = \{G; *\}$ de ordem $|\mathcal{G}|$ é dito ser o *produto direto* de seus subgrupos próprios $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_n$, respectivamente de ordens $|\mathcal{H}_1|, |\mathcal{H}_2|, \dots, |\mathcal{H}_n|$, se:

1. Os subgrupos não possuem nenhum elemento em comum, exceto a identidade.
2. Os elementos dos diferentes subgrupos comutam.
3. Todo elemento $g \in \mathcal{G}$ for esprimível de uma única maneira através dos produtos

$$g = h_1 * h_2 \cdots * h_n = h_2 * h_1 \cdots * h_n = \dots,$$

onde $h_1 \in \mathcal{H}_1, h_2 \in \mathcal{H}_2, \dots, h_n \in \mathcal{H}_n$.

Os subgrupos $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_n$ são denominados os *fatores diretos* de \mathcal{G} . Simbolicamente, escreve-se o produto direto como

$$\mathcal{G} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n,$$

sendo que este produto direto gera o grupo \mathcal{G} de ordem $|\mathcal{G}| = |\mathcal{H}_1| |\mathcal{H}_2| \cdots |\mathcal{H}_n|$.

Como consequência da definição acima, os subgrupos em questão são os subgrupos próprios invariantes de \mathcal{G} . Para exemplificar, supõe-se que \mathcal{G} possua dois subgrupos próprios invariantes: \mathcal{H}_1 de ordem $|\mathcal{H}_1|$ e \mathcal{H}_2 de ordem $|\mathcal{H}_2|$, *i. e.*,

$$\mathcal{H}_1 = \{h_{11} = I, h_{12}, \dots, h_{1|\mathcal{H}_1|}\}, \quad \mathcal{H}_2 = \{h_{21} = I, h_{22}, \dots, h_{2|\mathcal{H}_2|}\}.$$

Então, o produto direto $\mathcal{G} = \mathcal{H}_1 \otimes \mathcal{H}_2$ gera o conjunto composto pelas multiplicações de todos os elementos de \mathcal{H}_1 por todos os elementos de \mathcal{H}_2 :

$$\mathcal{G} = \{I, I * h_{22}, \dots, I * h_{2|\mathcal{H}_2|}, h_{12} * I, h_{12} * h_{22}, \dots, h_{12} * h_{2|\mathcal{H}_2|}, \dots, h_{1|\mathcal{H}_1|} * I, h_{1|\mathcal{H}_1|} * h_{22}, \dots, h_{1|\mathcal{H}_1|} * h_{2|\mathcal{H}_2|}\}.$$

Como os subgrupo são invariantes, ocorre que $h_{12} * h_{22} = h_{22} * h_{12}$, por exemplo, e, por isso, cada elemento de \mathcal{G} é obtido somente uma vez: $g_1 = I, g_2 = h_{22}, \dots, g_i = h_{12} * h_{22}, \dots, g_{|\mathcal{G}|} = h_{1|\mathcal{H}_1|} * h_{2|\mathcal{H}_2|} = h_{2|\mathcal{H}_2|} * h_{1|\mathcal{H}_1|}$, onde $|\mathcal{G}| = |\mathcal{H}_1| |\mathcal{H}_2|$.

Exemplo 3.7. Dado o grupo cíclico $\langle a \rangle$, de ordem 6, onde $a \in \langle a \rangle \mid a^6 = I$, os seus fatores diretos são

$$A = \{I, a^2, a^4\}, \quad B = \{I, a^3\}.$$

Então,

$$\langle a \rangle = A \otimes B = \{I, a^3, a^2, a^2 a^3 = a^5, a^4, a^4 a^3 = a\}.$$

Exemplo 3.8. Dentro do grupo \mathcal{C}_{4v} de isometrias pontuais de um quadrado, discutido no exercício 3.12, os subgrupos cíclicos $\{E, \sigma_{v57}\}$ e $\{E, \sigma_{v68}\}$ formam por produto direto o subgrupo

$$\{E, \sigma_{v57}\} \otimes \{E, \sigma_{v68}\} = \{E, C_4^2, \sigma_{v57}, \sigma_{v68}\}.$$

Se a estrutura completa de um grupo \mathcal{G} não for conhecida, mas se dois ou mais (supostos) subgrupos invariantes de \mathcal{G} o são, pelo fato de compartilharem o mesmo produto de grupo, a mesma identidade e de satisfazerem as condições da definição 3.21, então um novo subgrupo de \mathcal{G} (ou talvez o próprio) pode ser construído pelo produto direto dos subgrupos menores.

Também é possível partir-se de dois grupos distintos, \mathcal{G} e \mathcal{G}' , tais que os elementos do primeiro são independentes do segundo, e então construir o grupo $\mathcal{G} \otimes \mathcal{G}'$ superior, denominado *grupo do produto direto*. Isto é realizado pela formação de todos os pares

$$(a, a') \in \mathcal{G} \otimes \mathcal{G}', \text{ tais que } a \in \mathcal{G} \text{ e } a' \in \mathcal{G}'.$$

O produto de pares em $\mathcal{G} \otimes \mathcal{G}'$ é definido por

$$(a, a')(b, b') = (ab, a'b').$$

Como $ab \in \mathcal{G}$ e $a'b' \in \mathcal{G}'$, então $(ab, a'b') \in \mathcal{G} \otimes \mathcal{G}'$, de acordo com a exigência de clausura.

Se $I \in \mathcal{G}$ e $I' \in \mathcal{G}'$ são as respectivas identidades, então os pares (a, I') formam um subgrupo $\Gamma \subset \mathcal{G} \otimes \mathcal{G}'$ isomórfico a \mathcal{G} ,¹⁷ enquanto que os pares (I, a') formam um subgrupo $\Gamma' \in \mathcal{G} \otimes \mathcal{G}'$ isomórfico a \mathcal{G}' .

Exemplo 3.9 (*Grupo de isometrias pontuais de uma cunha*). Considere todas as operações de simetria de uma cunha cujas faces são triângulos equiláteros, mas que possui também uma certa espessura. Neste caso, além das operações de isometrias contidas no grupo \mathcal{D}_3 , o qual é isomorfo ao grupo \mathcal{C}_{3v} ,¹⁸ surge uma transformação isométrica adicional constituída pela reflexão em torno do plano perpendicular ao eixo de ordem 3, denotada pelo símbolo σ_h . Este operador é o gerador do grupo de ordem 2

$$\mathcal{C}_{1h} = \{E, \sigma_h\}.$$

Pode-se observar facilmente que qualquer rotação (própria ou imprópria) do triângulo seguida pela reflexão operada por σ_h terá o mesmo resultado que a ordem inversa de aplicação dos operadores. Neste caso, o grupo do produto direto $\mathcal{D}_3 \otimes \mathcal{C}_{1h}$ possui ordem 12 e este forma o grupo \mathcal{D}_{3h} , cujos elementos são

$$\mathcal{D}_{3h} = \mathcal{D}_3 \otimes \mathcal{C}_{1h} = \{E, C_3, C_3^2, \sigma_{v16}, \sigma_{v24}, \sigma_{v35}, \sigma_h, C_3\sigma_h, C_3^2\sigma_h, \sigma_{v16}\sigma_h, \sigma_{v24}\sigma_h, \sigma_{v35}\sigma_h\}.$$

3.6 MAPEAMENTOS ENTRE GRUPOS

Até este momento, os grupos abstratos e suas propriedades foram estudados de forma individual. Contudo, é possível estabelecer-se relações entre diferentes grupos que permitem generalizar o estudo das propriedades comuns entre os mesmos. Estas relações são estabelecidas a partir de *mapeamentos* entre os grupos, realizados de tal forma que existe um *homomorfismo* ou *isomorfismo* entre os mesmos. Alguns desses conceitos serão agora discutidos.

¹⁷A definição de isomorfia será apresentada na seção 3.6.

¹⁸Ver exercícios 3.10 e 3.11.

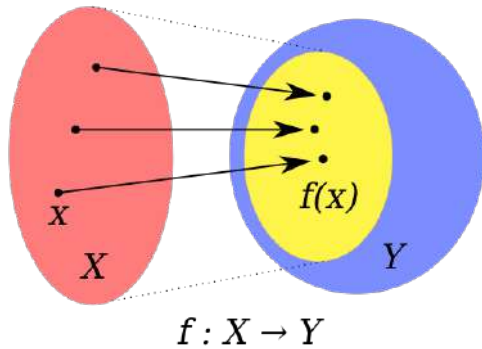


Figura 3.9: Representação do mapeamento entre os conjunto domínio (X) e contradomínio (Y) da função $f(x)$. O subconjunto amarelo de Y representa a imagem da função.

3.6.1 FUNÇÕES E MAPEAMENTOS

Um *mapeamento* entre grupos consiste em alguma relação entre um elemento de um dado grupo \mathcal{G} e um elemento de um outro grupo \mathcal{G}' . A forma como esta relação pode ser estabelecida é bastante ampla, e possui uma grande semelhança com o conceito de *função* em análise matemática. Por esta razão, será feita inicialmente a definição formal deste conceito, para depois discutir-se o mapeamento entre grupos.

Definição 3.22 (Função). Uma *função* é uma relação binária entre dois conjuntos, sendo um conjunto de entrada, denominado o *domínio* e um conjunto de saída, denominado o *contradomínio*. Se o conjunto domínio de uma função denominada f é identificado por X e o conjunto contradomínio é identificado por Y , então a relação binária entre os conjuntos é representada por

$$f: X \mapsto Y \text{ ou } X \xrightarrow{f} Y,$$

sendo dito que f é a *função de X em Y* .

A operação realizada por f sobre o elemento $x \in X$ é usualmente representada por $f(x)$, quando então x é denominado o *argumento* de f . Quando a função f estabelece a relação binária entre o argumento x e o elemento $y \in Y$, y é denominado o *valor de f em x* , e essa relação é usualmente representada por $y = f(x)$.

Para que a relação binária $f: X \mapsto Y$ seja uma função, é necessário que:

1. f seja *unívoca*: se $y = f(x)$ e $z = f(x)$, então $z = y$.
2. f seja *total*: para todo $x \in X$ existe um $y \in Y$ tal que $y = f(x)$.

Para cada $x \in X$ relacionado com $y \in Y \mid y = f(x)$ é estabelecido um *par ordenado* (x, y) . Sendo $\{x_1, x_2, \dots, x_n\} \subseteq X$ o conjunto de argumentos operados por f e $\{y_1, y_2, \dots, y_n\} \subseteq Y$, com $y_1 = f(x_1)$, etc, o conjunto de valores de f , o conjunto de pares ordenados

$$C = X \times Y = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$$

é denominado o *produto Cartesiano de X e Y* .¹⁹ Se $\{y_1, y_2, \dots, y_n\} \subset Y$, este conjunto é usualmente denominado a *imagem* da função.

A figura 3.9 ilustra o mapeamento entre conjuntos realizado pela função $f: X \mapsto Y$. Existem diversos tipos de funções, dependendo da regra que associa os elementos de X aos elementos de Y e da amplitude da ação de f sobre X . Alguns desses tipos serão citados:

Função injetora ou injetiva. Quando cada $y \in Y$ está associado a apenas um $x \in X$, *i. e.*, se para todo $x_1, x_2 \in X$ tais que $x_1 \neq x_2$, com os valores $y_1 = f(x_1)$ e $y_2 = f(x_2)$, então $y_1 \neq y_2$.

Função sobrejetora ou sobrejetiva. Quando todos os elementos do contradomínio estão associados a algum elemento do domínio, *i. e.*, o conjunto imagem se torna igual ao contradomínio inteiro.

Função bijetora ou bijetiva. Quando a função é ao mesmo tempo sobrejetora e injetora, *i. e.*, cada elemento do domínio está associado a um único elemento do contradomínio e vice-versa.

¹⁹Ver definição 3.27.

Mapeamento. Uma função $f : X \mapsto Y$ é denominada um *mapeamento* se o conjunto $\{x_1, x_2, \dots, x_n\}$ de elementos de X operados por f é o conjunto X inteiro. Um mapeamento pode ainda ser dividido em:

Mapeamento de X em Y . Quando a imagem do mapeamento está contida no contradomínio, i. e., $\{y_1, y_2, \dots, y_n\} \subset Y$. Este mapeamento pode ser representado por $f : X \rightarrow Y$.

Mapeamento de X sobre Y . Quando o conjunto de valores de f é igual ao contradomínio inteiro. Este mapeamento pode ser representado por $f : X \xrightarrow{\text{sobre}} Y$.

As definições recém apresentadas sobre funções e mapeamentos podem ser agora implementadas para o mapeamento entre dois grupos.

3.6.2 MAPEAMENTO ENTRE GRUPOS E HOMOMORFISMO

A definição e tipos de funções apresentados acima serão agora generalizados na definição de mapeamento entre grupos. Dentre todos os mapeamentos possíveis, os mais importantes são aqueles que estabelecem um *homomorfismo* entre dois grupos. Assim, nesta seção esses conceitos serão apresentados e discutidos.

Definição 3.23 (Mapeamento entre grupos). Dados os grupos $\mathcal{G} = \{G; *\}$ e $\mathcal{G}' = \{G'; \bullet\}$, o mapeamento Φ de \mathcal{G} em \mathcal{G}' , representado por

$$\Phi : \mathcal{G} \mapsto \mathcal{G}',$$

consiste em uma relação binária entre cada elemento $g \in G$ com um elemento $g' \in G'$, o qual é denominado o *valor de Φ em g* . Esta relação pode ser representada por $g' = \Phi(g)$.

Exemplo 3.10. Dados os grupos:

- $GL(n, \mathbb{R})$: grupo geral linear composto pelas matrizes $n \times n$ reais inversíveis, frente à multiplicação matricial.
- $(\mathbb{R}^*; \times)$: grupo formado pelo conjunto dos números reais, exceto o 0, frente ao produto algébrico.

Pode-se definir o mapeamento

$$\Phi : GL(n, \mathbb{R}) \mapsto (\mathbb{R}^*; \times)$$

tal que para cada matriz $A \in GL(n, \mathbb{R})$ corresponde o elemento $\Delta \in (\mathbb{R}^*; \times)$ determinado por

$$\Delta = \Phi(A) \equiv \det(A).$$

Diz-se então que cada matriz A é mapeada ao seu determinante Δ .

A partir da definição básica de mapeamento entre grupos, alguns tipos importantes de mapeamentos podem então ser definidos:

Homomorfismo. Dados os grupos $\mathcal{G} = \{G; *\}$ e $\mathcal{G}' = \{G'; \bullet\}$, o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}'$ será denominado um *homomorfismo do grupo \mathcal{G} para o grupo \mathcal{G}'* se cada elemento $g \in \mathcal{G}$ for mapeado a um único elemento $g' = \Phi(g) \in \mathcal{G}'$ de tal forma que

$$\Phi(g_1 * g_2) = \Phi(g_1) \bullet \Phi(g_2),$$

sendo $\{g_1, g_2\} \subset \mathcal{G}$ e $\{g'_1 = \Phi(g_1), g'_2 = \Phi(g_2)\} \subset \mathcal{G}'$. O elemento $g' = \Phi(g) \in \mathcal{G}'$ é denominado a *imagem* ou o *mapa do elemento g de \mathcal{G} sob o homomorfismo*.

Monomorfismo. Dados os grupos $\mathcal{G} = \{G; *\}$ e $\mathcal{G}' = \{G'; \bullet\}$, o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}'$ será denominado um *monomorfismo do grupo \mathcal{G} para o grupo \mathcal{G}'* se:

1. O mapeamento for homomórfico.
2. O mapeamento for *injetivo*, i. e., para todos $g_1, g_2 \in \mathcal{G}$ tais que $g_1 \neq g_2$, se $g'_1 = \Phi(g_1)$ e $g'_2 = \Phi(g_2)$, com $g'_1, g'_2 \in \mathcal{G}'$, então $g'_1 \neq g'_2$.

Epimorfismo. Dados os grupos $\mathcal{G} = \{G; *\}$ e $\mathcal{G}' = \{G'; \bullet\}$, o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}'$ será denominado um *epimorfismo do grupo \mathcal{G} para o grupo \mathcal{G}'* se:

1. O mapeamento for homomórfico.
2. O mapeamento for sobrejetivo, *i. e.*, todos os elementos de \mathcal{G}' forem mapeados a partir de algum elemento de \mathcal{G} .

Isomorfismo. Dados os grupos $\mathcal{G} = \{G; *\}$, de ordem $|\mathcal{G}|$, e $\mathcal{G}' = \{G'; \bullet\}$, de ordem $|\mathcal{G}'|$, o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}'$ será denominado um *isomorfismo* se:

1. Existir um homomorfismo do grupo \mathcal{G} para o grupo \mathcal{G}' .
2. Os grupos tiverem a mesma ordem; *i. e.*, $|\mathcal{G}'| = |\mathcal{G}|$.
3. Existir uma correspondência bijetora entre cada elemento $g \in \mathcal{G}$ com cada elemento $g' \in \mathcal{G}'$, *i. e.*, se $G = \{g_1, g_2, \dots, g_{|\mathcal{G}|}\}$ e $G' = \{g'_1, g'_2, \dots, g'_{|\mathcal{G}'|}\}$, então

$$g_1 \leftrightarrow g'_1, \quad g_2 \leftrightarrow g'_2, \dots, g_{|\mathcal{G}|} \leftrightarrow g'_{|\mathcal{G}'|}.$$

Os grupos \mathcal{G} e \mathcal{G}' são ditos então *isomórficos*, escrevendo-se $\mathcal{G} \cong \mathcal{G}'$.

Endomorfismo. Dado o grupo $\mathcal{G} = \{G; *\}$, o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}$ é denominado um *endomorfismo de \mathcal{G}* .

Automorfismo. Dado o grupo $\mathcal{G} = \{G; *\}$, o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}$ é denominado um *automorfismo de \mathcal{G}* se:

1. O mapeamento for endomórfico.
2. O mapeamento for isomórfico.

Exercício 3.13. Mostre que o mapeamento definido no exemplo 3.10 é um tipo de homomorfismo e classifique o mesmo.

Resolução: Dadas as matrizes $A_1, A_2 \in GL(n, \mathbb{R})$ e os números reais $\Delta_1, \Delta_2 \in (\mathbb{R}^*, \times)$ tais que $\Delta_1 = \det(A_1)$ e $\Delta_2 = \det(A_2)$. O mapeamento $\Phi : GL(n, \mathbb{R}) \mapsto (\mathbb{R}^*; \times) \mid \Phi(A) = \det(A)$ é um homomorfismo porque

$$\det(A_1 A_2) = \det(A_1) \det(A_2) = \Delta_1 \Delta_2.$$

Além disso, o mapeamento será um *epimorfismo* porque, uma vez que o grupo $GL(n, \mathbb{R})$ é de dimensão n^2 , sempre haverá uma combinação de parâmetros da matriz A cujo determinante será mapeado a qualquer número real Δ . Por outro lado, o mapeamento não é isomórfico porque a um dado Δ sempre haverá mais de uma matriz A mapeada ao mesmo.

Alguns exemplos de homomorfismo:

- No exercício 3.11 já foi demonstrado o isomorfismo entre os grupos \mathcal{C}_{3v} , \mathcal{D}_3 e \mathcal{S}_3 .
- $\Phi : \mathcal{C}_4 \mapsto \mathcal{Z}_4$, com $E \leftrightarrow 1$, $C_4 \leftrightarrow i$, $C_4^2 \leftrightarrow -1$ e $C_4^3 \leftrightarrow -i$. Observa-se o isomorfismo claramente nas tabelas de multiplicação:

	\mathcal{C}_4				\mathcal{Z}_4		
E	C_4	C_4^2	C_4^3	1	i	-1	$-i$
C_4	C_4^2	C_4^3	E	i	-1	$-i$	1
C_4^2	C_4^3	E	C_4	-1	$-i$	1	i
C_4^3	E	C_4	C_4^2	$-i$	1	i	-1

- Dados o grupo $(\mathbb{Z}; +)$ e o grupo cíclico $\mathcal{Z}/3 = \{0, 1, 2; +(\text{mod } 3)\}$, o mapeamento $\Phi : (\mathbb{Z}; +) \mapsto \mathcal{Z}/3$, onde para todo $n \in \mathbb{Z}$, $\Phi(n) = n \pmod{3}$ forma um epimorfismo.
- Dados $\mathcal{H} \subset GL(2, \mathbb{R})$, formado pelo conjunto

$$H = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mid a > 0, b \in \mathbb{R} \right\},$$

para qualquer $u \in \mathbb{C}$, define-se $f_u : \mathcal{H} \mapsto (\mathbb{C} \setminus \{0\}; \times)$ tal que

$$f_u \left[\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \right] = a^u.$$

Novamente este mapeamento consiste em um epimorfismo.

Algumas observações e consequências diretas das definições do homomorfismo $\Phi : \mathcal{G} \mapsto \mathcal{G}'$ acima:

- Se n elementos de \mathcal{G} forem mapeados ao mesmo elemento de \mathcal{G}' , diz-se que ocorre um mapeamento ou homomorfismo n -para-1 de \mathcal{G} para \mathcal{G}' . Claramente, se $n = 1$ para todo $g \in \mathcal{G}$, o mapeamento se reduz a um isomorfismo.
- Se $I \in \mathcal{G}$ e $I' \in \mathcal{G}'$ forem os respectivos elementos identidade, então, para todo $g \in \mathcal{G}$ e $g' = \Phi(g) \in \mathcal{G}'$,

$$g' = \Phi(I * g) = \Phi(I) \bullet \Phi(g) = \Phi(I) \bullet g' \implies I \mapsto I'. \quad (3.8a)$$

Ou seja, o elemento identidade em \mathcal{G} é sempre mapeado ao elemento identidade I' em \mathcal{G}' .

- Se $g, g^{-1} \in \mathcal{G}$ forem tais que $g * g^{-1} = I$ e $g', (g')^{-1} \in \mathcal{G}'$ forem tais que $g' \bullet (g')^{-1} = I'$ e $g' = \Phi(g)$, então

$$I' = \Phi(g * g^{-1}) = \Phi(g) \bullet \Phi(g^{-1}) = g' \bullet \Phi(g^{-1}) \implies (g')^{-1} = \Phi(g^{-1}). \quad (3.8b)$$

Ou seja, o elemento inverso em \mathcal{G} é mapeado ao respectivo elemento inverso em \mathcal{G}' .

- Se o elemento $g \in \mathcal{G}$ for de ordem n (finita), i. e., $g^n = I$, então

$$I' = \Phi(g^n) = \Phi(g) \bullet \Phi(g^{n-1}) = g' \bullet \Phi(g^{n-1}) = \dots = \underbrace{g' \bullet g' \bullet \dots \bullet g'}_{n \text{ vezes}}.$$

Ou seja,

$$(g')^n = I'. \quad (3.8c)$$

O teorema a seguir está relacionado com a propriedade (3.8c).

Teorema 3.5 (Ordem da imagem sob o homomorfismo). *Sejam os grupos $\mathcal{G} = \{G; *, I\}$ e $\mathcal{G}' = \{G'; \bullet, I'\}$. Seja o mapeamento $\Phi : \mathcal{G} \rightarrow \mathcal{G}'$, o qual forma um homomorfismo de \mathcal{G} em \mathcal{G}' . Seja $g \in \mathcal{G}$ um elemento de ordem finita. Então, $\text{ord}(\Phi(g))$ divide $\text{ord}(g)$.*

Demonstração. Seja $g' = \Phi(g)$; sejam $n, m \in \mathbb{Z}$ tais que $n = \text{ord}(g)$ e $m = \text{ord}(g')$. Então, elevando-se ambos os lados da identidade $(g')^m = I'$ a uma potência inteira positiva k resulta $(g')^{km} = I'$, o que implica, pela propriedade (3.8c), que

$$km = n \implies \frac{n}{m} = k = 1, 2, \dots$$

□

Se o mapeamento for monomórfico, então, necessariamente, $m = n$; caso contrário, é possível que $m < n$. A ordem da imagem do homomorfismo também é um divisor da ordem do grupo do domínio. A demonstração disto é dada pelo primeiro teorema do isomorfismo abaixo.

Mais algumas definições e consequências importantes:

Definição 3.24 (Núcleo e imagem do homomorfismo). Dado o homomorfismo $\Phi : \mathcal{G} \mapsto \mathcal{G}'$, o núcleo²⁰ de Φ é o conjunto de elementos de \mathcal{G} que são mapeados ao elemento identidade de \mathcal{G}' . Este conjunto é representado por

$$\ker(\Phi) = \{g \in \mathcal{G} \mid \Phi(g) = I'\}.$$

A imagem de Φ é o conjunto de imagens dos elementos de \mathcal{G} em \mathcal{G}' , i. e.,

$$\text{Im}(\Phi) \equiv \Phi(\mathcal{G}) = \{\Phi(g), \forall g \in \mathcal{G}\}.$$

A importância das definições acima está nos seguintes teoremas:

Teorema 3.6 (Primeiro teorema do isomorfismo). *Dados os grupos $\mathcal{G} = \{G; *, I\}$ e $\mathcal{G}' = \{G'; \bullet, I'\}$, se o mapeamento $\Phi : \mathcal{G} \mapsto \mathcal{G}'$ forma um homomorfismo de \mathcal{G} em \mathcal{G}' , então:*

1. O núcleo do homomorfismo forma um subgrupo invariante de \mathcal{G} .
2. A imagem do homomorfismo forma um subgrupo de \mathcal{G}' .

²⁰Em alemão: Kernel.

3. A imagem do homomorfismo é isomórfica ao grupo fator $\mathcal{G}/\ker(\Phi)$.

Demonstração. Será demonstrada aqui somente a primeira afirmação do teorema. Primeiro, demonstra-se que $\ker(\Phi)$ forma um subgrupo. Dados $g_1, g_2 \in \ker(\Phi)$, observa-se que

$$\Phi(g_1 * g_2) = \Phi(g_1) \bullet \Phi(g_2) = I',$$

portanto, $g_1 * g_2 \in \ker(\Phi)$ e a condição de clausura é satisfeita. Dado agora $g_1 \in \ker(\Phi)$ e $g_1^{-1} \in \mathcal{G}$. Então,

$$I' = \Phi(g_1 * g_1^{-1}) = \Phi(g_1) \bullet \Phi(g_1^{-1});$$

ou seja, se $g_1 \in \ker(\Phi)$, então necessariamente $g_1^{-1} \in \ker(\Phi)$, e os axiomas de grupo são satisfeitos.

Para demonstrar agora que $\ker(\Phi)$ é invariante, é necessário que para todo $g \in \mathcal{G}$, $g^{-1} \ker(\Phi) g = \ker(\Phi)$. Então, para todo $g_k \in \ker(\Phi)$,

$$\begin{aligned} \Phi(g^{-1} * g_k * g) &= \Phi(g^{-1}) \bullet \Phi(g_k) \bullet \Phi(g) = \Phi(g^{-1}) \bullet I' \bullet \Phi(g) \\ &= \Phi(g^{-1}) \bullet \Phi(g) = \Phi(g^{-1} * g) = I'. \end{aligned}$$

Portanto, para qualquer $g \in \mathcal{G}$, $g^{-1} * g_k * g \in \ker(\Phi)$ e o grupo é invariante. \square

Teorema 3.7 (Segundo teorema do isomorfismo). Dado o grupo \mathcal{G} , seja $\mathcal{H} \subseteq \mathcal{G}$ um subgrupo de \mathcal{G} e \mathcal{N} um subgrupo invariante de \mathcal{G} , então:

1. O produto $\mathcal{H}\mathcal{N}$ é um subgrupo de \mathcal{G} .
2. A intersecção $\mathcal{H} \cap \mathcal{N}$ é um subgrupo invariante de \mathcal{H} .
3. Os grupos fatores $(\mathcal{H}\mathcal{N})/\mathcal{N}$ e $\mathcal{H}/(\mathcal{H} \cap \mathcal{N})$ são isomórficos.

Teorema 3.8 (Terceiro teorema do isomorfismo). Dado o grupo \mathcal{G} com subgrupos invariantes \mathcal{N} e \mathcal{K} tais que

$$\mathcal{K} \subseteq \mathcal{N} \subseteq \mathcal{G},$$

então:

1. O fator \mathcal{N}/\mathcal{K} é um subgrupo normal do grupo fator \mathcal{G}/\mathcal{K} .
2. O grupo fator $(\mathcal{G}/\mathcal{K}) / (\mathcal{N}/\mathcal{K})$ é isomórfico a \mathcal{G}/\mathcal{N} .

Finalmente, o teorema a seguir afirma a singular importância do grupo simétrico S_n dentro da classe dos grupos finitos.

Teorema 3.9 (Teorema de Cayley²¹). Todo grupo \mathcal{G} é isomórfico a um subgrupo do grupo simétrico $\text{Sym}(\mathcal{G})$ atuando sobre \mathcal{G} .

A demonstração deste teorema pode ser realizada tanto para grupos finitos quanto para qualquer outro tipo de grupo.

Demonstração. (grupos finitos) Dado o grupo $\mathcal{G} = \{G; *\}$ de ordem n , onde $G = \{a_1, a_2, \dots, a_n\}$. Dado o elemento $b \in \mathcal{G}$, o produto

$$bG = \{b * a_1, b * a_2, \dots, b * a_n\}$$

contém todos os elementos de G , porém arranjados em uma ordem distinta da original, de acordo com o teorema 3.1. Dado agora o grupo S_n , existe sempre um elemento $\pi_b \in S_n$, cuja ação em G irá permutar os elementos do conjunto na mesma ordem de bG , ou seja,

$$\pi_b G = bG \implies \pi_b = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi_b(ba_1) & \pi_b(ba_2) & \dots & \pi_b(ba_n) \end{pmatrix},$$

onde $\pi_b(ba_i) = 1, 2, \dots, n$ é a operação que fornece a posição original do elemento a_i e cujo ordenamento em π_b indica a nova posição. Realiza-se então a associação $b \rightarrow \pi_b$.

²¹Em homenagem ao matemático britânico Arthur Cayley (1821 – 1895).

Toma-se agora um outro elemento $c \in \mathcal{G}$, a partir do qual é gerado o conjunto cG . O novo ordenamento dos elementos de G é descrito agora por um outro operador $\pi_c \in S_n$, dado por

$$\pi_c = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi_c(ca_1) & \pi_c(ca_2) & \dots & \pi_c(ca_n) \end{pmatrix}.$$

Realiza-se então a associação $c \rightarrow \pi_c$. Procedese desta forma para todos $a_i \in \mathcal{G}$ ($i = 1, \dots, n$), criando-se o subconjunto $P_{nn} \subseteq S_n$, composto por todas as permutações $\pi_{a_i} \leftarrow a_i$. Em particular, como $c * b * a_i \in \mathcal{G}$, deve existir um elemento $\pi_{cb} \in P_{nn}$ que descreve a permutação levada a cabo por $(c * b)G$, ou seja, se

$$\pi_{cb} = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi_c(cba_1) & \pi_c(cba_2) & \dots & \pi_c(cba_n) \end{pmatrix} \implies c * b \rightarrow \pi_{cb}.$$

Para provar que existe um isomorfismo entre \mathcal{G} e algum subgrupo de S_n , é necessário mostrar que o conjunto P_{nn} forma um subgrupo de S_n e que sua tabela de multiplicações é idêntica à tabela de \mathcal{G} , ou seja, assim como $d = c * b$, é necessário que $\pi_c \circ \pi_b = \pi_{cb}$.

Para tanto, observa-se que no operador π_c , por exemplo, a ordem numérica na primeira linha é irrelevante, uma vez que as posições na segunda linha serão sempre referentes às posições na primeira. Então, pode-se escrever

$$\pi_c = \begin{pmatrix} \pi_b(ba_1) & \pi_b(ba_2) & \dots & \pi_b(ba_n) \\ \pi_c(c(ba_1)) & \pi_c(c(ba_2)) & \dots & \pi_c(c(ba_n)) \end{pmatrix}.$$

Ou seja,

$$\begin{aligned} \pi_c \circ \pi_b &= \begin{pmatrix} \pi_b(ba_1) & \pi_b(ba_2) & \dots & \pi_b(ba_n) \\ \pi_c(c(ba_1)) & \pi_c(c(ba_2)) & \dots & \pi_c(c(ba_n)) \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & \dots & n \\ \pi_b(ba_1) & \pi_b(ba_2) & \dots & \pi_b(ba_n) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & \dots & n \\ \pi_c(c(ba_1)) & \pi_c(c(ba_2)) & \dots & \pi_c(c(ba_n)) \end{pmatrix} = \pi_{cb}. \end{aligned}$$

A associação $I \in \mathcal{G}$ com $I \in S_n$ também é respeitada, uma vez que $\pi_1 = I$ faz necessariamente parte de P_{nn} . Além disso, como $a_i^{-1} \in \mathcal{G}$, também é necessariamente feita a associação $a_i^{-1} \rightarrow \pi_{a_i^{-1}} \in P_{nn}$. Finalmente, como $|P_{nn}| = n$, o mapeamento é sobrejetivo e injetivo, ou seja, existe um isomorfismo entre \mathcal{G} e P_{nn} . \square

Exercício 3.14. Use o teorema de Cayley para demonstrar o isomorfismo entre os grupos \mathcal{C}_{3v} e S_3 .

Resolução: Representando o triângulo equilátero da figura 3.5a pela sequência de vértices $\{abc\}$, temos a seguinte associação entre os elementos de \mathcal{C}_{3v} e de S_3 , sendo estes dados na notação cíclica do exemplo 3.4:

$$\begin{aligned} C_3 \{abc\} = \{cab\} &\rightarrow \pi_5 \{abc\} = (132) \{abc\} = \{cab\} \implies C_3 \mapsto \pi_5 \\ C_3^2 \{abc\} = \{bca\} &\rightarrow \pi_4 \{abc\} = (123) \{abc\} = \{bca\} \implies C_3^2 \mapsto \pi_4 \\ \sigma_{v16} \{abc\} = \{acb\} &\rightarrow \pi_2 \{abc\} = (23) \{abc\} = \{acb\} \implies \sigma_{v16} \mapsto \pi_2 \\ \sigma_{v24} \{abc\} = \{cba\} &\rightarrow \pi_6 \{abc\} = (13) \{abc\} = \{cba\} \implies \sigma_{v24} \mapsto \pi_6 \\ \sigma_{v35} \{abc\} = \{bac\} &\rightarrow \pi_3 \{abc\} = (12) \{abc\} = \{bac\} \implies \sigma_{v35} \mapsto \pi_3. \end{aligned}$$

Assim, as tabelas de multiplicação ficam idênticas e o isomorfismo demonstrado, se uma das tabelas de multiplicação for alterada adequadamente, como por exemplo:

S_3						\mathcal{C}_{3v}					
I	π_2	π_3	π_4	π_5	π_6	E	σ_{v16}	σ_{v35}	C_3^2	C_3	σ_{v24}
π_2	I	π_4	π_3	π_6	π_5	σ_{v16}	E	C_3^2	σ_{v35}	σ_{v24}	C_3
π_3	π_5	I	π_6	π_2	π_4	σ_{v35}	C_3	E	σ_{v24}	σ_{v16}	C_3^2
π_4	π_6	π_2	π_5	I	π_3	C_3^2	σ_{v24}	σ_{v16}	C_3	E	σ_{v35}
π_5	π_3	π_6	I	π_4	π_2	C_3	σ_{v35}	σ_{v24}	E	C_3^2	σ_{v16}
π_6	π_4	π_5	π_2	π_3	I	σ_{v24}	C_3^2	C_3	σ_{v16}	σ_{v35}	E

3.7 ESTRUTURAS ALGÉBRICAS

Nesta seção será feita uma descrição sucinta de algumas *Estruturas Algébricas*, entre as quais os grupos estão contidos. A principal razão para tanto está na necessidade de definição formal de uma *álgebra*, onde serão discutidos os grupos e a álgebra de Lie.

Uma estrutura algébrica é composta por um conjunto de objetos, os quais se interrelacionam através de uma ou mais *operações finitárias* definidas sobre o conjunto. As definições dessas operações finitárias e de outros conceitos envolvendo relações entre objetos contidos em conjuntos serão apresentadas antes de se definir as estruturas propriamente ditas.

Definição 3.25 (Operação finitária). Uma *operação finitária* é uma operação (ou ação ou procedimento) aplicada sobre um número finito de objetos, produzindo um resultado.

Se a operação é realizada sobre um número infinito de objetos, está é denominada *operação infinitária*. Dentre as operações finitárias, as mais comuns são:

- Operações 1-árias ou *unárias*: operações executadas sobre um único objeto.
- Operações 2-árias ou *binárias*: operações executadas entre dois objetos. Adição e multiplicação são os exemplos mais comuns de operações binárias.

Definição 3.26 (Estrutura algébrica). Seja C um conjunto de objetos e \mathcal{F} uma coleção de operações sobre C e elementos especiais. A dupla $\langle C, \mathcal{F} \rangle$ é denominada uma *estrutura algébrica sobre C* .

Muitas vezes, por simplicidade e quando não houver ambiguidade, representa-se a estrutura $\langle C, \mathcal{F} \rangle$ simplesmente por C . Na sua maior generalidade possível, estruturas algébricas podem envolver um número arbitrário de conjuntos e uma vasta coleção de operações entre os mesmos. De interesse à física, serão definidas somente estruturas envolvendo, no máximo, dois conjuntos e com, no máximo, duas operações definidas sobre os mesmos.

Definição 3.27 (Produto Cartesiano de conjuntos). Dados m conjuntos C_1, \dots, C_n , o seu *produto Cartesiano* é o conjunto

$$C^n \equiv C_1 \times \dots \times C_n \doteq \{(c_1, \dots, c_n) \mid c_k \in C_k, \forall 1 \leq k \leq n\}.$$

O objeto (c_1, \dots, c_n) é denominado uma *n -upla ordenada*. Se $n = 2$, a 2-upla é também denominada de *par ordenado*.

Definição 3.28 (Comutatividade, associatividade e distributividade de operações binárias). Dado o conjunto C , denota-se por C^2 ou $C \times C$ o produto Cartesiano de todas as duplas ordenadas formadas em C e por $\chi : C^2 \rightarrow C$ a função que executa a operação binária entre os elementos de um par ordenado e com resultado aplicado sobre C . Dados $a, b, c \in C$, a operação binária entre quaisquer pares destes é denotada por $a\chi b$. Definem-se então as seguintes *leis de operações binárias*:

Lei da associatividade: Uma operação binária $\chi : C^2 \rightarrow C$ é dita *associativa* se

$$a\chi(b\chi c) = (a\chi b)\chi c.$$

Lei da comutatividade: Uma operação binária $\chi : C^2 \rightarrow C$ é dita *comutativa* ou *Abeliana* se

$$a\chi b = b\chi a.$$

Lei da distributividade: Dadas duas operações binárias $\chi_1, \chi_2 : C^2 \rightarrow C$, a operação χ_1 é dita *distributiva em relação a χ_2* se:

$$a\chi_1(b\chi_2 c) = (a\chi_1 b)\chi_2(a\chi_1 c).$$

Exemplo 3.11. A estrutura algébrica mais conhecida é $\langle \mathbb{R}, +, \times \rangle$, a qual usa as operações de adição (“+”) e multiplicação (“×”) usuais. De acordo com a definição acima, dados $a, b, c \in \mathbb{R}$:

- Adição é associativa: $a + (b + c) = (a + b) + c$.

- Multiplicação é associativa: $a \times (b \times c) = (a \times b) \times c$.
- Adição é comutativa: $a + b = b + a$.
- Multiplicação é comutativa: $a \times b = b \times a$.
- Multiplicação é distributiva em relação a adição: $a \times (b + c) = (a \times b) + (a \times c)$.
- Adição *não* é distributiva em relação a multiplicação: $a + (b \times c) \neq (a + b) \times (a + c) = a + (b \times c) + a \times [a + b + c - 1]$.

Dentre as diversas estruturas algébricas já definidas, as de interesse podem ser agrupadas da maneira apresentada a seguir.

3.7.1 ESTRUTURAS COMPOSTAS POR UM CONJUNTO COM OPERAÇÕES

Consistem nas estruturas com um único conjunto $\langle C, \mathcal{F} \rangle$. Dentre estas, destacam-se:

3.7.1.1 ESTRUTURAS DO TIPO GRUPO

Estruturas que contêm uma única operação binária.

Magma ou grupóide. Dado o conjunto C e a operação binária $\bullet : C^2 \rightarrow C$, a estrutura $\langle C, \bullet \rangle$ será um *grupóide* se os seus elementos satisfizerem:

Condição de clausura: $\forall a, b \in C, a \bullet b \in C$.

Semigrupo. Dado o conjunto C e a operação binária \bullet , denominada *produto*, a estrutura $\langle C, \bullet \rangle$ será um *semigrupo* se os seus elementos satisfizerem:

Condição de associatividade: $\forall a, c, b \in C, a \bullet (b \bullet c) = (a \bullet b) \bullet c$.

Nota-se aqui que a condição de clausura está ausente.

Subsemigrupo. Um *subsemigrupo* de $\langle C, \bullet \rangle$ é um subconjunto não vazio $D \subseteq C$ o qual satisfaz a condição de clausura sob o produto do semigrupo, *i. e.*, $\forall a, b \in D, a \bullet b \in D$.

Um subsemigrupo pode ser considerado um grupóide com a condição de associatividade.

Monóide. Dado o conjunto C e a operação binária \bullet , denominada *produto*, a estrutura $\langle C, \bullet \rangle$ será um *monóide* se os seus elementos satisfizerem:

Condição de associatividade: $\forall a, c, b \in C, a \bullet (b \bullet c) = (a \bullet b) \bullet c$.

Elemento identidade: $\exists I \in C$ tal que $\forall a \in C, a \bullet I = I \bullet a = a$.

Um monóide consiste em um semigrupo com elemento identidade.

Submonóide. Um *submonóide* de $\langle C, \bullet \rangle$ é um subconjunto não vazio $D \subseteq C$ o qual satisfaz a condição de clausura sob o produto do monóide, *i. e.*, $\forall a, b \in D, a \bullet b \in D$, e contém a identidade, *i. e.*, $I \in D$.

Grupo. Dado o conjunto C e a operação binária $\bullet : C^2 \rightarrow C$, denominada *produto*, a estrutura $\langle C, \bullet \rangle$ será um *grupo* se satisfizer os axiomas de grupo da definição 3.1.

Percebe-se nas definições acima que, partindo de um magma, cada estrutura posterior pode ser considerada, sob certas condições, uma extensão das estruturas anteriores. O diagrama na figura 3.10 representa esta evolução condicional. Cada estrutura ao final de uma seta “herda” as propriedades da estrutura anterior. A inversão no sentido das setas depende da adição de condições adicionais sobre a estrutura; por exemplo, todo grupo é um monóide, mas nem todo monóide é um grupo. É necessário também mencionar que nesta seção somente foram definidas as estruturas do lado direito do diagrama. As estruturas do lado esquerdo estão adicionalmente condicionadas à satisfação do axioma da divisão 3.2.

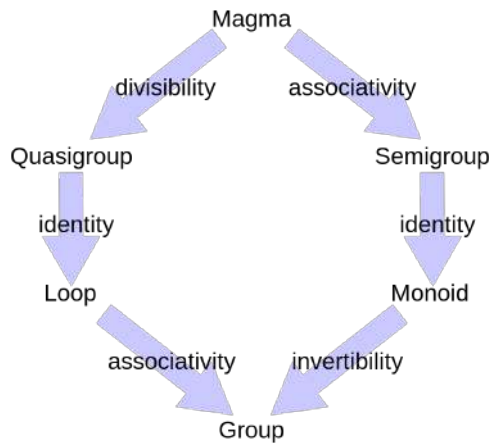


Figura 3.10: Diagrama representando a evolução na complexidade das estruturas algébricas do tipo grupo. A complexidade aumenta nos sentidos das setas. Estruturas no lado esquerdo do diagrama são equivalentes às estruturas do lado direito, porém com a inclusão do axioma da divisão.

3.7.1.2 ESTRUTURAS DO TIPO ANEL

Estruturas contendo duas operações binárias, denominadas *adição* e *produto*, sendo que o produto é distributivo em relação a adição.

Semianel. Dado o conjunto R e as operações binárias \oplus , denominada *adição* e \bullet , denominada *produto*, a estrutura $\langle R, \oplus, \bullet, 0, I \rangle$ será um *semianel* se, dados $a, b, c \in R$ estes satisfizerem:

1. **Frente a adição.** O semianel é um monóide comutativo, *i. e.:*
 - (a) **Associatividade:** $a \oplus (b \oplus c) = (a \oplus b) \oplus c$.
 - (b) **Comutatividade:** $a \oplus b = b \oplus a$.
 - (c) **Elemento identidade:** $\exists 0 \in R$ tal que $a \oplus 0 = 0 \oplus a = a$.
2. **Frente ao produto.** O semianel é um monóide, *i. e.:*
 - (a) **Associatividade:** $a \bullet (b \bullet c) = (a \bullet b) \bullet c$.
 - (b) **Elemento identidade:** $\exists I \in R$ tal que $a \bullet I = I \bullet a = a$.
3. **Distributividade.** O produto é distributivo frente a adição tanto *pela esquerda* quanto *pela direita*:
 - (a) **Distributividade pela esquerda:** $a \bullet (b \oplus c) = (a \bullet b) \oplus (a \bullet c)$.
 - (b) **Distributividade pela direita:** $(a \oplus b) \bullet c = (a \bullet c) \oplus (b \bullet c)$.

Uma operação que respeita distributividade por ambos os lados é também denominada *biaditiva*.
4. **Produto por 0:** $a \bullet 0 = 0 \bullet a = 0$.

Semianel comutativo. Consiste em um semianel o qual é comutativo frente ao produto.

Quase anel. Dado o conjunto R e as operações binárias \oplus , denominada *adição* e \bullet , denominada *produto*, a estrutura $\langle R, \oplus, \bullet, 0 \rangle$ será um *quase anel* se, dados $a, b, c \in R$ estes satisfizerem:

1. **Frente a adição.** O quase anel forma um grupo, sendo $0 \in R$ o elemento identidade.
2. **Frente a multiplicação.** O quase anel forma um semigrupo.
3. **Distributividade.** O produto é distributivo frente a adição *pela direita*: $(a \oplus b) \bullet c = (a \bullet c) \oplus (b \bullet c)$.

Anel não associativo. Dado o conjunto R e as operações binárias \oplus , denominada *adição* e \bullet , denominada *produto*, a estrutura $\langle R, \oplus, \bullet, 0 \rangle$ será um *anel não associativo* se, dados $a, b, c \in R$ estes satisfizerem:

1. **Frente a adição.** O anel forma um grupo Abelian, sendo $0 \in R$ o elemento identidade e, para todo $a \in R$, o elemento $-a \in R$ é a sua inversa aditiva: $a \oplus (-a) = -a \oplus a = 0$.
2. **Distributividade.** O produto é biaditivo frente a adição.

Anel de Lie. Um *anel de Lie* é um anel não associativo composto por um conjunto L com a operação de *adição*, denotada por “+”, e a operação de *colchete* (ou *colchete de Lie*), denotada por “[,]”. A estrutura $\langle L, +, [,], 0 \rangle$ será um anel de Lie se, dados $x, y, z \in L$ estes satisfizerem as seguintes condições:

1. **Frente ao colchete de Lie.** Os elementos do anel devem satisfazer:
 - (a) **Anticomutatividade:** $[x, y] = -[y, x]$. Por consequência, $[x, x] = 0$.
 - (b) **Identidade de Jacobi:** $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.
2. **Distributividade.** O colchete de Lie é biaditivo em relação a adição:

$$[x + y, z] = [x, z] + [y, z] \text{ e } [x, y + z] = [x, y] + [x, z].$$

É importante mencionar aqui que anéis de Lie não são *grupos de Lie* frente a adição.

Anel. Dado o conjunto R e as operações binárias \oplus , denominada *adição* e \bullet , denominada *produto*, a estrutura $\langle R, \oplus, \bullet, 0, I \rangle$ será um *anel* se, dados $a, b, c \in R$ estes satisfizerem os *axiomas de anel*:

1. **Frente a adição.** O anel forma um grupo Abelianos, sendo $0 \in R$ o elemento identidade e, para todo $a \in R$, o elemento $-a \in R$ é a sua inversa aditiva: $a \oplus (-a) = -a \oplus a = 0$.
2. **Frente ao produto.** O anel forma um monóide, sendo $I \in R$ o elemento identidade.
3. **Distributividade.** O produto é distributivo frente a adição tanto pela esquerda quanto pela direita.

Anel comutativo. Consiste em um anel o qual é comutativo frente ao produto.

Corpo. Dado o conjunto K e as operações binárias \oplus , denominada *adição* e \bullet , denominada *produto*, a estrutura $\mathbb{K} \doteq \langle K, \oplus, \bullet, 0, I \rangle$ será um *corpo*²² se, dados $a, b, c \in \mathbb{K}$ estes satisfizerem:

1. **Condições de clausura:** $a \oplus b \in \mathbb{K}$ e $a \bullet b \in \mathbb{K}$.
2. **Condições de associatividade:** $a \oplus (b \oplus c) = (a \oplus b) \oplus c$ e $a \bullet (b \bullet c) = (a \bullet b) \bullet c$.
3. **Condições de comutatividade:** $a \oplus b = b \oplus a$ e $a \bullet b = b \bullet a$.
4. **Elementos identidades:** $\exists 0 \in \mathbb{K}$ tal que $a \oplus 0 = 0 \oplus a = a$ e $\exists I \in \mathbb{K}$ tal que $a \bullet I = I \bullet a = a$.
5. **Elementos inversos:** para todo $a \in \mathbb{K}$ existe $-a \in \mathbb{K}$ tal que $a \oplus (-a) = -a \oplus a = 0$. Para todo $a \in \mathbb{K} \setminus \{0\}$, existe $a^{-1} \in \mathbb{K} \setminus \{0\}$ tal que $a \bullet a^{-1} = a^{-1} \bullet a = I$.
6. **Distributividade:** O produto é distributivo em relação a adição: $a \bullet (b \oplus c) = (a \bullet b) \oplus (a \bullet c)$.

Um corpo, portanto, consiste em dois grupos Abelianos frente a cada operação algébrica, com o embargo de que o elemento I possui inversa frente a adição, mas o elemento 0 não possui inversa frente ao produto.

Dentre as estruturas recém apresentadas, é interessante destacar aquelas que tipificam as definições básicas da teoria de conjuntos; quais sejam:

- O semianel comutativo $\langle \mathbb{N}, +, \times, 0, 1 \rangle$.
- O anel comutativo $\langle \mathbb{Z}, +, \times, 0, 1 \rangle$.
- Os corpos: $\mathbb{Q} \doteq \langle \mathbb{Q}, +, \times, 0, 1 \rangle$ dos números racionais e $\mathbb{I} \doteq \langle \mathbb{I}, +, \times, 0, 1 \rangle$ dos números irracionais.
- **O corpo real.** O corpo²³ $\mathbb{R} \doteq \langle \mathbb{R}, +, \times, 0, 1 \rangle$, formado pelos grupos Abelianos $\{\mathbb{R}; +\}$ e $\{\mathbb{R}^*; \times\}$, onde “+” e “ \times ” são as operações de adição e multiplicação algébricas, respectivamente.
- **O corpo complexo.** O corpo $\mathbb{C} \doteq \langle \mathbb{C}, +, \times, 0, 1 \rangle$, formado pelos grupos Abelianos $\{\mathbb{C}; +\}$ e $\{\mathbb{C}^*; \times\}$, onde “+” e “ \times ” são as operações de adição e multiplicação algébricas de números complexos,²⁴ respectivamente. Cabe mencionar aqui que o corpo \mathbb{R} está contido no corpo \mathbb{C} , em cuja situação é denominado um **subcorpo**.

²²O termo *corpo* empregado em português segue os termos em alemão *Körper* e em francês *corps*. Textos em inglês normalmente usam o termo *field*.

²³Note a diferença tipográfica entre o corpo \mathbb{R} e o conjunto \mathbb{R} .

²⁴Ver seção 2.2.

3.7.2 ESTRUTURAS DO TIPO MÓDULO

Consistem em sistemas de duas estruturas $\langle C_1, \mathcal{F}_1 \rangle$ e $\langle C_2, \mathcal{F}_2 \rangle$, muitas das quais definidas na seção anterior. Tratam-se de sistemas compostos por duas estruturas, cada uma envolvendo, no mínimo, duas operações binárias.

Dentre as estruturas compostas por dois conjuntos e por duas operações binárias, destacam-se as seguintes.

Grupo com operadores. A estrutura $\langle \mathcal{G}, \Omega \rangle$ é denominada um *grupo com operadores* ou Ω -grupo. Esta estrutura é composta por um grupo $\mathcal{G} = (G; \bullet)$ e um conjunto Ω cujos elementos operam sobre os elementos do grupo de uma determinada maneira. O Ω -grupo é definido com a ação de Ω sobre \mathcal{G} : $\Omega \times \mathcal{G} \rightarrow \mathcal{G} : (\omega, g) \mapsto g^\omega$, sendo $\omega \in \Omega$ e $g \in \mathcal{G}$, de tal forma que a operação é distributiva em relação ao produto de grupo:

$$(g \bullet h)^\omega = g^\omega \bullet h^\omega, \text{ onde } h \in \mathcal{G}.$$

Para todo $\omega \in \Omega$, o mapeamento $g \mapsto g^\omega$ consiste em um endomorfismo de \mathcal{G} . Portanto, o Ω -grupo pode ser visualizado como um grupo \mathcal{G} com uma família indexada $(u_\omega)_{\omega \in \Omega}$ de endomorfismos de \mathcal{G} .

Uma ação de \mathcal{G} sobre \mathcal{G} é um caso particular de um Ω -grupo quando $\Omega = \mathcal{G}$. Por exemplo, a ação esquerda de grupo é o endomorfismo $\mathcal{G} \times \mathcal{G} \rightarrow \mathcal{G}$ tal que para todos $g_1, g_2 \in \mathcal{G}$, $(g_1, g_2) \mapsto g_1 \bullet g_2$. Escrevendo simbolicamente esta ação como $g_1 \mathcal{G}$, observa-se que esta se trata simplesmente de uma classe lateral esquerda do grupo.

Espaço vetorial. Devido a sua importância, espaços vetoriais são discutidos em mais detalhes no capítulo 4.

Módulo. Um *módulo sobre um anel* consiste em uma generalização da noção de um espaço vetorial sobre um corpo, apresentado acima, sendo que os escalares são elementos de um anel, ao invés de um corpo. Além disso, os produtos por escalares ocorrem em ambos os lados.

Dado um grupo Abeliano $\mathcal{M} = \langle M, +, 0 \rangle$ frente a operação de *adição* “+” e com elemento identidade $0 \in \mathcal{M}$. Dado também o anel $A \doteq \langle A, \oplus, \bullet, 0, I \rangle$ e, finalmente, dada a operação $\cdot : A \times \mathcal{M} \rightarrow \mathcal{M}$, denominada *produto*, a estrutura $\langle \mathcal{M}, A, \cdot \rangle$ pode ser definida de duas maneiras:

A-Módulo à esquerda. O *A-módulo à esquerda* $\mathcal{M} \doteq \langle \mathcal{M}, A, \cdot \rangle$ é tal que o produto satisfaz as condições:

1. **Clausura no grupo.** Para todo $a \in A$ e para todo $m \in \mathcal{M}$, $a.m \in \mathcal{M}$.
2. **Associatividade.** Para cada $a, b \in A$ e $m \in \mathcal{M}$, $a.(b.m) = (a \bullet b).m$.
3. **Produto pela unidade de A.** Para todo $m \in \mathcal{M}$, $I.m = m$.
4. **Distributividade.** Para $a, b \in A$ e $m, n \in \mathcal{M}$,

$$\begin{aligned} a.(m + n) &= a.m + a.n \\ (a \oplus b).m &= a.m + b.m. \end{aligned}$$

A-Módulo à direita. O *A-módulo à direita* $\mathcal{M}_R \doteq \langle \mathcal{M}, A, \cdot_R \rangle$ possui produto $(\cdot_R \equiv \cdot)$ que satisfaz:

1. **Clausura no grupo.** Para todo $a \in A$ e para todo $m \in \mathcal{M}$, $m.a \in \mathcal{M}$.
2. **Associatividade.** Para cada $a, b \in A$ e $m \in \mathcal{M}$, $(m.a).b = m.(a \bullet b)$.
3. **Produto pela unidade de A.** Para todo $m \in \mathcal{M}$, $m.I = m$.
4. **Distributividade.** Para $a, b \in A$ e $m, n \in \mathcal{M}$,

$$\begin{aligned} (m + n).a &= m.a + n.a \\ m.(a \oplus b) &= m.a + m.b. \end{aligned}$$

3.7.3 ESTRUTURAS DO TIPO ÁLGEBRA

Sistemas compostos definidos sobre duas estruturas: um espaço vetorial e um corpo ou um anel e um A -módulo, equipados com uma operação adicional, o que totaliza no mínimo 5 operações binárias: duas operações no espaço vetorial/anel, duas no corpo/ A -módulo e uma envolvendo elementos de ambas as estruturas.

Álgebra sobre um corpo. Trata-se de um espaço vetorial equipado com um mapa bilinear.

Seja $\mathcal{A} = \langle \mathcal{A}, \mathbb{K}, +, \cdot \rangle$ um espaço vetorial sobre o corpo $\mathbb{K} \doteq \langle \mathbb{K}, \oplus, \bullet, 0, I \rangle$. Define-se a operação binária $\times : \mathcal{A} \times \mathcal{A} \mapsto \mathcal{A}$, denominada *produto de vetores*. Então \mathcal{A} será denominada uma *álgebra sobre \mathbb{K}* se, para todo $x, y, z \in \mathcal{A}$ e para todo $a, b \in \mathbb{K}$, as seguintes condições forem satisfeitas:

- **Distributividade pela esquerda:** $(x + y) \times z = x \times z + y \times z$.
- **Distributividade pela direita:** $x \times (y + z) = x \times y + x \times z$.
- **Compatibilidade com escalares:** $(a.x) \times (b.y) = (a \bullet b) \cdot (x \times y)$.

Álgebra sobre um anel comutativo. Esta álgebra generaliza o conceito de uma álgebra sobre um corpo, ao substituir este último por um anel comutativo.

Seja $A \doteq \langle A, \oplus, \bullet, 0, I \rangle$ um anel comutativo e \mathcal{M} o A -módulo definido sobre A . Uma *álgebra sobre um anel comutativo*, ou *A -álgebra*, é obtida a partir da definição da operação binária $[\cdot, \cdot] : \mathcal{M} \times \mathcal{M} \rightarrow \mathcal{M}$, denominada *A -produto*, o qual satisfaz a condição:

- **Bilinearidade:** para todo $x, y, z \in \mathcal{M}$ e $a, b \in A$,

$$\begin{aligned} [a.x + b.y, z] &= a.[x, y] + b.[y, z] \\ [x, a.y + b.z] &= a.[x, y] + b.[x, z]. \end{aligned}$$

Álgebra associativa. Dada uma A -álgebra, esta é dita *associativa* se o A -produto for associativo, i. e., para todo $x, y, z \in \mathcal{M}$,

$$[x, [y, z]] = [[x, y], z].$$

Álgebra não associativa. Quando o A -produto não for associativo. Neste caso, o resultado de $[x, [y, z]]$ pode ser definido de diferentes maneiras, o que gera álgebras especiais, como a álgebra de Lie.

Cabe mencionar que a classificação de álgebras entre associativas ou não associativas também é aplicável para álgebras sobre corpos.

Álgebra de Lie. Estas álgebras são não associativas, sendo que o produto da álgebra satisfaz a identidade de Jacobi. A sua importância reside no contexto de teoria de grupos contínuos (grupos de Lie) porque estão associadas às operações de isometria de um sistema físico. As álgebras de Lie surgem a partir dos geradores de grupos de Lie, geradores estes que são transformações infinitesimais. Esta álgebra será definida a partir de um espaço vetorial sobre um corpo, mas pode ser também generalizada no contexto de A -álgebras.

Uma *álgebra de Lie* é um espaço vetorial $\mathfrak{g} \doteq \langle \mathfrak{g}, \mathbb{K}, +, \cdot \rangle$ sobre o corpo $\mathbb{K} \doteq \langle \mathbb{K}, \oplus, \bullet, 0, I \rangle$ com a operação $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, denominada *colchete de Lie*, a qual, para todos $x, y, z \in \mathfrak{g}$ e $a, b \in \mathbb{K}$, satisfaz os axiomas:

- **Bilinearidade:** $[a.x + b.y, z] = a.[x, z] + b.[y, z]$, $[z, a.x + b.y] = a.[z, x] + b.[z, y]$.
- **Anticomutatividade:** $[x, y] = -[y, x]$. Por consequência, $[x, x] = 0$.
- **Identidade de Jacobi:** $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

A álgebra de Lie será discutida novamente na seção 5.10.5.

REFERÊNCIAS

- ALABISO, C.; WEISS, I. **A Primer on Hilbert Space Theory: Linear Spaces, Topological Spaces, Metric Spaces, Normed Spaces, and Topological Groups**. New York: Springer, 2014. (UNITEXT for Physics). 255 + xvii pp. ISBN 9783319037134. DOI: [10.1007/978-3-319-03713-4](https://doi.org/10.1007/978-3-319-03713-4).
- BARATA, João Carlos Alves. **Notas para um Curso de Física-Matemática**. São Paulo: USP. Versão de 28/02/2018. Disponível em: http://denebola.if.usp.br/~jbarata/Notas_de_aula/.
- BASSALO, José Maria F.; CATTANI, Mauro Sérgio D. **Teoria de Grupos**. Segunda edição. São Paulo: Livraria da Física, 2008. 286 + ix pp. ISBN 9788578610005.
- DRESSELHAUS, S.; DRESSELHAUS, G.; JORIO, Ado. **Group Theory: Application to the Physics of Condensed Matter**. Berlin: Springer, 2008. 582 + xv pp. ISBN 9783540328971. DOI: [10.1007/978-3-540-32899-5](https://doi.org/10.1007/978-3-540-32899-5).
- FAZZIO, A.; WATARI, K. **Introdução à Teoria de Grupos: com Aplicações em Moléculas e Sólidos**. Santa Maria: UFSM, 1998. 239 pp. ISBN 85-7391-010-0.
- HAMERMESH, Morton. **Group Theory and Its Application to Physical Problems**. New York: Dover Publications, 2012. (Dover Books on Physics). ISBN 9780486140391. Disponível em: <http://books.google.com.br/books?id=w6PDagAAQBAJ>.
- JOSHI, A. W. **Elements of Group Theory for Physicists**. Third Ed. New Delhi: John Wiley & Sons, 1982. 348 + xv pp. ISBN 0-470-27306-2.
- MESSIAH, Albert. **Quantum Mechanics. Volume I**. Amsterdam: North-Holland, 1961. 1136 + xxii pp. ISBN 0-486-40924-4.
- RILEY, K. F.; HOBSON, M. P.; BENCE, S. J. **Mathematical Methods for Physics and Engineering**. Third Ed. Cambridge: Cambridge University Press, 2006. 1333 + xxvii pp. ISBN 9781139450997. Disponível em: <http://books.google.com.br/books?id=Mq1nLEKhNcsC>.
- SZEKERES, Peter. **A Course in Modern Mathematical Physics: Groups, Hilbert Space, and Differential Geometry**. Cambridge: Cambridge University Press, 2004. ISBN 9780521536455. Disponível em: <http://books.google.com.br/books?id=pRgjkgEACAAJ>.
- TINKHAM, M. **Group Theory and Quantum Mechanics**. [S.l.]: Dover Publications, 2012. (Dover Books on Chemistry). 340 + xii pp. ISBN 9780486131665. Disponível em: <https://books.google.com.br/books?id=oe7DAgAAQBAJ>.

ESPAÇOS VETORIAIS



CONCEITO DE ESPAÇO VETORIAL possui importância singular para a física-matemática, pois todas as áreas da matemática aplicada, da física contemporânea e engenharia fazem uso de suas propriedades. Suas propriedades fornecem os fundamentos para o tratamento de sistemas de equações lineares, séries de Taylor e Fourier, na solução de equações diferenciais parciais, entre outras técnicas comumente empregadas em matemática aplicada. Espaços vetoriais fornecem também um ferramental abstrato e livre de coordenadas para o estudo das propriedades de objetos geométricos e físicos tais como tensores, a teoria quântica, a estrutura cristalina da matéria e a própria estrutura do espaço-tempo. Isso tudo é possível devido à possibilidade de se estender o conceito de um espaço vetorial, a partir de uma estrutura puramente algébrica e abstrata, para uma estrutura mais elevada, a qual contém elementos de análise matemática, geometria e topologia.

Historicamente, as primeiras concepções que levaram ao conceito de um espaço vetorial podem ser traçadas para o início do século XVII, quando foi desenvolvida a teoria da geometria afim com a introdução das coordenadas em espaços de duas ou três dimensões. Por volta do ano 1636, os matemáticos franceses René Descartes (1596 – 1650) e Pierre de Fermat (1607 – 1665) fundaram a disciplina da geometria analítica através do mapeamento das soluções de equação de duas variáveis a pontos sobre uma superfície. O desenvolvimento histórico que culminou com o conceito moderno de um espaço vetorial como uma estrutura algébrica (ver seção 3.7) ocorreu ao longo dos séculos XVIII – XX, contando com as contribuições de diversos matemáticos importantes, tais como Jean-Robert Argand (1768 – 1822), Bernard Bolzano (1781 – 1848), August Ferdinand Möbius (1790 – 1868), C. V. Mourey (1791? – 1830?), Giusto Bellavitis (1803 – 1880), William Rowan Hamilton (1805 – 1865), Hermann Günther Grassmann (1809 – 1877), Arthur Cayley (1821 – 1895), Edmond Laguerre (1834 – 1886), Giuseppe Peano (1858 – 1932), David Hilbert (1862 – 1943), Henri Léon Lebesgue (1875 – 1941) e Stefan Banach (1892 – 1945). Atualmente, a teoria das estruturas algébricas continua sendo desenvolvida como parte da *teoria das categorias*, criada por Samuel Eilenberg (1913 – 1998) e Saunders Mac Lane (1909 – 2005).

Neste capítulo serão apresentadas algumas das propriedades básicas de espaços vetoriais. Os principais objetivos deste capítulo são:

- Caracterizar um espaço vetorial como uma estrutura algébrica abstrata, hierarquicamente superior aos grupos discutidos no capítulo 3. A categorização de um espaço vetorial como uma estrutura algébrica foi realizada na seção 3.7.
- Prover conceitos e propriedades fundamentais que serão empregados posteriormente nos capítulos que tratam da teoria de representações de grupos (capítulo 5) e de tensores (capítulo 6).

Para atingir estes objetivos, na seção 4.12 o capítulo também abandona temporariamente a discussão de uma estrutura puramente algébrica para introduzir elementos de espaços métricos e topologia, os quais possibilitarão a definição de espaços vetoriais métricos ou topológicos, fundamentais para as teorias modernas da física.

4.1 ESPAÇO VETORIAL

O conceito de espaço vetorial é normalmente definido na disciplina de álgebra linear; porém, nesta seção o mesmo conceito será apresentado como um tipo de estrutura algébrica.

Definição 4.1 (Espaço vetorial). Um *espaço vetorial* sobre um corpo \mathbb{K} , também denominado um *espaço- \mathbb{K}* , e denotado por $\mathcal{V} \doteq \langle \mathcal{V}, \mathbb{K}, +, \cdot \rangle$, consiste em um conjunto V de elementos denominados *vetores* e dotado de uma operação $+$: $\mathcal{V}^2 \rightarrow \mathcal{V}$, denominada *soma vetorial*. O espaço é composto também por um corpo $\mathbb{K} \doteq \langle \mathbb{K}, \oplus, \bullet, 0, I \rangle$, cujos elementos são denominados *escalares*, e, finalmente, de uma operação \cdot : $\mathbb{K} \times \mathcal{V} \rightarrow \mathcal{V}$, denominada *produto por escalar*. Os conjuntos de vetores e escalares satisfazem as seguintes propriedades frente as operações definidas:

Frente a soma vetorial. A estrutura $\mathcal{V} = \langle V, +, 0 \rangle$ forma um grupo Abelianiano, sendo que o elemento identidade 0 é denominado *vetor nulo*.

Frente ao produto por escalar. Dado o corpo \mathbb{K} , para cada $\alpha \in \mathbb{K}$ e $u \in \mathcal{V}$, existe um vetor denotado por $\alpha.u \in \mathcal{V}$, denominado *produto de u por α* , o qual satisfaz as condições:

1. **Associatividade.** Para cada $\alpha, \beta \in \mathbb{K}$, $\alpha.(\beta.u) = (\alpha \bullet \beta).u$.
2. **Produto pela unidade de \mathbb{K} .** Para todo $u \in \mathcal{V}$, $I.u = u$.
3. **Distributividade.** Para cada $\alpha, \beta \in \mathbb{K}$ e $u, v \in \mathcal{V}$, o produto por escalares segue diferentes regras frente as operações “+” e “ \oplus ”:
 - (a) Distributivo pela esquerda em relação a soma vetorial: $\alpha.(u + v) = \alpha.u + \alpha.v$
 - (b) Distributividade mista pela direita em relação a “ \oplus ” e “+”: $(\alpha \oplus \beta).u = \alpha.u + \beta.u$.

Em física, espaços vetoriais são usualmente definidos sobre corpos reais ou complexos, quando então são também denominados *espaços vetoriais reais* ou *espaços vetoriais complexos*, respectivamente. Algumas propriedades dos espaços vetoriais são discutidas na seção a seguir.

4.2 SUBESPAÇOS VETORIAIS E SUBESPAÇOS COMPLEMENTARES

Um subespaço vetorial é um conjunto de vetores contidos no espaço vetorial, os quais satisfazem as condições de clausura frente a adição vetorial e produto por escalar.

Definição 4.2 (Subespaço vetorial). Seja \mathcal{V} um espaço vetorial. Dado um subconjunto $\mathcal{W} \subseteq \mathcal{V}$, este é denominado um *subespaço vetorial de \mathcal{V}* se satisfizer os axiomas de espaço vetorial. Adicionalmente, para todos os vetores $u, v \in \mathcal{W}$ e todos os escalares $\alpha, \beta \in \mathbb{K}$, a adição vetorial e o produto por escalar estão contidos no subespaço: $\alpha.u + \beta.v \in \mathcal{W}$.

Em particular, o subespaço $\mathcal{W} \subseteq \mathcal{V}$ sempre irá conter o vetor nulo. Dados dois subespaços $\mathcal{U}, \mathcal{W} \subseteq \mathcal{V}$, a sua intersecção $\mathcal{U} \cap \mathcal{W}$ também é subespaço de \mathcal{V} , pois se $u, v \in \mathcal{U} \cap \mathcal{W}$, então necessariamente $\alpha.u + \beta.v \in \mathcal{U} \cap \mathcal{W}$ também.

4.2.1 SUBESPAÇO SOMA

Sejam os subespaços $\mathcal{U}, \mathcal{W} \subseteq \mathcal{V}$. Embora a sua intersecção $\mathcal{U} \cap \mathcal{W}$ também seja um subespaço de \mathcal{V} , o mesmo não pode ser dito de sua união; $\mathcal{U} \cup \mathcal{W}$ em geral não é subespaço de \mathcal{V} . Contudo, é sempre possível definir-se o *subespaço soma*:

Definição 4.3 (Subespaço soma). Dados os subespaços $\mathcal{U}, \mathcal{W} \subseteq \mathcal{V}$, o conjunto

$$\mathcal{U} + \mathcal{W} \doteq \{u + w, \text{ tais que } u \in \mathcal{U} \text{ e } w \in \mathcal{W}\}$$

é *subespaço soma* de \mathcal{V} , pois se $u = u_1 + w_1 \in \mathcal{U} + \mathcal{W}$ e $w = u_2 + w_2 \in \mathcal{U} + \mathcal{W}$, então

$$\alpha.u + \beta.w = \alpha.(u_1 + w_1) + \beta.(u_2 + w_2) = (\alpha.u_1 + \beta.u_2) + (\alpha.w_1 + \beta.w_2) \in \mathcal{U} + \mathcal{W}.$$

4.2.2 SUBESPAÇOS COMPLEMENTARES

A definição do subespaço soma $\mathcal{U} + \mathcal{W}$ serve para determinar se \mathcal{U} e \mathcal{W} são também *subespaços complementares*.

Definição 4.4 (Subespaços complementares e soma direta de subespaços). Sejam os subespaços $\mathcal{U}, \mathcal{W} \subseteq \mathcal{V}$. Estes são ditos *subespaços complementares* se todo vetor $v \in \mathcal{V}$ possuir uma *decomposição única* $v = u + w$, sendo que $u \in \mathcal{U}$ e $w \in \mathcal{W}$.

Caso $\mathcal{U}, \mathcal{W} \subseteq \mathcal{V}$ sejam subespaços complementares, então \mathcal{V} é dito ser a *soma direta* destes subespaços. Esta operação é representada por

$$\mathcal{V} = \mathcal{U} \oplus \mathcal{W}.$$

O teorema a seguir estabelece as condições necessárias e suficientes para a determinação de subespaços complementares.

Teorema 4.1. Sejam os subespaços $\mathcal{U}, \mathcal{W} \subseteq \mathcal{V}$, sendo $\mathbf{0}$ o vetor nulo. Estes são subespaços complementares de \mathcal{V} se e somente se:

1. $\mathcal{V} = \mathcal{U} + \mathcal{W}$;
2. $\mathcal{U} \cap \mathcal{W} = \{\mathbf{0}\}$.

Um outro tipo importante de subespaço vetorial, um *subespaço invariante*, será apresentado na seção 4.6, após a definição de base de um espaço vetorial.

4.3 BASES DE UM ESPAÇO VETORIAL

Uma *base* de um espaço vetorial é um conjunto de vetores que reproduz todo o espaço a partir de combinações lineares entre os vetores do conjunto. Serão apresentadas uma série de definições para a construção do conceito de base e obtenção de suas propriedades.

Definição 4.5 (Varredura linear). Seja $C \doteq \{v_i\} \subset \mathcal{V}$ um subconjunto do espaço vetorial \mathcal{V} sobre o corpo \mathbb{K} . A *varredura linear*¹ de C , denotada por $\text{span}(C)$, é definida como o conjunto de todas as combinações lineares possíveis em C ; ou seja,

$$\text{span}(C) \doteq \left\{ \sum_i \alpha_i v_i, \forall \alpha_i \in \mathbb{K} \right\}.$$

Teorema 4.2 (Subespaço gerado por um conjunto). Seja $C \doteq \{v_i\} \subset \mathcal{V}$ um subconjunto do espaço vetorial \mathcal{V} sobre o corpo \mathbb{K} . Então,

1. $\text{span}(C)$ é um subespaço de \mathcal{V} .
2. Se $\mathcal{U} \subset \mathcal{V}$ é um subespaço tal que $C \subset \mathcal{U}$, então $\text{span}(C) \subset \mathcal{U}$.

Definição 4.6 (Espaço vetorial finito ou infinito e dimensão do espaço). Um espaço vetorial \mathcal{V} é dito ser de *dimensão finita* se este pode ser gerado a partir de um conjunto finito de vetores $C_n = \{v_1, \dots, v_n\} \subset \mathcal{V}$, sendo $n \in \mathbb{N}$:

$$\mathcal{V} = \text{span}(C_n).$$

Caso não exista um conjunto finito C_n capaz de gerar \mathcal{V} , então este é dito ser de *dimensão infinita*.

Quando \mathcal{V} tem dimensão finita, sua *dimensão* $\dim \mathcal{V}$ é definida como o menor número natural n tal que $\mathcal{V} = \text{span}(C_n)$.

Será definida agora uma noção fundamental para a teoria de espaços vetoriais: a independência (ou dependência) linear entre vetores.

Definição 4.7 (Vetores linearmente independentes). Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} . O conjunto de vetores $C \doteq \{v_i\}$, onde $i = 1, 2, \dots$ e $v_i \in \mathcal{V}$, é dito ser *linearmente independente* (identificado também por “LI”) se todos os subconjuntos finitos $\{v_j\} \subseteq C$ ($j = 1, 2, \dots, k$) forem tais que

$$\sum_{j=1}^k \alpha_j v_j = \mathbf{0} \iff \alpha_j = 0, \text{ para todo } j = 1, \dots, k,$$

sendo que $\{\alpha_j\} \subset \mathbb{K}$.

Um conjunto de vetores é dito ser *linearmente dependente* se este não for LI.

¹Do inglês: *linear span*.

Com as definições acima, torna-se agora possível definir-se a base de um espaço vetorial.

Definição 4.8 (Base de um espaço vetorial). Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} . Um subconjunto $B \subset \mathcal{V}$ é denominado uma *base de \mathcal{V}* se:

1. B for linearmente independente;
2. $\mathcal{V} = \text{span}(B)$; ou seja, se \mathcal{V} for gerado por B .

Teorema 4.3. *Todo espaço vetorial possui uma base.*

Observa-se que o teorema acima é válido tanto para espaços finitos quanto para infinitos. Os teoremas a seguir são válidos para espaços finitos.

Teorema 4.4. *Seja \mathcal{V} um espaço vetorial finito de dimensão n . Um subconjunto $B = \{e_1, \dots, e_n\}$ gera \mathcal{V} se e somente se B for linearmente independente.*

Corolário 4.1. *Seja \mathcal{V} um espaço vetorial finito. Se o subconjunto $B = \{e_1, \dots, e_n\}$ é uma base de \mathcal{V} , então $\dim \mathcal{V} = n$. Ou seja, a dimensão do espaço vetorial é idêntica à cardinalidade da base.*

Teorema 4.5 (Redução de base). *Seja \mathcal{V} um espaço vetorial finito de dimensão n . Se existe um subconjunto $C = \{v_1, \dots, v_m\} \subset \mathcal{V}$ com $m \geq n$ tal que $\mathcal{V} = \text{span}(C)$, então ou C é uma base de \mathcal{V} ou um número $m - n$ de vetores de C pode ser removido para formar uma base de \mathcal{V} .*

(Enunciado alternativo) *Seja \mathcal{V} um espaço vetorial finito de dimensão n . Seja $B = \{e_1, \dots, e_n\}$ uma base de \mathcal{V} . Qualquer conjunto de vetores $C = \{v_1, \dots, v_m\} \subset \mathcal{V}$ com $m > n$ é linearmente dependente.*

Teorema 4.6. *Seja \mathcal{V} um espaço vetorial finito. Sejam $B = \{e_1, \dots, e_n\} \subset \mathcal{V}$ e $B' = \{e'_1, \dots, e'_m\} \subset \mathcal{V}$ bases de \mathcal{V} . Então $m = n$.*

Demonstração. De acordo com o teorema 4.5, as desigualdades $n \leq m$ e $n \geq m$ devem ser ambas satisfeitas. Então, $n = m$. □

O importante teorema a seguir relaciona conjuntos LI com uma base de um espaço de dimensão finita.

Teorema 4.7 (Extensão de base). *Seja \mathcal{V} um espaço vetorial finito de dimensão n . Seja $A = \{v_1, v_2, \dots, v_m\} \subseteq \mathcal{V}$ um subconjunto de vetores linearmente independentes de \mathcal{V} , sendo que $m \leq n = \dim \mathcal{V}$. Então existe uma base $B = \{e_1, e_2, \dots, e_n\}$ de \mathcal{V} tal que*

$$e_1 = v_1, \quad e_2 = v_2, \quad \dots \quad e_m = v_m.$$

Em outras palavras, qualquer conjunto LI de \mathcal{V} pode ser estendido para formar uma base de \mathcal{V} . Apresenta-se agora um outro importante teorema, o qual versa sobre a decomposição de vetores.

Teorema 4.8 (Decomposição de vetores). *Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} . Se o conjunto $B = \{e_1, e_2, \dots\}$ é uma base de \mathcal{V} , então cada vetor $v \in \mathcal{V}$ pode ser decomposto de maneira unívoca pela combinação linear*

$$v = \sum_i v_i e_i, \quad (v_i \in \mathbb{K}).$$

Os escalares $\{v_i\}$ são denominados os componentes do vetor v com respeito à base B .

Os componentes de um vetor $v \in \mathcal{V}$ sobre uma base B com frequência são organizados na forma de uma matriz coluna, de tal forma que operações vetoriais possam ser realizadas por intermédio da álgebra vetorial. Este assunto é discutido na seção 4.9.4.

4.3.1 TRANSFORMAÇÃO DE BASE

Se para um dado espaço vetorial \mathcal{V} existirem pelos menos duas bases distintas (B e B' por exemplo), então é sempre possível decompor-se um vetor da base B por uma combinação linear dos vetores da base B' e *vice-versa*. Este processo é denominado a *transformação de base* $B \Rightarrow B'$.

Seja um espaço vetorial finito \mathcal{V} de dimensão n , o qual possui duas bases $B = \{e_1, \dots, e_n\}$ e $B' = \{e'_1, \dots, e'_n\}$. De acordo com o teorema 4.8, cada vetor da base B é decomposto de forma unívoca pelos vetores da base B' na forma

$$e_i = \sum_{j=1}^n A_{ji} \cdot e'_j, \quad (A_{ij} \in \mathbb{K}, i = 1, \dots, n), \quad (4.1)$$

sendo A_{ji} o j -ésimo componente do vetor e_i com respeito à base B' .

Seja agora um vetor $v \in \mathcal{V}$ qualquer. Novamente, de acordo com o 4.8, este pode ser decomposto pelas duas bases como

$$v = \sum_{i=1}^n v_i \cdot e_i = \sum_{i=1}^n v'_i \cdot e'_i,$$

sendo v_i e v'_i os componentes de v em respeito às bases B e B' , respectivamente. Mas, empregando-se a decomposição dos vetores em B em B' obtida acima, pode-se escrever

$$v = \sum_{j=1}^n \sum_{i=1}^n (v_i \cdot A_{ji}) \cdot e'_j = \sum_{j=1}^n v'_j \cdot e'_j,$$

de onde se identifica a lei de transformação das componentes de v como

$$v'_i = \sum_{j=1}^n A_{ij} \cdot v_j, \quad (i = 1, \dots, n). \quad (4.2)$$

A transformação das coordenadas de um vetor pode ser representada na forma matricial. Esta representação é realizada na seção 4.9.4.1

No capítulo 6 será realizada uma longa discussão a respeito de transformações de bases em espaços Euclidianos ou Riemannianos.

4.4 MAPEAMENTOS ENTRE ESPAÇOS VETORIAIS

Um *mapeamento* entre espaços vetoriais ou sobre um espaço vetorial é uma função que toma um vetor de um espaço e o mapeia sobre outro vetor, em um espaço distinto ou sobre o mesmo espaço, ou ainda sobre um escalar. Alguns tipos de mapeamentos praticados sobre ou entre espaços vetoriais são mencionados agora.

Definição 4.9 (Mapeamento entre espaços vetoriais). Sejam \mathcal{V} e \mathcal{W} dois espaços vetoriais sobre o mesmo corpo \mathbb{K} . Uma função $T : \mathcal{V} \mapsto \mathcal{W}$ é denominada um *mapeamento de \mathcal{V} sobre \mathcal{W}* se para todo $v \in \mathcal{V}$ existe um e somente um $w \in \mathcal{W}$ tal que

$$T(v) = w.$$

O vetor $w = T(u)$ é denominado o *valor de T em v* . O conjunto $V \subseteq \mathcal{V}$ que contém os vetores sobre o qual o mapeamento T atua é denominado o *domínio* de T , representado por $\text{Dom}(T)$. O conjunto $W \subseteq \mathcal{W} = \{w \mid T(v) = w, \forall v \in V\}$ é denominado a *imagem de V sob T* , representada por $\text{Img}(T)$.

Um *mapa linear*, também denominado *aplicação linear*, *mapeamento linear* ou *transformação linear*, é um mapeamento entre dois espaços vetoriais tal que as operações de soma vetorial e produto por escalar são preservadas.

Definição 4.10 (Mapa linear). Sejam \mathcal{V} e \mathcal{W} dois espaços vetoriais sobre o mesmo corpo \mathbb{K} . Uma função $L : \mathcal{V} \mapsto \mathcal{W}$ é denominada um *mapa linear* se para todo $u, v \in \mathcal{V}$ e para todo $\alpha \in \mathbb{K}$ as seguintes propriedades são satisfeitas:

1. **Aditividade:** $L(u + v) = L(u) + L(v)$.
2. **Homogeneidade de grau 1:** $L(\alpha \cdot u) = \alpha \cdot L(u)$.

O conjunto de todos os mapas lineares $L : \mathcal{V} \mapsto \mathcal{W}$ é representado por $\mathbb{L}(\mathcal{V}, \mathcal{W})$.

Alguns casos particulares de um mapa linear são:

Operador linear. Quando domínio e imagem coincidem ($\mathcal{W}=\mathcal{V}$), o mapa linear também é denominado *operador linear*. Um operador linear obviamente obedece as propriedades de aditividade e homogeneidade de um mapa linear. Adicionalmente, se $u = 0 \in \mathcal{V}$ é o vetor nulo de \mathcal{V} , então

$$L(0) = 0.$$

O conjunto de todos os operadores lineares $L : \mathcal{V} \mapsto \mathcal{V}$ é representado por $\mathbb{L}(\mathcal{V})$.

Forma linear. Quando o mapa linear é a função $f : \mathcal{V} \mapsto \mathbb{K}$, i. e., do espaço vetorial sobre o seu corpo escalar, este é denominado uma *forma linear* ou um *funcional linear*. Uma propriedade importante é que o mapeamento do funcional sobre o corpo é sempre ou trivial ou é sobrejetivo.

A partir da definição de um operador linear e do teorema sobre decomposição de vetores, é possível identificar a ação de um determinado operador linear atuando sobre um espaço vetorial com uma matriz que irá conter os seus componentes. Esta representação é realizada na seção 4.9.4.2.

Um *mapa bilinear*, também denominado *operador bilinear*, é uma função que combina elementos de dois espaços vetoriais para resultar em um elemento de um terceiro espaço vetorial que é linear em cada um dos argumentos.

Definição 4.11 (Mapa bilinear). Sejam \mathcal{V} , \mathcal{W} e \mathcal{X} três espaços vetoriais sobre o mesmo corpo \mathbb{K} . Um *mapa bilinear* B é definido como a função $B : \mathcal{V} \times \mathcal{W} \rightarrow \mathcal{X}$ tal que para todo $v \in \mathcal{V}$ e $w \in \mathcal{W}$:

1. O mapeamento $v \mapsto B(v, w)$ é um mapa linear de \mathcal{V} a \mathcal{X} .
2. O mapeamento $w \mapsto B(v, w)$ é um mapa linear de \mathcal{W} a \mathcal{X} .

Em outras palavras, mantendo o primeiro argumento do mapa fixo e permitindo o segundo variar, o resultado é um mapa linear, e vice-versa.

Forma bilinear. Definido no caso em que a imagem do mapeamento (\mathcal{X}) é o corpo base \mathbb{K} . Dado o espaço vetorial \mathcal{V} sobre o corpo \mathbb{K} , o mapa bilinear $B : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{K}$ é denominado uma *forma bilinear*, tal que, dados $u, v, w \in \mathcal{V}$ e $\alpha \in \mathbb{K}$:

1. $B(u + v, w) = B(u, w) + B(v, w)$
2. $B(u, v + w) = B(u, v) + B(u, w)$
3. $B(\alpha.u, v) = B(u, \alpha.v) = \alpha \bullet B(u, v)$.

Desta definição são definidos os *produto interno*, *produto escalar* e *forma quadrática*.

Forma sesquilinear. Trata-se de uma forma bilinear sobre um espaço vetorial complexo, a qual é linear em um dos argumentos, mas antilinear no outro. Existe uma liberdade em definir qual argumento é antilinear. Neste texto, o primeiro argumento terá tal propriedade. Seja \mathcal{V} um espaço linear complexo. Uma *forma sesquilinear sobre \mathcal{V}* é um mapa $\varphi : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{C}$ tal que, para todos $x, y, z \in \mathcal{V}$ e $\alpha, \beta \in \mathbb{C}$, as seguintes propriedades são satisfeitas:

1. $\varphi(x + y, z) = \varphi(x, z) + \varphi(y, z)$.
2. $\varphi(\alpha.x, \beta.y) = \alpha^* \beta \varphi(x, y)$, sendo $\alpha^* \in \mathbb{C}$ o complexo conjugado de α .

Definição 4.12 (Mapa multilinear). Trata-se de uma generalização de um mapa bilinear. Sejam $\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n$ ($n \geq 1$) e \mathcal{W} espaços vetoriais, todos sobre o mesmo corpo \mathbb{K} . Um mapa multilinear f é a função

$$f : \mathcal{V}_1 \times \mathcal{V}_2 \times \dots \times \mathcal{V}_n \mapsto \mathcal{W},$$

tal que para todo $v_i \in \mathcal{V}_i$ e para um dado $w \in \mathcal{W}$,

$$v_i \mapsto w = f(v_1, \dots, v_i, \dots, v_n)$$

é um mapa linear de \mathcal{V}_i a \mathcal{W} .

Claramente, mapas lineares ou bilineares são casos particulares de mapas multilineares.

Forma multilinear. Ocorre quando a imagem do mapa multilinear é o corpo substrato aos espaços vetoriais. Ou seja, uma *forma multilinear* é a função

$$f : \mathcal{V}_1 \times \mathcal{V}_2 \times \cdots \times \mathcal{V}_n \mapsto \mathbb{K},$$

a qual é separadamente linear em cada um de seus argumentos. Ou seja, que satisfaça:

- $\alpha \cdot f(v_1, \dots, v_i, \dots, v_n) = f(v_1, \dots, \alpha \cdot v_i, \dots, v_n)$, para todos $\alpha \in \mathbb{K}$ e $v_i \in \mathcal{V}_i$ e para todo $i = 1, \dots, n$.
- $f(v_1, \dots, v_i, \dots, v_n) + f(v_1, \dots, v'_i, \dots, v_n) = f(v_1, \dots, v_i + v'_i, \dots, v_n)$, para todos $v_i, v'_i \in \mathcal{V}_i$ e para todo $i = 1, \dots, n$.

Algumas definições e teoremas acerca de mapeamentos entre ou sobre espaços vetoriais são apresentados a seguir.

Definição 4.13 (Composição de mapeamentos). Dados os espaços vetoriais \mathcal{V} , \mathcal{W} e \mathcal{Z} , todos sobre o corpo \mathbb{K} e os mapeamentos $S : \mathcal{V} \mapsto \mathcal{W}$ e $T : \mathcal{W} \mapsto \mathcal{Z}$, a *composição de T com S* é o mapeamento $T \circ S : \mathcal{V} \mapsto \mathcal{Z}$ dado por

$$(T \circ S)(u) = T(S(u)),$$

para todo $u \in \mathcal{V}$.

A operação de composição \circ não é, em geral, comutativa, *i. e.*, $T \circ S \neq S \circ T$; porém é associativa e distributiva, como o teorema a seguir assegura.

Teorema 4.9. *Sejam $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3$ e \mathcal{V}_4 espaços vetoriais sobre \mathbb{K} . Sejam $S_1, S_2 : \mathcal{V}_1 \mapsto \mathcal{V}_2$, $T_1, T_2 : \mathcal{V}_2 \mapsto \mathcal{V}_3$ e $U_1 : \mathcal{V}_3 \mapsto \mathcal{V}_4$ mapeamentos. Finalmente, seja $\alpha \in \mathbb{K}$. Então, as seguintes propriedades são satisfeitas:*

1. **Associatividade:** $U_1 \circ (T_1 \circ S_1) = (U_1 \circ T_1) \circ S_1$.

2. **Distributividade:**

(a) $(T_1 + T_2) \circ S_1 = T_1 \circ S_1 + T_2 \circ S_1$;

(b) $T_1 \circ (S_1 + S_2) = T_1 \circ S_1 + T_1 \circ S_2$, se T_1 é linear;

(c) $(\alpha T_1) \circ S_1 = \alpha (T_1 \circ S_1)$;

(d) $T_1 \circ (\alpha S_1) = \alpha (T_1 \circ S_1)$, se T_1 é linear;

Teorema 4.10 (Composição de mapeamentos lineares). *Sejam $\mathcal{V}_1, \mathcal{V}_2$ e \mathcal{V}_3 espaços vetoriais sobre \mathbb{K} . Se $L_1 : \mathcal{V}_1 \mapsto \mathcal{V}_2$ e $L_2 : \mathcal{V}_2 \mapsto \mathcal{V}_3$ são mapas lineares, então a composição $L_2 \circ L_1 : \mathcal{V}_1 \mapsto \mathcal{V}_3$ também é linear.*

As definições e o teorema a seguir estão relacionados com o *mapeamento inverso* de um dado mapeamento. A noção de um mapeamentos e operadores inversos está ligada à definição de matrizes inversas e possibilita a obtenção de diversas outras propriedades importantes a respeito de operadores lineares atuando sobre espaços vetoriais.

Definição 4.14 (Mapeamento identidade). Seja \mathcal{V} um espaço vetorial e $I_{\mathcal{V}} : \mathcal{V} \mapsto \mathcal{V}$ o operador linear tal que para todo $v \in \mathcal{V}$, $I_{\mathcal{V}}(v) = v$. A função $I_{\mathcal{V}}$ é então denominada o *mapeamento identidade* ou o *operador identidade*.

Definição 4.15 (Mapeamento invertível). Sejam \mathcal{V} e \mathcal{W} espaços vetoriais, cujos vetores nulos são, respectivamente, $0_{\mathcal{V}}$ e $0_{\mathcal{W}}$. Seja $T : \mathcal{V} \mapsto \mathcal{W}$ um operador. O mapa T é dito *invertível* se existe um outro mapeamento $S : \mathcal{W} \mapsto \mathcal{V}$ tal que $S \circ T = I_{\mathcal{V}}$ e $T \circ S = I_{\mathcal{W}}$. Neste caso, S é denominado o *mapeamento inverso de T* , sendo representado por $S = T^{-1}$.

Teorema 4.11. *Se $T : \mathcal{V} \mapsto \mathcal{W}$ é um mapa invertível, então, para todos $v \in \mathcal{V}$ e $w \in \mathcal{W}$,*

$$\text{Im}(T^{-1}) = \text{Dom}(T) = \mathcal{V} \subseteq \mathcal{V} \text{ e } \text{Dom}(T^{-1}) = \text{Im}(T) = \mathcal{W} \subseteq \mathcal{W},$$

e

$$T(v) = w \iff T^{-1}(w) = v.$$

Teorema 4.12. *Se $S : \mathcal{V} \mapsto \mathcal{W}$ e $T : \mathcal{W} \mapsto \mathcal{Z}$ são mapeamentos invertíveis, então*

$$(T \circ S)^{-1} = S^{-1} \circ T^{-1}.$$

Teorema 4.13. *Se $L : \mathcal{V} \mapsto \mathcal{W}$ é um operador linear invertível, então o seu inverso $L^{-1} : \mathcal{W} \mapsto \mathcal{V}$ também é linear.*

4.5 ESPAÇO VETORIAL DUAL

Qualquer espaço vetorial possui um correspondente espaço dual, o qual consiste no conjunto de todos os funcionais lineares do primeiro, conjunto este que constitui um outro espaço vetorial.

Dado o espaço vetorial \mathcal{V} sobre o corpo \mathbb{K} , o conjunto de todos os seus funcionais lineares, *i. e.*, todos os mapas lineares do tipo $\varphi : \mathcal{V} \mapsto \mathbb{K}$, representado por $\text{Hom}(\mathcal{V}, \mathbb{K})$ ou $L(\mathcal{V}, \mathbb{K})$, forma um novo espaço vetorial sobre \mathbb{K} com o mesmo produto por escalar, denominado *espaço vetorial dual a \mathcal{V}* ou, simplesmente, *espaço dual*, quando as seguintes propriedades são satisfeitas. Para todos $x \in \mathcal{V}$, $\varphi, \psi \in \text{Hom}(\mathcal{V}, \mathbb{K})$ e $\alpha \in \mathbb{K}$:

1. **Aditividade:** $(\varphi + \psi)(x) = \varphi(x) + \psi(x)$.
2. **Homogeneidade:** $(\alpha \cdot \varphi)(x) = \alpha \cdot (\varphi(x))$.

O espaço dual a \mathcal{V} pode ser representado por \mathcal{V}^* ou \mathcal{V}' quando o corpo \mathbb{K} fica subentendido. Um dado elemento de \mathcal{V}^* é também denominado de *covetor* ou *forma-um* ou *1-forma*.

O produto Cartesiano entre o espaço \mathcal{V} e seu dual \mathcal{V}^* , resultando no mapeamento $[\cdot, \cdot] : \mathcal{V}^* \times \mathcal{V} \mapsto \mathbb{K}$, é um exemplo de uma forma bilinear. Um determinado par ordenado do domínio deste mapeamento, obtido a partir de $\varphi \in \mathcal{V}^*$ e $x \in \mathcal{V}$, pode ser representado por $[\varphi, x] \doteq \varphi(x)$ e ser denominado um *bracket*.

Existem dois tipos de espaços duais: o *espaço dual algébrico* e o *espaço dual contínuo*. O primeiro tipo é a definição acima apresentada, válida para qualquer espaço dual. O segundo tipo surge quando se consideram somente funcionais lineares contínuos.

4.6 SUBESPAÇOS INVARIANTES

O conceito de subespaços invariantes é fundamental para a teoria de representação de grupos, discutida no capítulo 5. A definição de um subespaço invariante depende dos operadores lineares que atuam sobre o espaço vetorial.

Definição 4.16 (Subespaço invariante). Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} . Seja $L : \mathcal{V} \mapsto \mathcal{V}$ um operador linear sobre \mathcal{V} . Um subespaço $\mathcal{U} \subseteq \mathcal{V}$ é denominado um *subespaço invariante sob o operador L* se

$$L(\mathcal{U}) = \{L(u), \forall u \in \mathcal{U}\} \subseteq \mathcal{U}.$$

Ou seja, a ação do operador linear L sobre qualquer vetor do subespaço gera um outro vetor que também pertence ao mesmo subespaço. O exemplo 4.1 ilustra a importância dos subespaços vetoriais invariantes.

A partir das definições de espaço vetorial e dos mapeamentos e formas aplicados aos vetores, foram caracterizados diversos tipos particulares de espaços vetoriais, alguns dos quais serão descritos a seguir.

4.7 ESPAÇO VETORIAL NORMADO

Um *espaço vetorial normado*, ou simplesmente *espaço normado*, ou ainda *espaço normalizado*, é um espaço vetorial sobre o corpo dos complexos² no qual os vetores possuem uma *norma* definida. Um espaço vetorial normado é a estrutura $\langle \mathcal{V}, \|\cdot\| \rangle$, onde \mathcal{V} é um espaço vetorial e $\|\cdot\|$ é uma norma definida sobre \mathcal{V} .

Uma *norma* é uma função que opera sobre os vetores de um espaço e que, como resultado, atribui aos mesmos uma quantidade real estritamente positiva que pode ser interpretada, quando o espaço vetorial é também um espaço métrico, como sendo a medida geométrica do comprimento ou da extensão dos mesmos. A exceção é o vetor nulo, o qual possui norma também nula.

Definição (Norma). Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} . Uma *norma sobre \mathcal{V}* é uma função $\|\cdot\| : \mathcal{V} \mapsto \mathbb{R}$, tal que para todos $x, y \in \mathcal{V}$ e $\alpha \in \mathbb{R}$, as seguintes propriedades são satisfeitas:

²Ou sobre qualquer um de seus subcorpos.

1. **Positividade:** $\|x\| \geq 0$, sendo que se $\|x\| = 0$, então x é o vetor nulo.
2. **Homogeneidade absoluta ou escalabilidade absoluta:** $\|\alpha \cdot x\| = |\alpha| \|x\|$.
3. **Inigualdade triangular ou subaditividade:** $\|x + y\| \leq \|x\| + \|y\|$.

4.8 ESPAÇO COM PRODUTO INTERNO

Trata-se de um espaço vetorial sobre os corpos dos reais ou complexos ou qualquer um de seus subcorpos, para o qual é definido um *produto interno*.

Definição 4.17 (Produto interno). Dado um espaço vetorial \mathcal{V} sobre um corpo \mathbb{K} ($= \mathbb{R}$ ou \mathbb{C}), um *produto interno sobre o espaço \mathcal{V}* é a forma bilinear

$$\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{K},$$

que associa cada par de vetores $(u, v) \in \mathcal{V} \times \mathcal{V}$ com um escalar $\langle u, v \rangle \in \mathbb{K}$.

Para todos $u, v, w \in \mathcal{V}$ e $\alpha \in \mathbb{K}$, os seguintes *axiomas do produto interno* são satisfeitos:

(EPI₁) **Conjugação complexa:** $\langle u, v \rangle = \langle v, u \rangle^*$.

(EPI₂) **Linearidade:** $\langle u, \alpha \cdot v \rangle = \alpha \cdot \langle u, v \rangle$.

(EPI₃) **Distributividade:** $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.

Desta forma, definem-se os espaços vetoriais em questão.

Definição 4.18 (Espaço vetorial com produto interno). Seja um espaço vetorial \mathcal{V} sobre um corpo \mathbb{K} , o qual é dotado de um produto interno $\langle \cdot, \cdot \rangle_{\mathcal{V}} : \mathcal{V} \times \mathcal{V} \mapsto \mathbb{K}$. A dupla $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ forma um *espaço vetorial com produto interno*.

Pode-se demonstrar como consequência das definições 4.17 e 4.18 que as seguintes propriedades são também válidas, para quaisquer $u, v, w \in \mathcal{V}$ e $\alpha \in \mathbb{K}$:

1. $\langle \mathbf{0}, u \rangle = \langle u, \mathbf{0} \rangle = 0$.
2. $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$.
3. $\langle \alpha \cdot u, v \rangle = \alpha^* \langle u, v \rangle$.
4. Se $\langle u, v \rangle = \langle u, w \rangle$ para todo $u \in \mathcal{V}$, então $v = w$.

Os conjuntos dos escalares em espaços com produto interno fazem parte dos corpos $\mathbb{K} = \mathbb{R}$ (reais) ou \mathbb{C} (complexos). Neste caso, estes espaços são consecutivamente denominados *espaço real ou complexo com produto interno*. Nestes casos, uma propriedade adicional é imposta, dependendo da natureza do espaço vetorial.

A definição do produto interno entre dois vetores de um espaço vetorial permite também introduzir-se a noção de *ortogonalidade* entre os vetores.

Definição 4.19 (Vetores ortogonais). Seja \mathcal{V} um espaço vetorial com produto interno. Dois vetores $x, y \in \mathcal{V}$ são ditos ser *ortogonais* ou *perpendiculares* se

$$\langle x, y \rangle = 0.$$

Neste caso, emprega-se a notação $x \perp y$.

Até este momento, o conceito de vetores ortogonais ou perpendiculares não possui necessariamente nenhum caráter geométrico, uma vez que a definição 4.19 é puramente algébrica. A concepção geométrica entre dois vetores perpendiculares é válida quando o espaço vetorial também é um espaço métrico. Este tipo de espaço será discutido mais adiante.

4.8.1 ESPAÇO REAL COM PRODUTO INTERNO

Neste caso, o corpo é real ($\mathbb{K} = \mathbb{R}$) e uma propriedade adicional é imposta, a qual distingue dois tipos de espaços reais.

Espaços pseudo-Euclidianos. Tratam-se de espaços reais com produto interno, os quais obedecem a propriedade adicional:

(EPI₄) Se $\langle x, y \rangle = 0$ para todo $y \in \mathcal{V}$, então $x = 0$.

Vetores de um espaço pseudo-Euclidiano não podem possuir norma, mas possuem *magnitude*.

Magnitude. Sejam \mathcal{V} um espaço real pseudo-Euclidiano e $x \in \mathcal{V}$. A *magnitude* de x é definida por $\langle x, x \rangle$.

Um vetor x de um espaço pseudo-Euclidiano que não seja o *vetor nulo* (i. e., a identidade aditiva do conjunto dos vetores) pode possuir uma magnitude nula ou negativa. Caso $\langle x, x \rangle = 0$, mas $x \neq 0$, este é dito **auto-ortogonal**.³

Espaços Euclidianos ou positivo-definidos. Tratam-se de espaços reais com produto interno aos quais é imposta uma condição adicional mais restritiva que (EPI₄):

(EPI'₄) **Positividade:** $\langle x, x \rangle \geq 0$, sendo que $\langle x, x \rangle = 0$ se e somente se $x = 0$.

A condição (EPI'₄) implica na condição (EPI₄), pois se existisse um vetor $x \neq 0$ tal que $\langle x, y \rangle = 0$ para todo $y \in \mathcal{V}$, então, fazendo $y = x$ teríamos $\langle x, x \rangle = 0$, o que viola (EPI'₄).

4.8.2 ESPAÇO COMPLEXO COM PRODUTO INTERNO

Neste caso, o corpo é complexo ($\mathbb{K} = \mathbb{C}$) e a condição adicional normalmente imposta é a positividade (EPI'₄).

Uma observação também é necessária em relação à definição 4.17 de linearidade (EPI₂). Esta foi aqui definida em relação ao segundo argumento, seguindo a convenção estabelecida na física pela descrição formal da mecânica quântica. Textos matemáticos usualmente definem a linearidade no primeiro argumento. Esta distinção é necessária porque o produto interno sobre o corpo \mathbb{C} é uma forma sesquilinear, como consequência da conjugação complexa.

4.8.3 NORMA DE UM VETOR INDUZIDA PELO PRODUTO INTERNO

Em espaços (reais ou complexos) nos quais a condição de positividade (EPI'₄) é obedecida, pode-se definir a *norma* de um vetor x , a qual é induzida pelo produto interno, por

$$\|x\| \doteq \sqrt{\langle x, x \rangle} \geq 0.$$

Verifica-se que esta definição satisfaz as condições de norma de um vetor.

Neste caso, o espaço vetorial com produto interno também é um espaço normado. Os seguintes teoremas são válidos nesses espaços.

Teorema 4.14 (Inigualdade de Cauchy-Schwarz). *Seja \mathcal{V} um espaço vetorial normado com produto interno. Para qualquer par de vetores $x, y \in \mathcal{V}$,*

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (4.3)$$

Corolário 4.2. *A igualdade em (4.3) somente irá ocorrer se os vetores x e y forem proporcionais entre si, isto é,*

$$\langle x, y \rangle = \|x\| \|y\| \iff y = \alpha x, \text{ para algum } \alpha \in \mathbb{K} \text{ (} \mathbb{R} \text{ ou } \mathbb{C} \text{)}.$$

Da definição de ortogonalidade 4.19 segue imediatamente a generalização do teorema de Pitágoras:

³Ver definição 4.19 de vetores ortogonais.

Teorema 4.15 (Teorema de Pitágoras). *Seja \mathcal{V} um espaço vetorial normado com produto interno. Se os vetores $x, y \in \mathcal{V}$ são ortogonais, então*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Demonstração. Basta empregar as definições:

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 = \|x\|^2 + \|y\|^2.$$

□

Dados dois vetores pertencentes a um espaço normado com produto interno: x e y , com $y \neq 0$. O vetor x sempre pode ser decomposto em duas partes: uma paralela a y e uma perpendicular. Este processo é denominado a **decomposição ortogonal** do vetor x ao longo de y .

Inicialmente escreve-se $x = x_{\parallel} + x_{\perp}$, sendo $x_{\parallel} = \alpha y$, para algum $\alpha \in \mathbb{K}$ a ser determinado, a componente de x paralela a y e x_{\perp} a componente perpendicular. Como x_{\perp} deve ser ortogonal a y ,

$$x_{\perp} = x - x_{\parallel} = x - \alpha y \implies 0 = \langle x_{\perp}, y \rangle = \langle x - \alpha y, y \rangle = \langle x, y \rangle - \alpha \|y\|^2.$$

Portanto, $\alpha = \langle x, y \rangle / \|y\|^2$ e o vetor x fica decomposto então por

$$x = \frac{\langle x, y \rangle}{\|y\|^2} y + \left(x - \frac{\langle x, y \rangle}{\|y\|^2} y \right).$$

4.8.4 BASES ORTONORMAIS EM ESPAÇOS NORMADOS COM PRODUTO INTERNO

Espaços nos quais a propriedade de positividade (EPI'_4) é obedecida são adicionalmente beneficiados por possuírem *bases ortonormais*. Bases ortonormais têm diversas propriedades especiais que levam a simplificações úteis em cálculos envolvendo a álgebra de vetores. Essas bases fazem uso constante do conceito de vetores ortogonais (definição 4.19).

Definição 4.20 (Conjunto ortogonal/ortonormal). *Seja \mathcal{V} um espaço vetorial com produto interno. Seja $C = \{x_1, \dots, x_n\} \subset \mathcal{V}$ um conjunto de vetores não nulos. O conjunto C é dito ser *ortogonal* se*

$$\langle x_i, x_j \rangle = 0,$$

ou seja, se $x_i \perp x_j$ para todos $1 \leq i \neq j \leq n$.

O mesmo conjunto C é dito ser *ortonormal* se

$$\langle x_i, x_j \rangle = \delta_{ij}, \text{ para todos } i, j = 1, \dots, n,$$

onde $\delta_{ij} = 1$ (se $i = j$) ou $\delta_{ij} = 0$ (se $i \neq j$) é denominada a *delta de Kronecker*.⁴

Teorema 4.16 (Independência linear de um conjunto ortogonal). *Seja \mathcal{V} um espaço vetorial com produto interno. Se $C = \{x_1, \dots, x_n\} \subset \mathcal{V}$, com $n \leq \dim \mathcal{V}$, é um conjunto ortogonal de vetores, então C é linearmente independente.*

Demonstração. Dado o conjunto C de vetores, suponha que exista um conjunto de escalares $\{\alpha_1, \dots, \alpha_n\} \subset \mathbb{K}$ tal que

$$x \doteq \sum_{i=1}^n \alpha_i x_i = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n = 0.$$

Então,

$$\langle x, x \rangle = \sum_{i,j=1}^n \alpha_i^* \alpha_j \langle x_i, x_j \rangle = \sum_{i=1}^n |\alpha_i|^2 \langle x_i, x_i \rangle + \sum_{\substack{i,j=1 \\ (i \neq j)}}^n \alpha_i^* \alpha_j \langle x_i, x_j \rangle \stackrel{0}{=} \sum_{i=1}^n |\alpha_i|^2 \langle x_i, x_i \rangle = 0.$$

Como $\langle x_i, x_i \rangle > 0$ para $x_i \in C$ e $|\alpha_i|^2 \geq 0$ para todo $\alpha_i \in \mathbb{K}$, a igualdade acima somente será satisfeita se $\alpha_i = 0$ ($i = 1, \dots, n$). □

⁴Ver também seção 6.1.2.

Teorema 4.17 (Base ortonormal). *Seja \mathcal{V} um espaço vetorial normado com produto interno de dimensão n . Se $B = \{b_1, \dots, b_n\} \subset \mathcal{V}$ é um conjunto ortonormal de vetores, então B é uma base de \mathcal{V} , denominada base ortonormal.*

Demonstração. Para que B seja uma base é necessário (definição 4.8) que: (i) B seja LI e que (ii) \mathcal{V} seja gerado por B . A condição (i) é satisfeita por B ser um conjunto ortogonal e pelo teorema 4.16. A condição (ii) é satisfeita porque a cardinalidade de B é igual a $\dim \mathcal{V}$ e pelos teoremas da seção 4.3. Portanto, B é base de \mathcal{V} . \square

O conceito de ortogonalidade de vetores também serve para definir um *subespaço complementar ortogonal*.

Definição 4.21 (Complemento ortogonal). *Seja \mathcal{V} um espaço vetorial de dimensão finita com produto interno $\langle \cdot, \cdot \rangle$ e $C \subset \mathcal{V}$ um subconjunto de \mathcal{V} . O complemento ortogonal de C , denotado por C^\perp , é o conjunto*

$$C^\perp \doteq \{v \in \mathcal{V} \mid \langle u, v \rangle = 0 \text{ para todo } u \in C\}.$$

Observa-se que, de acordo com o teorema 4.2, tanto C quanto C^\perp formam subespaços de \mathcal{V} por varredura. Da mesma maneira, $\{\mathbf{0}\}^\perp = \mathcal{V}$ e $\mathcal{V}^\perp = \{\mathbf{0}\}$. Adicionalmente, se $C_1 \subset C_2 \subset \mathcal{V}$, então $C_2^\perp \subset C_1^\perp$.

Com a definição de subespaço complementar ortogonal, resulta o seguinte importante teorema.

Teorema 4.18. *Se $\mathcal{C} \subset \mathcal{V}$ é subespaço de um espaço vetorial com produto interno de dimensão finita, então*

$$\mathcal{V} = \mathcal{C} \oplus \mathcal{C}^\perp.$$

Demonstração. De acordo com o teorema 4.1, para que \mathcal{C} e \mathcal{C}^\perp sejam subespaços complementares de \mathcal{V} , é necessário que (i) $\mathcal{V} = \mathcal{C} + \mathcal{C}^\perp$ e (ii) $\mathcal{C} \cap \mathcal{C}^\perp = \{\mathbf{0}\}$.

Para mostrar que a condição (i) é satisfeita, seja $n = \dim(\mathcal{C})$ e $B = \{e_1, \dots, e_n\}$ uma base ortonormal de \mathcal{C} . Então, para todo $v \in \mathcal{V}$, pode-se escrever

$$v = \underbrace{\left(\sum_{j=1}^n \langle v, e_j \rangle e_j \right)}_u + v - \underbrace{\left(\sum_{j=1}^n \langle v, e_j \rangle e_j \right)}_w.$$

Claramente $u \in \mathcal{C}$. Mas, em sendo $w = v - u$,

$$\langle w, e_j \rangle = \langle v, e_j \rangle - \langle u, e_j \rangle = 0, \quad (j = 1, \dots, n).$$

Ou seja, $w \in \mathcal{C}^\perp$, implicando que todo $v \in \mathcal{V}$ pode ser escrito como $v = u + w$.

Para provar a condição (ii), seja agora $v \in \mathcal{C} \cap \mathcal{C}^\perp$. Isto implica que v deve ser ortogonal a todos os vetores em \mathcal{C} , inclusive a si próprio. Portanto, necessariamente, $v = \mathbf{0}$, implicando em $\mathcal{C} \cap \mathcal{C}^\perp = \{\mathbf{0}\}$. \square

4.8.5 OPERADORES LINEARES EM ESPAÇOS COM PRODUTO INTERNO

A definição de um produto interno para o espaço vetorial possibilita a definição e subsequente obtenção de novas propriedades interessantes com relação a mapas e operadores lineares que atuam sobre esses espaços.

Definição 4.22 (Mapeamento adjunto). *Sejam $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ e $(\mathcal{W}, \langle \cdot, \cdot \rangle_{\mathcal{W}})$ espaços vetoriais com produto interno sobre o corpo \mathbb{K} . Seja $L : \mathcal{V} \mapsto \mathcal{W}$ um mapeamento linear de \mathcal{V} sobre \mathcal{W} . O adjunto de L é o mapeamento $L^* : \mathcal{W} \mapsto \mathcal{V}$, o qual satisfaz, para todos $v \in \mathcal{V}$ e $w \in \mathcal{W}$,*

$$\langle L(v), w \rangle_{\mathcal{W}} = \langle v, L^*(w) \rangle_{\mathcal{V}}.$$

Teorema 4.19. *Sejam $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ e $(\mathcal{W}, \langle \cdot, \cdot \rangle_{\mathcal{W}})$ espaços vetoriais com produto interno sobre o corpo \mathbb{K} . Para todo mapa linear $L : \mathcal{V} \mapsto \mathcal{W}$ existe um único adjunto $L^* : \mathcal{W} \mapsto \mathcal{V}$.*

Operador adjunto. A partir das definições de operador linear (definição 4.10) e de mapeamento adjunto (definição 4.22) e, de acordo com o teorema 4.19, dado um operador linear $L : \mathcal{V} \mapsto \mathcal{V}$, existe um único *operador adjunto* $L^* : \mathcal{V} \mapsto \mathcal{V}$ que satisfaz

$$\langle L(u), v \rangle = \langle u, L^*(v) \rangle,$$

para todos $u, v \in \mathcal{V}$.

Sendo L e \widehat{L} dois operadores lineares atuando sobre \mathcal{V} e dado $\alpha \in \mathbb{K}$, os respectivos operadores adjuntos satisfazem as propriedades:

1. $(\alpha L)^* = \alpha^* L^*$;
2. $(L^*)^* = L$;
3. $(L + \widehat{L})^* = L^* + \widehat{L}^*$;
4. $(L \circ \widehat{L})^* = \widehat{L}^* \circ L^*$.

Definição 4.23 (Operador autoadjunto ou Hermitiano). Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ um espaço vetorial com produto interno sobre o corpo \mathbb{K} . Seja $L : \mathcal{V} \mapsto \mathcal{V}$ um operador linear sobre \mathcal{V} . O operador L é denominado *autoadjunto* ou *Hermitiano* se

$$L^* = L.$$

O nome *operador Hermitiano* é usualmente atribuído quando $\mathbb{K} = \mathbb{C}$. Por outro lado, quando o corpo é real ($\mathbb{K} = \mathbb{R}$), o operador também pode ser denominado **simétrico**.

Teorema 4.20. Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ um espaço vetorial com produto interno sobre o corpo \mathbb{K} . Seja $L : \mathcal{V} \mapsto \mathcal{V}$ um operador linear sobre \mathcal{V} . O operador L é autoadjunto se e somente se

$$\langle L(u), v \rangle = \langle u, L(v) \rangle,$$

para todos $u, v \in \mathcal{V}$.

O teorema a seguir estabelece a relação entre operadores Hermitianos em espaços vetoriais de dimensão finita e matrizes Hermitianas, definidas na seção 4.9.3.2.

Teorema 4.21. Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ um espaço vetorial de dimensão finita com produto interno sobre o corpo \mathbb{K} com a base ortonormal ordenada B . Sejam também $L \in \mathbb{L}(\mathcal{V})$ e L_B a matriz de L em relação à base B . Então L é um operador Hermitiano se e somente se L_B for Hermitiana.

Um tipo especial de operadores lineares que atuam sobre espaços vetoriais com produto interno são os operadores unitários. A definição destes é largamente influenciada pela definição de uma matriz unitária, realizada na seção 4.9.3.2.

Definição 4.24 (Operador unitário). Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ um espaço vetorial com produto interno. Seja $U : \mathcal{V} \mapsto \mathcal{V}$ um operador linear que atua sobre os vetores de \mathcal{V} . Este operador é dito ser *unitário com respeito ao produto interno* se

$$U^* = U^{-1}.$$

Uma propriedade importante de operadores unitários é que estes preservam a norma de um vetor e o produto interno de dois vetores.

Teorema 4.22. Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ um espaço vetorial com produto interno. Seja $U : \mathcal{V} \mapsto \mathcal{V}$ um operador linear que atua sobre \mathcal{V} . Se U é um operador unitário com relação ao produto interno, então, para quaisquer $u, v \in \mathcal{V}$,

- $\|U(u)\| = \|u\|$.
- $\langle U(u), U(v) \rangle = \langle u, v \rangle$.

Definição 4.25 (Operador normal). Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle_{\mathcal{V}})$ um espaço com produto interno. Um operador linear $L \in \mathbb{L}(\mathcal{V})$ é dito *normal* se

$$L \circ L^* = L^* \circ L.$$

Alguns teoremas importantes referentes a operadores normais são apresentados a seguir.

Teorema 4.23. *Todos os operadores autoadjuntos ou unitários são operadores normais.*

Demonstração. Suponha primeiro que $L \in \mathbb{L}(\mathcal{V})$ seja autoadjunto. Então, $L = L^*$ por definição, o que imediatamente implica em $L \circ L^* = L^* \circ L$ e o operador é normal.

Suponha agora que L seja unitário. Então, $L^* = L^{-1}$ por definição, de tal forma que

$$L \circ L^* = L \circ L^{-1} = I_{\mathcal{V}} = L^{-1} \circ L = L^* \circ L,$$

sendo $I_{\mathcal{V}}$ o operador identidade. Portanto, L é normal. \square

Teorema 4.24. *Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ um espaço com produto interno sobre o corpo \mathbb{K} . Então, $L \in \mathbb{L}(\mathcal{V})$ é um operador normal se e somente se $\|L(v)\| = \|L^*(v)\|$ para todo $v \in \mathcal{V}$.*

4.9 EMPREGO DE MATRIZES EM ESPAÇOS VETORIAIS

Serão apresentadas algumas definições e propriedades com respeito a matrizes e suas aplicações em espaços vetoriais.

4.9.1 DEFINIÇÕES INICIAIS

Dados $n, m \in \mathbb{N}$ e o corpo \mathbb{K} . Uma **matriz** $m \times n$ **sobre** \mathbb{K} é um arranjo retangular de elementos de \mathbb{K} arranjados em m linhas e n colunas e representados por

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \text{ou} \quad \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (4.4)$$

onde $\{a_{ij}\} \subset \mathbb{K}$ ($1 \leq i \leq m$, $1 \leq j \leq n$). A quantidade $a_{ij} \in \mathbb{K}$ que está na i -ésima linha e j -ésima coluna é denominada o **elemento** ij da matriz.

De forma consistente com a definição de um espaço vetorial, o corpo \mathbb{K} que compõe os elementos da matriz é o mesmo corpo do conjunto dos escalares do espaço. Por esta razão, o elemento a_{ij} também é genericamente denominado como *escalar*.

Em aplicações na física, os corpos em questão são geralmente $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Neste caso, matrizes pertencentes a $\mathbb{R}^{m \times n}$ são denominadas **matrizes reais**, enquanto que matrizes em $\mathbb{C}^{m \times n}$ são denominadas **matrizes complexas**.

Uma matriz 1×1 consiste em um único escalar. Para $n \geq 2$, matrizes $n \times 1$ ou $1 \times n$ são chamadas **matrizes vetoriais** ou, quando não houver ambiguidade, simplesmente **vetores**. Um vetor

$$n \times 1 : \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

é denominado **vetor coluna**, enquanto que um vetor $1 \times n : (x_1 \ x_2 \ \cdots \ x_n)$ é denominado **vetor linha**.

Dado elemento $a_{ij} \in \mathbb{K}$ da matriz (4.4), os conjuntos $I_\ell \doteq \{1, \dots, m\}$ e $I_c \doteq \{1, \dots, n\}$ são denominados os **intervalos** dos índices i e j , respectivamente. Por sua vez, a matriz (4.4) pode ser representada pelos símbolos A , $A_{m,n}$, $[a_{ij}]_{m,n}$ ou simplesmente $[a_{ij}]$, caso os intervalos dos índices sejam irrelevantes. Por sua vez, o elemento a_{ij} também pode ser representado por $(A)_{ij}$ ou $(A)_{ij}$.

O símbolo,

$$\mathbb{K}^{m \times n} \doteq \left\{ [a_{ij}]_{m,n} : a_{ij} \in \mathbb{K}, \forall 1 \leq i \leq m, 1 \leq j \leq n \right\}$$

denota o conjunto de todas as matrizes $m \times n$ sobre o corpo \mathbb{K} . Em particular, o conjunto dos vetores colunas é denotado por $\mathbb{K}^n \equiv \mathbb{K}^{n \times 1}$.

Se $a_{ij} = 0$ para todos i e j nos seus intervalos, resulta a **matriz nula**

$$0_{m,n} = [0]_{m,n} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

4.9.2 ÁLGEBRA MATRICIAL

Serão agora definidas as operações fundamentais da álgebra matricial. Como os elementos das matrizes pertencem a um corpo \mathbb{K} qualquer, as operações denominadas *adição* ou *produto* deveriam ser representadas pelos símbolos genéricos das operações entre elementos do corpo:⁵ \oplus para a adição e \bullet para o produto. Contudo, sem haver ambiguidade, nesta seção estas operações serão representadas pelos símbolos adotados nas operações entre elementos dos corpos \mathbb{R} ou \mathbb{C} .

Sejam as matrizes $A, B, C \in \mathbb{K}^{m \times n}$ e o escalar $k \in \mathbb{K}$. A **soma** ou **adição matricial** de A por B resulta na matriz C , cujos elementos são

$$c_{ij} = a_{ij} + b_{ij}, \quad (i \in I_\ell, j \in I_c);$$

ou, na forma matricial,

$$\begin{aligned} C = A + B &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}. \end{aligned}$$

Claramente, a adição matricial é comutativa: $C = A + B = B + A$.

A **multiplicação** ou **produto matricial por escalar** de A por k resulta na matriz B , cujos elementos são

$$B \equiv [b_{ij}] = [ka_{ij}] = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}.$$

Estas operações foram explicitamente definidas de forma análoga às respectivas operações em espaços Euclidianos.⁶

O conjunto $\mathbb{K}^{m \times n}$ de todas as matrizes $m \times n$ forma um espaço vetorial sob as operações de adição matricial e multiplicação por escalar, sendo o vetor nulo a matriz nula $0_{m,n}$. Em particular, o conjunto $\mathbb{R}^{m \times n}$ forma um espaço vetorial real e o conjunto $\mathbb{C}^{m \times n}$ forma um espaço vetorial complexo.

O **produto** ou **multiplicação matricial** de uma matriz $A \in \mathbb{K}^{m \times p}$ por uma matriz $B \in \mathbb{K}^{p \times n}$ resulta na matriz $C \in \mathbb{K}^{m \times n}$, cujos elementos são

$$c_{ij} = \sum_{\ell=1}^p a_{i\ell} b_{\ell j}, \quad (i \in I_\ell, j \in I_c).$$

Esta operação é representada por

$$C = AB.$$

A multiplicação matricial da matriz A pela matriz B somente está definida se o intervalo $I_{c,A}$ (nas colunas) da primeira é idêntico ao intervalo $I_{\ell,B}$ (nas linhas) da segunda. A matriz

⁵Ver definição na página 124.

⁶Ver exemplo 164.

C resultante tem intervalos $I_{\ell,A}$ e $I_{c,B}$. Portanto, multiplicações matriciais do tipo BA quando $I_{c,B} \neq I_{\ell,A}$ não estão definidas.

Dadas as matrizes $A \in \mathbb{K}^{m \times n}$, $B, C \in \mathbb{K}^{n \times p}$ e $D \in \mathbb{K}^{p \times q}$ e o escalar $k \in \mathbb{K}$, verifica-se as propriedades da multiplicação matricial:

1. $A(kB) = k(AB)$.
2. **Distributividade:** $A(B + C) = AB + AC$.
3. **Associatividade:** $(AB)D = A(BD)$.

A **transposta** da matriz $A \in \mathbb{K}^{m \times n}$ é a matriz $\tilde{A} \in \mathbb{K}^{n \times m}$ (ou A^T) tal que

$$(\tilde{A})_{ij} = (A)_{ji}, \quad (1 \leq i \leq n, 1 \leq j \leq m);$$

$$\tilde{A} \equiv A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

Verifica-se facilmente que :

1. $(A^T)^T = A$.
2. $(kA)^T = kA^T$.
3. $(A + B)^T = A^T + B^T$.
4. $(AB)^T = B^T A^T$.

Para matrizes complexas, as seguintes operações são definidas. O **complexo conjugado** da matriz $A \in \mathbb{C}^{m \times n}$ é a matriz $A^* \in \mathbb{C}^{m \times n}$, cujos elementos são

$$(A^*)_{ij} = a_{ij}^*, \quad (i \in I_\ell, j \in I_c);$$

$$A^* = \begin{pmatrix} a_{11}^* & a_{12}^* & \cdots & a_{1n}^* \\ a_{21}^* & a_{22}^* & \cdots & a_{2n}^* \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}^* & a_{m2}^* & \cdots & a_{mn}^* \end{pmatrix}.$$

Claramente, $(A^*)^* \equiv A^{**} = A$. Além disso, se $A \in \mathbb{R}^{m \times n}$, então $A^* = A$.

O **Hermitiano conjugado** ou a **transposta conjugada** ou a **adjunta** da matriz $A \in \mathbb{C}^{m \times n}$ é a matriz $A^\dagger \in \mathbb{C}^{n \times m}$, cujos elementos são dados por

$$(A^\dagger)_{ij} = a_{ji}^*, \quad (1 \leq i \leq n, 1 \leq j \leq m);$$

$$A^\dagger = \begin{pmatrix} a_{11}^* & a_{21}^* & \cdots & a_{m1}^* \\ a_{12}^* & a_{22}^* & \cdots & a_{m2}^* \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}^* & a_{2n}^* & \cdots & a_{mn}^* \end{pmatrix}.$$

Dada a matriz $A \in \mathbb{C}^{m \times n}$ e o escalar $k \in \mathbb{C}$, as principais propriedades da conjugação Hermitiana são:

1. $(A^\dagger)^\dagger \equiv A^{\dagger\dagger} = A$.
2. $(kA)^\dagger = k^* A^\dagger$.
3. Dada $B \in \mathbb{C}^{m \times n}$, $(A + B)^\dagger = A^\dagger + B^\dagger$.
4. Dada $B \in \mathbb{C}^{n \times p}$, $(AB)^\dagger = B^\dagger A^\dagger$.

4.9.3 MATRIZES QUADRADAS

Se $m = n$, resulta a **matriz quadrada** $A_{n,n} \equiv A_n \equiv [a_{ij}]_{n,n} \equiv [a_{ij}]_n$. Os elementos a_{ij} com $j = i$ são chamados elementos da **diagonal principal** de A_n .

Algumas operações válidas para matrizes quadradas serão apresentadas agora. Dada a matriz quadrada $A \in \mathbb{K}^{n \times n}$, o seu **traço**, representado por $\text{Tr}(A)$, corresponde ao mapeamento $\text{Tr}(\cdot) : \mathbb{K}^{n \times n} \mapsto \mathbb{K}$, definido como a soma dos elementos na diagonal principal:

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii}.$$

Dadas as matrizes $A, B \in \mathbb{K}^{n \times n}$ e o escalar $k \in \mathbb{K}$, algumas propriedades do traço são:

1. $\text{Tr}(kA) = k\text{Tr}(A)$.
2. $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$.
3. $\text{Tr}(AB) = \text{Tr}(BA)$.
4. $\text{Tr}(\tilde{A}) = \text{Tr}(A)$.
5. $\text{Tr}(A^\dagger) = [\text{Tr}(A)]^*$.

4.9.3.1 DETERMINANTE

Dada a matriz quadrada $A \in \mathbb{K}^{n \times n}$, o seu **determinante** corresponde ao mapeamento (função) $\det_n(\cdot) \equiv \det(\cdot) : \mathbb{K}^{n \times n} \mapsto \mathbb{K}$. Este mapeamento pode ser definido de diferentes maneiras.

EXPANSÃO DE LAPLACE A expansão de Laplace calcula o determinante de uma matriz quadrada de forma recursiva. Inicialmente, o determinante da matriz $A \in \mathbb{K}^{1 \times 1} = [a]$ é definido como

$$\det_1(A) = a.$$

Em seguida, o determinante da matriz $A \in \mathbb{K}^{2 \times 2}$ é definido como

$$\det_2(A) = a_{11}a_{22} - a_{12}a_{21}.$$

Para generalizar o cálculo do determinante para $n \geq 2$, define-se as seguintes quantidades:

Menor. O menor $M_{ij} \in \mathbb{K}$ da matriz $A \in \mathbb{K}^{n \times n}$ é o determinante da matriz $(n-1) \times (n-1)$ obtida pela remoção dos elementos na i -ésima linha e j -ésima coluna.

Cofator. O cofator $C_{ij} \in \mathbb{K}$ é dado por $C_{ij} = (-1)^{i+j} M_{ij}$.

Então, para $n \geq 2$, o determinante da matriz $A \in \mathbb{K}^{n \times n}$ pode ser calculado de diferentes maneiras:

$$\det_n(A) = \sum_{j=1}^n a_{ij} C_{ij}, \quad (i \leq n)$$

é a *expansão de Laplace ao longo da i -ésima linha*, enquanto que

$$\det_n(A) = \sum_{i=1}^n a_{ij} C_{ij}, \quad (j \leq n)$$

é a *expansão de Laplace ao longo da j -ésima coluna*.

USANDO O SÍMBOLO DE LEVI-CIVITA Empregando o símbolo de Levi-Civita com n índices definido na seção 268, o determinante da matriz $A \in \mathbb{K}^{n \times n}$ pode ser calculado como

$$\det_n(A) = \sum_{i_1, i_2, \dots, i_n=1}^n \epsilon_{i_1 i_2 \dots i_n} a_{1i_1} a_{2i_2} \cdots a_{ni_n}.$$

PROPRIEDADES DE DETERMINANTES Dadas as matrizes $A, B \in \mathbb{K}^{n \times n}$ e o escalar $k \in \mathbb{K}$, as seguintes propriedades são válidas:

1. $\det(\tilde{A}) = \det(A)$.
2. $\det(A^\dagger) = [\det(A)]^*$.
3. (*troca de duas linhas ou colunas*) Seja A com determinante $\det(A)$ e seja B a matriz obtida trocando-se quaisquer duas linhas ou colunas de A . Então $\det(B) = -\det(A)$.
4. (*fatores em comum*) Se todos os elementos de uma dada linha ou coluna de A são proporcionais a k , então este fator em comum pode ser removido e $\det(A)$ será igual a k multiplicado pelo determinante da matriz restante.
Se $k = 0$, ou seja, os elementos de alguma linha ou coluna são nulos, então $\det(A) = 0$. Além disso, $\det(kA) = k^n \det(A)$.
5. (*linhas ou colunas idênticas*) Se quaisquer duas linhas ou colunas de A forem idênticas ou múltiplas entre si, então $\det(A) = 0$.
6. (*adicionando um múltiplo constante de uma linha ou coluna a outra*) O determinante de uma matriz não se modifica se os elementos de uma linha (coluna) forem somados a um múltiplo fixo dos elementos de outra linha (coluna).
7. (*determinante de um produto matricial*) $\det(AB) = \det(A) \det(B)$.

4.9.3.2 TIPOS ESPECIAIS DE MATRIZES QUADRADAS

Diversos tipos particulares de matrizes quadradas $A \in \mathbb{K}^{n \times n}$ são importantes. Alguns desses tipos serão definidos agora.

Matriz diagonal. É aquela matriz cujos únicos elementos não nulos estão na sua diagonal principal.

Matriz identidade (de ordem n). É a matriz diagonal $I_n \in \mathbb{K}^{n \times n}$ tal que $a_{ij} = \delta_{ij}$, sendo δ_{ij} o delta de Kronecker. Observa-se que $\det(I_n) = 1$.

Matriz diagonal superior ou inferior. A matriz $A \in \mathbb{K}^{n \times n}$ é uma *matriz diagonal superior* se os seus elementos abaixo da diagonal principal são todos nulos, ou uma *matriz diagonal inferior* se seus elementos acima da diagonal principal são nulos. Respectivamente,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2(n-1)} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{(m-1)(n-1)} & a_{(m-1)n} \\ 0 & 0 & \cdots & 0 & a_{mn} \end{pmatrix} \quad (\text{diagonal superior})$$

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{(m-1)1} & a_{(m-1)2} & \cdots & a_{(m-1)(n-1)} & 0 \\ a_{m1} & a_{m2} & \cdots & a_{m(n-1)} & a_{mn} \end{pmatrix} \quad (\text{diagonal inferior})$$

Matriz de cofatores (ou comatriz). Dada $A \in \mathbb{K}^{n \times n}$, a sua *comatriz* é a matriz $[C_{ij}]_{n,n}$ formada pelos cofatores de A .

Matriz adjunta (ou auxiliar). Dada $A \in \mathbb{K}^{n \times n}$, a sua *matriz adjunta* (*adjugate matrix*), denotada por $\text{adj}(A) \in \mathbb{K}^{n \times n}$, é a transposta da comatriz de A : $\text{adj}(A) = \widetilde{[C_{ij}]_{n,n}}$. Algumas de suas propriedades são

1. $\text{adj}(A) = I_n$.
2. Se $k \in \mathbb{K}$, $\text{adj}(kA) = k^{n-1} \text{adj}(A)$.

3. $\text{adj}(\tilde{A}) = \widetilde{\text{adj}(A)}$.
4. $\text{adj}(A^*) = [\text{adj}(A)]^*$.
5. $\text{adj}(A^\dagger) = [\text{adj}(A)]^\dagger$.
6. Se $B \in \mathbb{K}^{n \times n}$, $\text{adj}(AB) = \text{adj}(B) \text{adj}(A)$.
7. $\det[\text{adj}(A)] = [\det(A)]^{n-1}$.

Matriz inversa. Caso exista, a matriz $A^{-1} \in \mathbb{K}^{n \times n}$ será a *inversa* da matriz $A \in \mathbb{K}^{n \times n}$ se

$$AA^{-1} = A^{-1}A = I_n,$$

sendo $I_n \in \mathbb{K}^{n \times n}$ a matriz identidade.

Condições de existência da inversa e algumas de suas propriedades são:

1. Se $\det(A) \neq 0$, então A^{-1} existe e A é denominada **inversível** ou **não singular** ou **não degenerada**. Em caso contrário, A^{-1} não existe e A é denominada **singular** ou **degenerada**.
2. Se A for inversível, então

$$A^{-1} = \frac{\text{adj}(A)}{\det(A)}.$$

3. $(A^{-1})^{-1} = A$.
4. $(A^T)^{-1} = (A^{-1})^T$.
5. $(A^\dagger)^{-1} = (A^{-1})^\dagger$.
6. $(AB)^{-1} = B^{-1}A^{-1}$.
7. $(I_n)^{-1} = I_n$.
8. $\det(A^{-1}) = [\det(A)]^{-1}$.
9. $\text{adj}(A^{-1}) = [\text{adj}(A)]^{-1}$.

Matriz simétrica ou antissimétrica. A matriz A é dita *simétrica* se $A = \tilde{A}$. Por sua vez, a matriz será *antissimétrica* se $A = -\tilde{A}$. Se $A = \pm\tilde{A}$, então $A^{-1} = \pm\tilde{A}^{-1}$.

Matriz ortogonal. A matriz inversível A é *ortogonal* se $A^{-1} = A^T$. Neste caso, observa-se que $\det(A\tilde{A}) = \det(\tilde{A}A) = \det^2(A) = 1$. Ou seja, o determinante de uma matriz ortogonal é $\det(A) = \pm 1$.

Matriz Hermitiana ou anti-Hermitiana. A matriz A é dita *Hermitiana* se $A = A^\dagger$ ou *anti-Hermitiana* se $A = -A^\dagger$. Se $A = \pm A^\dagger$, então $A^{-1} = \pm (A^{-1})^\dagger$.

Matriz unitária. A matriz A é *unitária* se $A^{-1} = A^\dagger$. Neste caso, $\det(AA^\dagger) = \det(A)\det(A^\dagger) = |\det(A)|^2 = 1$. Ou seja, o determinante de uma matriz unitária é $\det(A) = e^{i\alpha}$ ($\alpha \in \mathbb{R}$).

É possível mostrar que se $A, B \in \mathbb{K}^{n \times n}$ são unitárias e $k \in \mathbb{K}$, então A^{-1} , kA , $A + B$ e AB também são unitárias.

Matriz normal. A matriz A é *normal* se

$$AA^\dagger = A^\dagger A.$$

Verifica-se facilmente que matrizes simétricas, ortogonais, Hermitianas ou unitárias são todas exemplos de matrizes normais. Da mesma forma, a matriz A^{-1} também é normal, pois $A^{-1}(A^{-1})^\dagger = A^{-1}(A^\dagger)^{-1} = (A^\dagger A)^{-1} = (AA^\dagger)^{-1} = (A^{-1})^\dagger A^{-1}$.

Matrizes similares. Dadas as matrizes $A, B \in \mathbb{K}^{n \times n}$, a matriz A é *similar* a B se existir uma matriz inversível $Q \in \mathbb{K}^{n \times n}$ tal que

$$B = Q^{-1}AQ. \quad (4.5)$$

Expressão (4.5) é denominada **transformação de similaridade**. Esta transformação serve para definir uma classe de equivalência⁷ entre as matrizes do espaço $\mathbb{K}^{n \times n}$. Outras propriedades de matrizes similares são:

⁷Definição 3.13.

1. A é inversível se e somente se B for inversível.
2. $\det(A) = \det(B)$.
3. $\text{Tr}(A) = \text{Tr}(B)$.

4.9.4 MATRIZES ASSOCIADAS A VETORES E OPERADORES

Nesta seção, as definições, operações e propriedades relacionadas a matrizes serão aplicadas a diversas operações envolvendo vetores e operadores lineares.

Será definido em primeiro lugar o mapeamento responsável pela representação de um vetor em uma matriz coluna, a partir de uma determinada base do espaço.

Definição 4.26 (Mapa coordenado da base). Dados: o espaço vetorial \mathcal{V}^n de dimensão n sobre o corpo \mathbb{K} , $B = \{e_1, \dots, e_n\}$ uma base de \mathcal{V}^n e o vetor $v = \sum_i v_i \cdot e_i \in \mathcal{V}^n$ decomposto sobre B de acordo com o teorema 4.8, o mapeamento $\varphi_B(v) : \mathcal{V}^n \mapsto \mathbb{K}^n$, denominado o *mapa coordenado de B* , é dado por

$$\varphi_B(v) \doteq [v]_B \equiv v_B \equiv v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

onde $v_B \in \mathbb{K}^n$ é uma matriz coluna.

Para evitar a ocorrência de ambiguidades e definir a matriz v_B de forma coerente, a base B deve rigorosamente ser uma **base ordenada**, na qual sempre haverá concordância entre o ordenamento dos índices dos componentes do vetor $\{v_i\}$ ($1 \leq i \leq n$) com os índices dos elemento da matriz $\{v_{B,i}\}$.

O teorema a seguir estabelece uma importante relação entre o produto interno e os mapas coordenados.

Teorema 4.25. *Seja $\langle \mathcal{V}, \langle, \rangle \rangle$ um espaço vetorial com produto interno sobre \mathbb{K} . Se B é uma base ortonormal ordenada para \mathcal{V} , então, para todos $u, v \in \mathcal{V}$,*

$$\langle u, v \rangle = u_B^\dagger v_B.$$

Demonstração. O enunciado do teorema é evidentemente correto se $\mathcal{V} = \{0\}$. Assumindo então que a dimensão do espaço é $n \geq 1$ e que $B = \{b_1, \dots, b_n\}$ é uma base ortonormal, então, de acordo com o teorema 4.8, existem $\{u_i\}, \{v_i\} \in \mathbb{K}$ tais que $u = u_1 b_1 + \dots + u_n b_n$ e $v = v_1 b_1 + \dots + v_n b_n$. Assim, de acordo com a definição 4.26 existem as matrizes coluna $u_B = (u_1 \dots u_n)^\top$ e $v_B = (v_1 \dots v_n)^\top$.

Agora, uma vez que B é ortonormal, $\langle b_i, b_j \rangle = \delta_{ij}$ ($\forall 1 \leq i, j \leq n$). Pelas definições 4.17 e 4.18 e suas consequências,

$$\langle u, v \rangle = \left\langle \sum_{i=1}^n u_i b_i, \sum_{i=1}^n v_i b_i \right\rangle = \sum_{i,j=1}^n u_i^* v_j \langle b_i, b_j \rangle = \sum_{i,j=1}^n u_i^* v_j \delta_{ij} = \sum_{i=1}^n u_i^* v_i = u_B^\dagger v_B.$$

□

4.9.4.1 TRANSFORMAÇÃO DE BASE

A transformação das coordenadas de um vetor discutida na seção 4.3.1 pode ser representada na forma matricial. Dado o espaço vetorial de dimensão n \mathcal{V}^n sobre o corpo \mathbb{K} com as bases ordenadas $B = \{e_1, \dots, e_n\}$ e $B' = \{e'_1, \dots, e'_n\}$ e o vetor $v \in \mathcal{V}^n$, a lei de transformação dos componentes de v sobre a base B para componentes sobre B' é dada por (4.2).

De acordo com a definição 4.26, o vetor v pode ser representado pelas matrizes coluna

$$v_B \doteq \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad v_{B'} \doteq \begin{pmatrix} v'_1 \\ \vdots \\ v'_n \end{pmatrix}.$$

Por outro lado, dado o conjunto $\{A_{ij}\} \subset \mathbb{K}$, onde A_{ij} é o i -ésimo componente do vetor $e_j \in B$ com respeito à base B' , de acordo com (4.2), estes componentes podem ser organizados na matriz $A_{BB'} \equiv [A_{ij}]_{n,n} \in \mathbb{K}^{n \times n}$.

Desta maneira, a lei de transformação (4.2) pode ser expressa em termos de multiplicação matricial como

$$v_{B'} = A_{BB'} v_B,$$

ficando novamente subentendido que as multiplicações ocorrem via a operação de produto “•” do corpo \mathbb{K} .

Pode ser demonstrado que a matriz de transformação $A_{BB'}$ é inversível e a sua inversa, a matriz $A_{BB'}^{-1} \in \mathbb{K}^{n \times n}$, composta por elementos $(A_{BB'}^{-1})_{ij} \in \mathbb{K}$, corresponde à transformação inversa $B' \rightarrow B$; ou seja, $A_{B'B}^{-1} = A_{B'B}$, tal que

$$v_B = A_{B'B} v_{B'} \implies v_i = \sum_{j=1}^n (A_{B'B})_{ij} \cdot v'_j.$$

4.9.4.2 MATRIZ DE UM OPERADOR LINEAR

A ação de um operador linear também pode ser representada por meio de uma matriz. Esta propriedade é crucial, por exemplo, para a teoria de representação de grupos, apresentada no capítulo 5.

Definição 4.27 (Matriz de um operador linear). Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} ; seja $T : \mathcal{V} \rightarrow \mathcal{V}$ um operador linear atuando sobre \mathcal{V} e seja $B = \{b_1, b_2, \dots\}$ uma base de \mathcal{V} . De acordo com o teorema 4.8, a decomposição

$$T(b_i) = \sum_j T_{ji} \cdot b_j, \quad (\text{para todo } b_i \in B \text{ e com } T_{ij} \in \mathbb{K}),$$

existe e é única. Os escalares $\{T_{ij}\}$ são denominados os *componentes do operador linear T com respeito à base B* . Estes componentes podem ser organizados como os elementos de uma matriz quadrada $T_B \doteq [T_{ij}]$, denominada a *matriz do operador T com respeito à base B* .

A definição da matriz associada a uma operador linear é útil porque permite sempre descrever a ação de um operador abstrato sobre vetores de um espaço também abstrato (desde que exista uma base) através de objetos “concretos” que são as matrizes, através da álgebra matricial.

Dados os objetos empregados no teorema 4.8 e na definição 4.27, se o vetor v é decomposto por

$$v = \sum_i v_i \cdot b_i,$$

então o seu vetor imagem $\tilde{v} = T(v)$ é decomposto e expresso por

$$\tilde{v} = T(v) = T\left(\sum_j v_j \cdot b_j\right) = \sum_j v_j \cdot [T(b_j)] = \sum_{i,j} (T_{ij} \cdot v_j) \cdot b_i \equiv \sum_i \tilde{v}_i \cdot b_i.$$

Ou seja, os componentes de \tilde{v} são dados por

$$\tilde{v}_i = \sum_j T_{ij} \cdot v_j.$$

Com isto, se os vetores v e \tilde{v} forem representados em termos de seus componentes na base B através das matrizes coluna

$$v_B \doteq \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix}, \quad \tilde{v}_B \doteq \begin{pmatrix} \tilde{v}_1 \\ \tilde{v}_2 \\ \vdots \end{pmatrix},$$

então os componentes de \tilde{v} são obtidos a partir dos componentes de v por meio da multiplicação matricial

$$\tilde{v}_B = T_B v_B, \quad \text{sendo } T_B \doteq [T_{ij}]_{n,n}$$

a matriz do operador T na base B . No resultado acima, fica subentendido que os produtos entre os elementos da matriz T_B e os elementos da matriz v ocorrem via a operação “ \bullet ” de produto entre elementos do corpo \mathbb{K} .

Dados agora dois operadores, T e S , que atuam sobre \mathcal{V} , a sua composição $ST \doteq S \circ T$ é representada na notação matricial por

$$\begin{aligned} ST(b_i) &= \sum_j S(T_{ji}.b_j) = \sum_j T_{ji} \cdot [S(b_j)] = \sum_{j,k} (S_{kj} \bullet T_{ji}) \cdot b_k \\ &= \sum_k (ST)_{ki} \cdot b_k, \end{aligned}$$

sendo

$$(ST)_{ij} = \sum_k S_{ik} \bullet T_{kj} \implies S_B T_B = [(ST)_{ij}].$$

Ou seja, o elemento ij da matriz que representa a composição dos operadores S e T é exatamente o mesmo elemento da matriz resultante da multiplicação de S_B por T_B .

O teorema a seguir se aplica a mapeamentos lineares em geral.

Teorema 4.26. *Sejam \mathcal{V} e \mathcal{W} espaços vetoriais sobre o corpo \mathbb{K} de dimensões n e m , respectivamente; seja $L : \mathcal{V} \mapsto \mathcal{W}$ um mapa linear atuando entre \mathcal{V} e \mathcal{W} e sejam $B = \{b_1, \dots, b_n\}$ uma base ordenada de \mathcal{V} e $C = \{c_1, \dots, c_m\}$ uma base ordenada de \mathcal{W} . A **matriz correspondente a $L \in \mathbb{L}(\mathcal{V}, \mathcal{W})$ com respeito a B e C , ou a matriz BC** , é dada por*

$$[L]_{BC} = [[L(b_1)]_C [L(b_2)]_C \cdots [L(b_n)]_C],$$

de tal forma que, para todo $v \in \mathcal{V}$, existe $w \in \mathcal{W}$, dado por $w = L(v)$, sendo que seus componentes estão relacionados por

$$w_C = [L]_{BC} v_B.$$

Demonstração. Para todo $b_j \in B$, a decomposição de $L(b_j) \in \mathcal{W}$ com a base C , de acordo com o teorema 4.8, é dada por

$$L(b_j) = \sum_{i=1}^m L_{ij} \cdot c_i \implies [L(b_j)]_C = [L_{1j} L_{2j} \cdots L_{mj}]^T \quad (1 \leq j \leq n),$$

onde $[L(b_j)]_C$ é a matriz coluna dos componentes de $L(b_j)$ com respeito a C , de acordo com a definição 4.26.

Seja agora um vetor $v \in \mathcal{V}$, o qual é decomposto como

$$v = \sum_{i=1}^n v_i \cdot b_i \implies v_B = [v_1 v_2 \cdots v_n]^T.$$

Então, se $w = L(v) \in \mathcal{W}$, resulta

$$w = L(v) = L\left(\sum_{j=1}^n v_j \cdot b_j\right) = \sum_{j=1}^n v_j \cdot L(b_j).$$

Portanto, a matriz que contém os componentes de $L(v)$ com respeito a C pode ser escrita

$$\begin{aligned} [L(v)]_C &= \left[\sum_{j=1}^n v_j \cdot L(b_j) \right]_C = \sum_{j=1}^n v_j \bullet [L(b_j)]_C = v_1 \bullet [L(b_1)]_C + \cdots + v_n \bullet [L(b_n)]_C \\ &= v_1 \bullet \begin{bmatrix} L_{11} \\ L_{21} \\ \vdots \\ L_{m1} \end{bmatrix} + \cdots + v_n \bullet \begin{bmatrix} L_{1n} \\ L_{2n} \\ \vdots \\ L_{mn} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n v_j \bullet L_{1j} \\ \sum_{j=1}^n v_j \bullet L_{2j} \\ \vdots \\ \sum_{j=1}^n v_j \bullet L_{mj} \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{m1} & L_{m2} & \cdots & L_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}. \end{aligned}$$

Ou seja,

$$[L(v)]_C = w_C = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{m1} & L_{m2} & \cdots & L_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = [L]_{BC} v_B,$$

sendo

$$[L]_{BC} = [[L(b_1)]_C [L(b_2)]_C \cdots [L(b_n)]_C] \implies ([L]_{BC})_{ij} = L_{ij} = ([L(b_j)]_C)_i.$$

□

Um corolário do teorema 4.26 leva à matriz de um operador linear T atuando em um espaço vetorial \mathcal{V} considerando somente a base B . Neste caso, a matriz do operador linear $[T]_{BB} \equiv [T]_B \equiv \mathbb{T}_B$ é a *matriz correspondente a T com respeito a B* , ou a *matriz B de T* .

Corolário 4.3. *Seja \mathcal{V} um espaço vetorial sobre \mathbb{K} de dimensão n e $B = \{b_1, \dots, b_n\}$ uma base ordenada de \mathcal{V} . Se $T \in \mathbb{L}(\mathcal{V})$, então a matriz B de T é*

$$\mathbb{T}_B = [[T(b_1)]_B [T(b_2)]_B \cdots [T(b_n)]_B] \iff (\mathbb{T}_B)_{ij} = ([T(b_j)]_B)_i \quad (1 \leq i, j \leq n),$$

de tal forma que $\tilde{v}_B = \mathbb{T}_B v_B$, para todo $v \in \mathcal{V}$.

Demonstração. Dado o enunciado do teorema 4.26, se $\mathcal{W} = \mathcal{V}$ e $C = B$, então para todo operador linear $T : \mathcal{V} \mapsto \mathcal{V}$ resulta a matriz \mathbb{T}_B . □

O resultado do teorema 4.26 também pode ser utilizado para a construção da matriz de transformação de bases.

Corolário 4.4. *Seja \mathcal{V} um espaço de dimensão n e $B = \{b_1, \dots, b_n\}$ e $E = \{e_1, \dots, e_n\}$ duas bases para \mathcal{V} . Se $\mathbf{b}_i^E = [b_{1i}^E \ b_{2i}^E \ \cdots \ b_{ni}^E]^T$ ($1 \leq i \leq n$) for o vetor coluna contendo os componentes de $b_i \in B$ em respeito a E , então a matriz de transformação de bases A_{BE} é dada por*

$$A_{BE} = [b_1^E \ b_2^E \ \cdots \ b_n^E] \iff (A_{BE})_{ij} = (b_j^E)_i = b_{ij}^E,$$

tal que $v_E = A_{BE} v_B$.

Demonstração. Dado o espaço \mathcal{V} e o vetor $b_i \in B$, considera-se o mapeamento identidade (definição 4.14) tal que $I_{\mathcal{V}}(b_i) = b_i$, a matriz de $I_{\mathcal{V}}$ na base E , segundo o teorema 4.26, é

$$[I_{\mathcal{V}}]_{BE} \equiv A_{BE} = [[I_{\mathcal{V}}(b_1)]_E [I_{\mathcal{V}}(b_2)]_E \cdots [I_{\mathcal{V}}(b_n)]_E] = [b_1^E \ b_2^E \ \cdots \ b_n^E].$$

Neste caso, para todo $v \in \mathcal{V}$, o qual possui componentes dados pelas matrizes coluna v_B e v_E , resulta

$$v_E = A_{BE} v_B.$$

□

A representação matricial da ação do operador T também pode ser expressa entre distintas bases, de acordo com o teorema a seguir.

Teorema 4.27. *Seja \mathcal{V} um espaço vetorial finito de dimensão n sobre \mathbb{K} , com $B = \{b_i\}$ e $E = \{e_i\}$ duas bases ordenadas distintas. Seja $T \in \mathbb{L}(\mathcal{V})$ o qual tem representações matriciais \mathbb{T}_B e \mathbb{T}_E sobre as distintas bases. Se $A_{BE} \in \mathbb{K}^{n \times n}$ é a matriz da transformação da base B para a base E , então*

$$\mathbb{T}_E A_{BE} = A_{BE} \mathbb{T}_B = \mathbb{T}_{BE},$$

sendo \mathbb{T}_{BE} a matriz que representa T entre as bases. Ou seja, as matrizes \mathbb{T}_B e \mathbb{T}_E são equivalentes, pois

$$\mathbb{T}_B = A_{BE}^{-1} \mathbb{T}_E A_{BE}.$$

Demonstração. Sejam $v, \tilde{v} \in \mathcal{V}$ tais que $\tilde{v} = T(v)$, então $\tilde{v}_E = \mathbb{T}_E v_E$. Mas $v_E = A_{BE} v_B$; portanto,

$$\tilde{v}_E = \mathbb{T}_E A_{BE} v_B = \mathbb{T}_{BE} v_B.$$

Por outro lado, $\tilde{v}_B = \mathbb{T}_B v_B$. Mas, como $\tilde{v}_B = A_{EB} \tilde{v}_E$, resulta $\tilde{v}_E = A_{BE} \mathbb{T}_B v_B$. Portanto,

$$\mathbb{T}_E A_{BE} = A_{BE} \mathbb{T}_B.$$

□

O teorema a seguir relaciona a matriz de transformação de bases em um espaço com produto interno com matrizes unitárias.

Teorema 4.28. Se B e E são duas bases ortonormais ordenadas para um espaço com produto interno $\langle \mathcal{V}, \langle, \rangle \rangle$ sobre \mathbb{K} , então a matriz de transformação de bases A_{BE} é uma matriz unitária.

Demonstração. Dadas as bases ordenadas ortonormais de \mathcal{V} $B = \{b_1, \dots, b_n\}$ e $E = \{e_1, \dots, e_n\}$, então $\langle b_i, b_j \rangle = \delta_{ij}$ ($1 \leq i, j \leq n$) e, para todos $u, v \in \mathcal{V}$, pelo teorema 4.25, $\langle u, v \rangle = u^\dagger_E v_E$.

Agora, de acordo com o corolário 4.4, a matriz de transformação de bases é

$$A_{BE} = [b_1^E \ b_2^E \ \cdots \ b_n^E] \iff (A_{BE})_{ij} = (b_j^E)_i.$$

Chamando então $A_{BE} \equiv A$ e calculando o elemento $(A^\dagger A)_{ij}$,

$$(A^\dagger A)_{ij} = \sum_{k=1}^n (b_i^{E*})_k (b_j^E)_k = b_i^{E\dagger} b_j^E = \langle b_i, b_j \rangle = \delta_{ij}.$$

Ou seja,

$$A^\dagger A = AA^\dagger = I_n.$$

□

O exemplo a seguir ilustra a ação da matriz de um operador linear atuando sobre um espaço invariante.

Exemplo 4.1 (Matriz de um operador atuando sobre um espaço invariante). Dado um subespaço $\mathcal{U} \subset \mathcal{V}$, de dimensão $\dim \mathcal{U} = m < n = \dim \mathcal{V}$, seja $B_{\mathcal{U}} = \{b_1, b_2, \dots, b_m\}$ uma base de \mathcal{U} . Se \mathcal{U} é invariante frente ao operador L , então

$$L(b_i) = \sum_{j=1}^m L_{ji} b_j, \quad (i = 1, \dots, m),$$

sendo $\{L_{ji}\}$ os elementos da matriz do operador L com respeito à base $B_{\mathcal{V}}$ (definição 4.27). De acordo com o teorema 4.7, a base $B_{\mathcal{U}}$ pode ser estendida para uma base $B_{\mathcal{V}} = \{b_1, \dots, b_n\}$ de \mathcal{V} .

Portanto, se a ação $L(b_i) \in \mathcal{U}$ for descrita na forma matricial,

$$\tilde{b}L = (b_1 \ b_2 \ \cdots \ b_m \ \cdots \ b_n) \begin{pmatrix} L_{11} & L_{12} & \cdots & L_{1m} & \cdots & L_{1n} \\ L_{21} & L_{22} & \cdots & L_{2m} & \cdots & L_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ L_{m1} & L_{m2} & \cdots & L_{mm} & \cdots & L_{mn} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \cdots & L_{nm} & \cdots & L_{nn} \end{pmatrix},$$

isto implica que $L_{ij} = 0$ para $j \leq m$ e $i > m$, isto é, a matriz L deve estar na forma *bloco-diagonal superior*

$$L = \begin{pmatrix} L_1 & & L_3 \\ - & - & - \\ 0 & & L_2 \end{pmatrix},$$

onde L_1 é uma submatriz $m \times m$, a qual contém os elementos de $[L(\mathcal{U})]$ na base $B_{\mathcal{U}}$, enquanto que L_2 e L_3 são submatrizes de tamanhos $(n-m) \times (n-m)$ e $m \times (n-m)$, respectivamente. A submatriz 0 é a matriz nula de tamanho $(n-m) \times m$. Estas submatrizes são elementos essenciais para a discussão a respeito de representações redutíveis realizada no capítulo 5.

Se $\mathcal{U}, \mathcal{W} \subset \mathcal{V}$ forem subespaços complementares de \mathcal{V} ,⁸ i. e., $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, sendo que $\dim(\mathcal{U}) = m$ e $\dim(\mathcal{W}) = n-m$, então a base $B_{\mathcal{V}} = \{b_1, \dots, b_m, b_{m+1}, \dots, b_n\}$ pode ser organizada de tal forma que $B_{\mathcal{U}} = \{b_1, \dots, b_m\}$ e $B_{\mathcal{W}} = \{b_{m+1}, \dots, b_n\}$. Neste caso, se ambos \mathcal{U} e \mathcal{W} forem invariantes frente a um operador L , então a matriz deste terá a forma *bloco-diagonal*

$$L = \begin{pmatrix} L_1 & 0 \\ - & - & - \\ 0 & L_2 \end{pmatrix}.$$

⁸Ver seção 4.2.2.

4.10 AUTOVETORES E AUTOVALORES DE UM OPERADOR LINEAR

Operadores lineares e suas matrizes associadas podem formar problemas de autovalores. Algumas propriedades desses problemas de autovalores serão apresentadas nesta seção. Serão assumidos espaços vetoriais de dimensão finita $n \geq 1$.

Definição 4.28. Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} e $L : \mathcal{V} \mapsto \mathcal{V}$ um operador linear. Um **autovetor** de L é um vetor não nulo $v \in \mathcal{V}$ tal que

$$L(v) = \lambda.v,$$

para algum $\lambda \in \mathbb{K}$, denominado o **autovalor** de L . O conjunto

$$\mathcal{E}_L(\lambda) \doteq \{v \in \mathcal{V} : L(v) = \lambda v\}$$

é denominado o **autoespaço** de L correspondente a λ . Por fim, o símbolo $\sigma(L)$ denota o conjunto de autovalores possuídos pelo operador L .

Nota-se que o escalar $0 \in \mathbb{R}$ ou \mathbb{C} pode ser um autovalor. Por sua vez, o vetor nulo $0 \in \mathcal{V}$ não pode ser um autovalor, mas é elemento de $\mathcal{E}_L(\lambda)$, pois

$$L(0) = 0 = \lambda.0$$

vale para qualquer operador L .

Teorema 4.29. *Seja \mathcal{V} um espaço vetorial. Se $L : \mathcal{V} \mapsto \mathcal{V}$ é um operador linear com autovalor λ , então $\mathcal{E}_L(\lambda)$ é um subespaço de \mathcal{V} .*

Demonstração. Já foi estabelecido que o vetor nulo pertence a $\mathcal{E}_L(\lambda)$. Agora, dados os vetores $u, v \in \mathcal{E}_L(\lambda)$ e o escalar $\mu \in \mathbb{K}$, observa-se que

$$L(u + v) = L(u) + L(v) = \lambda.u + \lambda.v = \lambda.(u + v),$$

e

$$L(\mu.v) = \mu.L(v) = \mu.(\lambda.v) = \lambda.(\mu.v).$$

Portanto, $u + v \in \mathcal{E}_L(\lambda)$ e $\mu.v \in \mathcal{E}_L(\lambda)$. □

Teorema 4.30. *Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} e seja $L \in \mathbb{L}(\mathcal{V})$ um operador linear com autovalores $\lambda_1, \lambda_2 \in \mathbb{K}$ e com correspondentes autovetores $v_1, v_2 \in \mathcal{V}$, respectivamente.*

1. Se $\lambda_1 \neq \lambda_2$, então $v_1 \neq v_2$.
2. $\mathcal{E}_L(\lambda_1) \cap \mathcal{E}_L(\lambda_2) = \{0\}$ se e somente se $\lambda_1 \neq \lambda_2$.

Teorema 4.31. *Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} e seja $L \in \mathbb{L}(\mathcal{V})$ um operador linear com autovalores distintos $\lambda_1, \dots, \lambda_n \in \mathbb{K}$. Se $v_1, \dots, v_n \in \mathcal{V}$ são os correspondentes autovetores de $\lambda_1, \dots, \lambda_n$ respectivamente, então o conjunto $\{v_1, \dots, v_n\}$ é linearmente independente.*

4.10.1 O TEOREMA ESPECTRAL

Operadores lineares em espaços com produto interno com frequência têm associados aos mesmos autovalores e autovetores. Nesta seção serão apresentados alguns teoremas importantes, em particular o teorema espectral.

Teorema 4.32. *Seja $(\mathcal{V}, \langle \cdot, \cdot \rangle)$ um espaço com produto interno sobre o corpo \mathbb{K} de dimensão $n \in \mathbb{N}$ e seja $L \in \mathbb{L}(\mathcal{V})$.*

1. Se $\mathbb{K} = \mathbb{C}$, então L possui um autovetor.
2. Se $\mathbb{K} = \mathbb{R}$ e L é autoadjunto com respeito a algum produto interno em \mathcal{V} , então L possui um autovetor.

Teorema 4.33. *Seja $(\mathcal{V}, \langle, \rangle)$ um espaço com produto interno sobre o corpo \mathbb{K} e seja $L \in \mathbb{L}(\mathcal{V})$ um operador normal.*

1. *Se $v \in \mathcal{V}$ é um autovetor de L com correspondente autovalor $\lambda \in \mathbb{K}$, então v é um autovetor de L^* com autovalor λ^* .*
2. *Se $v_1, v_2 \in \mathcal{V}$ são autovetores de L com correspondentes autovalores $\lambda_1, \lambda_2 \in \mathbb{K}$ tais que $\lambda_1 \neq \lambda_2$, então $v_1 \perp v_2$.*

Teorema 4.34. *Seja $(\mathcal{V}, \langle, \rangle)$ um espaço com produto interno sobre o corpo \mathbb{K} e seja $L \in \mathbb{L}(\mathcal{V})$ um operador autoadjunto.*

1. *Todos autovalores de L são reais.*
2. *Se $v \in \mathcal{E}_L(\lambda)$ e $u \in \mathcal{V}$ é ortogonal a v , então $L(u) \perp v$.*

Demonstração. Parte (1). Seja λ um autovalor de L . Então existe $v \in \mathcal{V}$ tal que $v \neq 0$ e $L(v) = \lambda v$. Como L é autoadjunto, resulta que

$$\langle L(v), v \rangle = \langle v, L(v) \rangle \iff \langle \lambda v, v \rangle = \langle v, \lambda v \rangle \stackrel{\text{def. 4.17}}{\implies} \lambda \langle v, v \rangle = \lambda^* \langle v, v \rangle \stackrel{\langle v, v \rangle \neq 0}{\implies} \lambda = \lambda^*.$$

Parte (2). Seja v um autovetor de L e $u \in \mathcal{V}$ tal que $u \perp v$. Então existe $\lambda \in \mathbb{K}$ tal que $L(v) = \lambda v$ e $\langle u, v \rangle = 0$. Portanto, como L é autoadjunto e pela propriedade (EPI₂),

$$0 = \lambda \langle u, v \rangle = \langle u, L(v) \rangle = \langle L(u), v \rangle,$$

o que mostra que $L(u) \perp v$. □

Teorema 4.35 (Teorema espectral). *Seja $(\mathcal{V}, \langle, \rangle)$ um espaço com produto interno sobre o corpo \mathbb{K} de dimensão $n \in \mathbb{N}$.*

1. *Seja $\mathbb{K} = \mathbb{R}$. Então $L \in \mathbb{L}(\mathcal{V})$ é um operador autoadjunto se e somente se \mathcal{V} possui uma base ortonormal formada pelos autovetores de L .*
2. *Seja $\mathbb{K} = \mathbb{C}$. Então $L \in \mathbb{L}(\mathcal{V})$ é um operador normal se e somente se \mathcal{V} possui uma base ortonormal formada pelos autovetores de L .*

O importante teorema a seguir se refere à versão do teorema espectral para matrizes Hermitianas.

Teorema 4.36. *Se $A \in \mathbb{K}^{m \times n}$ é uma matriz Hermitiana, então existe uma matriz U unitária tal que $U^{-1}AU$ é uma matriz diagonal, cujos elementos são os autovalores de A .*

Demonstração. Seja $(\mathcal{V}, \langle, \rangle)$ um espaço vetorial com produto interno de dimensão n . Seja $L \in \mathbb{L}(\mathcal{V})$ um operador Hermitiano. De acordo o teorema 4.23, o operador é também normal. Então, de acordo com o teorema espectral, o espaço \mathcal{V} possui uma base ortonormal O formada pelos autovetores de L .

Sejam então os autovetores e autovalores de L determinados por $L(x_j) = \lambda_j x_j$ ($j = 1, \dots, n$). De acordo com o teorema 4.21, dada uma base ortonormal B de \mathcal{V} , a matriz de L em relação à base B também é Hermitiana. Sejam então as matrizes coluna x_j^B ($1 \leq j \leq n$) que contém os componentes do autovetor x_j em relação a B . Neste caso, matriz que representa o operador na base B satisfaz

$$[L(x_j)]_B = L_B x_j^B = \lambda_j x_j^B,$$

de acordo com o corolário 4.3, sendo que $L_B \equiv A$ é Hermitiana.

Construindo agora a matriz

$$U = [x_1^B \ x_2^B \ \dots \ x_n^B] \iff (U)_{ij} = (x_j^B)_i,$$

de acordo com o corolário 4.4 a matriz $U = A_{OB}$ é a matriz de transformação de bases ($O \rightarrow B$). Portanto, de acordo com o teorema 4.28, a matriz U é unitária.

Calculando então a transformação de similaridade (4.5) sobre A com U , $U^{-1}AU = U^\dagger AU$, resulta

$$(U^\dagger AU)_{ij} = \sum_{k, \ell=1}^n (U^\dagger)_{ik} (A)_{k\ell} (U)_{\ell j} = \sum_{k, \ell=1}^n (x_i^{B*})_k (A)_{k\ell} (x_j^B)_\ell$$

$$= \lambda_j \sum_{k=1}^n (x_i^{B^*})_k (x_j^B)_k = \lambda_j \langle x_i, x_j \rangle = \lambda_j \delta_{ij},$$

de acordo com o teorema 4.25.

Ou seja, dada a matriz Hermitiana A , existe uma matriz unitária U tal que

$$A' = U^{-1}AU \iff (A')_{ij} = \lambda_i \delta_{ij} \quad (1 \leq i, j \leq n)$$

é diagonal e composta pelos autovalores de A . □

4.11 ESPAÇOS TENSORIAIS

Na seção 4.5 mostrou-se que qualquer espaço vetorial \mathcal{V} possui um espaço dual $\mathcal{V}^* = \text{Hom}(\mathcal{V}, \mathbb{K})$. A partir de \mathcal{V} e \mathcal{V}^* desenvolve-se agora o conceito de uma classe mais geral de espaços denominados *espaços tensoriais*. O conceito de tensores⁹ pode também ser generalizado com base nas definições e propriedades que serão aqui apresentadas.

Antes de se apresentar a definição de um espaço tensorial propriamente dito, serão realizadas as definições de *função característica*, *soma formal finita* e de um *espaço vetorial livre*.

Definição 4.29 (Função característica). Dado um conjunto S , se U é um subconjunto de S , o mapeamento $\chi_U : S \rightarrow \{0, 1\}$, definido por

$$\chi_U(s) = \begin{cases} 1, & s \in U \\ 0, & s \notin U, \end{cases}$$

sendo que $s \in S$, é denominado a *função característica de U* .

Definição 4.30 (Soma formal finita). Dados um conjunto S e um corpo \mathbb{K} , o conjunto de todas as funções

$$\sum_{i=1}^n k_i s_i = k_1 s_1 + k_2 s_2 + \cdots + k_n s_n, \quad (n = 0, 1, 2, \dots) \text{ e onde } k_i \in \mathbb{K}, s_i \in S,$$

é denominado uma *soma formal finita*.

A definição acima é dita “formal” porque se S já for uma estrutura algébrica com operações de adição e/ou produto por escalar definidos, então os produtos e somas na “soma formal” não estão necessariamente relacionados com os primeiros.

Definição 4.31 (Espaço vetorial livre). Dados um conjunto S e um corpo \mathbb{K} , o conjunto de todos os mapeamentos lineares $F(S, \mathbb{K}) : S \rightarrow \mathbb{K}$, que são nulos exceto para um número finito de elementos de S é denominado um *espaço vetorial livre*, identificado por $F(S, \mathbb{K})$.

Como $F(S, \mathbb{K})$ é um mapa linear, então dados $\{f, g\} \subset F(S, \mathbb{K})$, $s \in S$ e $k \in \mathbb{K}$, as condições

$$(f + g)(s) = f(s) + g(s), \quad (kf)(s) = kf(s)$$

são satisfeitas, como exigido para que $F(S, \mathbb{K})$ seja um espaço vetorial. Além disso, tomando-se qualquer conjunto finito $\{s_1, \dots, s_n\} \subset S$ e a função característica $\chi_t \doteq \chi_{\{t\}}$, para $t \in S$, a soma formal finita

$$f = \sum_{i=1}^n f(s_i) \chi_{s_i} = f(s_1) \chi_{s_1} + \cdots + f(s_n) \chi_{s_n}$$

sempre pertencerá a $F(S, \mathbb{K})$.

Para definir o espaço tensorial, toma-se dois espaços vetoriais \mathcal{V} e \mathcal{W} sobre o mesmo corpo \mathbb{K} e define-se um mapeamento bilinear denominado o *produto tensorial* e denotado por \otimes , o qual toma o produto cartesiano dos espaços e o organiza de uma maneira que elimina certas ambiguidades que poderiam surgir. O espaço tensorial assim definido é um tipo de espaço vetorial livre, mas definido de tal forma que produtos cartesianos distintos não possam ser interpretados como equivalentes neste contexto.

⁹Conforme discutido no capítulo 6.

Definição 4.32 (Produto tensorial entre espaços vetoriais). Sejam os espaços vetoriais \mathcal{V} e \mathcal{W} , ambos sobre o mesmo corpo \mathbb{K} . Dados $\{v, v'\} \subset \mathcal{V}$, $\{w, w'\} \subset \mathcal{W}$ e $\{a, b\} \subset \mathbb{K}$, o mapeamento bilinear $\otimes : \mathcal{V} \times \mathcal{W} \mapsto F(\mathcal{V} \times \mathcal{W}, \mathbb{K})$, definido por

$$(a.v + b.v') \otimes w = a.v \otimes w + b.v' \otimes w, \quad v \otimes (a.w + b.w') = a.v \otimes w + b.v \otimes w',$$

de tal forma que as imagens do mapeamento pertencem ao espaço vetorial livre $F(\mathcal{V} \times \mathcal{W}, \mathbb{K})$, é denominado *produto tensorial entre os espaços \mathcal{V} e \mathcal{W}* .

Para a definição do espaço tensorial, um cuidado adicional é necessário pois no produto Cartesiano de \mathcal{V} com \mathcal{W} que constrói os elementos de $F(\mathcal{V} \times \mathcal{W}, \mathbb{K})$, pares tais como $a(v, w)$ e $(a.v, w)$ ou $(v + v', w)$ e $(v, w) + (v', w)$ que são elementos distintos em um espaço livre deverão ser considerados como equivalentes. Para evitar essa ambiguidade, a definição de um espaço tensorial emprega também os conceitos de relação de equivalência e classes de equivalência e de espaços fatores, definidos na seção 3.3.3.

Definição 4.33 (Espaço tensorial). O produto tensorial entre os espaços vetoriais \mathcal{V} e \mathcal{W} sobre o corpo comum \mathbb{K} consiste no espaço fator

$$\mathcal{V} \otimes \mathcal{W} \doteq F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim,$$

onde \sim é a relação de equivalência entre todos os pares ordenados $(v, w) \in F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim$, sendo $\{v, v'\} \in \mathcal{V}$, $\{w, w'\} \in \mathcal{W}$ e $c \in \mathbb{K}$, tal que as seguintes propriedades são satisfeitas:

1. **Identidade.** $(v, w) \sim (v, w)$.
2. **Distributividade.** $(v, w) + (v', w) \sim (v + v', w)$ e $(v, w) + (v, w') \sim (v, w + w')$.
3. **Múltiplos escalares.** $c(v, w) \sim (c.v, w)$ e $c(v, w) \sim (v, c.w)$.

Se os espaços \mathcal{V} e \mathcal{W} são ambos de dimensão finita, sejam as bases $E = \{e_i \mid i = 1, \dots, n\} \subset \mathcal{V}$ e $G = \{g_j \mid j = 1, \dots, m\} \subset \mathcal{W}$. Dados $v = \sum_i v_i \cdot e_i \in \mathcal{V}$ e $w = \sum_j w_j \cdot g_j \in \mathcal{W}$, o produto tensorial $v \otimes w \in F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim$ pode ser decomposto como

$$v \otimes w = \left(\sum_{i=1}^n v_i \cdot e_i \right) \otimes \left(\sum_{j=1}^m w_j \cdot g_j \right) = \sum_{i=1}^n \sum_{j=1}^m v_i \bullet w_j (e_i \otimes g_j).$$

Neste caso, o conjunto $EG \doteq \{e_i \otimes g_j \mid i = 1, \dots, n; j = 1, \dots, m\}$ é uma base do espaço tensorial, o qual possui dimensão $\dim(F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim) = \dim(\mathcal{V}) \dim(\mathcal{W}) = nm$.

Usando o termo **tensor** para denominar um elemento qualquer de um espaço tensorial $F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim$ finito, então, se $T \in F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim$ e EG é uma base de $F(\mathcal{V} \times \mathcal{W}, \mathbb{K}) / \sim$, o tensor T pode ser decomposto como

$$T = \sum_{i=1}^n \sum_{j=1}^m T_{ij} (e_i \otimes g_j),$$

sendo $\{T_{ij}\}$ os componentes do tensor. Se os componentes de T forem obtidos a partir dos vetores v e w , então $T_{ij} = v_i \bullet w_j$.

4.12 ELEMENTOS DE ESPAÇOS MÉTRICOS E TOPOLOGIA

Antes de introduzir estruturas algébricas mais complexas, será necessário realizar algumas definições adicionais, as quais introduzem conceitos que estabelecem relações entre três importantes ramos da matemática: álgebra abstrata, geometria e topologia. Uma definição importante, que será realizada em seguida, introduz o conceito de uma *métrica* ou *função distância* ou simplesmente *distância* entre os diferentes elementos de um conjunto.

Definição 4.34 (Métrica). Dado o conjunto C , uma *métrica em C* é uma função $d : C \times C \mapsto \mathbb{R}$, denominada *função distância* ou, simplesmente, *distância*, tal que, para todos $x, y, z \in C$ as seguintes propriedades são satisfeitas:

1. **Positividade:** $d(x, y) \geq 0$, sendo que se $d(x, y) = 0$, então $x = y$.
2. **Simetria:** $d(x, y) = d(y, x)$.
3. **Inigualdade triangular ou subaditividade:** $d(x, z) \leq d(x, y) + d(y, z)$.

Na sua concepção mais geral, um *espaço métrico* é um conjunto no qual se define a noção de distância, denominada a *métrica do espaço*, entre os seus elementos.

Definição 4.35 (Espaço métrico). Dado um conjunto C e uma métrica $d: C \times C \rightarrow \mathbb{R}$ atuando sobre C , a estrutura $\langle C, d \rangle$ é denominada um *espaço métrico*.

A partir desta definição é possível caracterizar-se um espaço métrico tanto do ponto de vista da geometria quanto da álgebra. Um exemplo de espaço métrico é o espaço Euclidiano,¹⁰ o qual tem sido categorizado desde o período clássico da Grécia antiga e no qual as noções cotidianas e intuitivas de posição e distância estão definidas. A definição aqui apresentada é aquela realizada a partir de estruturas algébricas abstratas.

Definição 4.36 (Sequência de Cauchy). Uma *sequência de Cauchy* é aquela sequência cujos elementos se tornam arbitrariamente próximos entre si à medida que a sequência progride. Alternativamente, uma sequência de Cauchy pode ser definida quando, dada uma quantidade positiva ϵ arbitrariamente pequena, todos os elementos da mesma distam entre si por uma distância menor que ϵ , exceto possivelmente por um subconjunto finito de elementos da sequência.

A partir das definições de espaço métrico e de uma sequência de Cauchy, um outro conceito útil, um *espaço métrico completo*, pode ser definido.

Definição 4.37 (Espaço métrico completo). Um espaço métrico \mathbb{M} é denominado *espaço métrico completo* ou *espaço de Cauchy* se todas as sequências de Cauchy compostas por pontos em \mathbb{M} possuem um limite que também está contido em \mathbb{M} . Em outras palavras, se todas as sequências de Cauchy em \mathbb{M} convergem em \mathbb{M} .

De forma coloquial, um espaço métrico completo é aquele em que não existem “pontos ausentes” no seu interior ou sobre o contorno. Exemplos de espaços métricos completos são os corpos dos números reais (\mathbb{R}) ou dos complexos (\mathbb{C}), nos quais a métrica é definida como o valor absoluto da distância entre os números: $d(x, y) = |x - y|$. Já um espaço métrico não completo é o corpo dos números racionais (\mathbb{Q}) com a mesma métrica, pois a sequência

$$x_1 = 1, \quad x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n} \quad (n = 1, 2, \dots),$$

embora composta por pontos em \mathbb{Q} , converge para $\sqrt{2}$, o qual não pertence ao corpo.

Neste ponto é importante ressaltar que as definições de métrica (definição 4.34) e espaço métrico (definição 4.35) recém realizadas são bastante restritivas e excluiriam diversas estruturas geométricas e algébricas importantes na matemática e física atuais. Um exemplo imediato é o espaço de Minkowski,¹¹ cuja métrica não é positivo-definida e não respeita a condição de positividade. O espaço de Minkowski é um exemplo de um espaço *pseudo-Riemanniano*. Espaços (pseudo-)Riemannianos são uma generalização de espaços métricos quando a métrica é determinada pela forma bilinear¹² (6.26). Em um espaço pseudo-Riemanniano, a métrica não é necessariamente positivo-definida.

As observações e definições abaixo referem-se a conceitos introduzidos nas áreas de geometria e topologia. Estes conceitos não são discutidos em maiores detalhes neste texto, mas são necessários para categorizar de uma forma adequada certas estruturas como grupos contínuos.¹³ As definições apresentadas não são rigorosas, sendo, outrossim, realizadas de uma forma coloquial, porém compreensiva.

Topologia é um ramo da matemática que estuda formas e espaços topológicos (ver definição abaixo), sendo considerada uma extensão da geometria. A topologia estuda as propriedades matemáticas do espaço e como estas são ou não preservadas sob deformações contínuas, tais como o alongamento ou a flexão do espaço. Dentre as propriedades consideradas, estão a conexão, continuidade e os contornos do espaço.

¹⁰Ver exemplo 4.2.

¹¹Discutido na seção 6.15.1.

¹²Definição 4.11.

¹³Seção 5.10.

Definição 4.38 (Topologia). Dado um conjunto C , uma *topologia em C* consiste em uma coleção \mathcal{O} de subconjuntos, denominados os *conjuntos abertos da topologia*, ou simplesmente os *abertos da topologia*, os quais satisfazem as condições:

1. O conjunto vazio \emptyset é aberto e o conjunto C também é aberto. Além disso, $\{\emptyset, C\} \subset \mathcal{O}$.
2. Se U e V são conjuntos abertos, então também o é a sua intersecção $U \cap V$,

$$U \in \mathcal{O} \text{ e } V \in \mathcal{O} \implies U \cap V \in \mathcal{O}.$$

3. Se $\{V_i \mid i \in \mathbb{N}\}$ é uma coleção qualquer de conjuntos abertos, então a sua união $\bigcup_{i \in \mathbb{N}} V_i$ também é um conjunto aberto.

Definição 4.39 (Espaço topológico). Dado um conjunto C e um topologia \mathcal{O} sobre C , a estrutura $\langle C, \mathcal{O} \rangle$ é denominada um *espaço topológico*. Os elementos de C são usualmente denominados de *pontos* e a topologia \mathcal{O} define a *vizinhança* desses pontos.

Pode-se mostrar que todo espaço métrico é um espaço topológico induzido pela sua métrica; contudo, a recíproca não é verdadeira: nem todo espaço topológico deve, necessariamente, possuir uma métrica.

Definição 4.40 (Homeomorfismo). Um *homeomorfismo* ou *isomorfismo topológico* ou ainda uma *função bi-contínua* é uma função contínua entre espaços topológicos que possui uma função inversa também contínua. Homeomorfismos são os isomorfismos (seção 3.6) entre espaços topológicos, i. e., são mapeamentos que preservam todas as propriedades topológicas de um dado espaço.

Definição 4.41 (Variedade). Uma *variedade*¹⁴ é um espaço topológico que se assemelha ao espaço Euclidiano nas vizinhanças de cada ponto. De forma mais precisa, cada ponto de uma variedade n -dimensional possui uma vizinhança que é homeomórfica ao espaço Euclidiano de dimensão n .

Definição 4.42 (Variedade diferenciável). Uma *variedade diferenciável* é um tipo de variedade que é localmente similar a um espaço vetorial para que as operações de cálculo definidas na análise matemática sejam realizáveis.

Será apresentado agora o importante exemplo do espaço Euclidiano como um espaço métrico.

Exemplo 4.2 (O espaço métrico Euclidiano E^n). Um dos espaços métricos mais importantes e empregados na física-matemática é o espaço Euclidiano de n dimensões E^n , o qual contém o espaço de 3 dimensões da geometria euclidiana, ao qual também é atribuído um espaço vetorial normado e dotado de produto interno.¹⁵

Como um espaço métrico, o espaço Euclidiano é um conjunto de pontos que satisfazem certas relações entre si, em particular a métrica (ou distância) entre os mesmos. Esta métrica permite estabelecer-se uma relação geométrica entre os pontos do espaço, usualmente representada pelo *sistema Cartesiano de coordenadas*. Dessa maneira, em uma dimensão, o espaço E^1 é visualizado por meio da reta real; em duas dimensões, o espaço E^2 é visualizado através do plano Cartesiano e em três ou mais dimensões o espaço E^n ($n \geq 3$) é visualizado por um sistema Cartesiano de n coordenadas ortogonais.

Um ponto no espaço E^3 é ilustrado na figura 4.1. Um ponto P neste espaço é dado pela terna ordenada $P \doteq (x_1, x_2, x_3)$ ou, como também costuma-se escrever, $P = (x, y, z)$. Ao se representar este ponto geometricamente em um sistema Cartesiano de 3 coordenadas, neste caso, os números x , y e z são as *coordenadas* do ponto P ao longo de cada eixo do sistema. Estas coordenadas são obtidas traçando-se retas que partem de P e que são perpendiculares a cada eixo, conforme ilustrado na figura.

Para se definir o espaço Euclidiano, é necessário primeiro estabelecer a métrica do mesmo.

Definição 4.43 (Métrica Euclidiana). Seja $\mathbb{R} \doteq \langle \mathbb{R}, +, \times, 0, 1 \rangle$ o corpo dos números reais.¹⁶ Seja

$$\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ vezes}} = \{(x_1, \dots, x_n) \mid x_j \in \mathbb{R}, j = 1, \dots, n\}$$

¹⁴Em inglês: *manifold*.

¹⁵Discutido na seção 4.1.

¹⁶Ver definição na página 124.

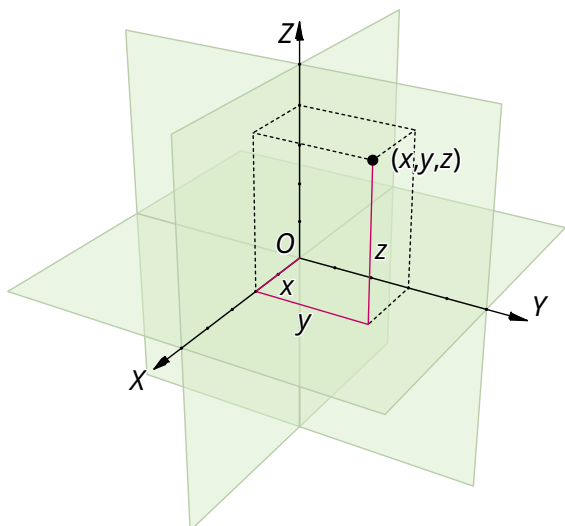


Figura 4.1: No espaço E^3 , todo ponto é determinado pela terna $(x_1, x_2, x_3) \equiv (x, y, z)$ e possui uma representação geométrica em termos de um sistema Cartesiano.

o conjunto de todas as n -uplas ordenadas obtidas a partir do produto Cartesiano¹⁷ do corpo \mathbb{R} por si mesmo. Dados então $p, q \in \mathbb{R}^n$, a métrica ou *distância Euclidiana* entre p e q é a extensão do segmento de reta \overline{pq} conectando ambos os pontos, ou seja,

$$d(p, q) = d(p, q) = \sqrt{\sum_{j=1}^n (q_j - p_j)^2} = \sqrt{(q_1 - p_1)^2 + \cdots + (q_n - p_n)^2}.$$

A partir da definição da métrica Euclidiana, pode-se realizar a definição de um espaço Euclidiano.

Definição 4.44 (Espaço métrico Euclidiano E^n). Seja

$$\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ vezes}} = \{(x_1, \dots, x_n) \mid x_j \in \mathbb{R}, j = 1, \dots, n\}$$

o conjunto de todas as n -uplas ordenadas obtidas a partir do produto Cartesiano do corpo dos números reais. Sejam $d, q \in \mathbb{R}^n$ tais que a distância entre estes pontos é definida pela métrica Euclidiana $d \equiv d(p, q)$. A estrutura $E^n \doteq \langle \mathbb{R}^n, d \rangle$ forma o *Espaço Euclidiano de dimensão n* .

É importante observar aqui que na definição de um espaço métrico em geral, e no espaço E^n em particular, não se define nenhuma outra operação ou relação entre seus elementos (pontos) além da métrica. Por isso, um espaço métrico não é, por exemplo, um espaço vetorial, salvo se este for dotado de estruturas adicionais.

4.13 ESPAÇO VETORIAL MÉTRICO

Um espaço vetorial métrico é um espaço vetorial no qual é definida a noção de métrica, ou distância (ver definição 4.34), entre seus elementos. Um espaço vetorial métrico, como uma estrutura algébrica, é também um tipo de espaço métrico (definição 4.35), visto como uma estrutura geométrica. Este é um exemplo típico de uma estrutura matemática híbrida (algébrica e geométrica).

4.13.1 ESPAÇO VETORIAL MÉTRICO COM NORMA.

Dado um espaço vetorial normado $\langle \mathcal{V}, \|\cdot\| \rangle$, este é um *espaço vetorial métrico normado* quando é possível definir-se uma *métrica sobre \mathcal{V}* pela norma da seguinte maneira: para todos $x, y \in \mathcal{V}$,

$$d(x, y) \doteq \|x - y\| = \|x + (-y)\|,$$

¹⁷Definição 3.27.

sendo $-y \in \mathcal{V}$ o elemento inverso de y em \mathcal{V} . Diz-se então que a métrica d é *induzida* pela norma $\|\cdot\|$.

A métrica acima definida, além de apresentar a propriedade de homogeneidade como consequência da definição de norma, possui também a propriedade de *invariância translacional*, i. e.,

$$d(x+a, y+a) = d(x, y),$$

para todos $x, y, a \in \mathcal{V}$.

De forma recíproca, se a métrica d sobre o espaço \mathcal{V} satisfaz as condições de homogeneidade e invariância translacional, então a norma é *induzida pela métrica* por $\|x\| \doteq d(x, 0)$, $\forall x \in \mathcal{V}$, sendo 0 o vetor nulo.

4.13.2 ESPAÇO DE BANACH.

Um espaço vetorial métrico pode ser completo ou não, de acordo com a definição 4.37. Um espaço vetorial métrico completo é denominado um *espaço de Banach*.

4.13.3 ESPAÇO DE HILBERT

Um espaço de Hilbert é uma generalização tanto de um espaço com produto interno quanto de um espaço Euclidiano, de tal forma que o espaço vetorial pode ter uma dimensão finita arbitrária ou mesmo infinita.

Um *espaço de Hilbert* \mathcal{H} é um espaço vetorial complexo com produto interno, o qual é também um espaço métrico completo, sendo assim um exemplo de um espaço de Banach. No espaço de Hilbert, a métrica do espaço é induzida pelo produto interno, sendo que este é uma forma sesquilinear. Para todos $x, y \in \mathcal{H}$ e $\alpha, \beta \in \mathbb{C}$, o produto interno satisfaz as propriedades de conjugação complexa, distributividade e positividade. Além disso, por ser uma forma sesquilinear:

- $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$ (linear no segundo argumento).
- $\langle \alpha x, y \rangle = \alpha^* \langle x, y \rangle$ (antilinear no primeiro argumento).

A norma e a métrica sobre \mathcal{H} satisfazem as propriedades:

- $\|x\| = \sqrt{\langle x, x \rangle}$.
- $d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}$.

ESPAÇO DE HILBERT DUAL.

A cada espaço de Hilbert \mathcal{H} é possível associar-se um espaço dual \mathcal{H}^* , composto pelos funcionais lineares que atuam sobre os vetores de \mathcal{H} . Como o resultado de um funcional linear deste tipo resulta em um elemento do corpo, este resultado pode ser usado para a implementação do produto interno em \mathcal{H} . Este é o procedimento empregado no desenvolvimento formal da mecânica quântica.

No caso de um espaço de Hilbert, existe uma correspondência bijetiva com o seu espaço dual, garantida pelo *teorema de representação de Fréchet-Riesz*.

Teorema 4.37 (Teorema de representação de Fréchet-Riesz). *Seja \mathcal{H} um espaço de Hilbert e \mathcal{H}^* o seu espaço dual, composto por todos os funcionais lineares contínuos $\varphi : \mathcal{H} \mapsto \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}). Se o vetor $x \in \mathcal{H}$, então existe um e somente um funcional φ_x , dado por*

$$\varphi_x(y) \doteq \langle x, y \rangle \quad (\forall y \in \mathcal{H}),$$

o qual é um elemento de \mathcal{H}^* .

Portanto, a função $\varphi_x(y)$ determina o produto interno entre os vetores $x, y \in \mathcal{H}$. A correspondência entre o vetor $x \in \mathcal{H}$ e o funcional $\varphi_x \in \mathcal{H}^*$ estabelece um mapeamento $\Phi : \mathcal{H} \mapsto \mathcal{H}^*$ o qual é um (anti-) isomorfismo isométrico, ou seja, para todos os vetores $x, y \in \mathcal{H}$ e para todo $\alpha \in \mathbb{K}$:

- Φ é bijetivo.

- Φ preserva a norma: $\|x\| = \|\varphi_x\|$.
- Φ é aditivo: $\Phi(x + y) = \Phi(x) + \Phi(y)$.
- Se $\mathbb{K} = \mathbb{R}$, então $\Phi(\alpha x) = \alpha\Phi(x)$. Se $\mathbb{K} = \mathbb{C}$, então $\Phi(\alpha x) = \alpha^*\Phi(x)$.

4.13.4 ESPAÇO AFIM

Um espaço afim¹⁸ é uma estrutura geométrica que generaliza certas propriedades de linhas paralelas em um espaço Euclidiano (exemplo 4.2) atribuindo assim ao mesmo a estrutura adicional de um espaço vetorial. A definição de um espaço afim serve para definir de forma rigorosa as noções geométricas de “somas vetoriais” como composições de setas direcionadas sobre o plano ou espaço Cartesianos. Em um espaço afim não existe um ponto particular que serve como uma origem. Dessa forma, nenhum vetor possui uma origem fixa nem pode ser associado univocamente a um ponto do espaço. Ao invés disso, um espaço afim é composto por *vetores deslocamento* entre dois pontos do espaço. Esta definição permite atribuir um vetor à subtração de dois pontos do espaço métrico, mas não atribui sentido à soma dos mesmos. Pela mesma concepção, é concebível adicionar-se um vetor a um ponto do espaço afim, tendo como resultado um novo ponto deslocado a partir do primeiro pelo vetor deslocamento.

Definição 4.45 (Espaço afim). Seja um espaço vetorial \mathcal{V} sobre um corpo \mathbb{K} (sendo “ $+$ ” o operador de soma vetorial), juntamente com um conjunto $M \neq \emptyset$, cujos elementos são denominados *pontos*. A estrutura $\langle M, \mathcal{V}, + \rangle$ forma um *espaço afim sobre \mathcal{V}* ou *espaço afim de direção \mathcal{V}* se para todo ponto $p \in M$ e vetor $v \in \mathcal{V}$ existir um mapeamento

$$+ : M \times \mathcal{V} \mapsto M, \quad (v, p) \mapsto v + p$$

que satisfaça as propriedades:

(EA₁) **Identidade à esquerda.** Se $0 \in \mathcal{V}$ for o vetor nulo, para todo $p \in M$, $0 + p = p$.

(EA₂) **Associatividade.** Para todos $v, w \in \mathcal{V}$ e $p \in M$, $v + (w + p) = (v + w) + p$.

(EA₃) **Unicidade.** Para todos $p, q \in M$, existe um único vetor $v \in \mathcal{V}$ tal que $v + p = q$.

O mapeamento “ $+$ ” é uma ação de grupo livre e transitiva¹⁹ de \mathcal{V} sobre M . A ação é livre pela propriedade (EA₁); ao passo que a unicidade (EA₃) assegura que a ação é transitiva, pois, dados $p, q \in M$ quaisquer, se $q = p + u = p + u'$, com $u, u' \in \mathcal{V}$, resulta que $p + u - u' = p$ e, portanto, pela propriedade (EA₁), necessariamente $u' = u$.

Como o grupo $\mathcal{V} \in \mathcal{V}$ é Abeliano, não há diferença entre ações pela esquerda ou pela direita; assim, a identidade $v + p = p + v$ é válida.

Definindo-se uma origem $O \in M$, pode-se transformar o conjunto M em um espaço vetorial. Reciprocamente, qualquer espaço vetorial \mathcal{V} é um espaço afim sobre si próprio.

Denotando-se o vetor único v provido por (EA₃) como \vec{pq} , escreve-se $q = p + \vec{pq}$. De (EA₁) resulta que $\vec{pp} = 0$. Então, de (EA₂) segue que o mapeamento

$$M \times M \mapsto \mathcal{V}, \quad (q, p) \mapsto \vec{pq}$$

satisfaz a equação

$$\vec{pq} +_{\mathcal{V}} \vec{qr} = \vec{pr}.$$

Em particular, este resultado implica que $\vec{pq} = -\vec{qp}$.

A partir da definição de um espaço afim, as seguintes propriedades da geometria analítica podem ser rigorosamente definidas.

Translação. Para cada $v \in \mathcal{V}$, a função

$$\tau_v : M \mapsto M, \quad p \mapsto p + v$$

é uma *translação* do ponto $p \in M$ pelo vetor v . Escrevendo $q = p + v$, então $v \equiv \vec{pq}$ é o vetor que conecta (translada) o ponto p ao ponto $q \in M$.

¹⁸Do inglês: *affine space*.

¹⁹Definição 3.20.

Origem e vetor posição. Seja $O \in M$ um ponto em particular denominado a *origem* do espaço afim. A função

$$\mathcal{V} \mapsto M, \quad v \mapsto O + v$$

é bijetiva, com inversa

$$M \mapsto \mathcal{V}, \quad p \mapsto \overrightarrow{Op}.$$

O vetor \overrightarrow{Op} é denominado o *vetor posição* de p com respeito a O .

Coordenadas e sistema coordenado. Se $\dim(\mathcal{V}) = m$ e se $\{b_1, \dots, b_m\}$ é uma base em \mathcal{V} , existe uma única m -upla $(x_1, \dots, x_m) \in \mathbb{K}^m$ tal que

$$\overrightarrow{Op} = \sum_{j=1}^m x_j b_j. \quad (4.6)$$

Neste caso, os números $x_1, \dots, x_m \in \mathbb{K}$ são denominados as *coordenadas (afins)* do ponto p com respeito ao *sistema de coordenadas* $(O; b_1, \dots, b_m)$. A função bijetiva

$$M \mapsto \mathbb{K}^m, \quad p \mapsto (x_1, \dots, x_m),$$

a qual mapeia cada ponto $p \in M$ a suas coordenadas, é denominada a *função coordenada de M com respeito a $(O; b_1, \dots, b_m)$* .

A dimensão do espaço afim é a dimensão do espaço vetorial que o compõe. A figura 4.2 mostra a construção de um sistema Cartesiano de coordenadas no espaço como um espaço afim de dimensão 3 sobre o corpo \mathbb{R} , no qual a cada ponto P do espaço é atribuído um vetor \overrightarrow{OP} , expresso pela combinação linear (4.6) dos vetores de uma base ortonormal de \mathcal{V} .

IDENTIDADE DE CHASLES E A REGRA DO PARALELOGRAMO. Dados três pontos quaisquer $a, b, c \in M$, conforme visualizado no painel esquerdo da figura 4.3, resulta que $c = a + \overrightarrow{ac}$, $b = a + \overrightarrow{ab}$ e $c = b + \overrightarrow{bc}$. Então, de (EA₂), $c = b + \overrightarrow{bc} = (a + \overrightarrow{ab}) + \overrightarrow{bc} = a + (\overrightarrow{ab} + \overrightarrow{bc})$. E de (EA₃) resulta que

$$c = a + \overrightarrow{ac} = a + (\overrightarrow{ab} + \overrightarrow{bc}) \implies \overrightarrow{ac} = \overrightarrow{ab} + \overrightarrow{bc},$$

a qual é a *identidade de Chasles*. Se $b = a$, resulta de (EA₁), $a = a + \overrightarrow{aa} \implies \overrightarrow{aa} = 0$. Ou, fazendo $c = a$, resulta $\overrightarrow{ab} = -\overrightarrow{ba}$.

Finalmente, dados agora 4 pontos $a, b, c, d \in M$, da identidade de Chasles resulta

$$\overrightarrow{ab} + \overrightarrow{bc} = \overrightarrow{ad} + \overrightarrow{dc} = \overrightarrow{ac},$$

de onde vem a *regra do paralelogramo*

$$\overrightarrow{ab} = \overrightarrow{dc} \text{ se e somente se } \overrightarrow{bc} = \overrightarrow{ad}.$$

Para finalizar esta seção, serão apresentados a seguir alguns exemplos importantes de espaços vetoriais.

Exemplo 4.3 (O espaço vetorial real \mathcal{R}^3). O espaço vetorial mais comum na física é aquele formado pelo grupo \mathcal{R}^3 definido na seção 3.1.2.3 e pelo corpo dos números reais \mathbb{R} , definido na página 124. Esta estrutura é definida como

$$\mathcal{R}^3 \doteq \langle \mathcal{R}^3, \mathbb{R}, +, \cdot \rangle$$

e é corriqueiramente denominada o “espaço Cartesiano” ou o “espaço Euclideo,” embora ambas as designações não sejam rigorosamente corretas, uma vez que um espaço Euclideo é um espaço métrico.²⁰

O produto por escalar neste espaço vetorial é definido da seguinte maneira. Dados o vetor $\mathbf{a} = (a_1, a_2, a_3) \in \mathcal{R}^3$ e o escalar $\alpha \in \mathbb{R}$, o produto de \mathbf{a} por α resulta no vetor $\mathbf{b} \in \mathcal{R}^3$ dado por

$$\mathbf{b} = \alpha \cdot \mathbf{a} = (\alpha a_1, \alpha a_2, \alpha a_3).$$

²⁰Exemplo 4.2.

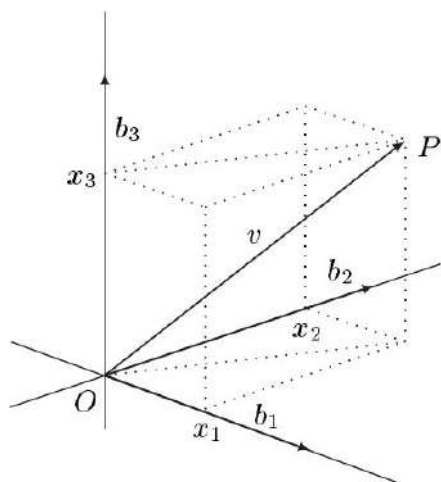


Figura 4.2: O sistema de coordenadas $(O; b_1, b_2, b_3)$ de um espaço afim de dimensão 3 sobre o corpo \mathbb{R} . A base empregada é ortogonal, com $b_j = \overrightarrow{O \|b_j\|}$ ($j = 1, 2, 3$), onde $\|b_j\|$ é a norma do vetor. Observa-se também o vetor $v \equiv \overrightarrow{OP}$ dado por (4.6), bem como as suas coordenadas $\{x_1, x_2, x_3\}$.

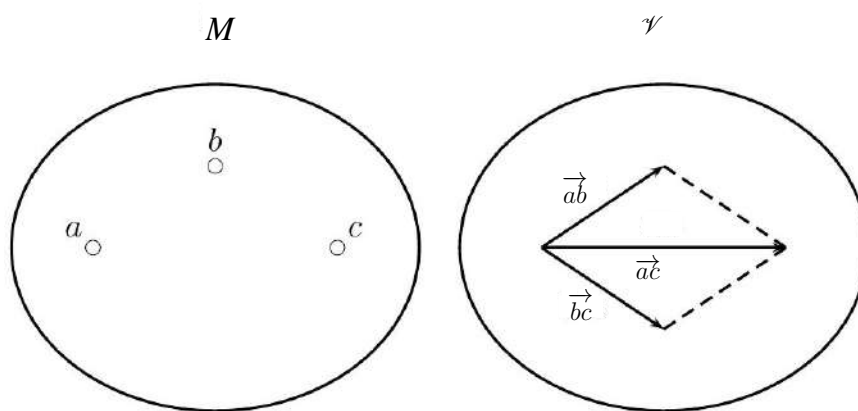


Figura 4.3: Espaço afim $\langle M, \mathcal{V}, + \rangle$ visualizado tanto do ponto de vista do conjunto M (painel esquerdo), com os pontos a, b, c , quanto do ponto de vista do espaço \mathcal{V} (painel direito), com os vetores \vec{ab} , \vec{bc} e \vec{ac} . A regra do paralelogramo é satisfeita no espaço afim.

Verifica-se facilmente que esta operação satisfaz as condições de associatividade, produto pela unidade e distributividade.

Adicionalmente, o espaço \mathcal{R}^3 é sempre considerado com uma ou mais definições de produto interno, sendo o *produto escalar* o mais corriqueiro. O produto escalar do vetor a pelo vetor b é definido como

$$\langle a, b \rangle \equiv a \cdot b \doteq a_1 b_1 + a_2 b_2 + a_3 b_3 \in \mathbb{R}.$$

Verifica-se facilmente que o produto escalar satisfaz os requisitos de um produto interno. Além disso, o produto escalar no \mathcal{R}^3 induz a norma dos vetores do espaço: para todo $a \in \mathcal{R}^3$,

$$\|a\| \equiv |a| \doteq \sqrt{a \cdot a} = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Esta definição em particular da norma do vetor a é denominada de *norma Euclideana no \mathcal{R}^3* .

Em suma, o espaço real \mathcal{R}^3 é um espaço vetorial normado com produto interno. Uma generalização imediata é realizada para o espaço real \mathcal{R}^n , formado pelas n -uplas (a_1, a_2, \dots, a_n) ($n = 1, 2, \dots$).

Exemplo 4.4 (O espaço vetorial complexo \mathcal{C}^3). O espaço vetorial \mathcal{C}^3 é uma generalização do espaço real \mathcal{R}^3 . Este espaço é formado pelo grupo Abelian aditivo \mathcal{C}^3 definido na seção 3.1.2.3 e pelo corpo dos números complexos \mathbb{C} , definido na página 124. Esta estrutura é definida como

$$\mathcal{C}^3 \doteq \langle \mathcal{C}^3, \mathbb{C}, +, \cdot \rangle.$$

O produto por escalar neste espaço vetorial é definido da forma usual. Dados o vetor $a = (a_1, a_2, a_3) \in \mathcal{C}^3$ e o escalar $\alpha \in \mathbb{C}$, o produto de a por α resulta no vetor $b \in \mathcal{C}^3$ dado por

$$b = \alpha \cdot a = (\alpha a_1, \alpha a_2, \alpha a_3).$$

O produto interno no \mathcal{C}^3 é uma forma sesquilinear. Dados $\mathbf{a}, \mathbf{b} \in \mathcal{C}^3$,

$$\langle \mathbf{a}, \mathbf{b} \rangle = a_1^* b_1 + a_2^* b_2 + a_3^* b_3.$$

Claramente, $\langle \mathbf{a}, \mathbf{b} \rangle \in \mathbb{C}$. Finalmente, a norma no \mathcal{C}^3 é a generalização da norma Euclideana. Dado $\mathbf{a} \in \mathcal{C}^3$,

$$\|\mathbf{a}\| = \sqrt{\langle \mathbf{a}, \mathbf{a} \rangle} = \sqrt{|a_1|^2 + |a_2|^2 + |a_3|^2}.$$

Em suma, o espaço vetorial complexo \mathcal{C}^3 é um espaço vetorial normado com produto interno. Uma generalização imediata é realizada para o espaço complexo \mathcal{C}^n , formado pelas n -uplas (a_1, a_2, \dots, a_n) ($n = 1, 2, \dots$).

Exemplo 4.5 (O espaço vetorial Euclidiano \mathbb{E}^3). Neste ponto é possível definir-se o *espaço vetorial Euclidiano \mathbb{E}^3* , também denominado o *espaço real coordenado \mathbb{R}^3* . A definição aqui realizada concorda com as noções usuais empregadas em geometria analítica sobre o plano Euclidiano em 2 dimensões ou o espaço tridimensional da geometria Euclidiana, no qual os pontos do espaço são identificados através de um sistema Cartesiano de coordenadas. Adicionalmente, os vetores neste espaço são representados no sistema Cartesiano como setas orientadas, as quais conectam dois pontos no mesmo: ou a origem O ao ponto P ou o ponto P ao ponto Q . A operação algébrica de adição vetorial é representada no sistema como uma composição das setas orientadas.

A construção parte da definição do espaço (métrico) \mathbb{E}^3 realizada no exemplo 4.2. Em seguida, toma-se a definição do espaço vetorial real \mathcal{R}^3 no exemplo 4.3. Então, de acordo com a definição 4.45, o par $\mathbb{E}^3 \doteq (\mathbb{E}^3, \mathcal{R}^3)$ constitui um espaço (vetorial) afim, no qual é usualmente empregada a base canônica ortonormal do \mathcal{R}^3 formada pelos vetores

$$\hat{e}_1 \equiv \hat{x}_1 \equiv \hat{i} \doteq (1, 0, 0) \quad \hat{e}_2 \equiv \hat{x}_2 \equiv \hat{j} \doteq (0, 1, 0) \quad \hat{e}_3 \equiv \hat{x}_3 \equiv \hat{k} \doteq (0, 0, 1).$$

Sendo agora os pontos do \mathbb{E}^3 $O \doteq (0, 0, 0)$ (a origem), $1_x \equiv (1, 0, 0)$, $1_y \doteq (0, 1, 0)$ e $1_z \doteq (0, 0, 1)$, resulta que os vetores de base do \mathcal{R}^3 são também os vetores posição $\hat{i} = \overrightarrow{O1_x}$, $\hat{j} = \overrightarrow{O1_y}$ e $\hat{k} = \overrightarrow{O1_z}$ do \mathbb{E}^3 . Qualquer outro ponto $P \doteq (p_1, p_2, p_3) \in \mathbb{E}^3$ é localizado no sistema Cartesiano pela seta orientada \overrightarrow{OP} a qual é simplesmente o vetor posição

$$\overrightarrow{OP} \equiv \mathbf{r}_P = \sum_{i=1}^3 p_i \hat{e}_i, \quad (\overrightarrow{OP} \in \mathbb{E}^3).$$

Dado um outro ponto $Q \doteq (q_1, q_2, q_3) \in \mathbb{E}^3$, o vetor deslocamento de P a Q ($\overrightarrow{PQ} \equiv \mathbf{r}_{PQ}$) é

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = \sum_{i=1}^3 (q_i - p_i) \hat{e}_i.$$

Além disso, pela definição do produto escalar entre os vetores \overrightarrow{OQ} e \overrightarrow{OP} no \mathcal{R}^3 ,

$$\langle \overrightarrow{OQ}, \overrightarrow{OP} \rangle \equiv \overrightarrow{OQ} \cdot \overrightarrow{OP} = \sqrt{\sum_{i=1}^3 q_i p_i}.$$

Adicionalmente, pela definição das normas destes vetores e também pela distância entre os pontos P e Q dados pela métrica Cartesiana no \mathbb{E}^3 , observa-se que

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|\overrightarrow{OQ} - \overrightarrow{OP}\| = \sqrt{\sum_{i=1}^3 (q_i - p_i)^2}.$$

Ou seja, as medidas da métrica e da norma no \mathbb{E}^3 concordam entre si. Consequentemente, a norma do \mathcal{R}^3 induz a métrica do \mathbb{E}^3 e vice-versa. Por esta razão, o espaço \mathbb{E}^3 , além de ser um espaço afim, também é um espaço vetorial métrico normado.

Todas as definições e propriedades consideradas neste exemplo são automaticamente generalizadas para o espaço vetorial Euclidiano de n dimensões \mathbb{E}^n , com $n = 1, 2, 3, 4, \dots$

Exemplo 4.6 (A álgebra do espaço \mathcal{R}^3). No exemplo 4.3, mostrou-se que o \mathcal{R}^3 é um espaço vetorial normado com produto interno. As definições contidas nesse exemplo abrangem quase todas as operações algébricas normalmente apresentadas em disciplinas de álgebra linear, cálculo e física básica: (i) adição vetorial, (ii) produto de vetor por escalar e (iii) produto escalar e norma. A única operação que não foi definida no exemplo é aquela denominada *produto vetorial*. Isto porque a definição do produto vetorial no \mathcal{R}^3 promove este a uma álgebra,²¹ sendo que então o conjunto completo de operações é comumente denominado *álgebra vetorial*.

O produto vetorial é definido então como o mapa bilinear $\times : \mathcal{R}^3 \times \mathcal{R}^3 \mapsto \mathcal{R}^3$ tal que, dados $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{R}^3$,

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \implies c_i \doteq \sum_{j,k=1}^3 \epsilon_{ijk} a_j b_k \quad (i = 1, 2, 3), \quad (4.7)$$

onde ϵ_{ijk} é o símbolo de Levi-Civita definido na seção 6.1.2. O produto vetorial satisfaz a condição de bilinearidade, pois para $\alpha, \beta \in \mathbb{R}$,

$$\begin{aligned} [(\alpha\mathbf{a}) + (\beta\mathbf{b})] \times \mathbf{c} &= (\alpha\mathbf{a}) \times \mathbf{c} + (\beta\mathbf{b}) \times \mathbf{c} = \alpha(\mathbf{a} \times \mathbf{c}) + \beta(\mathbf{b} \times \mathbf{c}) \\ \mathbf{c} \times [(\alpha\mathbf{a}) + (\beta\mathbf{b})] &= \alpha(\mathbf{c} \times \mathbf{a}) + \beta(\mathbf{c} \times \mathbf{b}). \end{aligned}$$

Além disso, como é bem sabido, o produto vetorial é anticomutativo: $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$. Como consequência, a álgebra vetorial é não associativa: $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$, pois²²

$$\begin{aligned} [\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \epsilon_{ijk} \epsilon_{klm} a_j b_l c_m \stackrel{(6.3)}{=} a_j c_j b_i - a_j b_j c_i \\ [(\mathbf{a} \times \mathbf{b}) \times \mathbf{c}]_i &= \epsilon_{ijk} \epsilon_{jlm} a_l b_m c_k = a_k c_k b_i - b_k c_k a_i, \end{aligned}$$

ou seja, $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a}$. Contudo, a identidade de Jacobi é satisfeita, *i. e.*,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0.$$

Portanto, talvez de uma forma até surpreendente, a álgebra vetorial é uma *álgebra de Lie*.

REFERÊNCIAS

- ALABISO, C.; WEISS, I. **A Primer on Hilbert Space Theory: Linear Spaces, Topological Spaces, Metric Spaces, Normed Spaces, and Topological Groups**. New York: Springer, 2014. (UNITEXT for Physics). 255 + xvii pp. ISBN 9783319037134. DOI: [10.1007/978-3-319-03713-4](https://doi.org/10.1007/978-3-319-03713-4).
- AMANN, H.; ESCHER, J. **Analysis I**. Basel: Birkhäuser, 2013. (Grundstudium Mathematik). 435 + xv pp. ISBN 9783034877947.
- ARFKEN, G. B.; WEBER, H. J.; HARRIS, F. E. **Mathematical Methods for Physicists: A Comprehensive Guide**. New York: Elsevier, 2013. ISBN 9780123846549.
- BARATA, João Carlos Alves. **Notas para um Curso de Física-Matemática**. São Paulo: USP. Versão de 28/02/2018. Disponível em: http://denebola.if.usp.br/~jbarata/Notas_de_aula/.
- LANG, S. **Undergraduate Algebra**. [S.l.]: Springer New York, 2013. (Undergraduate Texts in Mathematics). ISBN 9781475768985. Disponível em: <https://books.google.com.br/books?id=SqrqBwAAQBAJ>.
- RILEY, K. F.; HOBSON, M. P.; BENCE, S. J. **Mathematical Methods for Physics and Engineering**. Third Ed. Cambridge: Cambridge University Press, 2006. 1333 + xxvii pp. ISBN 9781139450997. Disponível em: <http://books.google.com.br/books?id=Mq1n1EKhNcsC>.
- SCHILLING, A.; LANKHAM, I.; NACHTERGAELE, B. **Linear Algebra as an Introduction to Abstract Mathematics**. [S.l.]: WORLD SCIENTIFIC Publishing Company Incorporated, 2015. ISBN 9789814730358. Disponível em: <https://books.google.com.br/books?id=J3SSjgEACAAJ>.

²¹Ver seção 3.7.3.

²²Usando a convenção de Einstein para somas implícitas (seção 6.1.1).

SZEKERES, Peter. **A Course in Modern Mathematical Physics: Groups, Hilbert Space, and Differential Geometry**. Cambridge: Cambridge University Press, 2004. ISBN 9780521536455. Disponível em: <<http://books.google.com.br/books?id=pRgjkgEACAAJ>>.

TEORIA DE REPRESENTAÇÕES DE GRUPOS

A TEORIA DE REPRESENTAÇÃO é um ramo da matemática que estuda *estruturas algébricas abstratas* através da representação de seus elementos como *transformações lineares de espaços vetoriais* e estuda os *módulos* sobre essas estruturas algébricas abstratas. As diferentes estruturas algébricas abstratas são discutidas na seção 3.7.

De forma resumida, uma *representação* torna uma estrutura algébrica abstrata mais “concreta” através da descrição de seus elementos por matrizes e da conversão das *operações algébricas* entre os elementos das estruturas às operações usuais de soma e multiplicação matriciais. As estruturas algébricas suscetíveis a tais representações incluem grupos, álgebras associativas e álgebras de Lie. A aplicação mais importante da teoria de representação ocorre para grupos abstratos.

A **Teoria de Representações de Grupos** descreve grupos abstratos em termos de transformações lineares de espaços vetoriais; em particular, os elementos do grupo são representados como matrizes e, assim, a operação de multiplicação de grupo se torna a multiplicação matricial usual. Nisto reside a importância da teoria de representações: ela permite que a formulação teórica de sistemas físicos abstratos se reduza às propriedades da álgebra linear, a qual é bem compreendida.

Ao longo do desenvolvimento dos novos modelos físicos que surgiram a partir da formulação teórica da mecânica quântica e da teoria quântica de campos, podem ser destacados vários exemplos nos quais a teoria de grupos e a teoria de representações possibilitaram a compreensão de objetos físicos abstratos e suas interrelações. Um destes exemplos é a proposição do uso das *matrizes de Pauli*,¹ as quais pertencem ao grupo $GL(2, \mathbb{C})$ e apresentam simetria de isospin, formando uma base ao $SU(2)$. Neste caso, tratou-se da construção formal de um modelo matemático ao recém-descoberto spin do elétron, o qual é propriedade física fundamental dos férmions e não possui contrapartida clássica, sendo, neste sentido, uma propriedade “abstrata” das partículas. Posteriormente, com a proposição de uma equação da onda relativística pelo físico inglês Paul Adrien Maurice Dirac (1902–1984), o estado quântico de um elétron passou a ser descrito por um *espinor*, o qual é um elemento de um campo vetorial abstrato, construído de tal forma que satisfaz as isometrias do grupo de Lorentz homogêneo. Posteriormente, ao longo da década de 1960 e com as contribuições principais do físico paquistanês Mohammad Abdus Salam (1926–1996) e dos físicos norte-americanos Sheldon Lee Glashow (1932–) e Steven Weinberg (1933–2021), ocorreu a construção da Lagrangiana da interação eletrofraca, a qual envolve interações entre os campos eletromagnéticos e os bósons vetoriais da interação fraca, e cuja unificação foi possibilitada pela construção do *grupo de gauge* $SU(2) \otimes U(1)$. Ainda mais recente, a construção do *modelo padrão da física de partículas* envolveu o acréscimo do setor da cromodinâmica quântica à Lagrangiana da teoria eletrofraca, bem como de suas interações mútuas. Esta proeza foi possibilitada pela representação do grupo total de isometrias das partículas fundamentais no espaço-tempo de Minkowski: o supergrupo $SU(3) \otimes SU(2) \otimes U(1)$. Estes são apenas alguns exemplos nos quais as representações de grupos foram importantes na exploração das simetrias dos sistemas físicos e na consequente determinação e soluções das equações que descrevem a evolução dos mesmos.

O termo *representação de um grupo* também é empregado de uma forma mais geral ao propor uma “descrição” de um grupo abstrato como um grupo de transformações de algum objeto

¹Geralmente atribuídas ao físico austríaco Wolfgang Pauli (1900–1958).

matemático. De uma maneira mais formal, uma representação é um homomorfismo do grupo abstrato a um grupo de automorfismos do objeto. Se o objeto em questão é um espaço vetorial, a representação é linear. Quando o objeto matemático possui outra natureza, o termo *realização* costuma também ser empregado. Na matemática, a teoria da representação permeia diversas disciplinas tais como a álgebra, a análise, a geometria e topologia e a teoria de números. Tais desenvolvimentos possibilitaram, ao longo do século XX, a criação da *Teoria das Categorias*, a qual unifica a descrição de diferentes estruturas matemáticas oriundas de diversas áreas e de suas interrelações.

5.1 PRIMEIRAS DEFINIÇÕES E REPRESENTAÇÕES

Nesta seção, apresenta-se as definições básicas da teoria de representação de grupos e alguns exemplos imediatos, aplicados aos grupos finitos.

A representação de um determinado grupo abstrato \mathcal{G} é realizada usando o conceito de homomorfismo introduzido na seção 3.6. De uma forma coloquial, pode-se dizer que a representação de um grupo abstrato é semelhante ao que é realizado quando se decompõe um vetor. Se A é um vetor que possui sentido físico e matemático independente do sistema de coordenadas adotado, a “representação prática” deste vetor é obtida escolhendo-se um determinado sistema de coordenadas (Cartesiano, esférico, cilíndrico, etc), o qual possui um conjunto de coordenadas $\{q_i\}$ e de vetores de base $\{\hat{e}_i\}$. Então, pode-se escrever $A = \sum_i \alpha_i \hat{e}_i$, sendo $\{\alpha_i = \alpha_i(q)\}$ as componentes do vetor A neste sistema de coordenadas. A partir deste momento, as informações sobre o vetor A serão fornecidas pelas suas coordenadas.

De forma semelhante, a teoria de representação de grupos toma um grupo abstrato \mathcal{G} , cujos elementos podem não possuir uma descrição matemática conhecida, mas cujas propriedades são idênticas a de um outro grupo D conhecido. Então, a representação de \mathcal{G} consiste na escolha de um “sistema de coordenadas” bem comportado (o grupo D), o qual irá então fornecer as informações desejadas sobre \mathcal{G} .

Para a física, as representações desejadas serão estabelecidas por matrizes não singulares que formam um grupo frente a multiplicação matricial; especificamente, um homomorfismo será estabelecido entre \mathcal{G} e o grupo geral linear. Como a dimensão do grupo \mathcal{G} pode ser tanto uma quantidade finita quanto infinita, a definição mais geral de uma representação deve ser realizada de forma independente da mesma, mencionando somente todos os automorfismos do espaço vetorial considerado. Posteriormente, a definição geral será reduzida aos casos mais usuais de grupos finitos, grupos compactos, etc.

Definição 5.1 (Grupo geral linear de um espaço vetorial). Sendo \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} , o grupo geral linear de \mathcal{V} , denotado por $GL(\mathcal{V})$ ou $Aut(\mathcal{V})$, é o grupo de todos os automorfismos de \mathcal{V} , i. e., o conjunto de todas as transformações lineares bijetivas $\mathcal{V} \rightarrow \mathcal{V}$, em conjunto com a regra de composição funcional como a operação de grupo.

A definição de grupo geral acima, portanto, faz uso do conceito de espaço vetorial (capítulo 4) e de transformação ou mapa linear, definidos na seção 4.4. Se o espaço vetorial \mathcal{V} possui dimensão n finita, então o grupo $GL(\mathcal{V})$ se torna isomórfico ao grupo $GL(n, \mathbb{K})$ definido na seção 3.1.2.3. Contudo, o isomorfismo irá depender da escolha de uma base, como discutido mais adiante.

Definição 5.2 (Representação de grupo (geral)). Uma representação de um grupo \mathcal{G} em um espaço vetorial \mathcal{V} sobre um corpo \mathbb{K} é um homomorfismo de grupo de $\mathcal{G} = \{G; *\}$ ao grupo geral linear $GL(\mathcal{V})$, i. e., consiste no mapeamento

$$\rho : \mathcal{G} \mapsto GL(\mathcal{V}),$$

tal que, para todos os elementos $g_1, g_2 \in \mathcal{G}$,

$$\rho(g_1 * g_2) = \rho(g_1) \rho(g_2),$$

onde o produto $\rho(g_1) \rho(g_2)$ é realizado com o produto do grupo $GL(\mathcal{V})$.

O espaço vetorial \mathcal{V} é denominado o **espaço de representação** e a dimensão do mesmo é a **dimensão da representação**.

Esta definição é completamente geral, por não pressupor que o grupo tenha uma dimensão finita, ou que sua ordem seja finita. A definição a seguir é mais restritiva, porém mais adequada, uma vez que as representações de interesse aplicam-se a esses grupos em particular e serão realizadas por intermédio de matrizes.

Definição 5.3 (Representação de grupo (dimensão finita)). Dado o grupo $\mathcal{G} = \{G; *\}$ e um subgrupo $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$, uma *representação do grupo \mathcal{G}* é um homomorfismo de \mathcal{G} para $GL(n, \mathbb{K})$, i. e.,

$$\rho : \mathcal{G} \longrightarrow GL(n, \mathbb{K}),$$

tal que:

1. Para cada elemento $g \in \mathcal{G}$ existe uma única matriz $D(g) \in D(\mathcal{G})$, a qual é a imagem de g sob o homomorfismo.
2. Para todos $g_1, g_2 \in \mathcal{G}$, com suas respectivas imagens $D(g_1), D(g_2) \in D(\mathcal{G})$,

$$D(g_1 * g_2) = D(g_1) D(g_2). \quad (5.1)$$

O grupo $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$ é a imagem do homomorfismo, i. e., o conjunto de todas as matrizes $D(g), \forall g \in \mathcal{G}$.

O espaço vetorial sobre o corpo \mathbb{K} no qual as matrizes $D \in GL(n, \mathbb{K})$ estão contidas é o **espaço de representação** e n , a ordem de D , é denominado a **dimensão da representação**. Se $D_{ij}(g)$ é o elemento da matriz $D(g)$ na i -ésima linha e na j -ésima coluna, então a multiplicação matricial em (5.1) implica em

$$D_{ij}(g_1 * g_2) = \sum_{k=1}^n D_{ik}(g_1) D_{kj}(g_2), \quad (i, j = 1, \dots, n). \quad (5.2)$$

Todas as propriedades de um homomorfismo que foram discutidas na seção 3.6 continuam sendo válidas para uma representação; em particular, as coincidências entre os elementos identidade e inverso entre os dois grupos, consequências diretas da definição de homomorfismo.

É necessário enfatizar também que a dimensão da representação (a ordem n das matrizes) não corresponde necessariamente à ordem de \mathcal{G} ou à sua dimensão, caso este seja um grupo contínuo. Portanto, em princípio é possível estabelecer homomorfismos entre \mathcal{G} e $GL(n, \mathbb{K})$ para todos $n \geq 1$. Assim, em geral existe mais de uma representação, em diferentes ordens, para o mesmo grupo. Distintas representações do mesmo grupo \mathcal{G} serão distinguidas pelo índice superescrito (μ), onde $\mu = 1, 2, \dots, m$, sendo m o número de diferentes representações de \mathcal{G} . Sendo também n_μ ou $[\mu]$ a dimensão da μ -ésima representação, então esta será identificada por

$$D^{(\mu)}(\mathcal{G}) \subseteq GL(n_\mu, \mathbb{K}), \quad (\mu = 1, 2, \dots, m),$$

enquanto que a imagem de $g \in \mathcal{G}$ nesta particular representação será identificada por

$$D^{(\mu)}(g) \in D^{(\mu)}(\mathcal{G}).$$

Nas seções a seguir serão apresentadas algumas técnicas usuais para se determinar o homomorfismo $\rho : \mathcal{G} \mapsto GL(n, \mathbb{K})$, ou em outras palavras, para se obter as matrizes $D(\mathcal{G}) \subset GL(n, \mathbb{K})$.

5.1.1 VETORES E FUNÇÕES DE BASE E REPRESENTAÇÕES REGULARES

Como o grupo $GL(n, \mathbb{K})$ é composto pelas matrizes não singulares de ordem n pertencentes a um espaço vetorial \mathcal{V} sobre o corpo \mathbb{K} , uma maneira prática de se estabelecer a representação consiste no emprego de **vetores de base**² pertencentes a \mathcal{V} e adequadamente escolhidos. Um vetor de base pode ser escrito como uma matriz coluna $u = (u_1 \ u_2 \ \dots \ u_n)^T \in \mathcal{V}$, sendo seus componentes $\{u_i\}$ ($i = 1, \dots, n$) denominados as **funções de base**, as quais possuem as mais variadas origens (números, coordenadas, rótulos, etc).

²Ressalta-se que um vetor de base não é, necessariamente, um elemento da base do espaço vetorial.

Uma vez determinados os vetores de base, a *representação regular* do grupo $\mathcal{G} = \{G; * \} = \{g_1, g_2, \dots, g_n; *\}$ (de ordem n) é construída a partir do homomorfismo

$$\rho: g\mathcal{G} \mapsto D(g) \subseteq GL(n, \mathbb{K}),$$

i. e., por todas ações de grupo de \mathcal{G} sobre si mesmo.³

Para todos $g \in \mathcal{G}$, a matriz $D(g)$ irá executar a mesma ação sobre as funções de base do vetor u que é realizada pela ação $g\mathcal{G}$, respeitando a condição imposta a um homomorfismo, isto é, a relação (5.1).

Este mapeamento é implementado fazendo uso dos vetores de base da seguinte maneira. Escreve-se inicialmente todos os elementos de \mathcal{G} na forma de uma matriz linha $[\mathcal{G}]$, onde

$$[\mathcal{G}] = (g_1 \ g_2 \ \dots \ g_n).$$

A cada ação de grupo⁴ de \mathcal{G} sobre os elementos de u atribui-se um vetor linha através dos mapeamentos: se $\mathcal{G} \mapsto [\mathcal{G}]$, então, para todo $g_i \in \mathcal{G}$, $g_i\mathcal{G} \mapsto [g_i\mathcal{G}] = (g_i * g_1 \ g_i * g_2 \ \dots \ g_i * g_n)$. Escrevendo $u = [\widetilde{\mathcal{G}}]$, segue que a ação de g_i sobre \mathcal{G} pode ser escrita como uma matriz coluna u_i , a qual é o resultado da ação de uma matriz $M(g_i) \in GL(n, \mathbb{K})$ sobre u , de acordo com

$$u_i = M(g_i) u. \quad (5.3a)$$

Mas, obviamente, $u_i = [\widetilde{g_i\mathcal{G}}]$, de onde se conclui que

$$[g_i\mathcal{G}] = [\mathcal{G}] \widetilde{M(g_i)}.$$

O resultado acima sugere que a representação de g_i em $D(\mathcal{G})$ deve ser

$$D(g_i) = \widetilde{M(g_i)}, \quad (5.3b)$$

sendo que as matrizes $D(g_i)$ satisfazem a condição de homomorfismo (5.1).

De fato, dados $g_1, g_2, g_3 \in \mathcal{G}$ tais que $g_3 = g_1 * g_2$ e a relação (5.3a), as multiplicações sucessivas

$$u_1 = M(g_1) u \longrightarrow M(g_2) u_1 = M(g_2) M(g_1) u$$

devem resultar em $u_3 = M(g_3) u$, sendo $M(g_3) = M(g_2) M(g_1)$, pois isto implica em que $D(g_3) = D(g_1) D(g_2)$, de acordo com (5.1). O exemplo a seguir exemplifica este processo.

Exemplo 5.1 (Representação regular do grupo S_3). Uma representação regular de dimensão 6 para o grupo S_3 é obtida a partir da metodologia descrita acima.

Dado o grupo $S_3 = \{I, \pi_2, \dots, \pi_6; \circ\}$, cujos elementos foram definidos no exemplo 3.2, deseja-se uma representação do mesmo com o grupo $D^{(4)}(S_3) \subset GL(6, \mathbb{R})$. Definindo então o vetor de base $u = (I \ \pi_2 \ \dots \ \pi_6)^T$, em primeiro lugar, obviamente,

$$I \mapsto M(I) = D^{(4)}(I) = I_6, \text{ pois } I_6 u = u,$$

sendo I_6 a matriz identidade de ordem 6. Para os demais elementos, a ação de grupo $\pi_i S_3$ pode ser obtida a partir das linhas da tabela de multiplicação derivada no exercício 3.4. Então, de acordo com (5.3a),

$$\begin{aligned} u_2 = M(\pi_2) u &\rightsquigarrow \begin{pmatrix} \pi_2 \\ I \\ \pi_4 \\ \pi_3 \\ \pi_6 \\ \pi_5 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{pmatrix} \implies M(\pi_2) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \\ u_3 = M(\pi_3) u &\rightsquigarrow \begin{pmatrix} \pi_3 \\ \pi_5 \\ I \\ \pi_6 \\ \pi_2 \\ \pi_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{pmatrix} \implies M(\pi_3) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \\ u_4 = M(\pi_4) u &\rightsquigarrow \begin{pmatrix} \pi_4 \\ \pi_6 \\ \pi_2 \\ \pi_5 \\ I \\ \pi_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{pmatrix} \implies M(\pi_4) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}, \end{aligned}$$

³Ver definição de grupo com operadores na seção 3.7.2.

⁴Definição 3.20.

$$u_5 = M(\pi_5) u \rightsquigarrow \begin{pmatrix} \pi_5 \\ \pi_3 \\ \pi_6 \\ I \\ \pi_4 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{pmatrix} \implies M(\pi_5) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$u_6 = M(\pi_6) u \rightsquigarrow \begin{pmatrix} \pi_6 \\ \pi_4 \\ \pi_5 \\ \pi_2 \\ \pi_3 \\ I \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{pmatrix} \implies M(\pi_6) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Portanto, de acordo com (5.3b), $D^{(4)}(\pi_i) = \widetilde{M}(\pi_i)$, ou seja,

$$D^{(4)}(I) = I_6 \qquad D^{(4)}(\pi_2) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$D^{(4)}(\pi_3) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \qquad D^{(4)}(\pi_4) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$D^{(4)}(\pi_5) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \qquad D^{(4)}(\pi_6) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Pode-se verificar que

I_6	$D^{(4)}(\pi_2)$	$D^{(4)}(\pi_3)$	$D^{(4)}(\pi_4)$	$D^{(4)}(\pi_5)$	$D^{(4)}(\pi_6)$
$D^{(4)}(\pi_2)$	I_6	$D^{(4)}(\pi_4)$	$D^{(4)}(\pi_3)$	$D^{(4)}(\pi_6)$	$D^{(4)}(\pi_5)$
$D^{(4)}(\pi_3)$	$D^{(4)}(\pi_5)$	I_6	$D^{(4)}(\pi_6)$	$D^{(4)}(\pi_2)$	$D^{(4)}(\pi_4)$
$D^{(4)}(\pi_4)$	$D^{(4)}(\pi_6)$	$D^{(4)}(\pi_2)$	$D^{(4)}(\pi_5)$	I_6	$D^{(4)}(\pi_3)$
$D^{(4)}(\pi_5)$	$D^{(4)}(\pi_3)$	$D^{(4)}(\pi_6)$	I_6	$D^{(4)}(\pi_4)$	$D^{(4)}(\pi_2)$
$D^{(4)}(\pi_6)$	$D^{(4)}(\pi_4)$	$D^{(4)}(\pi_5)$	$D^{(4)}(\pi_2)$	$D^{(4)}(\pi_3)$	I_6

a qual é idêntica à tabela de S_3 . Portanto, esta representação é isomórfica.

O exemplo acima obteve uma representação do grupo S_3 que consiste em um isomorfismo do mesmo em $D^{(4)}(S_3) \subset GL(6, \mathbb{R})$. Neste caso, aplica-se a definição a seguir.

Definição 5.4 (Representação fiel). Dada a representação $\rho : \mathcal{G} \mapsto GL(n, \mathbb{K})$, quando esta é isomórfica à sua imagem, diz-se que a representação é *fiel*. Quando esta condição não é satisfeita, a representação é dita *não fiel* ou *degenerada*.

A representação obtida no exemplo 5.1 para o grupo S_3 é fiel. Contudo, uma representação fiel não necessita possuir a dimensão igual à ordem do grupo. Com base no exemplo 5.1, pode-se fazer a seguinte definição para uma representação regular.

Definição 5.5 (Representação regular). Dado um grupo finito $\mathcal{G} = \{G; *\}$ de ordem n , a sua representação regular $\rho : g\mathcal{G} \mapsto D(\mathcal{G}) \subseteq GL(n, \mathbb{R})$ é uma representação fiel tal que para todo $g_k \in \mathcal{G}$ ($k = 1, \dots, n$), os elementos de sua representação serão dados por

$$D_{ij}(g_k) = \delta_{ik,j}, \quad \text{onde } \delta_{ik,j} = \begin{cases} 1, & \text{se } g_i * g_k = g_j \\ 0, & \text{se } g_i * g_k \neq g_j. \end{cases} \quad (5.4)$$

É fácil verificar que a regra (5.4) realmente fornece uma representação. Supondo que os elementos $g_m, g_n, g_l \in \mathcal{G}$ sejam tais que $g_m * g_n = g_l$, então é necessário que $D(g_m)D(g_n) = D(g_l)$. Escrevendo o elemento ij de ambos os lados da equação,

$$\begin{cases} [D(g_m)D(g_n)]_{ij} = \sum_{r=1}^n D_{ir}(g_m)D_{rj}(g_n) \stackrel{(5.4)}{=} \sum_{r=1}^n \delta_{im,r}\delta_{rn,j} \implies \sum_{r=1}^n \delta_{im,r}\delta_{rn,j} = \delta_{il,j}. \\ D_{ij}(g_l) = \delta_{il,j} \end{cases}$$

O resultado desta soma é um dígito binário, para o qual basta que um dos símbolos $\delta_{im,r}$ ou $\delta_{rn,j}$ seja nulo (para todos os índices r) para que o elemento ij resulte igual a zero. Este é o resultado usual. Para o elemento ser unitário, é necessário que ambos os símbolos sejam iguais a um para um dado r . No lado esquerdo da equação isto implica que, simultaneamente, $g_i * g_m = g_r$ e $g_r * g_n = g_j$. Por consequência, $(g_i * g_m) * g_n = g_i * (g_m * g_n) = g_i * g_l = g_j$, o que garante que o elemento do lado direito também é igual a um. Portanto, a representação regular realmente constitui uma representação do grupo \mathcal{G} .

5.1.2 REPRESENTAÇÃO NATURAL

Uma outra maneira de se obter uma representação, particularmente útil quando o grupo é de isometrias, consiste na *representação natural*. Nesta representação, as matrizes $D(\mathcal{G})$ descrevem o que ocorre ao objeto quando se aplica ao mesmo uma transformação de simetria, em relação a um referencial fixo.

O exemplo a seguir obtém uma representação natural $D^{(5)}(S_3)$ de dimensão 3, baseado no isomorfismo de S_3 com o grupo \mathcal{C}_{3v} de isometrias pontuais de um triângulo equilátero.

Exemplo 5.2. Uma representação de dimensão 3 para o grupo S_3 pode ser obtida, baseada no seu isomorfismo com o grupo de operações de isometria de um triângulo equilátero.

Deseja-se construir o grupo $D^{(5)}(S_3) \subset GL(3, \mathbb{R})$. Para tanto, inicialmente retorna-se aos exercícios 3.10 e 3.14, os quais obtiveram o grupo \mathcal{C}_{3v} e o seu isomorfismo com S_3 . Se o vetor de base for definido agora como a sequência dos vértices do triângulo, $u = (a \ b \ c)^T$, então a ação de cada elemento de \mathcal{C}_{3v} sobre o mesmo irá provocar uma alteração nas posições dos vértices, em relação ao referencial fixo determinado pelos números externos ao triângulo, como pode ser visto na figura 3.5a. Estas operações são mapeadas isomorficamente através das multiplicações das matrizes de $D^{(5)}(S_3)$ sobre o vetor de base, empregando-se o isomorfismo entre \mathcal{C}_{3v} e S_3 . Desta maneira,

$$\begin{array}{llll}
 I \mapsto E & \implies & & \implies D^{(5)}(I) = I_3 \\
 \pi_2 \mapsto \sigma_{v16} & \implies & \begin{pmatrix} a \\ c \\ b \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} & \implies D^{(5)}(\pi_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
 \pi_3 \mapsto \sigma_{v35} & \implies & \begin{pmatrix} b \\ a \\ c \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} & \implies D^{(5)}(\pi_3) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
 \pi_4 \mapsto C_3^2 & \implies & \begin{pmatrix} b \\ c \\ a \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} & \implies D^{(5)}(\pi_4) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\
 \pi_5 \mapsto C_3 & \implies & \begin{pmatrix} c \\ a \\ b \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} & \implies D^{(5)}(\pi_5) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\
 \pi_6 \mapsto \sigma_{v24} & \implies & \begin{pmatrix} c \\ b \\ a \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} & \implies D^{(5)}(\pi_6) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}
 \end{array}$$

Verifica-se então que a tabela de multiplicação

I_3	$D^{(5)}(\pi_2)$	$D^{(5)}(\pi_3)$	$D^{(5)}(\pi_4)$	$D^{(5)}(\pi_5)$	$D^{(5)}(\pi_6)$
$D^{(5)}(\pi_2)$	I_3	$D^{(5)}(\pi_4)$	$D^{(5)}(\pi_3)$	$D^{(5)}(\pi_6)$	$D^{(5)}(\pi_5)$
$D^{(5)}(\pi_3)$	$D^{(5)}(\pi_5)$	I_3	$D^{(5)}(\pi_6)$	$D^{(5)}(\pi_2)$	$D^{(5)}(\pi_4)$
$D^{(5)}(\pi_4)$	$D^{(5)}(\pi_6)$	$D^{(5)}(\pi_2)$	$D^{(5)}(\pi_5)$	I_3	$D^{(5)}(\pi_3)$
$D^{(5)}(\pi_5)$	$D^{(5)}(\pi_3)$	$D^{(5)}(\pi_6)$	I_3	$D^{(5)}(\pi_4)$	$D^{(5)}(\pi_2)$
$D^{(5)}(\pi_6)$	$D^{(5)}(\pi_4)$	$D^{(5)}(\pi_5)$	$D^{(5)}(\pi_2)$	$D^{(5)}(\pi_3)$	I_3

é exatamente a tabela do grupo S_3 obtida no exercício 3.4. Esta representação também é um isomorfismo e, portanto, a mesma também é uma representação fiel.

5.2 REPRESENTAÇÕES DE GRUPOS DE TRANSFORMAÇÕES LINEARES

As definições e exemplos das seções anteriores enfatizam o conceito de representação de grupo usualmente empregado na física, consistindo este na descrição das propriedades e evolução de sistemas compostos por partículas e campos e suas interações em termos de operadores que atuam sobre os componentes do sistema e que são descritos por grupos de transformação. Neste contexto, os grupos de representação de interesse, e que são constituídos pelas operações de transformação aplicadas aos sistemas físicos, são usualmente formados por matrizes.

Nesta formulação, o estado de um sistema físico é usualmente expresso como um vetor que pertence a um determinado espaço vetorial. Qualquer transformação atuando sobre esse vetor deve gerar um novo vetor, o qual pertence ao mesmo espaço vetorial e que descreve um outro estado físico do sistema. Os operadores responsáveis pelas possíveis transformações às quais o

sistema está sujeito pertencem a determinados grupos de transformação, os quais podem então ser representados por um grupo de matrizes não singulares contido no $GL(n, \mathbb{K})$.

Nesta seção será discutida a formulação da teoria de representação de grupos usualmente empregada na base formal da mecânica quântica e da teoria de quântica de campos.

5.2.1 ESPAÇOS VETORIAIS E OPERADORES NA MECÂNICA QUÂNTICA

De acordo com a descrição formal da mecânica quântica, proposta inicialmente por Paul Adrien Maurice Dirac (1902 – 1984), a cada estado físico do sistema associa-se um **vetor ket** representado pelo símbolo $|u\rangle$, o qual pertence a um espaço de Hilbert \mathcal{H} sobre o corpo \mathbb{C} .⁵ Associado ao espaço \mathcal{H} , existe também o seu espaço dual \mathcal{H}^* , composto por todos os funcionais lineares $\chi : \mathcal{H} \mapsto \mathbb{C}$. Assim, se $\chi(|u\rangle)$ é uma forma linear aplicada ao vetor $|u\rangle \in \mathcal{H}$, este funcional em particular será um vetor pertencente ao espaço dual \mathcal{H}^* . Estes são denominados **vetores bra**, sendo representados pelo símbolo “ $\langle \cdot |$ ”. Desta maneira, o funcional linear $\chi(|u\rangle)$ define o vetor bra $\langle \chi | \in \mathcal{H}^*$.

Como o mapeamento de um funcional linear é sobrejetivo, o resultado da operação $\chi(|u\rangle)$ irá sempre resultar em um número (usualmente) complexo, o qual é representado pelo símbolo $\langle \chi | u \rangle \doteq \chi(|u\rangle)$ e é denominado um **bracket**. Adicionalmente, para um espaço de Hilbert a correspondência entre o vetor ket $|u\rangle$ e o vetor bra $\langle \chi |$ é única, pois existe um único número $\chi(|u\rangle) = \langle \chi | u \rangle \in \mathbb{C}$ resultante da aplicação da forma linear; ou seja, a função $\chi(|u\rangle)$ é de fato bijetora. Como consequência, a quantidade $\langle \chi | u \rangle$ satisfaz as condições impostas ao produto interno entre dois vetores de um espaço vetorial. Além disso, sempre existe um único vetor ket $|\chi\rangle \in \mathcal{H}$ para o qual o resultado da aplicação $\langle \chi | \chi \rangle = \chi(|\chi\rangle)$ irá resultar em um número que satisfaz as condições impostas à sua norma. Por isso, as operações $\langle \chi | \chi \rangle$ e $\langle \chi | u \rangle$ definem, no espaço de Hilbert, respectivamente a norma do vetor $|\chi\rangle$ e o seu produto interno com o vetor $|u\rangle$.

O mapeamento isomórfico $|\chi\rangle \leftrightarrow \langle \chi |$ é assegurado pelo *teorema de representação de Fréchet-Riesz* (teorema 4.37). Uma das importantes consequências deste teorema é que a dimensão do espaço dual \mathcal{H}^* é igual à dimensão de \mathcal{H} . Vetores bra e ket associados por esta correspondência unívoca são ditos **conjugados** entre si e são rotulados pelo mesmo símbolo. Assim, dado o vetor ket $|u\rangle \in \mathcal{H}$, o seu vetor bra conjugado é representado por $\langle u | \in \mathcal{H}^*$.

Assume-se agora que para um dado ket $|u\rangle$ do espaço vetorial existe um outro ket $|v\rangle$ resultante da ação de um certo operador linear O sobre o primeiro. Esta ação é representada por

$$|v\rangle = O|u\rangle.$$

Na mecânica quântica, o operador O é um agente abstrato, mas que está associado a uma determinada variável dinâmica mensurável, representada por O . Estas quantidades mensuráveis são denominadas **observáveis**. Para que o operador O possa ser associado a um observável, é usualmente assumido que este seja um *operador Hermitiano*.⁶

De acordo com o formalismo da mecânica quântica, cada processo de medida do observável O em um sistema que está inicialmente no estado $|u\rangle$ (suposto normalizado) irá resultar, em geral, em diferentes valores para o observável. Imaginando um *ensemble* de sistemas físicos, isto é um conjunto hipotético de sistemas físicos idênticos entre si, todos inicialmente no mesmo estado físico $|u\rangle$, o conjunto de medidas do observável O irá resultar em uma distribuição de valores, cujo valor médio (denominado *valor esperado*) é dado por

$$\langle O \rangle = \langle u | O | u \rangle.$$

Dada agora uma base ortonormal $\{|n\rangle\}$ do espaço de Hilbert, um sistema físico no estado $|u\rangle$ pode ser descrito⁷ pela decomposição vetorial $|u\rangle = \sum_n \langle n | u \rangle |n\rangle$, onde a quantidade $u_n \doteq \langle n | u \rangle$ é a componente do vetor $|u\rangle$ sobre o vetor de base $|n\rangle$. Esta quantidade pode ser considerada como a n -ésima linha de um vetor coluna. Ou seja, o estado do sistema sempre pode ser descrito em termos da base ortonormal $\{|n\rangle\}$. Por sua vez, o operador O também pode ser escrito em termos desta base como

$$O = \sum_{m,n} \langle m | O | n \rangle |m\rangle \langle n|.$$

⁵Ver seção 4.13.3.

⁶Definição 4.23.

⁷Empregando-se operadores de projeção, os quais não serão discutidos aqui.

Nesta expressão, a quantidade $\langle m|O|n\rangle \doteq O_{mn}$ é o elemento mn da matriz quadrada O que representa o operador O na representação do grupo ao qual o mesmo faz parte.⁸

Portanto, a ação do operador O sobre o sistema no estado $|u\rangle$ leva este ao estado $|v\rangle$ dado por

$$|v\rangle = O|u\rangle = \sum_{m,n} \langle m|O|n\rangle \langle n|u\rangle |m\rangle.$$

Contudo, como $|v\rangle$ sempre pode ser escrito em termos da base $\{|n\rangle\}$ como $|v\rangle = \sum_n \langle n|v\rangle |n\rangle$, resulta que a n -ésima linha do vetor coluna que representa o ket $|v\rangle$ é determinada a partir da matriz O como

$$\langle n|v\rangle = \sum_m \langle n|O|m\rangle \langle m|u\rangle.$$

Por outro lado, se $|u\rangle$ pertence à base do espaço de Hilbert ($|u\rangle = |\ell\rangle \in \{|n\rangle\}$), observa-se que a ação do operador O sobre o mesmo é igual a

$$O|\ell\rangle = \sum_{m,n} \langle m|O|n\rangle \langle n|\ell\rangle |m\rangle = \sum_m \langle m|O|\ell\rangle |m\rangle,$$

ou seja, $O|\ell\rangle$ sempre pode ser escrito em termos da base $\{|n\rangle\}$ através da matriz O e, portanto, pertence ao mesmo espaço de Hilbert.

Finalmente, o valor médio obtido ao ser realizada uma medida da variável dinâmica O quando o sistema está no estado $|u\rangle$ resulta ser igual a

$$\langle O\rangle = \langle u|O|u\rangle = \sum_m \sum_n \langle m|O|n\rangle \langle m|u\rangle^* \langle n|u\rangle.$$

Ou seja, a matriz O , a qual é a representação do operador O , está diretamente associada ao processo de medida da quantidade física O . Este resultado mostra a importância da teoria de representação de grupos para a mecânica quântica.

5.2.2 ESPAÇOS VETORIAIS E SUAS REPRESENTAÇÕES

Os argumentos da seção anterior podem ser apresentados novamente de uma maneira ligeiramente mais abstrata, para ilustrar suas aplicações a outros formalismos empregados na física. Dado um grupo de transformações $\mathcal{G} = \{G; *\}$, cuja representação é dada pelo grupo de matrizes $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$, as transformações que compõe \mathcal{G} atuam sobre um sistema físico cujo estado pode ser descrito por um vetor pertencente a um espaço vetorial complexo \mathcal{V}^n , de dimensão n , no qual estão definidos tanto o produto interno quanto a norma dos vetores. Para certos sistemas físicos, é necessário que este seja um espaço métrico completo de dimensão arbitrária, ou seja, um espaço de Hilbert.

O estado do sistema físico é suposto completamente descrito por um vetor $u \in \mathcal{V}^n$. Dada uma base ortonormal $\Psi \doteq \{\psi_i\}$ em \mathcal{V}^n , o vetor u pode ser desenvolvido em termos dos vetores de base como

$$u = \sum_{i=1}^n u_i \psi_i,$$

onde $\{u_i\}$ são os componentes de u na base Ψ . Estes componentes são obtidos então a partir do produto interno:

$$u_i = \langle \psi_i, u \rangle.$$

Observa-se aqui que a escolha alternativa $u_i = \langle u, \psi_i \rangle$ também poderia ser feita, uma vez que ambas estão relacionadas por $\langle u, \psi_i \rangle = \langle \psi_i, u \rangle^*$. Como $\langle u, u \rangle \geq 0$, resulta a relação de Parseval

$$\sum_i |\langle \psi_i, u \rangle|^2 = \langle u, u \rangle.$$

A expressão obtida acima para a componente u_i permite realizar uma correspondência entre o vetor abstrato u com a matriz coluna $u = (u_1 \ u_2 \ \dots \ u_n)^T$, a qual pode ser considerada uma representação matricial para o vetor na base Ψ .

⁸Ver seção 4.9.4.2.

Dado agora um operador linear $A \in \mathcal{G}$, a ação deste sobre o vetor u consiste em um mapeamento linear $A : \mathcal{V}^n \mapsto \mathcal{V}^n$ que irá associar u com outro vetor $v \in \mathcal{V}^n$ através da relação

$$v = Au.$$

Quando $u = \psi_i \in \Psi$, ou seja, um vetor da base, resulta $v = A\psi_i$. Mas o vetor v , por sua vez, também pode ser expresso em termos da base Ψ como $v = \sum_j \langle \psi_j, v \rangle \psi_j$, ou seja,

$$\begin{aligned} v = A\psi_i &= \sum_j \langle \psi_j, A\psi_i \rangle \psi_j \\ &= \sum_j D_{ji}(A) \psi_j, \end{aligned}$$

onde $D_{ji}(A) \doteq \langle \psi_j, A\psi_i \rangle$. Isto implica em que ao operador linear abstrato A é possível atribuir-se uma *representação matricial na base Ψ* através da associação

$$A \longleftrightarrow D(A),$$

onde $D(A)$ é uma matriz $n \times n$ cujos elementos são $D_{ij}(A)$ ($i, j = 1, \dots, n$).

Para verificar que o conjunto de matrizes $D(\mathcal{G}) \doteq \{D(A)\} (\forall A \in \mathcal{G})$ realmente forma uma representação de \mathcal{G} , considera-se a aplicação consecutiva dos operadores $A, B \in \mathcal{G}$ sobre o vetor de base $\psi_i \in \Psi$. Por um lado, esta operação resulta em

$$(A * B) \psi_i = A(B\psi_i) = A \left(\sum_k \psi_k D_{ki}(B) \right) = \sum_k (A\psi_k) D_{ki}(B) = \sum_{j,k} \psi_j D_{jk}(A) D_{ki}(B),$$

enquanto que, por outro lado,

$$(A * B) \psi_i = \sum_j \psi_j D_{ji}(A * B).$$

Ou seja,

$$D_{ij}(A * B) = \sum_k D_{ik}(A) D_{kj}(B) \implies D(A * B) = D(A) D(B),$$

satisfazendo as condições (5.1) e (5.2).

Em resumo, mostrou-se que mediante a escolha de uma base ortonormal adequada, tanto os vetores quanto os operadores de um espaço vetorial podem ser representados por meio de matrizes. Esta argumentação ilustra a importância tanto da teoria de representação de grupo para a física quanto a preferência para a mesma de grupos de representações compostos por matrizes quadradas pertencentes ao grupo geral linear.

5.3 REPRESENTAÇÕES EQUIVALENTES E CARACTERES

Como foi visto nas seções anteriores, a representação de um determinado grupo pode, em princípio, ser realizada por mais de um grupo de representações de diferentes dimensões, dependendo da escolha feita para o conjunto de vetores de base. Por outro lado, para uma dada dimensão da representação, a qual foi obtida com uma determinada base do espaço vetorial, pode sempre ser realizada uma mudança de base,⁹ com a qual obtém-se uma outra representação de mesma dimensão. As duas representações assim obtidas são denominadas equivalentes.

Definição 5.6 (Representações equivalentes). Dado o grupo $\mathcal{G} = \{G; *\}$, duas representações $D(\mathcal{G})$ e $D'(\mathcal{G})$, ambas de dimensão n , i. e., $D(\mathcal{G}), D'(\mathcal{G}) \subseteq GL(n, \mathbb{K})$, são ditas *representações equivalentes* se para todo $g \in \mathcal{G}$ existe uma matriz não singular $S \in GL(n, \mathbb{K})$ tal que

$$D'(g) = S^{-1} D(g) S. \tag{5.5}$$

Esta transformação é denominada *transformação de similaridade*.

A transformação de similaridade entre duas matrizes já havia sido definida por (4.5). É possível verificar que a matriz $D'(g)$ obtida a partir da definição acima realmente constitui uma representação de g , de acordo com a definição 5.3:

⁹Ver seção 4.3.1.

1. **Clausura:** Dados $g_1, g_2 \in \mathcal{G}$ e suas respectivas representações $D(g_1), D(g_2) \in D(\mathcal{G})$,

$$\begin{aligned} D'(g_1) D'(g_2) &= S^{-1} D(g_1) S S^{-1} D(g_2) S \\ &= S^{-1} D(g_1) D(g_2) S = S^{-1} D(g_1 * g_2) S = D'(g_1 * g_2). \end{aligned}$$

2. **Unicidade da representação do elemento unitário:** como $D(I) = I_n$,

$$D'(I) = S^{-1} D(I) S = S^{-1} I_n S = I_n.$$

3. **Unicidade da representação do elemento inverso:** se $g^{-1} \in \mathcal{G}$ é o elemento inverso de g , então $D(g^{-1}) = D^{-1}(g)$ e

$$D'(g^{-1}) = S^{-1} D(g^{-1}) S = S^{-1} D^{-1}(g) S = [S^{-1} D(g) S]^{-1} = [D'(g)]^{-1}.$$

Portanto, uma dada representação de dimensão n em geral não é única. Deseja-se então obter representações em particular ou quantidades derivadas das representações que sejam invariantes frente a uma transformação de similaridade. Este problema será abordado na próxima seção. Uma quantidade relacionada à representação e que é invariante será definida agora.

Definição 5.7 (Caractere da representação). Seja \mathcal{V} um espaço vetorial de dimensão finita sobre o corpo \mathbb{K} , seja $\mathcal{G} = \{G; *\}$ um grupo finito e seja $\rho : \mathcal{G} \rightarrow GL(n, \mathbb{K})$ uma representação de \mathcal{G} em \mathcal{V} . O *caractere da representação* ρ é o mapeamento

$$\chi_\rho : \mathcal{G} \rightarrow \mathbb{K}$$

tal que para todo $g \in \mathcal{G}$ com representação $D(g) \in GL(n, \mathbb{K})$,

$$\chi_\rho(g) = \text{Tr}[D(g)] = \sum_{i=1}^n D_{ii}(g), \quad (5.6)$$

sendo $\chi_\rho(g)$ denominado o *caractere do elemento g na representação ρ* .

Verifica-se facilmente que o caractere de uma representação é uma quantidade invariante frente a uma transformação de similaridade. Dados $g \in \mathcal{G}$ e $D(g), D'(g) \in GL(n, \mathbb{K})$ duas representações equivalentes obtidas pela transformação (5.5), então

$$\begin{aligned} \chi_{\rho'}(g) &= \sum_{i=1}^n D'_{ii}(g) = \sum_{k=1}^n \sum_{m=1}^n D_{km}(g) \left[\sum_{i=1}^n S_{mi}(S^{-1})_{ik} \right] \\ &= \sum_{k=1}^n \sum_{m=1}^n D_{km}(g) \delta_{mk} = \sum_{k=1}^n D_{kk}(g) = \chi_\rho(g). \end{aligned}$$

A mesma demonstração pode ser realizada de uma forma mais compacta empregando a seguinte propriedade do traço de uma matriz,

$$\chi_{\rho'}(g) = \text{Tr}[D'(g)] = \text{Tr}[S^{-1} D(g) S] = \text{Tr}[D(g) S^{-1} S] = \text{Tr}[D(g)] = \chi_\rho(g).$$

Observa-se portanto que o caractere da representação depende somente da dimensão da mesma, e não da base adotada.

5.4 SOMA E PRODUTO DIRETOS DE MATRIZES E REPRESENTAÇÕES

Nas seções anteriores verificou-se que existe, em geral, mais de uma representação para um determinado grupo. Essas distintas representações podem ser de diferentes dimensões ou, caso tenham a mesma dimensão, distinguem-se por serem equivalentes ou não equivalentes. Representações de dimensão $n > 1$ podem também ser construídas a partir de representações de dimensões mais baixas através das técnicas de *soma direta* ou *produto direto de representações*.

5.4.2 SOMA DIRETA DE REPRESENTAÇÕES

A questão que se coloca agora é a seguinte: a soma direta de matrizes pode ser empregada para gerar novas representações? Dado o grupo $\mathcal{G} = \{G; *\}$ e duas representações conhecidas, $D^{(1)}(\mathcal{G}) \subseteq GL(n_1, \mathbb{K})$ e $D^{(2)}(\mathcal{G}) \subseteq GL(n_2, \mathbb{K})$, a soma direta

$$D^{(3)}(\mathcal{G}) \doteq D^{(1)}(\mathcal{G}) \oplus D^{(2)}(\mathcal{G}),$$

com $D^{(3)}(\mathcal{G}) \subseteq GL(n_1 + n_2, \mathbb{K})$, será também uma representação de \mathcal{G}

Dados $g_1, g_2, g_3 \in \mathcal{G}$, de acordo com a definição 5.3 e as expressões (5.1) e (5.2), se $D^{(3)}(g_1)$, $D^{(3)}(g_2)$ e $D^{(3)}(g_3)$ são representações, então

$$D^{(3)}(g_3) = D^{(3)}(g_1 * g_2) = D^{(3)}(g_1) D^{(3)}(g_2),$$

ou seja,

$$D_{ij}^{(3)}(g_3) = \sum_{k=1}^{n_1+n_2} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2), \quad (i, j = 1, \dots, n_1 + n_2).$$

Mas,

$$D_{ij}^{(3)}(g_i) = \begin{cases} D_{ij}^{(1)}(g_i), & i \leq n_1, j \leq n_1 \\ D_{(i-n_1)(j-n_1)}^{(2)}(g_i), & n_1 + 1 \leq i \leq n_1 + n_2, n_1 + 1 \leq j \leq n_1 + n_2 \\ 0, & \text{no restante.} \end{cases}$$

Portanto, é possível realizar a seguinte partição:

$$\begin{aligned} \binom{1 \leq i \leq n_1}{1 \leq j \leq n_1} D_{ij}^{(3)}(g_3) &= \sum_{k=1}^{n_1} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) + \sum_{k=n_1+1}^{n_1+n_2} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) \\ \binom{1 \leq i \leq n_1}{n_1+1 \leq j \leq n_1+n_2} D_{ij}^{(3)}(g_3) &= \sum_{k=1}^{n_1} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) + \sum_{k=n_1+1}^{n_1+n_2} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) \\ \binom{n_1+1 \leq i \leq n_1+n_2}{1 \leq j \leq n_1} D_{ij}^{(3)}(g_3) &= \sum_{k=1}^{n_1} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) + \sum_{k=n_1+1}^{n_1+n_2} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) \\ \binom{n_1+1 \leq i \leq n_1+n_2}{n_1+1 \leq j \leq n_1+n_2} D_{ij}^{(3)}(g_3) &= \sum_{k=1}^{n_1} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2) + \sum_{k=n_1+1}^{n_1+n_2} D_{ik}^{(3)}(g_1) D_{kj}^{(3)}(g_2), \end{aligned}$$

ou seja,

$$\begin{aligned} \binom{1 \leq i \leq n_1}{1 \leq j \leq n_1} D_{ij}^{(3)}(g_3) &= \sum_{k=1}^{n_1} D_{ik}^{(1)}(g_1) D_{kj}^{(1)}(g_2) = D_{ij}^{(1)}(g_3) \\ \binom{1 \leq i \leq n_1}{n_1+1 \leq j \leq n_1+n_2} D_{ij}^{(3)}(g_3) &= 0 \\ \binom{n_1+1 \leq i \leq n_1+n_2}{1 \leq j \leq n_1} D_{ij}^{(3)}(g_3) &= 0 \\ \binom{n_1+1 \leq i \leq n_1+n_2}{n_1+1 \leq j \leq n_1+n_2} D_{ij}^{(3)}(g_3) &= \sum_{k=n_1+1}^{n_1+n_2} D_{(i-n_1)(k-n_1)}^{(2)}(g_1) D_{(k-n_1)(j-n_1)}^{(2)}(g_2) = D_{(i-n_1)(j-n_1)}^{(2)}(g_3), \end{aligned}$$

o que satisfaz a condição de clausura da representação.

Desta forma, mostrou-se que para todo $g \in \mathcal{G}$,

$$D^{(3)}(g) = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & D^{(1)}(g) & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & D^{(2)}(g) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Em particular,

$$D^{(3)}(I) = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & D^{(1)}(I) & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & D^{(2)}(I) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & I_{n_1} & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots & I_{n_2} & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} = I_{n_1+n_2}.$$

Observa-se também que, como a soma direta não é comutativa, uma outra representação $D^{(4)}(\mathcal{G})$, distinta de $D^{(3)}(\mathcal{G})$, pode ser obtida através de $D^{(2)}(\mathcal{G}) \oplus D^{(1)}(\mathcal{G})$.

De uma forma geral, a partir de N representações distintas do grupo \mathcal{G} , identificadas por $D^{(1)}(\mathcal{G}), \dots, D^{(N)}(\mathcal{G})$ e de dimensões n_1, \dots, n_N , respectivamente, é sempre possível para todo $g \in \mathcal{G}$ construir uma representação $D(g)$ de dimensão $n_1 + \dots + n_N$ pela soma direta

$$D(g) = \bigoplus_{\mu=1}^N D^{(\mu)}(g) = D^{(1)}(g) \oplus D^{(2)}(g) \oplus \dots \oplus D^{(N)}(g), \quad (5.9a)$$

cujas matrizes bloco-diagonais são

$$D(g) = \begin{pmatrix} \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & D^{(1)}(g) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} \\ \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & D^{(2)}(g) & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} \\ \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} \\ \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} & \begin{matrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{matrix} \end{pmatrix}.$$

Além disso, se $\chi(g) = \text{Tr}[D(g)]$ é o caractere de $g \in \mathcal{G}$ e $\chi_\mu(g) = \text{Tr}[D^{(\mu)}(g)]$ é o caractere de g na μ -ésima representação ($\mu = 1, \dots, N$), então a propriedade (5.8b) implica que

$$\chi(g) = \sum_{\mu=1}^N \chi_\mu(g) = \chi_1(g) + \dots + \chi_N(g). \quad (5.9b)$$

O artifício da soma direta de matrizes será explorado na seção 5.5 para a construção dos conceitos de representações redutíveis ou irredutíveis.

5.4.3 PRODUTO DIRETO DE MATRIZES

Dadas novamente a matriz A de tamanho $m \times n$ e a matriz B de tamanho $p \times q$, o *produto direto* ou o *produto de Kronecker* de ambas, executado pela operação binária representada pelo símbolo “ \otimes ”, resulta em uma matriz C de tamanho $mp \times nq$ tal que

$$C = A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix},$$

sendo que cada “elemento” $a_{ij}B$ consiste na verdade na matriz de tamanho $p \times q$ dada por

$$a_{ij}B = \begin{pmatrix} a_{ij}b_{11} & a_{ij}b_{12} & \cdots & a_{ij}b_{1q} \\ a_{ij}b_{21} & a_{ij}b_{22} & \cdots & a_{ij}b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{ij}b_{p1} & a_{ij}b_{p2} & \cdots & a_{ij}b_{pq} \end{pmatrix}.$$

Para se obter um elemento de C em termos dos elementos de A e B , pode-se usar a notação $c_{ik,j\ell}$, onde uma linha de C é denotada pelo símbolo dual “ ik ” e uma coluna por “ $j\ell$ ”, de tal forma que

$$c_{ik,j\ell} = a_{ij}b_{k\ell}, \quad (1 \leq i \leq m, 1 \leq j \leq n, 1 \leq k \leq p, 1 \leq \ell \leq q).$$

Como um exemplo, o produto direto das matrizes

$$A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} h & r \\ s & t \\ u & v \end{pmatrix}$$

é a matriz 6×6

$$C = A \otimes B = \begin{pmatrix} aB & bB & cB \\ dB & eB & fB \end{pmatrix} = \begin{pmatrix} ah & ar & bh & br & ch & cr \\ as & at & bs & bt & cs & ct \\ au & av & bu & bv & cu & cv \\ dh & dr & eh & er & fh & fr \\ ds & dt & es & et & fs & ft \\ du & dv & eu & ev & fu & fv \end{pmatrix}.$$

Se $\{A_i\}$ ($i = 1, \dots, p$) é um conjunto de matrizes, então o produto direto das mesmas pode ser escrito como

$$B = A_1 \otimes A_2 \otimes \dots \otimes A_p \equiv \bigotimes_{i=1}^p A_i. \quad (5.10)$$

Propriedade 5.2 (Produto direto de matrizes). O produto direto de matrizes possui as seguintes propriedades:

- O produto direto é associativo: $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.

- O produto direto não é comutativo: $A \otimes B \neq B \otimes A$.

- O produto direto é distributivo com respeito à soma matricial:

$$A \otimes (B + C) = A \otimes B + A \otimes C.$$

- Distributividade com respeito à soma direta:

$$(A \oplus B) \otimes C = A \otimes C \oplus B \otimes C.$$

- Dadas as matrizes A_1, A_2, B_1 e B_2 , se os produtos matriciais $A_1 A_2$ e $B_1 B_2$ existem, então

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2). \quad (5.11a)$$

- Devido à propriedade acima,

$$(AB)^{[p]} = (A)^{[p]} (B)^{[p]},$$

onde

$$A^{[p]} = \bigotimes_{i=1}^p A = \underbrace{A \otimes A \otimes \dots \otimes A}_{p \text{ vezes}}.$$

- A inversa do produto direto é o produto direto das inversas, caso estas existam:

$$B^{-1} = \left[\bigotimes_{i=1}^p A_i \right]^{-1} = \bigotimes_{i=1}^p A_i^{-1}.$$

- O Hermitiano conjugado do produto direto é o produto direto dos Hermitianos conjugados:

$$B^\dagger = \left[\bigotimes_{i=1}^p A_i \right]^\dagger = \bigotimes_{i=1}^p A_i^\dagger.$$

- Seja A uma matriz $m \times m$ e B uma matriz $p \times p$. O determinante do produto direto das mesmas é dado por:

$$\det(A \otimes B) = [\det(A)]^p [\det(B)]^m.$$

- Se a matriz B é dada por (5.10), então

$$\text{Tr}(B) = \prod_{i=1}^p \text{Tr}(A_i) = \text{Tr}(A_1) \text{Tr}(A_2) \dots \text{Tr}(A_p). \quad (5.11b)$$

- Se A e B são matrizes quadradas com autovalores e autovetores $\{\lambda_i\}, \{x_i\}$ e $\{\mu_j\}, \{y_j\}$, respectivamente, os autovalores de $A \otimes B$ são $\{\lambda_i \mu_j\}$ e seus autovetores são $\{x_i \otimes y_j\}$, i. e., se $Ax_i = \lambda_i x_i$ e $By_j = \mu_j y_j$, então

$$(A \otimes B)(x_i \otimes y_j) = \lambda_i \mu_j (x_i \otimes y_j).$$

5.4.4 PRODUTO DIRETO DE REPRESENTAÇÕES

Dado agora o grupo $\mathcal{G} = \{G; *\}$ e duas representações conhecidas, $D^{(\mu)}(\mathcal{G})$ e $D^{(\nu)}(\mathcal{G})$, considere-se, para cada elemento $g \in \mathcal{G}$, o seguinte produto direto:

$$D^{(\mu \times \nu)}(g) \doteq D^{(\mu)}(g) \otimes D^{(\nu)}(g).$$

A questão que se coloca é se a matriz $D^{(\mu \times \nu)}(g)$ consiste em uma representação do elemento g . Se isto for verdade, então, para todos os produtos $g_1 * g_2 = g_3$, resulta

$$D(g_1) D(g_2) = \left[D^{(\mu)}(g_1) \otimes D^{(\nu)}(g_1) \right] \left[D^{(\mu)}(g_2) \otimes D^{(\nu)}(g_2) \right].$$

Pelas propriedades do produto direto, resulta que

$$\begin{aligned} D(g_1) D(g_2) &= \left[D^{(\mu)}(g_1) D^{(\mu)}(g_2) \right] \otimes \left[D^{(\nu)}(g_1) D^{(\nu)}(g_2) \right] \\ &= D^{(\mu)}(g_3) \otimes D^{(\nu)}(g_3) \\ &= D^{(\mu \times \nu)}(g_3). \end{aligned}$$

Ou seja, o produto direto de duas representações também é uma representação.

Genericamente, a operação

$$D^{(\mu \times \nu)}(\mathcal{G}) = D^{(\mu)}(\mathcal{G}) \otimes D^{(\nu)}(\mathcal{G}) \tag{5.12}$$

é denominada o *produto direto das representações* $D^{(\mu)}(\mathcal{G})$ e $D^{(\nu)}(\mathcal{G})$ e a representação $D^{(\mu \times \nu)}(\mathcal{G})$ resultante é denominado o **grupo do produto direto**.

A propriedade (5.11b) possui também uma consequência importante no que concerne aos caracteres de uma representação. Dadas as representações $D^{(\mu)}(g)$ e $D^{(\nu)}(g)$ de qualquer elemento $g \in \mathcal{G}$ e o seu produto direto $D^{(\mu \times \nu)}(g)$, os caracteres destas representações estão relacionados, a partir de (5.6), por

$$\chi^{(\mu \times \nu)}(g) = \chi^{(\mu)}(g) \chi^{(\nu)}(g). \tag{5.13}$$

Estas propriedades serão exploradas nas próximas seções.

5.5 REPRESENTAÇÕES REDUTÍVEIS OU IRREDUTÍVEIS

Chega-se então a um ponto crucial na teoria de representações de grupos: a obtenção de uma *representação irreduzível* de uma determinada dimensão. Os conceitos de representações redutíveis ou irreduzíveis são centrais na teoria de representação de grupos. Na seção anterior mostrou-se que uma representação de dimensão maior pode ser construída a partir da soma direta de representações de dimensões mais baixas. Considera-se agora a questão oposta: dada uma representação $D(\mathcal{G})$ de uma dada dimensão, é possível descrevê-la em termos de representações mais “simples”, *i. e.*, de mais baixas dimensões?

Um procedimento sistemático que pode ser adotado para responder a esta questão consiste em realizar uma ou mais transformações de similaridade (5.5) sobre $D(\mathcal{G})$ até que as matrizes da representação fiquem próximas (mas não necessariamente idênticas) à forma bloco-diagonal. Caso isto seja possível, esta representação é dita *redutível*.

Definição 5.8 (Representação redutível). Dado o grupo $\mathcal{G} = \{G; *\}$ e uma representação de dimensão n $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$. Esta representação é denominada *representação redutível* se para todo $g \in \mathcal{G}$ com representação $D(g) \in D(\mathcal{G})$ existir uma transformação de similaridade (5.5) tal que $D(g)$ resulte na forma *bloco-escalonada*

$$D(g) = \left(\begin{array}{c|c} D^{(1)}(g) & A(g) \\ \hline 0 & D^{(2)}(g) \end{array} \right), \tag{5.14}$$

onde $D^{(1)}(g) \in D^{(1)}(\mathcal{G})$ e $D^{(2)}(g) \in D^{(2)}(\mathcal{G})$ (com $D^{(1)}(\mathcal{G}), D^{(2)}(\mathcal{G}) \subset D(\mathcal{G})$) são também representações de \mathcal{G} , respectivamente de dimensões n_1 e n_2 tais que $n_1 + n_2 = n$. Por sua vez, $A(g)$ é uma matriz retangular de tamanho $n_1 \times n_2$.

Verifica-se facilmente que a matriz $D(g)$ expressa em (5.14) é uma representação do elemento g . Dados $g_1, g_2 \in \mathcal{G}$ com respectivas representações $D(g_1)$ e $D(g_2)$ dadas na forma (5.14), então

$$\begin{aligned} D(g_1 * g_2) &= D(g_1) D(g_2) = \begin{pmatrix} D^{(1)}(g_1) & A(g_1) \\ \hline 0 & D^{(2)}(g_1) \end{pmatrix} \begin{pmatrix} D^{(1)}(g_2) & A(g_2) \\ \hline 0 & D^{(2)}(g_2) \end{pmatrix} \\ &= \begin{pmatrix} D^{(1)}(g_1) D^{(1)}(g_2) & D^{(1)}(g_1) A(g_2) + A(g_1) D^{(2)}(g_2) \\ \hline 0 & D^{(2)}(g_1) D^{(2)}(g_2) \end{pmatrix}. \end{aligned}$$

O resultado acima está claramente na forma (5.14), uma vez que $D^{(1)}(g_1) D^{(1)}(g_2)$ é uma representação de dimensão n_1 de $g_1 * g_2$, enquanto que $D^{(2)}(g_1) D^{(2)}(g_2)$ é uma representação de dimensão n_2 do mesmo elemento e $D^{(1)}(g_1) A(g_2) + A(g_1) D^{(2)}(g_2)$ é uma matriz retangular de tamanho $n_1 \times n_2$. Portanto, a condição de homomorfismo (5.1) está satisfeita.

A condição para que uma representação seja redutível está ligada à teoria de espaços vectoriais. Retornando ao espaço vectorial \mathcal{V}^n , de dimensão n , considerado na seção 5.2.2,¹⁰ a redutibilidade de uma representação depende da existência de um subespaço próprio invariante $\mathcal{V}^{n_1} \subset \mathcal{V}^n$, de dimensão $n_1 < n$.¹¹ Se $u = (u_1 \dots u_{n_1} u_{n_1+1} \dots u_n)^T$ é um vetor de \mathcal{V}^n , então a ação da representação $D(g)$ na forma reduzida (5.14) sobre o mesmo irá resultar no vetor u' dado por $u' = D(g)u$. Se o espaço vectorial possui um subespaço invariante de dimensão n_1 , então um vetor contido neste subespaço pode ser escrito, mediante uma escolha conveniente de base, como $u = (u_1 \dots u_{n_1} 0 \dots 0)^T = (\begin{smallmatrix} \vdots & \vdots & \vdots & \vdots \\ u_{n_1} & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots \end{smallmatrix})^T$ e o vetor u' resulta

$$\begin{pmatrix} D^{(1)}(g) & A(g) \\ \hline 0 & D^{(2)}(g) \end{pmatrix} \begin{pmatrix} u_{n_1} \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} u'_{n_1} \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} D^{(1)}(g)u_{n_1} \\ \vdots \\ 0 \end{pmatrix}.$$

o mesmo ocorrendo para todos os outros elementos de \mathcal{G} . Ou seja, os vetores u' resultantes da ação de $D(\mathcal{G})$ sobre os vetores contidos no subespaço n_1 também estão contidos no mesmo subespaço, caracterizando o mesmo como invariante.

Por outro lado, para os vetores contidos no subespaço complementar de dimensão $n_2 = n - n_1$, os quais podem ser escritos como $u = (0 \dots 0 u_{n_1+1} \dots u_n)^T = (\begin{smallmatrix} \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \vdots & \vdots & \vdots \end{smallmatrix})^T$, a aplicação do mesmo operador resulta em

$$u'' = \begin{pmatrix} D^{(1)}(g) & A(g) \\ \hline 0 & D^{(2)}(g) \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ u_{n_2} \end{pmatrix} = \begin{pmatrix} A(g)u_{n_2} \\ \vdots \\ D^{(2)}(g)u_{n_2} \end{pmatrix}.$$

Ou seja, o subespaço complementar não é invariante.

Agora surge a questão da redutibilidade das representações $D^{(1)}(g)$ e $D^{(2)}(g)$ em (5.14). Se existir uma matriz $S_1 \in GL(n_1, \mathbb{K})$ tal que a transformação de similaridade (5.5) sobre $D^{(1)}(g)$ a coloque também na forma (5.14), então

$$D^{(1)}(g) = \begin{pmatrix} D^{(3)}(g) & A'(g) \\ \hline 0 & D^{(4)}(g) \end{pmatrix},$$

o mesmo ocorrendo com $D^{(2)}(g)$. Novamente, a condição imposta para a redutibilidade destas representações consiste na existência de subespaços invariantes.

Procedendo desta forma, chega-se a um ponto onde não mais existem subespaços invariantes para a subsequente redução da representação e a matriz $D(g)$ original fica então expressa na

¹⁰Ver também capítulo 4.

¹¹Subespaços invariantes foram considerados na seção 4.6.

forma

$$D(g) = \begin{pmatrix} \begin{array}{c|c} D^{(1)}(g) & A^{(1)}(g) \\ \hline 0 & \dots \end{array} \\ \begin{array}{c|c} D^{(2)}(g) & A^{(2)}(g) \\ \hline 0 & \dots \end{array} \\ \begin{array}{c|c} D^{(3)}(g) & A^{(3)}(g) \\ \hline 0 & \dots \end{array} \\ \vdots \\ \begin{array}{c|c} D^{(k-1)}(g) & A^{(k-1)}(g) \\ \hline 0 & \dots \end{array} \\ \begin{array}{c|c} D^{(k)}(g) & A^{(k)}(g) \\ \hline 0 & \dots \end{array} \end{pmatrix}.$$

Por outro lado, se for possível encontrar uma base ortonormal para \mathcal{V}^n na qual todas as matrizes da representação $D(\mathcal{G})$ podem ser reduzidas na forma (5.14), porém com $A(g) = 0$, então esta representação é dita completamente redutível.

Definição 5.9 (Representação completamente redutível). Dado o grupo $\mathcal{G} = \{G; *\}$ e uma representação $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$. Esta representação é denominada *representação completamente redutível* se para todo $g \in \mathcal{G}$ com representação $D(g) \in D(\mathcal{G})$ existir uma transformação de similaridade (5.5) tal que $D(g)$ resulte na forma bloco-diagonal $D(g) = D^{(1)}(g) \oplus D^{(2)}(g)$, ou seja,

$$D(g) = \begin{pmatrix} \begin{array}{c|c} \vdots & 0 \\ \hline 0 & \vdots \end{array} \\ \vdots \\ \begin{array}{c|c} \vdots & 0 \\ \hline 0 & \vdots \end{array} \end{pmatrix}, \tag{5.15}$$

onde $D^{(1)}(g) \in D^{(1)}(\mathcal{G}) \subset D(\mathcal{G})$ e $D^{(2)}(g) \in D^{(2)}(\mathcal{G}) \subset D(\mathcal{G})$ são também representações de \mathcal{G} , respectivamente de dimensões n_1 e n_2 tais que $n_1 + n_2 = n$.

Se esta condição for satisfeita, ambos os subespaços \mathcal{V}^{n_1} e \mathcal{V}^{n_2} são invariantes e o espaço vetorial \mathcal{V}^n pode ser *decomposto* pela soma direta

$$\mathcal{V}^n = \mathcal{V}^{n_1} \oplus \mathcal{V}^{n_2},$$

como foi discutido na seção 4.2.2.

Aplicando agora novas transformações de similaridade tanto sobre $D^{(1)}(g)$ quanto sobre $D^{(2)}(g)$ em (5.15), estas podem, por sua vez, também ser colocadas na forma bloco-diagonal, dependendo da existência de subespaços invariantes de \mathcal{V}^{n_1} e \mathcal{V}^{n_2} . Nota-se aqui que sempre é possível construir uma transformação de similaridade do tipo (5.5) a qual irá transformar somente o bloco $D^{(1)}(g)$ em (5.15), mantendo $D^{(2)}(g)$ invariante. Uma matriz deste tipo seria

$$S = \begin{pmatrix} \begin{array}{c|c} \vdots & 0 \\ \hline C_1 & \vdots \end{array} \\ \begin{array}{c|c} 0 & \vdots \\ \hline \vdots & I_{n_2} \end{array} \end{pmatrix},$$

sendo $C_1 \in GL(n_1, \mathbb{K})$, pois

$$\begin{aligned} D'(g) = S^{-1}D(g)S &= \begin{pmatrix} \begin{array}{c|c} \vdots & 0 \\ \hline C_1^{-1} & \vdots \end{array} \\ \begin{array}{c|c} 0 & \vdots \\ \hline \vdots & I_{n_2} \end{array} \end{pmatrix} \begin{pmatrix} \begin{array}{c|c} \vdots & 0 \\ \hline D^{(1)}(g) & \vdots \end{array} \\ \begin{array}{c|c} 0 & \vdots \\ \hline \vdots & D^{(2)}(g) \end{array} \end{pmatrix} \begin{pmatrix} \begin{array}{c|c} \vdots & 0 \\ \hline C_1 & \vdots \end{array} \\ \begin{array}{c|c} 0 & \vdots \\ \hline \vdots & I_{n_2} \end{array} \end{pmatrix} \\ &= \begin{pmatrix} \begin{array}{c|c} \vdots & 0 \\ \hline C_1^{-1}D^{(1)}(g)C_1 & \vdots \end{array} \\ \begin{array}{c|c} 0 & \vdots \\ \hline \vdots & D^{(2)}(g) \end{array} \end{pmatrix}. \end{aligned}$$

Repetindo este processo, chega-se a um ponto em que não mais existem subespaços invariantes para a representação e a matriz $D(g)$ resulta na forma bloco-diagonal

$$D(g) = \begin{pmatrix} \begin{array}{c|c} \vdots & \mathbf{0} \\ \hline \hat{D}^{(1)}(g) & \vdots \end{array} \\ \begin{array}{c|c} \vdots & \vdots \\ \hline \mathbf{0} & \hat{D}^{(2)}(g) \end{array} \\ \vdots \\ \begin{array}{c|c} \vdots & \vdots \\ \hline \mathbf{0} & \hat{D}^{(N)}(g) \end{array} \end{pmatrix}, \tag{5.16a}$$

onde as matrizes $\hat{D}^{(1)}(g), \dots, \hat{D}^{(N)}(g)$ não podem mais ser reduzidas à forma bloco-diagonal por nenhuma transformação de similaridade. Estas representações são denominadas *irredutíveis*.

Definição 5.10 (Representação irredutível). Dado o grupo $\mathcal{G} = \{G; *\}$ e uma representação $\hat{D}(\mathcal{G}) \subseteq GL(n, \mathbb{K})$. Esta representação é denominada *representação irredutível (Irrep)* se, para todo $g \in \mathcal{G}$ com representação $\hat{D}(g) \in \hat{D}(\mathcal{G})$, não existir uma transformação de similaridade (5.5) tal que $\hat{D}(g)$ resulte na forma bloco-diagonal (5.15).

A definição acima implica em que não existem subespaços próprios invariantes que decompõem o espaço \mathcal{V}^n . Quando uma representação $D(\mathcal{G})$ puder ser escrita na forma bloco-diagonal (5.16a) em termos das matrizes $\hat{D}^{(\mu)}(g)$, estas são denominadas as *representações irredutíveis* ou *irreps* de \mathcal{G} e a forma da representação $D(g)$ em (5.16a) é dita **completamente reduzida**.

Dentre as N irreps em (5.16a) pode ocorrer que diversas sejam equivalentes entre si, todas com a mesma dimensão. Representações irredutíveis equivalentes não são consideradas distintas entre si e se pode usar a mesma designação para estas. Neste caso, dada a irrep $\hat{D}^{(\mu)}(g)$ de dimensão n_μ , se houver $m_\mu = 0, 1, \dots$ irreps equivalentes entre si, a representação $D(g)$ em (5.16a) pode ser escrita como

$$D(g) = \bigoplus_{\mu=1}^N m_\mu \hat{D}^{(\mu)}(g) = m_1 \hat{D}^{(1)}(g) \oplus m_2 \hat{D}^{(2)}(g) \oplus \dots \oplus m_N \hat{D}^{(N)}(g). \quad (5.16b)$$

Portanto, se para um dado grupo \mathcal{G} for possível escrever a sua representação na forma completamente reduzida (5.16), as suas irreps conterão toda a informação carregada pelos operadores de transformação. A vantagem de empregar as irreps é evidente, uma vez que estas em geral são matrizes de baixa dimensão, fáceis de serem manuseadas.

A questão que ainda resta é se qualquer representação pode ser escrita em termos de irreps. O teorema a seguir oferece um primeiro passo em sentido à resposta a esta questão ao mostrar que uma representação qualquer de um grupo finito é sempre equivalente a uma representação unitária.

Definição 5.11 (Representação unitária). Dado o grupo \mathcal{G} e uma representação $\Gamma(\mathcal{G}) \subseteq GL(n, \mathbb{K})$. Se para todo $g \in \mathcal{G}$ a sua representação $\Gamma(g) \in \Gamma(\mathcal{G})$ for uma matriz unitária, *i. e.*,

$$\Gamma^\dagger(g) \Gamma(g) = \Gamma(g) \Gamma^\dagger(g) = I_n,$$

então $\Gamma(\mathcal{G})$ é denominada uma *representação unitária* de \mathcal{G} .

Teorema 5.1. Dado o grupo $\mathcal{G} = \{G; *\}$ finito e uma representação $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$. Esta representação é sempre equivalente a uma representação unitária $\Gamma(\mathcal{G}) \subseteq GL(n, \mathbb{K})$.

Demonstração. Define-se a matriz Hermitiana

$$H = \sum_{g \in \mathcal{G}} D(g) D^\dagger(g),$$

ou seja, a qual satisfaz $H = H^\dagger$. De acordo com o teorema 4.36, qualquer matriz Hermitiana sempre pode ser diagonalizada por uma transformação de similaridade (5.5) através do uso de uma matriz unitária U . Então,

$$U^{-1} H U = H_d,$$

onde H_d é uma matriz diagonal cujos elementos são os autovalores (reais) de H . Mas,

$$\begin{aligned} H_d &= U^{-1} \left(\sum_{g \in \mathcal{G}} D(g) D^\dagger(g) \right) U \\ &= \sum_{g \in \mathcal{G}} U^{-1} D(g) U U^{-1} D^\dagger(g) U \\ &= \sum_{g \in \mathcal{G}} D'(g) D'^\dagger(g), \end{aligned}$$

onde $D'(g) = U^{-1} D(g) U$. Tomando agora o k -ésimo elemento de H_d , observa-se que

$$d_k = (H_d)_{kk} = \sum_{g \in \mathcal{G}} \sum_j D'_{kj}(g) D'^\dagger_{jk}(g) = \sum_{g \in \mathcal{G}} \sum_j D'_{kj}(g) D'^*_{kj}(g) = \sum_{g \in \mathcal{G}} \sum_j |D'_{kj}(g)|^2,$$

ou seja, $d_k \geq 0$. Para que d_k seja nulo, é necessário que $D'_{kj}(g) = 0, \forall j, g$, o que implica que $\det [D(g)] = 0, \forall g$, o que implicaria que $D(g) \notin GL(n, \mathbb{K})$. Portanto, necessariamente, $d_k > 0$ e $\det(H_d) > 0$.

Pode-se definir agora a matriz Hermitiana diagonal E , cujos elementos são $E_{ij} = \sqrt{d_i} \delta_{ij}$, tal que $E^2 = H_d$. Definindo então a matriz de transformação $V = UE$, para todo $D(g) \in D(\mathcal{G})$, resulta que a partir da representação $D(g)$ é sempre possível obter uma representação unitária equivalente $\Gamma(g)$ através da transformação de similaridade

$$\Gamma(g) = V^{-1}D(g)V = E^{-1}U^{-1}D(g)UE = E^{-1}D'(g)E.$$

Para verificar que as matrizes $\Gamma(g)$ realmente são unitárias, calcula-se

$$\begin{aligned} \Gamma(g)\Gamma^\dagger(g) &= E^{-1}D'(g)E[E^{-1}D'(g)E]^\dagger = E^{-1}D'(g)EE^\dagger D'^\dagger(g)(E^{-1})^\dagger \\ &= E^{-1}D'(g)H_d D'^\dagger(g)E^{-1} = E^{-1}D'(g)\left(\sum_{g' \in \mathcal{G}} D'(g')D'^\dagger(g')\right)D'^\dagger(g)E^{-1} \\ &= E^{-1}\sum_{g' \in \mathcal{G}} D'(g)D'(g')D'^\dagger(g')D'^\dagger(g)E^{-1} = E^{-1}\sum_{g' \in \mathcal{G}} D'(g * g')D'^\dagger(g * g')E^{-1} \\ &= E^{-1}H_d E^{-1}. \end{aligned}$$

Mas, como $(E^{-1})_{ij} = d_i^{-1/2} \delta_{ij}$, resulta que

$$[\Gamma(g)\Gamma^\dagger(g)]_{ij} = \sum_{k, \ell} (E^{-1})_{i\ell} (H_d)_{\ell k} (E^{-1})_{kj} = \sum_{k, \ell} d_i^{-1/2} d_\ell d_k^{-1/2} \delta_{i\ell} \delta_{\ell k} \delta_{kj} = d_i^{-1/2} d_j^{1/2} \delta_{ij} = \delta_{ij}.$$

Ou seja, $\Gamma(g)\Gamma^\dagger(g) = I_n$ para todo $g \in \mathcal{G}$ e a representação $\Gamma(\mathcal{G})$ é unitária. \square

A importância deste teorema é percebida retornando a (5.14). Se a representação $D(\mathcal{G})$ é redutível, então realizando a transformação de similaridade com a matriz $V = UE$ obtém-se a representação $\Gamma(\mathcal{G})$, a qual deve estar na forma (5.14). Contudo, como $\Gamma(g)$ é unitária para qualquer $g \in \mathcal{G}$, observa-se que, se

$$\Gamma(g) = \begin{pmatrix} \Gamma^{(1)}(g) & A(g) \\ 0 & \Gamma^{(2)}(g) \end{pmatrix},$$

isto é, se

$$\Gamma(g) = \begin{pmatrix} \Gamma_{11}^{(1)} & \cdots & \Gamma_{1n_1}^{(1)} & A_{11} & \cdots & A_{1n_2} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Gamma_{n_1 1}^{(1)} & \cdots & \Gamma_{n_1 n_1}^{(1)} & A_{n_1 1} & \cdots & A_{n_1 n_2} \\ 0 & \cdots & 0 & \Gamma_{11}^{(2)} & \cdots & \Gamma_{1n_2}^{(2)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \Gamma_{n_2 1}^{(2)} & \cdots & \Gamma_{n_2 n_2}^{(2)} \end{pmatrix}, \text{ então } \Gamma^\dagger(g) = \begin{pmatrix} \Gamma_{11}^{(1)*} & \cdots & \Gamma_{n_1 1}^{(1)*} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Gamma_{1n_1}^{(1)*} & \cdots & \Gamma_{n_1 n_1}^{(1)*} & 0 & \cdots & 0 \\ A_{11}^* & \cdots & A_{n_1 1}^* & \Gamma_{11}^{(2)*} & \cdots & \Gamma_{n_2 1}^{(2)*} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{1n_2}^* & \cdots & A_{n_1 n_2}^* & \Gamma_{1n_2}^{(2)*} & \cdots & \Gamma_{n_2 n_2}^{(2)*} \end{pmatrix},$$

ou seja,

$$\Gamma^\dagger(g) = \begin{pmatrix} \Gamma^{(1)\dagger}(g) & 0 \\ A^\dagger(g) & \Gamma^{(2)\dagger}(g) \end{pmatrix}.$$

Isto significa que se $\Gamma(g)$ for unitária, então, necessariamente, $A(g) = 0$, e a representação é sempre totalmente redutível.

A conclusão, portanto, é que para grupos finitos sempre é possível encontrar uma representação completamente reduzida em termos das representações irredutíveis unitárias.

5.6 TEOREMAS FUNDAMENTAIS SOBRE REPRESENTAÇÕES DE GRUPOS E CARACTERES

Serão apresentados agora alguns teoremas adicionais sobre representações irredutíveis de grupos e seus caracteres. Esses teoremas oferecem diversas informações importantes, como o número total de irreps de um determinado grupo e suas dimensões.

5.6.1 TEOREMAS SOBRE REPRESENTAÇÕES

Alguns dos principais teoremas sobre representações de grupos são apresentados nesta seção.

Lema 5.1 (Lema de Schur 1). *Se $\hat{D}(\mathcal{G})$ é uma representação irredutível de um grupo \mathcal{G} e se uma matriz P comuta com todas as matrizes $\hat{D}(a) \in \hat{D}(\mathcal{G})$ ($\forall a \in \mathcal{G}$), então P deve ser uma matriz constante, i. e., $P = \lambda I$, sendo λ um escalar.*

Demonstração. Será mostrado inicialmente que qualquer matriz Hermitiana que comuta com todas as matrizes de uma representação irredutível é uma matriz constante da matriz identidade. De acordo com o teorema 5.1, a representação $\hat{D}(\mathcal{G})$ é sempre equivalente a uma representação unitária $\hat{\Gamma}(\mathcal{G})$. Portanto, se n é a dimensão da representação, existe uma matriz $S \in GL(n, \mathbb{K})$ que realiza a transformação de similaridade $\hat{D}(\mathcal{G}) \rightarrow \hat{\Gamma}(\mathcal{G})$.

Seja então H uma matriz Hermitiana que comuta com todas as matrizes da irrep unitária $\hat{\Gamma}(\mathcal{G})$, i.e.,

$$H\hat{\Gamma}(a) = \hat{\Gamma}(a)H, \forall a \in \mathcal{G}.$$

De acordo com o teorema 4.36, existe uma matriz unitária U que transforma a matriz H na matriz diagonal H_d , i.e.,

$$U^{-1}HU = H_d \implies H = UH_dU^{-1},$$

sendo que os elementos de H_d são os autovalores de H , doravante chamados de λ_i ($1 \leq i \leq n$), sendo n a dimensão da representação.

Supõe-se agora que nem todos os autovalores são idênticos entre si. Assumindo que um certo autovalor é repetido k vezes ($1 \leq k \leq n$), escolhe-se a matriz U de forma a colocar estes autovalores nas primeiras k posições de H_d ; isto é, $\lambda_1 = \dots = \lambda_k \neq \lambda_\ell$ ($k+1 \leq \ell \leq n$).

Realizando então a transformação de similaridade com a matriz U ,

$$U^{-1}HUU^{-1}\hat{\Gamma}(a)U = U^{-1}\hat{\Gamma}(a)UU^{-1}HU \implies H_d\hat{\Gamma}'(a) = \hat{\Gamma}'(a)H_d, \forall a \in \mathcal{G}.$$

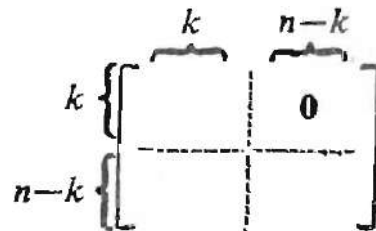
Tomando o elemento ij , resulta

$$\lambda_i\hat{\Gamma}'_{ij}(a) = \hat{\Gamma}'_{ij}(a)\lambda_j \implies (\lambda_i - \lambda_j)\hat{\Gamma}'_{ij}(a) = 0, \forall a \in \mathcal{G}.$$

Ou seja,

$$\hat{\Gamma}'_{ij}(a) = 0, \text{ para } \begin{cases} 1 \leq i \leq k \\ k+1 \leq j \leq n \\ \text{todo } a \in \mathcal{G}. \end{cases}$$

Isto significa que as matrizes $\hat{\Gamma}'(a)$ resultam na forma bloco-escalonada ilustrada na figura abaixo:



Mas, como $\hat{\Gamma}'(a)$ é unitária para todo $a \in \mathcal{G}$, isto implica que a representação é redutível, de acordo com a discussão realizada na página 186, e, portanto, $\hat{D}(\mathcal{G})$ é redutível, por ser equivalente a $\hat{\Gamma}(\mathcal{G})$.

Assim, como $\hat{D}(\mathcal{G})$ é irredutível por hipótese, é necessário que $k = n$, ou seja, todos os autovalores de H são idênticos entre si ($\lambda_i = \lambda, \forall i$) e, portanto, $H = \lambda I$.

Seja agora P uma matriz que comuta com todas as matrizes de $\hat{\Gamma}(\mathcal{G})$, i.e., $P\hat{\Gamma}(a) = \hat{\Gamma}(a)P$ ($\forall a \in \mathcal{G}$). Tomando o Hermitiano conjugado de ambos os lados,

$$P^\dagger\hat{\Gamma}^\dagger(a) = \hat{\Gamma}^\dagger(a)P^\dagger \implies P^\dagger\hat{\Gamma}^{-1}(a) = \hat{\Gamma}^{-1}(a)P^\dagger \implies P^\dagger\hat{\Gamma}(a^{-1}) = \hat{\Gamma}(a^{-1})P^\dagger,$$

para todo $a \in \mathcal{G}$. Ou seja, P^\dagger também comuta com todas as matrizes de $\hat{\Gamma}(\mathcal{G})$. Definindo então as matrizes Hermitianas

$$H_1 = \frac{1}{2}(P + P^\dagger), \quad H_2 = \frac{i}{2}(P - P^\dagger),$$

escreve-se

$$P = H_1 + iH_2, \quad P^\dagger = H_1 - iH_2.$$

Assim,

$$\begin{aligned} P\hat{\Gamma}(a) = \hat{\Gamma}(a)P &\implies H_1\hat{\Gamma}(a) + iH_2\hat{\Gamma}(a) = \hat{\Gamma}(a)H_1 + i\hat{\Gamma}(a)H_2 \\ P^\dagger\hat{\Gamma}(a) = \hat{\Gamma}(a)P^\dagger &\implies H_1\hat{\Gamma}(a) - iH_2\hat{\Gamma}(a) = \hat{\Gamma}(a)H_1 - i\hat{\Gamma}(a)H_2. \end{aligned}$$

Somando ambos os resultados, resulta que $H_1\hat{\Gamma}(a) = \hat{\Gamma}(a)H_1$ e subtraindo resulta $H_2\hat{\Gamma}(a) = \hat{\Gamma}(a)H_2$.

Ou seja, como P e P^\dagger comutam com todas as matrizes de $\hat{\Gamma}(\mathcal{G})$, isto implica que as matrizes Hermitianas H_1 e H_2 também devem comutar. Contudo, de acordo com a conclusão da primeira parte da demonstração, $H_1 = \lambda_1 I$ e $H_2 = \lambda_2 I$. Portanto, $P = (\lambda_1 + i\lambda_2)I = \lambda I$ e P deve ser uma matriz constante.

Finalmente, tomando a matriz $R = S^{-1}$ para realizar a transformação inversa $\hat{\Gamma}(\mathcal{G}) \longrightarrow \hat{D}(\mathcal{G})$, como $P = \lambda I$, então $P' = R^{-1}PR = P$ e o lema está demonstrado.

A recíproca do enunciado do Lema de Schur 1 também é verdadeira: se somente uma matriz constante comuta com todas as matrizes de uma representação, então essa representação deve ser irredutível. \square

Lema 5.2 (Lema de Schur 2). *Se $\hat{D}^{(\mu)}(\mathcal{G})$ e $\hat{D}^{(\nu)}(\mathcal{G})$ são duas representações irredutíveis de um grupo \mathcal{G} , de dimensões n_μ e n_ν respectivamente, e se uma matriz M de tamanho $(n_\mu \times n_\nu)$ satisfaz a relação*

$$\hat{D}^{(\mu)}(a)M = M\hat{D}^{(\nu)}(a),$$

para todo $a \in \mathcal{G}$, então ou:

- (a) $M = 0$, i. e., M é a matriz nula, ou
- (b) $\det(M) \neq 0$, em cujo caso $\hat{D}^{(\mu)}$ e $\hat{D}^{(\nu)}$ são representações equivalentes (com $n_\mu = n_\nu$).

Demonstração. De acordo com o teorema 5.1, existem matrizes $S_1 \in GL(n_\mu, \mathbb{K})$ e $S_2 \in GL(n_\nu, \mathbb{K})$ tais que realizam as transformações de similaridade $\hat{D}^{(\mu)}(\mathcal{G}) \longrightarrow \hat{\Gamma}^{(\mu)}(\mathcal{G})$ e $\hat{D}^{(\nu)}(\mathcal{G}) \longrightarrow \hat{\Gamma}^{(\nu)}(\mathcal{G})$. Então, parte-se de

$$\hat{\Gamma}^{(\mu)}(a)M = M\hat{\Gamma}^{(\nu)}(a) \quad (\forall a \in \mathcal{G}). \quad (5.17a)$$

Tomando agora o Hermitiano conjugado de ambos os lados, obtém-se

$$M^\dagger \hat{\Gamma}^{(\mu)\dagger}(a) = \hat{\Gamma}^{(\nu)\dagger}(a)M^\dagger \implies M^\dagger \hat{\Gamma}^{(\mu)}(a^{-1}) = \hat{\Gamma}^{(\nu)}(a^{-1})M^\dagger,$$

para todo $a \in \mathcal{G}$. Multiplicando na direita por M ,

$$M^\dagger \hat{\Gamma}^{(\mu)}(a^{-1})M = \hat{\Gamma}^{(\nu)}(a^{-1})M^\dagger M \implies M^\dagger M \hat{\Gamma}^{(\nu)}(a) = \hat{\Gamma}^{(\nu)}(a^{-1})M^\dagger M.$$

Ou seja, a matriz $M^\dagger M$ comuta com $\hat{\Gamma}^{(\nu)}(a)$ para todos $a \in \mathcal{G}$. Portanto, de acordo com o Lema de Schur 1, a mesma deve ser uma matriz constante,

$$M^\dagger M = \lambda I_{n_\nu}. \quad (5.17b)$$

Considerando primeiro o caso $n_\mu = n_\nu = n$, de acordo com as propriedades de matrizes apresentadas na seção 4.9,

$$\det(M^\dagger M) = \det(M^\dagger) \det(M) = |\det(M)|^2 = \lambda^n.$$

Se $\lambda \neq 0$, então $\det(M) \neq 0$ e, portanto, M^{-1} existe. Neste caso, de (5.17a) conclui-se que

$$\hat{\Gamma}^{(\nu)}(a) = M^{-1} \hat{\Gamma}^{(\mu)}(a) M \quad (\forall a \in \mathcal{G}),$$

o que mostra que $\hat{\Gamma}^{(\mu)}(\mathcal{G})$ e $\hat{\Gamma}^{(\nu)}(\mathcal{G})$ são representações equivalentes.

Por outro lado, se $\lambda = 0$, toma-se o elemento ii de (5.17b),

$$(M^\dagger M)_{ii} = \sum_{k=1}^n |M_{ki}|^2 = 0 \implies M_{ki} = 0, \text{ para todo } 1 \leq k \leq n.$$

Contudo, como i assume qualquer valor $1 \leq i \leq n$, isto implica que $M = 0$; ou seja, M é nula.

Considerando agora o caso $n_\mu \neq n_\nu$, assume-se, sem perda de generalidade, que $n_\mu < n_\nu$. Será criada uma nova matriz M' a qual é a matriz M acrescida com um bloco inferior que consiste na matriz nula com $n_\nu - n_\mu$ linhas e n_ν colunas:

$$M' = \left[\begin{array}{c} \overbrace{M}^{n_\mu} \\ \underbrace{0}_{n_\nu - n_\mu} \end{array} \right] \quad M'_{ij} = \begin{cases} M_{ij}, & 1 \leq i \leq n_\mu \\ 0, & n_{\mu+1} + 1 \leq i \leq n_\nu. \end{cases}$$

Com esta definição,

$$M'^{\dagger} = \left[\begin{array}{c} \overbrace{M^{\dagger}}^{n_\mu} \\ \underbrace{0}_{n_\nu - n_\mu} \end{array} \right] n_\nu.$$

Verifica-se facilmente por multiplicação matricial que $M'^{\dagger}M' = M^{\dagger}M$ e, portanto, de (5.17b) conclui-se que

$$\det(M'^{\dagger}M') = \det(M^{\dagger}M) = \lambda^{n_\nu}.$$

Contudo, observa-se que $\det(M'^{\dagger}) = \det(M') = 0$. Portanto, $\lambda = 0$ novamente, o que leva a $M'^{\dagger}M' = 0$. Analisando novamente elemento a elemento de $M'^{\dagger}M'$, conclui-se finalmente que $M = 0$. \square

Chega-se então ao *grande teorema da ortogonalidade* para as representações irredutíveis de um grupo. Este teorema ocupa uma posição central na teoria de representações de grupos.

Teorema 5.2 (Teorema da ortogonalidade). *Seja \mathcal{G} um grupo de ordem g e sejam $\Gamma^{(\mu)}(\mathcal{G})$ e $\Gamma^{(\nu)}(\mathcal{G})$ representações irredutíveis unitárias não equivalentes de \mathcal{G} , de dimensões n_μ e n_ν , respectivamente. Então, para todo $a_i \in \mathcal{G}$,*

$$\sum_{a_i \in \mathcal{G}} \Gamma_{jm}^{(\mu)}(a_i) \Gamma_{nk}^{(\nu)}(a_i^{-1}) = \sum_{a_i \in \mathcal{G}} \Gamma_{jm}^{(\mu)}(a_i) \Gamma_{kn}^{(\nu)*}(a_i) = \frac{g}{n_\mu} \delta_{\mu\nu} \delta_{jk} \delta_{mn}. \quad (5.18)$$

Demonstração. Será construída a matriz

$$M = \sum_{a \in \mathcal{G}} \Gamma^{(\mu)}(a) X \Gamma^{(\nu)}(a^{-1}),$$

sendo $\Gamma^{(\mu)}(\mathcal{G})$ e $\Gamma^{(\nu)}(\mathcal{G})$ duas representações irredutíveis unitárias não equivalentes de \mathcal{G} ($\mu \neq \nu$) e X uma matriz ($n_\mu \times n_\mu$) arbitrária e independente dos elementos do grupo. Dado agora o elemento $b \in \mathcal{G}$, existe $c \in \mathcal{G}$ tal que $c = b * a$. Então, multiplicando ambos os lados à esquerda por $\Gamma^{(\mu)}(b)$,

$$\begin{aligned} \Gamma^{(\mu)}(b) M &= \Gamma^{(\mu)}(b) \sum_{a \in \mathcal{G}} \Gamma^{(\mu)}(a) X \Gamma^{(\nu)}(a^{-1}) \\ &= \sum_{a \in \mathcal{G}} \Gamma^{(\mu)}(b * a) X \Gamma^{(\nu)}(a^{-1}) \\ &= \sum_{c \in \mathcal{G}} \Gamma^{(\mu)}(c) X \Gamma^{(\nu)}(c^{-1} * b) \\ &= \sum_{c \in \mathcal{G}} \Gamma^{(\mu)}(c) X \Gamma^{(\nu)}(c^{-1}) \Gamma^{(\nu)}(b) \\ &= M \Gamma^{(\nu)}(b), \end{aligned}$$

para todo $b \in \mathcal{G}$. Portanto, pelo Lema de Schur 2, $M = 0$.

Tomando o elemento ij de M ,

$$(M)_{ij} = \sum_{a \in \mathcal{G}} \sum_{k=1}^{n_\mu} \sum_{\ell=1}^{n_\nu} \Gamma_{ik}^{(\mu)}(a) (X)_{k\ell} \Gamma_{\ell j}^{(\nu)}(a^{-1}) = 0, \quad \begin{pmatrix} 1 \leq i \leq n_\mu \\ 1 \leq j \leq n_\nu \end{pmatrix}.$$

A matriz X é arbitrária, mas por conveniência, esta será agora tomada como a matriz cujos elementos são todos nulos, exceto pelo elemento mn , o qual é igual a um; ou seja, $(X)_{k\ell} = \delta_{km} \delta_{\ell n}$. Assim,

$$\sum_{a \in \mathcal{G}} \Gamma_{im}^{(\mu)}(a) \Gamma_{nj}^{(\nu)}(a^{-1}) = 0, \quad (\mu \neq \nu).$$

Seja agora a matriz N definida como

$$N = \sum_{a \in \mathcal{G}} \Gamma^{(\mu)}(a) X \Gamma^{(\mu)}(a^{-1}).$$

Esta matriz se distingue de M pelo fato que $\nu = \mu$. Mas, pela mesma cadeia de operações realizadas acima, conclui-se que

$$\Gamma^{(\mu)}(a) N = N \Gamma^{(\mu)}(a), \text{ para todo } a \in \mathcal{G}.$$

Portanto, pelo Lema de Schur 1, a matriz N deve ser uma matriz constante $N = \lambda 1_{n_\mu}$.

Tomando então o elemento ij da matriz, obtém-se

$$(N)_{ij} = \sum_{a \in \mathcal{G}} \sum_{k, \ell=1}^{n_\mu} \Gamma_{ik}^{(\mu)}(a) (X)_{k\ell} \Gamma_{\ell j}^{(\mu)}(a^{-1}) = \lambda \delta_{ij}.$$

Mas, com a escolha já feita para a matriz X , tal que $(X)_{ij} = \delta_{im} \delta_{jn}$, resulta

$$\sum_{a \in \mathcal{G}} \Gamma_{in}^{(\mu)}(a) \Gamma_{nj}^{(\mu)}(a^{-1}) = \lambda \delta_{ij}.$$

Para se obter o valor da constante λ , calcula-se o traço de N . De acordo com as propriedades do traço apresentadas na seção 4.9.3,

$$\begin{aligned} \text{Tr}(N) &= n_\mu \lambda = \sum_{a \in \mathcal{G}} \text{Tr} \left[\Gamma^{(\mu)}(a) X \Gamma^{(\mu)}(a^{-1}) \right] \\ &= \sum_{a \in \mathcal{G}} \text{Tr} \left[\Gamma^{(\mu)}(a) \Gamma^{(\mu)}(a^{-1}) X \right] \\ &= \sum_{a \in \mathcal{G}} \text{Tr} (1_{n_\mu} X) = g \text{Tr}(X). \end{aligned}$$

Ou seja, $\lambda = g \text{Tr}(X) / n_\mu$.

Mas, com a escolha feita para a matriz X , observa-se que $\text{Tr}(X) = \delta_{mn}$. Portanto,

$$\sum_{a \in \mathcal{G}} \Gamma_{im}^{(\mu)}(a) \Gamma_{nj}^{(\mu)}(a^{-1}) = \frac{g}{n_\mu} \delta_{ij} \delta_{mn}.$$

Combinando ambos os resultados, chega-se a (5.18). □

O teorema da ortogonalidade possui uma interpretação interessante dentro da teoria de espaços vetoriais. Essa interpretação será apresentada na seção 5.6.2.

Uma expressão complementar ao teorema da ortogonalidade é a *relação de completeza para as representações*, apresentada a seguir.

Teorema 5.3 (Relação de completeza). *Seja \mathcal{G} um grupo de ordem g e*

$$\Gamma \doteq \left\{ \Gamma^{(\mu)}(\mathcal{G}) \mid \mu = 1, \dots, N \right\}$$

o conjunto de suas representações irredutíveis unitárias. Então, dados $a_i, a_j \in \mathcal{G}$,

$$\sum_{\mu=1}^N \sum_{k, \ell=1}^{n_\mu} \sqrt{\frac{n_\mu}{g}} \Gamma_{k\ell}^{(\mu)}(a_i) \sqrt{\frac{n_\mu}{g}} \Gamma_{k\ell}^{(\mu)*}(a_j) = \delta_{ij},$$

onde n_μ é a dimensão da μ -ésima representação.

5.6.2 INTERPRETAÇÃO DO TEOREMA DA ORTOGONALIDADE

Considera-se um grupo finito \mathcal{G} de ordem g . Assume-se agora que o número total de representações irredutíveis não equivalentes de \mathcal{G} seja igual a N . Dada a representação irredutível unitária $\Gamma^{(\mu)}(\mathcal{G}) \subseteq GL(n_\mu, \mathbb{K})$ de dimensão n_μ , cada matriz $\Gamma^{(\mu)}(a_i) \in \Gamma^{(\mu)}(\mathcal{G})$ na representação

contém um total de n_μ^2 elementos. A quantidade $\Gamma_{jk}^{(\mu)}(a_i)$ é um destes elementos, com $a_i \in \mathcal{G}$, $1 \leq \mu \leq N$ e $1 \leq j, k \leq n_\mu$.

Define-se agora a g -upla

$$\Gamma_{jk}^{(\mu)} \doteq \left(\Gamma_{jk}^{(\mu)}(a_1), \Gamma_{jk}^{(\mu)}(a_2), \dots, \Gamma_{jk}^{(\mu)}(a_g) \right).$$

As diferentes g -uplas que podem ser construídas a partir de todas as irreps unitárias de \mathcal{G} são distinguidas pelos índices μ, j, k . Assim, pode-se formar o conjunto

$$\Gamma \doteq \left\{ \Gamma_{jk}^{(\mu)} \mid 1 \leq \mu \leq N, 1 \leq j, k \leq n_\mu \right\}, \quad (5.19)$$

composto por todas as g -uplas possíveis. O número total de g -uplas neste conjunto, ou seja, a sua cardinalidade, é igual a

$$|\Gamma| = \sum_{\mu=1}^N n_\mu^2.$$

Será demonstrado agora que o conjunto Γ é composto por vetores, denominados **vetores de representação**, de um espaço vetorial normado e com produto interno, de dimensão g , o qual possui propriedades semelhantes ao espaço vetorial \mathcal{R}^g , porém com algumas diferenças. Cada componente de $\Gamma_{jk}^{(\mu)}(\mathcal{G})$ é um elemento do corpo \mathbb{K} . Os corpos de interesse para a física são, em sua grande maioria, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Será assumido então um destes corpos em particular.

Retornando então à definição de um espaço vetorial, apresentada na seção 4.1, seja o espaço vetorial complexo $\mathcal{K}^g \doteq \langle \mathcal{K}, \mathbb{K} \rangle$ sobre o corpo \mathbb{K} , onde o conjunto \mathcal{K} é o conjunto de vetores em \mathcal{K}^g , formado pelas g -uplas

$$\mathcal{K} = \{ \kappa \mid \kappa = (\kappa_1, \kappa_2, \dots, \kappa_g), \text{ onde } \kappa_i \in \mathbb{K} \}.$$

Este conjunto forma um grupo Abeliano frente a operação de adição vetorial “+”:

- *Clausura*: se $\kappa, \eta \in \mathcal{K}$, então $\kappa + \eta = (\kappa_1 + \eta_1, \kappa_2 + \eta_2, \dots, \kappa_g + \eta_g) \in \mathcal{K}$.
- *Associatividade*: se $\kappa, \eta, \zeta \in \mathcal{K}$, então $\kappa + (\eta + \zeta) = (\kappa + \eta) + \zeta$.
- *Elemento identidade*: existe $0 = (0, 0, \dots, 0) \in \mathcal{K}$ tal que para todo $\kappa \in \mathcal{K}$ resulta $0 + \kappa = \kappa + 0 = \kappa$.
- *Elemento inverso*: para todo $\kappa \in \mathcal{K}$ existe o elemento $(-\kappa) = (-\kappa_1, -\kappa_2, \dots, -\kappa_g)$ tal que $\kappa + (-\kappa) = (-\kappa) + \kappa = 0$.
- *Comutatividade*: para todos $\kappa, \eta \in \mathcal{K}$, resulta $\kappa + \eta = \eta + \kappa$.

Se o corpo \mathbb{K} forma o conjunto dos escalares de \mathcal{K}^g , então o produto por escalar é definido, para todo $\kappa \in \mathcal{K}$ e todo $k \in \mathbb{K}$, como

$$k \cdot \kappa = (k\kappa_1, k\kappa_2, \dots, k\kappa_g).$$

O produto interno em \mathcal{K}^g será definido como a forma sesquilinear $\langle \cdot, \cdot \rangle : \mathcal{K}^g \times \mathcal{K}^g \mapsto \mathbb{K}$ tal que, dados $\kappa, \eta \in \mathcal{K}^g$, o produto interno de κ por η será a quantidade

$$\langle \kappa, \eta \rangle = \sum_{i=1}^g \kappa_i^* \eta_i.$$

Esta definição satisfaz todos os requisitos de um produto interno; em particular, $\langle \kappa, \kappa \rangle = \sum_i |\kappa_i|^2 \geq 0$, sendo que $\langle \kappa, \kappa \rangle = 0$ somente se $\kappa = 0$.

Retornando agora ao conjunto Γ , é fácil verificar que $\Gamma \subset \mathcal{K}^g$. Considera-se então o produto interno entre os vetores $\Gamma_{kn}^{(\nu)}, \Gamma_{jm}^{(\mu)} \in \Gamma$. Este resulta igual a

$$\left\langle \Gamma_{kn}^{(\nu)}, \Gamma_{jm}^{(\mu)} \right\rangle = \sum_{i=1}^g \Gamma_{jm}^{(\mu)}(a_i) \Gamma_{kn}^{(\nu)*}(a_i) = \sum_{i=1}^g \Gamma_{jm}^{(\mu)}(a_i) \Gamma_{nk}^{(\nu)}(a_i^{-1}).$$

O último resultado ocorre porque as matrizes $\Gamma^{(\mu)}(a)$ são unitárias, ou seja, $\Gamma^{(\mu)}(a)\Gamma^{(\mu)\dagger}(a) = \Gamma^{(\mu)\dagger}(a)\Gamma^{(\mu)}(a) = I_{n_\mu}$, de onde resulta que

$$\Gamma^{(\mu)\dagger}(a) = \left[\Gamma^{(\mu)}(a) \right]^{-1} = \Gamma^{(\mu)}(a^{-1}),$$

sendo que o último resultado é uma das propriedades do grupo de representação. Por isso, dados a matriz unitária $\Gamma^{(\nu)}(a_i)$ e o elemento $\Gamma_{nk}^{(\nu)}(a_i)$ da mesma em particular, ao se tomar o conjugado Hermitiano da matriz, resulta que $\Gamma_{nk}^{(\nu)\dagger}(a_i) = \Gamma_{kn}^{(\nu)*}(a_i) = \left[\Gamma^{(\nu)}(a_i) \right]_{nk}^{-1} = \Gamma_{nk}^{(\nu)}(a_i^{-1})$.

Porém, do teorema da ortogonalidade (5.18) obtém-se que

$$\left\langle \Gamma_{kn}^{(\nu)}, \Gamma_{jm}^{(\mu)} \right\rangle = 0, \text{ se } \begin{matrix} \mu \neq \nu \\ j \neq k \\ m \neq n \end{matrix} \text{ ou } \left\langle \Gamma_{jm}^{(\mu)}, \Gamma_{jm}^{(\mu)} \right\rangle = \frac{g}{n_\mu} \implies \left\langle \Gamma_{kn}^{(\nu)}, \Gamma_{jm}^{(\mu)} \right\rangle = \frac{g}{n_\mu} \delta_{\mu\nu} \delta_{jk} \delta_{mn}, \quad (5.20)$$

sendo que o último resultado serve para definir a norma de $\Gamma_{jm}^{(\mu)}(g)$. Portanto, conclui-se que o conjunto Γ é formado por vetores mutuamente ortogonais frente ao produto interno definido. Este resultado é muito interessante porque a cardinalidade de qualquer conjunto de vetores mutuamente ortogonais nunca pode exceder a dimensão do espaço vetorial,¹² ou seja,

$$|\Gamma| = \sum_{\mu=1}^N n_\mu^2 \leq g. \text{ Na verdade, } |\Gamma| = g. \quad (5.21)$$

Posteriormente será demonstrado que, de fato, $|\Gamma| = g$.¹³ Este resultado é importante porque fornece uma informação adicional sobre o número e as dimensões das irreps de um grupo finito.

Os exercícios a seguir ilustram a construção das irreps do grupo S_3 e aplicações do teorema da ortogonalidade às mesmas.

Exercício 5.1. Encontre o conjunto de representações irredutíveis do grupo S_3 .

Resolução: De acordo com a condição (5.21), o número de irreps unitárias deste grupo e suas dimensões deve satisfazer $|\Gamma| = |S_3| = 6$. Existem duas possibilidades: ou todas as irreps têm dimensão 1, existindo então 6 irreps não equivalentes, ou há 2 irreps de dimensão 1 e uma de dimensão 2, pois $1^2 + 1^2 + 2^2 = 6$. Será mostrado na seção 5.6.5 que a segunda opção é a correta. *Irreps de dimensão 1:* de acordo com a tabela de multiplicações do S_3 , obtida no exercício 3.4, $\pi_2^2 = \pi_3^2 = \pi_6^2 = I$. Então, a representação $\Gamma(S_3)$ de dimensão 1 deve ser tal que

$$\Gamma(\pi_2^2) = \Gamma(\pi_2)\Gamma(\pi_2) = [\Gamma(\pi_2)]^2 = \Gamma(I) = 1,$$

ou seja, $\Gamma(\pi_2) = \pm 1$, com o mesmo ocorrendo para π_3 e π_6 . Por outro lado, $\pi_4^3 = \pi_5^3 = I$, de onde se conclui que

$$[\Gamma(\pi_4)]^3 = [\Gamma(\pi_5)]^3 = 1 \implies \Gamma(\pi_4), \Gamma(\pi_5) = 1, t, t^2, \text{ sendo } t = -\frac{1}{2}(1 - i\sqrt{3}).$$

Mas, como $\pi_2 \circ \pi_3 = \pi_4$, segue que $\Gamma(\pi_2)\Gamma(\pi_3) = \Gamma(\pi_4)$. Da mesma forma, como $\pi_3 \circ \pi_2 = \pi_5$, segue que $\Gamma(\pi_3)\Gamma(\pi_2) = \Gamma(\pi_5)$. Portanto, existem somente duas escolhas possíveis de irreps de dimensão 1 que satisfazem a tabela de S_3 :

$$\begin{aligned} \Gamma^{(1)}(S_3) : & \Gamma^{(1)}(g) = 1, \forall g \in S_3 \\ \Gamma^{(2)}(S_3) : & \begin{cases} \Gamma^{(2)}(I) = \Gamma^{(2)}(\pi_4) = \Gamma^{(2)}(\pi_5) = 1 \\ \Gamma^{(2)}(\pi_2) = \Gamma^{(2)}(\pi_3) = \Gamma^{(2)}(\pi_6) = -1. \end{cases} \end{aligned}$$

Ambas são representações degeneradas.

Irrep de dimensão 2: busca-se agora a representação $\Gamma^{(3)}(S_3) \subset GL(2, \mathbb{R})$. Já de início, $\Gamma^{(3)}(I) = I_2$. Para os demais elementos, usa-se novamente a informação contida na tabela do grupo para obter as matrizes na forma

$$\Gamma^{(3)}(\pi_i) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \text{ para } i = 2, \dots, 6 \text{ e com } a, b, c, d \in \mathbb{R}.$$

¹²Ver teoremas e definições na seção 4.3.

¹³A demonstração será realizada na página 200.

Novamente, como $\pi_2^2 = \pi_3^2 = \pi_6^2 = I$, então, para $i = 2, 3, 6$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \implies \begin{matrix} a^2 + bc = 1 & ab + bd = 0 \\ ac + cd = 0 & bc + d^2 = 1. \end{matrix}$$

Dentre as equações acima,

$$ab + bd = b(a + d) = 0 \implies d = -a \text{ ou } \boxed{b = 0}.$$

Toma-se a opção destacada. Da mesma forma,

$$ac + cd = 0 \implies d = -a \text{ ou } \boxed{c = 0}.$$

Das equações restantes, resulta então que

$$a^2 + bc = 1 \implies \boxed{a = \pm 1} \quad bc + d^2 = 1 \implies \boxed{d = \pm 1}.$$

Há 4 combinações de sinais, com 3 disponíveis para uma representação fiel. Contudo, somente para um elemento a escolha é adequada. Realizando as outras escolhas para os demais elementos, o resultado não seguirá a tabela do grupo. Assim, escolhe-se somente

$$\boxed{\Gamma^{(3)}(\pi_2) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}},$$

uma vez que para essa escolha a matriz é ortogonal: $\Gamma^{(3)}(\pi_2) \widetilde{\Gamma^{(3)}(\pi_2)} = I_2$.

Buscam-se então matrizes também ortogonais para os demais elementos, ou seja, para $i = 3, \dots, 6$, matrizes $\Gamma^{(3)}(\pi_i)$ tais que

$$\left[\Gamma^{(3)}(\pi_i) \right]^{-1} = \widetilde{\Gamma^{(3)}(\pi_i)}.$$

Então, se

$$\Gamma^{(3)}(\pi_i) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \left[\Gamma^{(3)}(\pi_i) \right]^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix},$$

resultando em

$$\begin{matrix} \frac{d}{ad - bc} = a & \frac{b}{ad - bc} = -c \\ \frac{c}{ad - bc} = -b & \frac{a}{ad - bc} = d. \end{matrix}$$

Então,

$$\frac{d}{ad - bc} = d(ad - bc) \implies ad - bc = \pm 1.$$

Escolhendo $ad - bc = 1$, $a = d$ e $b = -c$, escreve-se

$$\Gamma^{(3)}(\pi_i) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \text{ com } \boxed{a^2 + b^2 = 1}.$$

Como $\pi_4^3 = I$, resulta

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^3 = \begin{pmatrix} a^3 - 3ab^2 & 3a^2b - b^3 \\ -3a^2b + b^3 & a^3 - 3ab^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \implies \begin{matrix} a^3 - 3ab^2 = 1 \\ 3a^2b - b^3 = 0 \end{matrix} \implies \left\{ \begin{matrix} b = 0 \\ 3a^2 - b^2 = 0 \end{matrix} \right\}.$$

Então,

$$3a^2 = 1 - a^2 \implies 4a^2 = 1 \implies \boxed{a = -\frac{1}{2}} \text{ e } b^2 = 1 - a^2 \implies \boxed{b = \frac{\sqrt{3}}{2}}.$$

Portanto,

$$\boxed{\Gamma^{(3)}(\pi_4) = \frac{1}{2} \begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}}.$$

Tabela 5.1: Representações irreduzíveis ortogonais do grupo S_3 .

	Irreps de S_3		
	$\Gamma(1)$	$\Gamma(2)$	$\Gamma(3)$
n_μ	1	1	2
π_i			
I	1	1	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
π_2	1	-1	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
π_3	1	-1	$\frac{1}{2} \begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$
π_4	1	1	$\frac{1}{2} \begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$
π_5	1	1	$\frac{1}{2} \begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$
π_6	1	-1	$\frac{1}{2} \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$

Agora, como $\pi_2 \circ \pi_3 = \pi_4$,

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} \Rightarrow \begin{matrix} a = 1/2 & b = -\sqrt{3}/2 \\ c = -\sqrt{3}/2 & d = -1/2 \end{matrix} \Rightarrow \Gamma^{(3)}(\pi_3) = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}.$$

Por outro lado, como $\pi_5 = \pi_3 \circ \pi_2$,

$$\Gamma^{(3)}(\pi_5) = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \Gamma^{(3)}(\pi_5) = \frac{1}{2} \begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

Finalmente, como $\pi_6 = \pi_2 \circ \pi_5$,

$$\Gamma^{(3)}(\pi_6) = \frac{1}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix} \Rightarrow \Gamma^{(3)}(\pi_6) = \frac{1}{2} \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

Verifica-se facilmente que todas as matrizes obtidas são ortogonais e que esta representação é fiel. A tabela 5.1 destaca as irreps do grupo S_3 .

Exercício 5.2. Verifique o teorema da ortogonalidade para as representações irreduzíveis do grupo S_3 .

Resolução. Dadas as irreps na tabela 5.1 e o teorema (5.18), é possível realizar várias combinações distintas. Por exemplo:

(a) $\mu = \nu = 1$. Neste caso, $j = m = n = k = 1$ e $\Gamma_{jm}^{(1)}(a_i) = \Gamma_{nk}^{(1)}(a_i^{-1}) = 1$, resultando em

$$1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2 = \frac{6}{1}.$$

(b) $\mu = \nu = 2$. Neste caso,

$$1.1 + 1.1 + (-1)(-1) + (-1)(-1) + 1.1 + 1.1 + (-1)(-1) = \frac{6}{1}.$$

(c) $\mu = 1, \nu = 2$. Neste caso,

$$1.1 + 1.1 + 1.(-1) + 1.(-1) + 1.1 + 1.(-1) = 0.$$

(d) $\mu = 1, \nu = 3$. Tomando $j = m = 1$ e $k = 1, n = 2$,

$$1.0 + 1.0 + 1. \left(-\frac{\sqrt{3}}{2} \right) + 1. \frac{\sqrt{3}}{2} + 1. \left(-\frac{\sqrt{3}}{2} \right) + 1. \frac{\sqrt{3}}{2} = 0.$$

(e) $\mu = \nu = 3$. Tomando $j = m = k = n = 2$:

$$1.1 + 1.1 + \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) + \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) + \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) + \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) = \frac{6}{2}.$$

(f) $\mu = \nu = 3$. Tomando $j = 1, m = 2, k = 1, n = 2$:

$$0.0 + 0.0 + \left(-\frac{\sqrt{3}}{2}\right)\left(-\frac{\sqrt{3}}{2}\right) + \frac{\sqrt{3}\sqrt{3}}{2} + \left(-\frac{\sqrt{3}}{2}\right)\left(-\frac{\sqrt{3}}{2}\right) + \frac{\sqrt{3}\sqrt{3}}{2} = \frac{6}{2}.$$

(g) $\mu = \nu = 3$. Tomando $j = 1, m = 2, k = 2, n = 1$:

$$0.0 + 0.0 + \left(-\frac{\sqrt{3}}{2}\right)\left(-\frac{\sqrt{3}}{2}\right) + \frac{\sqrt{3}}{2}\left(-\frac{\sqrt{3}}{2}\right) + \left(-\frac{\sqrt{3}}{2}\right)\frac{\sqrt{3}}{2} + \frac{\sqrt{3}\sqrt{3}}{2} = 0.$$

Observa-se que todos os resultados acima satisfazem o teorema da ortogonalidade. Diversas outras combinações são possíveis, mas que não serão apresentadas aqui.

Observa-se aqui que qualquer grupo possui sempre a representação irreduzível trivial

$$\Gamma(a) = 1, \text{ para todo } a \in \mathcal{G}.$$

Esta representação não é fiel, mas permite a obtenção de um conjunto de relações entre os elementos de quaisquer outra representação a partir do teorema da ortogonalidade. Colocando, em (5.18), $\Gamma_{kn}^{(\nu)*}(a_i) = 1$ e assumindo que $\mu \neq \nu$, resulta

$$\sum_{a_i \in \mathcal{G}} \Gamma_{jm}^{(\mu)}(a_i) = 0, \quad \text{ou} \quad \sum_{\substack{a_i \in \mathcal{G} \\ a_i \neq I}} \Gamma_{jm}^{(\mu)}(a_i) = -\delta_{jm},$$

uma vez que a representação da identidade é sempre I_{n_μ} . Esta expressão fornece um sistema de g equações que os elementos da μ -ésima representação devem satisfazer. Estas equações são úteis, por exemplo, para verificar se uma determinada representação é realmente irreduzível e unitária. Verifica-se facilmente que os elementos das representações $\Gamma^{(2)}(S_3)$ e $\Gamma^{(3)}(S_3)$, apresentadas na tabela 5.1, satisfazem estas relações.

5.6.3 TEOREMAS SOBRE CARACTERES

Serão apresentados agora teoremas importantes concernentes aos caracteres de uma representação.

Teorema 5.4. *Dado o grupo $\mathcal{G} = \{G; *\}$ e uma representação $\rho : \mathcal{G} \mapsto D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$, os elementos de \mathcal{G} pertencentes a uma dada classe de conjugação possuem os mesmos caracteres em ρ .*

Demonstração. Retornando à definição 3.16 de uma classe de conjugação, dado o elemento $a \in \mathcal{G}$, a sua classe de conjugação é dada por

$$C_a = \{b^{-1} * a * b, \forall b \in \mathcal{G}\}.$$

Então, para qualquer $c \in C_a$, existe um elemento $d \in \mathcal{G}$ tal que $c = d^{-1} * a * d$. Neste caso, as respectivas representações destes elementos devem satisfazer

$$D(c) = D(d^{-1} * a * d) = D(d^{-1}) D(a) D(d) = [D(d)]^{-1} D(a) D(d).$$

Em consequência, o caractere do elemento c na representação ρ é dado por (5.6) e deve satisfazer

$$\begin{aligned} \chi_\rho(c) &= \text{Tr}[D(c)] = \text{Tr}\left\{[D(d)]^{-1} D(a) D(d)\right\} \\ &= \text{Tr}\left\{D(a) [D(d)]^{-1} D(d)\right\} = \text{Tr}\{D(a)\} = \chi_\rho(a). \end{aligned}$$

Portanto, todos os elementos da classe C_a possuem o mesmo caractere na representação ρ . \square

Dado o grupo \mathcal{G} , suas diferentes representações são distinguidas pelo índice “ μ ” em $D^{(\mu)}(\mathcal{G})$. Como o caractere um determinado elemento de \mathcal{G} depende da representação empregada em particular, é conveniente então distinguir os caracteres também com o mesmo índice empregado para a representação; ou seja, se a representação de $g \in \mathcal{G}$ é $D^{(\mu)}(g)$, então o seu caractere nesta representação pode ser identificado por $\chi^{(\mu)}(g)$. Quando o grupo possui mais de um classe de conjugação, é útil também distinguir as mesmas para uma dada representação. Neste caso, o símbolo $\chi_a^{(\mu)}$ indica o caractere dos elementos da classe C_a na representação μ .

Para um determinado grupo finito, muitas vezes é útil construir a sua *tabela de caracteres*, a qual indica tanto a partição de seus elementos em classes de conjugação quanto os valores dos caracteres para cada representação irredutível. Os exemplo a seguir obtém a tabela de caracteres do grupo S_3 .

Exemplo 5.3 (Tabela de caracteres do grupo S_3). Dado o grupo S_3 , cuja tabela de multiplicações foi obtida no exercício 3.4, as suas classes de conjugação são: a classe $C_I = \{I\}$, a classe C_{π_2} , composta por $\pi_i^{-1} \circ \pi_2 \circ \pi_i$, para $i = 1, \dots, 6$:

$$\left\{ \begin{array}{c} I \\ \pi_2 \\ \pi_3 \\ \pi_5 \\ \pi_4 \\ \pi_6 \end{array} \right\} \circ \pi_2 \circ \left\{ \begin{array}{c} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{array} \right\} = \left\{ \begin{array}{c} \pi_2 \\ \pi_2 \\ \pi_6 \\ \pi_6 \\ \pi_3 \\ \pi_3 \end{array} \right\} \implies C_{\pi_2} = \{\pi_2, \pi_3, \pi_6\} = C_{\pi_3} = C_{\pi_6},$$

e a classe C_{π_4} , composta por

$$\left\{ \begin{array}{c} I \\ \pi_2 \\ \pi_3 \\ \pi_5 \\ \pi_4 \\ \pi_6 \end{array} \right\} \circ \pi_4 \circ \left\{ \begin{array}{c} I \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \pi_5 \\ \pi_6 \end{array} \right\} = \left\{ \begin{array}{c} \pi_4 \\ \pi_5 \\ \pi_5 \\ \pi_4 \\ \pi_4 \\ \pi_5 \end{array} \right\} \implies C_{\pi_4} = \{\pi_4, \pi_5\} = C_{\pi_5}.$$

Portanto, de acordo com a tabela 5.1 de irreps do S_3 , obtém-se a tabela 5.2 de caracteres do grupo.

Tabela 5.2: Tabela de caracteres do grupo S_3 . Os caracteres de cada representação são identificados por índices que seguem a classificação dos irreps do S_3 apresentados na tabela 5.1.

Caracteres	Classes	C_I	C_{π_2}	C_{π_4}
		$\{I\}$	$\{\pi_2, \pi_3, \pi_6\}$	$\{\pi_4, \pi_5\}$
$\chi^{(1)}$		+1	+1	+1
$\chi^{(2)}$		+1	-1	+1
$\chi^{(3)}$		+2	0	-1

Obtém-se agora o teorema de ortogonalidade dos caracteres a partir do teorema (5.18).

Teorema 5.5 (Ortogonalidade dos caracteres). *Seja \mathcal{G} um grupo de ordem g , com representações irredutíveis unitárias $\Gamma^{(\mu)}(\mathcal{G})$ e $\Gamma^{(\nu)}(\mathcal{G})$, de dimensões n_μ e n_ν , respectivamente. Se C_a é a classe de conjugação do elemento $a \in \mathcal{G}$, c_a é a cardinalidade de C_a , e n_C é o número de classes de conjugação de \mathcal{G} , então*

$$\sum_{k=1}^{n_C} \sqrt{\frac{c_k}{g}} \chi_k^{(\mu)} \sqrt{\frac{c_k}{g}} \chi_k^{(\nu)*} = \delta_{\mu\nu}, \quad (5.22a)$$

onde $\chi_k^{(\mu)}$ é o caractere de qualquer elemento na classe C_k na representação $\Gamma^{(\mu)}(\mathcal{G})$ e a soma é realizada sobre todas as classes de conjugação distintas do grupo \mathcal{G} .

Demonstração. Partindo do teorema de ortogonalidade (5.18), coloca-se $j = m$ e $n = k$ e soma-se sobre m e n , resultando

$$\sum_{m=1}^{n_\mu} \sum_{n=1}^{n_\nu} \sum_{a_i \in \mathcal{G}} \Gamma_{mm}^{(\mu)}(a_i) \Gamma_{nn}^{(\nu)*}(a_i) = \frac{g}{n_\mu} \delta_{\mu\nu} \sum_{m=1}^{n_\mu} \sum_{n=1}^{n_\nu} \delta_{mn}.$$

Mas,

$$\sum_{m=1}^{n_\mu} \Gamma_{mm}^{(\mu)}(a_i) = \chi^{(\mu)}(a_i), \quad \sum_{n=1}^{n_\nu} \Gamma_{nn}^{(\nu)*}(a_i) = \chi^{(\nu)*}(a_i) \quad \text{e} \quad \delta_{\mu\nu} \sum_{m=1}^{n_\mu} \sum_{n=1}^{n_\nu} \delta_{mn} = \delta_{\mu\nu} n_\mu,$$

resultando

$$\sum_{a_i \in \mathcal{G}} \chi^{(\mu)}(a_i) \chi^{(\nu)*}(a_i) = g \delta_{\mu\nu}. \quad (5.22b)$$

A soma sobre os elementos a_i pode ser agora particionada em conjuntos contendo os elementos das classes de conjugação de \mathcal{G} , os quais possuem todos os mesmos caracteres. Desta maneira, o resultado acima pode ser escrito

$$\sum_{k=1}^{n_C} c_k \chi_k^{(\mu)} \chi_k^{(\nu)*} = g \delta_{\mu\nu},$$

sendo que a soma agora é sobre as classes de \mathcal{G} , $\chi_k^{(\mu)}$ é o caractere dos elementos na classe C_k e na representação μ e c_k é o número de elementos na classe. Obtém-se assim o resultado (5.22). \square

Como sempre existe, para qualquer grupo, uma irrep unitária ou trivial $\Gamma^{(1)}(a) = +1$ para todo $a \in \mathcal{G}$, os caracteres desta irrep serão os mesmos em qualquer classe, ou seja, $\chi^{(1)}(a) = \chi_k^{(1)} = +1$ ($\forall a, k$). Colocando então $\nu = 1$ na condição de ortogonalidade (5.22a), os caracteres de qualquer outra representação $\mu \neq 1$ devem satisfazer

$$\sum_{k=1}^{n_C} c_k \chi_k^{(\mu)} = 0, \quad (\mu \neq 1). \quad (5.22c)$$

O teorema de ortogonalidade dos caracteres possui também uma interpretação geométrica com base na teoria de espaços vetoriais, de forma semelhante à interpretação realizada para o teorema da ortogonalidade das representações.

5.6.4 INTERPRETAÇÃO DO TEOREMA DA ORTOGONALIDADE DOS CARACTERES

Dado o grupo finito \mathcal{G} de ordem g , se o mesmo possui n_C classes de conjugação, sendo $c_k = |C_k|$ a cardinalidade da k -ésima classe, então pode-se definir a tupla

$$\hat{\chi}^{(\mu)} \doteq \left(\sqrt{\frac{c_1}{g}} \chi_1^{(\mu)}, \sqrt{\frac{c_2}{g}} \chi_2^{(\mu)}, \dots, \sqrt{\frac{c_{n_C}}{g}} \chi_{n_C}^{(\mu)} \right), \quad (5.23a)$$

a qual é composta pelos caracteres de cada classe na representação μ . Assim, cada tupla $\hat{\chi}^{(\mu)}$ é distinguida pelo índice que identifica a representação do grupo, existindo tantas tuplas quantas representações. Restringindo às representações irredutíveis, assume-se que existem, ao todo, N irreps de \mathcal{G} . Neste caso, o conjunto

$$\chi \doteq \left\{ \hat{\chi}^{(1)}, \hat{\chi}^{(2)}, \dots, \hat{\chi}^{(N)} \right\} \quad (5.23b)$$

é composto por $N = |\chi|$ tuplas.

Como a quantidade $\chi_k^{(\mu)} \in \mathbb{K}$ ($\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$), a tupla $\hat{\chi}^{(\mu)}$ é um vetor, denominado *vetor de caracteres*, do espaço vetorial \mathcal{K}^{n_C} sobre o corpo \mathbb{K} , com dimensão n_C . Definindo então a produto interno do vetor ψ pelo vetor ζ (com $\psi, \zeta \in \mathcal{K}^{n_C}$) por

$$\langle \psi, \zeta \rangle = \sum_{i=1}^{n_C} \psi_i \zeta_i^*,$$

o teorema (5.22) mostra que

$$\left\langle \hat{\chi}^{(\mu)}, \hat{\chi}^{(\nu)} \right\rangle = \sum_{i=1}^{n_C} \hat{\chi}_i^{(\mu)} \hat{\chi}_i^{(\nu)*} = \sum_{i=1}^{n_C} \sqrt{\frac{c_i}{g}} \chi_i^{(\mu)} \sqrt{\frac{c_i}{g}} \chi_i^{(\nu)*} = \delta_{\mu\nu}. \quad (5.23c)$$

Portanto, o conjunto $\chi \subset \mathcal{K}^{n_C}$ é composto por vetores ortonormais do espaço vetorial.

Como a cardinalidade de um conjunto contendo vetores ortogonais não pode exceder a dimensão do espaço vetorial, resulta que

$$|\chi| = N \leq n_C. \quad (5.24)$$

Será mostrado posteriormente que, na verdade, $N = n_C$, ou seja, o número de representações irredutíveis de \mathcal{G} é igual ao número de suas classes de conjugação. Este resultado também é útil para a determinação do número total de representações irredutíveis de um grupo finito.

Exemplo 5.4 (Número de Irreps de S_3). O número de classes de conjugação do grupo S_3 é $n_C = 3$, como pode ser visto na tabela 5.2. Portanto,

$$N_{S_3} = n_C = 3,$$

sendo N_{S_3} o número de representações irredutíveis do grupo. Combinando esta informação com a expressão (5.21), conclui-se que realmente o único conjunto possível de Irreps do S_3 é aquele obtido no exercício 5.1.

Exemplo 5.5 (Tabela de caracteres do grupo C_{1h}). O grupo C_{1h} , dado por (3.7), é composto somente pela identidade E e pela reflexão σ_h . Por ser um grupo Abeliano, ambos elementos formam classes de conjugação por si próprios e, por conseguinte, $N = n_C = |C_{1h}| = 2$. Isto implica que há somente duas representações irredutíveis, ambas unidimensionais. Como $\Gamma^{(1)}(C_{1h}) = \chi^{(1)}(C_{1h}) = +1$, a outra irrep, para satisfazer a tabela de multiplicações do grupo, deve ser $\Gamma^{(2)}(E) = \chi^{(2)}(E) = +1$ e $\Gamma^{(2)}(\sigma_h) = \chi^{(2)}(\sigma_h) = -1$. Resulta então a tabela 5.3.

Tabela 5.3: Tabela de caracteres do grupo C_{1h} .

Caracteres	Classes	C_E	C_{σ_h}
		$\{I\}$	$\{\sigma_h\}$
$\chi^{(1)}$		+1	+1
$\chi^{(2)}$		+1	-1

Finalmente, existe também uma relação de completeza para os caracteres.

Teorema 5.6 (Relação de completeza de caracteres). Seja \mathcal{G} um grupo de ordem g e $\chi_k^{(\mu)}$ o caractere de qualquer elemento na classe de conjugação C_k e na representação irredutível unitária $\Gamma^{(\mu)}(\mathcal{G})$. Então,

$$\sum_{\mu=1}^N \sqrt{\frac{c_k}{g}} \chi_k^{(\mu)} \sqrt{\frac{c_\ell}{g}} \chi_\ell^{(\mu)*} = \delta_{k\ell}, \quad (5.25)$$

sendo c_k a cardinalidade da classe C_k .

5.6.5 DECOMPOSIÇÃO DE UMA REPRESENTAÇÃO EM IRREPS

Uma vez estabelecidos os principais teoremas envolvendo representações de grupos e seus caracteres, pode-se retornar à questão inicial acerca da expressão ou decomposição de uma representação qualquer em termos de representações irredutíveis de mais baixa dimensionalidade.

Sendo \mathcal{G} um grupo finito, supõe-se a existência de uma representação $D(\mathcal{G}) \subseteq GL(n, \mathbb{K})$, a qual pode ser uma representação regular ou natural, por exemplo, sendo assumido de $D(\mathcal{G})$ não é uma irrep. Os teoremas apresentados anteriormente e suas consequências permitem estabelecer, para qualquer grupo finito, o número (N) e as dimensões ($\{n_\mu\}$) de suas irreps. Então, sempre é possível expressar a matriz $D(a_i)$ (para todo $a_i \in \mathcal{G}$) como a soma direta das irreps de \mathcal{G} dada na expressão (5.16b), ou seja,

$$D(a_i) = \bigoplus_{\mu=1}^N m_\mu \Gamma^{(\mu)}(a_i) = m_1 \Gamma^{(1)}(a_i) \oplus m_2 \Gamma^{(2)}(a_i) \oplus \dots \oplus m_N \Gamma^{(N)}(a_i), \quad (5.26)$$

sendo que as irreps são supostas unitárias. Os únicos parâmetros ainda indeterminados na expressão acima são os índices $\{m_\mu\}$, os quais indicam quantas vezes uma determinada irrep irá aparecer na soma direta. Para determinar estes índices, calcula-se o traço das matrizes envolvidas, o que significa a obtenção dos caracteres das representações. Com o auxílio da propriedade (5.8b) da soma direta, resulta

$$\chi(a_i) = \sum_{\nu=1}^N m_\nu \chi^{(\nu)}(a_i), \quad (5.27)$$

sendo $\chi(a_i)$ o caractere de a_i na representação $D(a_i)$.

Os caracteres $\{\chi^{(\nu)}(a_i)\}$ das irreps são denominados **caracteres primitivos** ou **caracteres simples**, ao passo que os caracteres $\chi(a_i)$ de uma representação redutível são denominados **caracteres compostos**, como o resultado recém obtido indica. Multiplicando ambos os lados por $\chi^{(\mu)*}(a_i)$ e somando sobre todos os elemento de \mathcal{G} , resulta

$$\sum_{i=1}^g \chi(a_i) \chi^{(\mu)*}(a_i) = \sum_{\nu=1}^N m_{\nu} \sum_{i=1}^g \chi^{(\nu)}(a_i) \chi^{(\mu)*}(a_i) \implies \sum_{k=1}^{n_C} c_k \chi_k \chi_k^{(\mu)*} = \sum_{\nu=1}^N m_{\nu} \sum_{k=1}^{n_C} c_k \chi_k^{(\nu)} \chi_k^{(\mu)*}.$$

Na última expressão acima, os caracteres dos elementos foram agrupados por classes de conjugação, sendo c_k o número de elementos na k -ésima classe, n_C o número total de classes de conjugação em \mathcal{G} e χ_k o caractere da k -ésima classe na representação $D(a_i)$. Usando então o teorema (5.22), resulta

$$m_{\mu} = \frac{1}{g} \sum_{k=1}^{n_C} c_k \chi_k \chi_k^{(\mu)*}. \tag{5.28}$$

Este resultado permite finalmente a decomposição completa de uma representação em termos das irreps do grupo.

Um outro resultado importante envolvendo os índices $\{m_{\mu}\}$ pode ser obtido. Multiplicando (5.27) pelo seu complexo conjugado e somando sobre os elementos de \mathcal{G} , obtém-se

$$\begin{aligned} \sum_{i=1}^g \chi(a_i) \chi^*(a_i) &= \sum_{\mu, \nu=1}^N m_{\mu} m_{\nu} \sum_{i=1}^g \chi^{(\nu)}(a_i) \chi^{(\mu)*}(a_i), \\ \sum_{k=1}^{n_C} c_k |\chi_k|^2 &= \sum_{\mu, \nu=1}^N m_{\mu} m_{\nu} \sum_{k=1}^{n_C} c_k \chi_k^{(\nu)} \chi_k^{(\mu)*} = g \sum_{\mu=1}^N m_{\mu}^2. \end{aligned}$$

Se a representação $D(a_i)$ já era irredutível, então todos os índices m_{μ} são nulos, exceto por um que é igual à unidade. Portanto, uma condição necessária e suficiente para que uma dada representação seja irredutível é

$$\sum_{i=1}^g |\chi(a_i)|^2 = \sum_{k=1}^{n_C} c_k |\chi_k|^2 = g. \tag{5.29}$$

Verifica-se na tabela 5.2 que os caracteres das irreps do grupo S_3 satisfazem este critério.

Por outro lado, se a representação for redutível, então a expressão

$$\frac{1}{g} \sum_{k=1}^{n_C} c_k |\chi_k|^2 = \sum_{\mu=1}^N m_{\mu}^2$$

fornece informações sobre a decomposição da mesma em irreps.

Exercício 5.3. Estude a decomposição da representação natural do grupo S_3 em irreps.

Resolução. A representação natural $D^{\text{nat}}(S_3)$ foi obtida no exemplo 5.2. Verifica-se que os caracteres da representação são os seguintes:

$$\chi_I = 3, \quad \chi_{\pi_2} = 1, \quad \chi_{\pi_4} = 0.$$

Por sua vez, os caracteres das irreps são dados na tabela 5.2. Assim, da expressão (5.28), com $k = I, \pi_2, \pi_4$, resulta

$$\begin{aligned} m_1 &= \frac{1}{6} \sum_{k=1}^3 c_k \chi_k \chi_k^{(1)*} = \frac{1}{6} (1.3.1 + 3.1.1 + 2.0.1) = 1 \\ m_2 &= \frac{1}{6} \sum_{k=1}^3 c_k \chi_k \chi_k^{(2)*} = \frac{1}{6} (1.3.1 + 3.1.(-1) + 2.0.1) = 0 \\ m_3 &= \frac{1}{6} \sum_{k=1}^3 c_k \chi_k \chi_k^{(3)*} = \frac{1}{6} (1.3.2 + 3.1.0 + 2.0.(-1)) = 1. \end{aligned}$$

Portanto,

$$D^{\text{nat}}(S_3) = \Gamma^{(1)}(S_3) \oplus \Gamma^{(3)}(S_3).$$

Será apresentado agora um teorema relacionado com a decomposição de uma representação regular em irreps.

Teorema 5.7. *Dado um grupo \mathcal{G} , a sua representação regular contém todas as suas representações irreduzíveis, com multiplicidades iguais às dimensões destas irreps.*

Demonstração. Seja a representação regular (5.4) de um grupo de ordem g . A representação do elemento I será sempre $D^{\text{reg}}(I) = I_g$ e, portanto, $\chi^{\text{reg}}(I) = g$. Por outro lado, para os demais elementos $g_k \in \mathcal{G}$ as diagonais principais das respectivas representações serão dadas por

$$D_{ii}^{\text{reg}}(g_k) = \begin{cases} 1, & \text{se } g_k * g_i = g_i \\ 0, & \text{se } g_k * g_i \neq g_i, \end{cases} \quad (i = 1, \dots, g).$$

Como $g_k \neq I$, então $g_k * g_i \neq g_i$ sempre; ou seja, $D_{ii}^{\text{reg}}(g_k) = 0$ ($i = 1, \dots, g$). Portanto, $\chi^{\text{reg}}(g_k) = 0$.

Usando agora a relação (5.27), considera-se a decomposição do caractere composto do elemento identidade em termos dos caracteres do mesmo elemento nas irreps,

$$\chi^{\text{reg}}(I) = g = \sum_{\nu=1}^N m_{\nu} \chi^{(\nu)}(I) = \sum_{\nu=1}^N m_{\nu} n_{\nu}, \quad (5.30)$$

uma vez que $D^{(\nu)}(I) = I_{n_{\nu}}$ sempre. Por outro lado, os índices $\{m_{\nu}\}$ são dados por (5.28),

$$m_{\nu} = \frac{1}{g} \sum_{k=1}^{n_C} c_k \chi_k^{\text{reg}} \chi_k^{(\nu)*}.$$

Contudo, como $\chi_I^{\text{reg}} = g$, $\chi_I^{(\nu)} = n_{\nu}$, $c_I = 1$ e $\chi_k^{\text{reg}} = 0$ para $k \neq I$, resulta que

$$m_{\nu} = n_{\nu}. \quad (5.31)$$

□

Exemplo 5.6 (Decomposição da representação regular do grupo S_3). A representação regular do grupo S_3 foi obtida no exemplo 5.1, enquanto que suas irreps foram obtidas no exercício 5.1. De acordo com (5.31), esta representação é decomposta pelas irreps de acordo com

$$D^{\text{reg}}(S_3) = \Gamma^{(1)}(S_3) \oplus \Gamma^{(2)}(S_3) \oplus \Gamma^{(3)}(S_3) \oplus \Gamma^{(3)}(S_3).$$

O teorema 5.7 será empregado agora para demonstrar a igualdade na expressão (5.21).

Demonstração da expressão (5.21). É possível agora demonstrar a expressão $|\Gamma| = g$. Para tanto, emprega-se a representação regular (5.4) de um grupo de ordem g . A propriedade (5.30) mostra que

$$\sum_{\nu=1}^N m_{\nu} n_{\nu} = g.$$

Mas, levando em conta que para a representação regular a multiplicidade da ν -ésima irrep é dada por (5.31), resulta então que

$$g = \sum_{\nu=1}^N n_{\nu}^2,$$

o que completa a demonstração. □

Uma vez demonstrado que o conjunto (5.19) é uma base do espaço vetorial \mathcal{K}^g , conclui-se que não existe nenhum outro vetor de representação que seja ortogonal aos vetores em Γ . Isto, por sua vez, implica que não existe nenhum outro vetor de caractere que seja ortogonal aos vetores no conjunto χ em (5.23b). Portanto, os vetores de caracteres em χ também devem formar um conjunto completo de vetores de base para o espaço \mathcal{K}^{n_C} e a cardinalidade de χ deve ser assim igual à dimensão do espaço. Isto completa a demonstração da igualdade em (5.24).

Uma última propriedade importante dos caracteres de uma representação pode ser deduzida. Sendo $\Gamma^{(\mu)}(\mathcal{G})$ uma irrep unitária de dimensão n_{μ} , uma irrep equivalente pode ser obtida a partir da transformação de similaridade

$$\Gamma^{(\mu)'}(\mathcal{G}) = S^{-1} \Gamma^{(\mu)}(\mathcal{G}) S,$$

onde S é uma matriz $n_\mu \times n_\mu$ não singular. Agora, para qualquer elemento $a \in \mathcal{G}$, a sua representação $\Gamma^{(\mu)}(a)$ possui um conjunto de autovalores $\{\lambda_j\}$ ($j = 1, \dots, n_\mu$), sendo cada autovalor obtido a partir do autovetor $x_j = (x_{j1} \ x_{j2} \ \dots \ x_{jn_\mu})^T$ pela equação secular

$$\Gamma^{(\mu)}(a)x_j = \lambda_j^{(\mu)}x_j \implies \det \left[\Gamma^{(\mu)}(a) - \lambda^{(\mu)}I_{n_\mu} \right] = 0.$$

Pode-se mostrar que as raízes da equação secular, *i. e.*, os n_μ autovalores de $\Gamma^{(\mu)}(a)$ são unimodulares, ou seja, $|\lambda_j^{(\mu)}| = 1, \forall j = 1, \dots, n_\mu$.

Construindo então a matriz S tal que suas colunas são os autovetores de $\Gamma^{(\mu)}(a)$, *i. e.*,

$$S = \begin{pmatrix} x_1 & x_2 & \dots & x_{n_\mu} \\ \downarrow & \downarrow & \dots & \downarrow \end{pmatrix} = \begin{pmatrix} x_{11} & x_{21} & \dots & x_{n_\mu 1} \\ x_{12} & x_{22} & \dots & x_{n_\mu 2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n_\mu} & x_{2n_\mu} & \dots & x_{n_\mu n_\mu} \end{pmatrix} \implies S_{ij} = x_{ji},$$

resulta que

$$\begin{aligned} \left[S^{-1} \Gamma^{(\mu)}(a) S \right]_{ij} &= \sum_{k, \ell=1}^{n_\mu} (S^{-1})_{ik} \Gamma_{k\ell}^{(\mu)}(a) S_{\ell j} = \sum_{k, \ell=1}^{n_\mu} (S^{-1})_{ik} \Gamma_{k\ell}^{(\mu)}(a) x_{j\ell} \\ &= \sum_{k=1}^{n_\mu} (S^{-1})_{ik} \lambda_j^{(\mu)} x_{jk} = \lambda_j^{(\mu)} \sum_{k=1}^{n_\mu} (S^{-1})_{ik} S_{kj} = \lambda_j^{(\mu)} \delta_{ij}. \end{aligned}$$

Ou seja,

$$\Gamma^{(\mu)'}(a) = \begin{pmatrix} \lambda_1^{(\mu)} & 0 & \dots & 0 \\ 0 & \lambda_2^{(\mu)} & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n_\mu}^{(\mu)} \end{pmatrix}.$$

Como uma transformação de similaridade não altera o valor do caractere, resulta que o caractere do elemento a pode ser calculado como a soma dos autovalores de sua representação, *i. e.*,

$$\chi^{(\mu)}(a) = \chi_a^{(\mu)} = \sum_{j=1}^{n_\mu} \lambda_j^{(\mu)}, \tag{5.32a}$$

sendo $\chi_a^{(\mu)}$ o caractere de todos os elementos da classe de conjugação C_a .

Agora, se o elemento a é tal que $a^m = I$, isto implica que $[\Gamma^{(\mu)}(a)]^m = I_{n_\mu}$ (propriedade 3.8c). Além disso, todos os elementos na classe de conjugação C_a possuem a mesma ordem. Então,

$$\left[\Gamma^{(\mu)'}(a) \right]^m = \left[S^{-1} \Gamma^{(\mu)}(a) S \right]^m = S^{-1} \Gamma^{(\mu)}(a) S S^{-1} \Gamma^{(\mu)}(a) S \dots S^{-1} \Gamma^{(\mu)}(a) S = S^{-1} \left[\Gamma^{(\mu)}(a) \right]^m S,$$

ou seja,

$$\left[\Gamma^{(\mu)'}(a) \right]^m = \begin{pmatrix} \lambda_1^m & 0 & \dots & 0 \\ 0 & \lambda_2^m & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{n_\mu}^m \end{pmatrix} = I_{n_\mu} \implies \left[\lambda_j^{(\mu)} \right]^m = 1 \quad (j = 1, \dots, n_\mu). \tag{5.32b}$$

Ou seja, os autovalores da representação do elemento a são as m raízes da unidade, sendo m a ordem do elemento a .

Embora este último resultado tenha sido obtido com o emprego de uma irrep, o mesmo vale para qualquer representação do grupo.

5.6.6 CONSTRUÇÃO DE UMA TABELA DE CARACTERES

Para que se possa realizar a decomposição de uma dada representação em irreps, é necessário o conhecimento da tabela de caracteres do grupo \mathcal{G} de ordem g . Os resultados obtidos nesta seção possibilitam a construção desta última de uma forma sistemática, bastando para tanto a determinação das seguintes propriedades:

1. O número de irreps é igual ao número de classes de conjugação de \mathcal{G} : $N = n_C$ (propriedade 5.24).
2. As dimensões das irreps de \mathcal{G} satisfazem a relação (propriedade 5.21)

$$\sum_{\mu=1}^N n_{\mu}^2 = g.$$

3. A irrep *identidade* ou *trivial* $\Gamma^{(1)}(a) = +1$ ($n_1 = 1$), para todo $a \in \mathcal{G}$, está sempre presente para todo grupo. Portanto, $\chi^{(1)}(a) = \chi_k^{(1)} = +1$.
4. Para qualquer representação, a irrep da identidade é sempre $\Gamma^{(\mu)}(I) = 1_{n_{\mu}}$. Portanto, $\chi^{(\mu)}(I) = \chi_I^{(\mu)} = n_{\mu}$.
5. Os caracteres da irrep $\Gamma^{(\mu)}(\mathcal{G})$ satisfazem a relação (propriedade 5.29)

$$\sum_{i=1}^g \left| \chi^{(\mu)}(a_i) \right|^2 = \sum_{k=1}^{n_C} c_k \left| \chi_k^{(\mu)} \right|^2 = g \xrightarrow{\text{Passo (4)}_{c_I=1}} \sum_{k=2}^{n_C} c_k \left| \chi_k^{(\mu)} \right|^2 = g - n_{\mu}^2.$$

6. O caractere do elemento a de ordem m na μ -ésima irrep pode ser expresso como a soma dos autovalores de $\Gamma^{(\mu)}(a)$, sendo que cada autovalor é a m -ésima raiz da unidade (propriedade 5.32).
7. O vetor formado pelos caracteres de uma dada irrep é ortogonal ao vetor formado pelos caracteres de uma outra irrep qualquer (propriedades 5.22 e 5.23).
8. O vetor formado pelos caracteres de uma dada classe de conjugação (ao longo de todas as irreps) é ortogonal ao vetor de uma outra classe (propriedade 5.25).

Exemplo 5.7. As classes de conjugação do grupo S_3 são $C_1 = \{I\}$ ($c_1 = 1$), $C_2 = \{\pi_2, \pi_3, \pi_6\}$ ($c_2 = 3$) e $C_3 = \{\pi_4, \pi_5\}$ ($c_3 = 2$). Aplicando sistematicamente as informações contidas no algoritmo acima, resulta:

1. O número de classes de conjugação é $n_C = 3$. Então, o número de irreps do S_3 é $N = n_C = 3$.
2. A ordem do grupo é $g = |S_3| = 6$. Então, as dimensões das irreps devem satisfazer a relação

$$\sum_{\mu=1}^3 n_{\mu}^2 = n_1^2 + n_2^2 + n_3^2 = 6.$$

O único conjunto de números naturais que satisfazem esta equação é $\{1, 1, 2\}$. Portanto, $n_1 = n_2 = 1$ e $n_3 = 2$.

3. A irrep trivial do S_3 é, portanto, $\chi_k^{(1)} = +1$ ($k = 1, 2, 3$). Com isso, fica preenchida a primeira linha da tabela 5.2.
4. Os caracteres das irreps $\mu = 2, 3$ satisfazem as equações

$$\sum_{k=1}^3 c_k \left| \chi_k^{(\mu)} \right|^2 = \left| \chi_1^{(\mu)} \right|^2 + 3 \left| \chi_2^{(\mu)} \right|^2 + 2 \left| \chi_3^{(\mu)} \right|^2 = 6 \implies 3 \left| \chi_2^{(\mu)} \right|^2 + 2 \left| \chi_3^{(\mu)} \right|^2 = \begin{cases} 5, & \mu = 2 \\ 2, & \mu = 3. \end{cases}$$

5. Cada caractere é igual à soma de uma raiz da unidade. Todos os elementos de uma dada classe têm a mesma ordem. Sendo m_k a ordem dos elementos da k -ésima classe, então $m_1 = 1$, $m_2 = 2$ e $m_3 = 3$. Chamando $z_{kp} = \lambda_j^{(\mu)}$, esta quantidade é uma das raízes de $z^{m_k} = 1$, ou seja,¹⁴ $z_{kp} = \exp(2p\pi i/m_k)$ ($k = 1, \dots, n_C$, $p = 0, \dots, m_k - 1$).
6. Os caracteres das irreps $\mu = 2, 3$ também devem satisfazer

$$\sum_{k=1}^3 c_k \chi_k^{(\mu)} = \chi_1^{(\mu)} + 3\chi_2^{(\mu)} + 2\chi_3^{(\mu)} = 0 \implies 3\chi_2^{(\mu)} + 2\chi_3^{(\mu)} = \begin{cases} -1, & \mu = 2 \\ -2, & \mu = 3. \end{cases}$$

¹⁴Ver fórmula (2.9).

Para obter os caracteres das representações $\mu = 2, 3$ deve-se realizar escolhas que satisfazem as condições impostas pelos passos (4) a (6) do algoritmo.

- $\mu = 2$. Esta representação também é unidimensional ($n_2 = 1$). Então, $\chi_1^{(2)} = +1$. Para satisfazer as condições dos passos (4) – (6):

$$3 \left| \chi_2^{(2)} \right|^2 + 2 \left| \chi_3^{(2)} \right|^2 = 5 \quad 3\chi_2^{(2)} + 2\chi_3^{(2)} = -1,$$

$$\chi_2^{(2)} = \lambda_2^{(2)}, \left[\lambda_2^{(2)} \right]^2 = 1 \Rightarrow \lambda_2^{(2)} = \pm 1; \quad \chi_3^{(2)} = \lambda_3^{(2)}, \left[\lambda_3^{(2)} \right]^3 = 1 \Rightarrow \lambda_3^{(2)} = 1, e^{2\pi i/3} \text{ ou } e^{4\pi i/3}.$$

As únicas escolhas possíveis são $\chi_2^{(2)} = -1$ e $\chi_3^{(2)} = +1$, resultando preenchida a segunda linha da tabela 5.2.

- $\mu = 3$. Agora, $n_3 = 2$, implicando em $\chi_1^{(3)} = +2$. O passo (7) demanda que

$$\chi_k^{(2)} \chi_\ell^{(2)*} + \chi_k^{(3)} \chi_\ell^{(3)*} = \frac{g}{c_k} \delta_{k\ell} - 1 \Rightarrow \begin{cases} \chi_2^{(2)} \chi_3^{(2)*} + \chi_2^{(3)} \chi_3^{(3)*} = -1 \\ \chi_2^{(2)} \chi_2^{(2)*} + \chi_2^{(3)} \chi_2^{(3)*} = 1 \\ \chi_3^{(2)} \chi_3^{(2)*} + \chi_3^{(3)} \chi_3^{(3)*} = 2 \end{cases} \Rightarrow \begin{cases} \chi_2^{(3)} \chi_3^{(3)*} = 0 \\ \left| \chi_2^{(3)} \right|^2 = 0 \\ \left| \chi_3^{(3)} \right|^2 = 1. \end{cases}$$

Ou seja, $\chi_2^{(3)} = 0$. O valor de $\left| \chi_3^{(3)} \right|^2$ também poderia ter sido obtido com o passo (4). Já o valor de $\chi_3^{(3)}$ deve ser consistente com o passo (6), ou seja,

$$3\chi_2^{(3)} + 2\chi_3^{(3)} = -2 \implies \chi_3^{(3)} = -1,$$

o que completa a terceira linha da tabela 5.2. Nota-se que este último resultado está consistente com o passo (5), o qual exige que

$$\chi_3^{(3)} = \sum_{j=1}^2 \lambda_{3p_j}, \text{ sendo } \lambda_{3p} = e^{2p\pi i/3}, (p = 0, 1, 2).$$

Fazendo a escolha $p_1 = 1$ e $p_2 = 2$, resulta

$$\chi_3^{(3)} = e^{2\pi i/3} + e^{4\pi i/3} = e^{i\pi} \left(e^{-i\pi/3} + e^{i\pi/3} \right) = -2 \cos(\pi/3) = -1.$$

As tabelas de caracteres de grupos de ordens mais altas podem ser sistematicamente construídas seguindo-se o mesmo algoritmo.

5.7 BASES SIMETRIZADAS PARA REPRESENTAÇÕES IRREDUTÍVEIS

Chega-se então ao problema prático de como se obter e reduzir uma representação em geral. Como já foi mencionado na seção 5.2, em muitos problemas na física o estado do sistema é descrito por um vetor em um espaço vetorial. Operadores neste espaço (lineares ou não) correspondem a transformações realizadas sobre o sistema físico, levando o mesmo a um outro estado, descrito por um outro vetor no mesmo espaço. Estas transformações podem ser divididas em grupos abstratos os quais, dada uma base ortonormal adequada, serão representados por matrizes do $GL(n, \mathbb{K})$.

O aspecto inconveniente no procedimento recém mencionado está no fato de que os espaços vetoriais em questão muitas vezes são de dimensão alta (em muitos casos infinita) e as representações assim obtidas devem ser descritas por matrizes de ordem alta. É neste ponto que se manifesta a possibilidade de reduzir essas representações em termos das irreps do grupo, como foi discutido ao longo das seções 5.5 e 5.6. Partindo de uma base ortonormal genérica, deseja-se reduzir a representação através de uma escolha adequada de subconjuntos de vetores da base, sendo que cada subconjunto forma a base em um subespaço invariante sob as operações dos

elementos do grupo. Em consequência, a representação genérica inicial será reduzida em termos das irreps de baixa dimensão. Nesta seção será apresentado um método que possibilita a obtenção destes subconjuntos e alguns exemplos serão apresentados.

Retornando à discussão realizada na seção 5.2.2, dado o espaço vetorial \mathcal{V}^n e uma base ortonormal $\Phi \doteq \{\phi_1, \phi_2, \dots, \phi_n\}$ do mesmo e dado o grupo de transformações $\mathcal{T} = \{T; *\}$, cuja representação na base Φ é composta pelas matrizes $D(\mathcal{T})$, qualquer transformação $A \in \mathcal{T}$ terá sua ação sobre o vetor de base ϕ_i dada por

$$A\phi_i = \sum_{j=1}^n \phi_j D_{ji}(A), \quad (5.33)$$

sendo assumido que $D(A)$ é uma representação redutível em termos de irreps de dimensões mais baixas.

Para reduzir a representação $D(\mathcal{T})$, deseja-se encontrar uma matriz unitária U adequada, tal que realizando a transformação de similaridade com a mesma, obtém-se

$$D^{\text{red}}(A) = U^{-1}D(A)U, \quad (\forall A \in \mathcal{T}),$$

onde as matrizes $D^{\text{red}}(\mathcal{T})$ estão todas na forma bloco-diagonal (5.16a), escritas em termos das irreps unitárias de \mathcal{T} . Para tanto, reescreve-se a expressão (5.33) para descrever a ação do operador A sobre todos os vetores da base Φ na forma matricial como

$$A\Phi = \Phi D(A),$$

onde $\Phi \doteq (\phi_1 \phi_2 \dots \phi_n)$ é a matriz linha composta por todos os vetores da base e $A\Phi \doteq (A\phi_1 A\phi_2 \dots A\phi_n)$.

Se U é a matriz desejada, então

$$A\Phi U = \Phi U U^{-1} D(A) U \implies A(\Phi U) = (\Phi U) D^{\text{red}}(A).$$

Este resultado sugere que a base ortonormal adequada para os fins desejados é o conjunto Ψ definido por

$$\Psi \doteq \Phi U = \{\psi_1, \psi_2, \dots, \psi_n\},$$

com vetores $\{\psi_i\}$ tais que

$$\psi_i = (\Phi U)_i = \sum_{j=1}^n \phi_j U_{ji}, \quad (i = 1, \dots, n). \quad (5.34a)$$

Com o propósito de encontrar os elementos de matriz $\{U_{ij}\}$, considera-se agora a derradeira redução de $D(\mathcal{T})$. De (5.26),

$$D(\mathcal{T}) = \bigoplus_{\mu=1}^N m_\mu \Gamma^{(\mu)}(\mathcal{T}),$$

sendo $\Gamma(\mathcal{T}) \doteq \{\Gamma^{(\mu)}(\mathcal{T})\}$ o conjunto das N irreps unitárias do grupo \mathcal{T} . Dentro do conjunto $\Gamma(\mathcal{T})$, a μ -ésima irrep possui dimensão $n_\mu < n$ e, portanto, representa o grupo \mathcal{T} dentro de um subespaço invariante $\mathcal{V}^{n_\mu} \subset \mathcal{V}^n$. É possível então particionar o conjunto Ψ em subconjuntos

$$\Psi_p^{(\mu)} \subset \Psi, \text{ tais que } \sum_{\mu} m_\mu n_\mu = n.$$

O índice p em $\Psi_p^{(\mu)}$ indica a p -ésima ocorrência da irrep $\Gamma^{(\mu)}$ em $D(\mathcal{T})$; ou seja, $1 \leq p \leq m_\mu$, sendo que a irrep somente irá ocorrer na redução se $m_\mu > 0$.

Denotando os vetores da base $\Psi_p^{(\mu)}$ do subespaço invariante $\mathcal{V}_p^{n_\mu} \subset \mathcal{V}^n$ por $\{\psi_{pi}^{(\mu)}\}$, ou seja,

$$\Psi_p^{(\mu)} = \{\psi_{p1}^{(\mu)}, \psi_{p2}^{(\mu)}, \dots, \psi_{pn_\mu}^{(\mu)}\},$$

a ação do operador A sobre o vetor $\psi_{pi}^{(\mu)}$ é dada por

$$A\psi_{pi}^{(\mu)} = \sum_{k=1}^{n_\mu} \psi_{pk}^{(\mu)} \Gamma_{ki}^{(\mu)}(A). \quad (5.34b)$$

Esta é uma terceira relação que ajuda a determinar os elementos da matriz unitária U , os quais, por sua vez, determinam os vetores de base $\{\psi_{pi}^{(\mu)}\}$.

As três expressões que determinam a matriz U são rerepresentadas abaixo:

$$\sum_{k=1}^n U_{k(\mu pi)}^* U_{k(\nu qj)} = \delta_{\mu\nu} \delta_{pq} \delta_{ij} \quad (5.35a)$$

$$\sum_{\mu, p, k} U_{i(\mu pk)}^* U_{j(\mu pk)} = \delta_{ij} \quad (5.35b)$$

$$\sum_{k=1}^{n_\mu} U_{\ell(\mu pk)} \Gamma_{ki}^{(\mu)}(A) = \sum_{j=1}^n D_{\ell j}(A) U_{j(\mu pi)}. \quad (5.35c)$$

As expressões (5.35) serão empregadas para a obtenção da base simetrizada de uma representação regular. Dado um grupo $\mathcal{G} = \{A, B, C, \dots; *\}$ de ordem g , os elementos das matrizes de sua representação regular $D^{\text{reg}}(\mathcal{G})$ são dados por (5.4). Como a dimensão das matrizes de $D^{\text{reg}}(\mathcal{G})$ é igual à ordem do grupo ($n = g$), os índices de $D^{\text{reg}}(A)$ em (5.35c) indicam também os elementos de \mathcal{G} . Ou seja, a variação do índice j mantendo ℓ fixo corresponde à varredura da linha do elemento A na tabela de multiplicações de grupo. Chamando $j \rightarrow C$ e $\ell \rightarrow B$, sendo $B, C \in \mathcal{G}$, a relação (5.35c) pode ser escrita como

$$\sum_{k=1}^{n_\mu} U_{B,(\mu pk)} \Gamma_{ki}^{(\mu)}(A) = \sum_{C \in \mathcal{G}} D_{B,C}(A) U_{C,(\mu pi)} = U_{BA,(\mu pi)}.$$

pois

$$D_{B,C}(A) = \begin{cases} 1, & \text{se } B * A = C \\ 0, & \text{se } B * A \neq C. \end{cases}$$

Este resultado vale para todos $A, B \in \mathcal{G}$ e $1 \leq i \leq n_\mu$. Se for feita a escolha $B = I$ com a convenção de que $I \leftrightarrow 1$ sempre, então

$$U_{A,(\mu pi)} = \sum_{k=1}^{n_\mu} U_{1(\mu pk)} \Gamma_{ki}^{(\mu)}(A). \quad (5.36)$$

Esta última relação, em conjunto com (5.35a,b) determinam a matriz U .

5.8 BASES PARA REPRESENTAÇÕES DE GRUPOS DE PRODUTO DIRETO

Na seção 3.5 mostrou-se que um dado grupo \mathcal{G} sempre pode ser escrito como o produto direto de seus subgrupos invariantes, i. e., se $\mathcal{H}_1, \mathcal{H}_2 \subset \mathcal{G}$ são invariantes, então

$$\mathcal{G}' = \mathcal{H}_1 \otimes \mathcal{H}_2.$$

Se $\mathcal{H}_1 + \mathcal{H}_2 = \mathcal{G}$, então $\mathcal{G}' = \mathcal{G}$. Os subgrupos invariantes de \mathcal{C}_{4v} constituem em um exemplo simples deste conceito. Os subgrupos $\mathcal{H}_1 = \{E, \sigma_{vA}\}$ e $\mathcal{H}_2 = \{E, \sigma_{vB}\}$ formam o grupo

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \{E, \sigma_{vA}, \sigma_{vB}, C_4^2\}.$$

Contudo mencionou-se também que um grupo superior pode ser criado a partir de dois grupos menores. Tomando por exemplo o grupo das rotações próprias \mathcal{C}_n , se a operação de simetria denominada *inversão espacial* (J) for acrescentada às rotações próprias, esta forma o grupo cíclico $\langle J \rangle = \{E, J\}$. Foi mencionado na seção 3.4.1 que a operação de inversão comuta com qualquer outra operação de simetria. Portanto, é possível criar um grupo contendo tanto rotações próprias quanto inversões através do produto $\mathcal{C}_n \otimes \langle J \rangle$.

Posteriormente, na seção 5.4.4 mostrou-se que o produto direto de duas representações também é uma representação (expressão 5.12), sendo que o caractere do grupo produto direto é o produto dos caracteres (expressão 5.13).

5.8.1 REDUÇÃO DA REPRESENTAÇÃO DO PRODUTO DIRETO. SÉRIE DE CLEBSCH-GORDAN

Dadas agora duas representações irreduzíveis do grupo \mathcal{G} , $\Gamma^{(\mu)}(\mathcal{G})$ e $\Gamma^{(\nu)}(\mathcal{G})$, o grupo produto direto $D^{(\mu \times \nu)}(\mathcal{G}) = \Gamma^{(\mu)}(\mathcal{G}) \otimes \Gamma^{(\nu)}(\mathcal{G})$ também é representação. Porém, em princípio esta representação é redutível. Assim, a redução do grupo produto direto em irreps de \mathcal{G} consiste na obtenção dos coeficientes naturais $\{(\mu\nu\eta)\}$ tais que

$$D^{(\mu \times \nu)}(\mathcal{G}) = \Gamma^{(\mu)}(\mathcal{G}) \otimes \Gamma^{(\nu)}(\mathcal{G}) = \bigoplus_{\eta=1}^N (\mu\nu\eta) \Gamma^{(\eta)}(\mathcal{G}). \quad (5.37a)$$

A expansão (5.37a) é denominada uma *série de Clebsch-Gordan*.

Claramente, de (5.27) resulta que

$$\chi^{(\mu \times \nu)}(a) = \chi^{(\mu)}(a) \chi^{(\nu)}(a) = \sum_{\eta=1}^N (\mu\nu\eta) \chi^{(\eta)}(a), \quad (\forall a \in \mathcal{G}). \quad (5.37b)$$

Dadas as expressões (5.22b) e (5.28), os coeficientes $(\mu\nu\eta)$ são obtidos por

$$(\mu\nu\eta) = \frac{1}{g} \sum_{a \in \mathcal{G}} \chi^{(\mu \times \nu)}(a) \chi^{(\eta)*}(a) = \frac{1}{g} \sum_{a \in \mathcal{G}} \chi^{(\mu)}(a) \chi^{(\nu)}(a) \chi^{(\eta)*}(a) = \frac{1}{g} \sum_{k=1}^{nc} c_k \chi_k^{(\mu)} \chi_k^{(\nu)} \chi_k^{(\eta)*}.$$

Obviamente, $(\mu\nu\eta) = (\nu\mu\eta)$.

Diz-se que um grupo \mathcal{G} é *simplesmente redutível* se o produto direto de duas irreps quaisquer de \mathcal{G} não contém cada uma de suas irreps mais de uma vez; isto é, se $(\mu\nu\eta) \leq 1$.

Em geral, dadas duas representações redutíveis de \mathcal{G} , $D^{(\mu)}(\mathcal{G})$ e $D^{(\nu)}(\mathcal{G})$, o seu produto direto $D^{(\mu \times \nu)}(\mathcal{G}) = D^{(\mu)}(\mathcal{G}) \otimes D^{(\nu)}(\mathcal{G})$ sempre pode ser reduzido, uma vez que

$$D^{(\mu)}(\mathcal{G}) = \bigoplus_{\eta=1}^N m_{\eta}^{\mu} \Gamma^{(\eta)}(\mathcal{G}) \quad D^{(\nu)}(\mathcal{G}) = \bigoplus_{\kappa=1}^N m_{\kappa}^{\nu} \Gamma^{(\kappa)}(\mathcal{G})$$

e, portanto,

$$\begin{aligned} D^{(\mu \times \nu)}(\mathcal{G}) &= D^{(\mu)}(\mathcal{G}) \otimes D^{(\nu)}(\mathcal{G}) = \left(\bigoplus_{\eta=1}^N m_{\eta}^{\mu} \Gamma^{(\eta)}(\mathcal{G}) \right) \otimes \left(\bigoplus_{\kappa=1}^N m_{\kappa}^{\nu} \Gamma^{(\kappa)}(\mathcal{G}) \right) \\ &= \bigoplus_{\eta=1}^N \bigoplus_{\kappa=1}^N m_{\eta}^{\mu} m_{\kappa}^{\nu} \Gamma^{(\eta)}(\mathcal{G}) \otimes \Gamma^{(\kappa)}(\mathcal{G}) = \bigoplus_{\tau=1}^N \left[\sum_{\eta, \kappa=1}^N m_{\eta}^{\mu} m_{\kappa}^{\nu} (\eta\kappa\tau) \right] \Gamma^{(\tau)}(\mathcal{G}). \end{aligned}$$

Esta expressão pode sempre ser estendida ao produto direto de mais de duas representações.

5.8.2 BASES PARA REPRESENTAÇÕES DE PRODUTOS DIRETOS. COEFICIENTES DE CLEBSCH-GORDAN

A base para uma representação de produto direto pode ser obtida conhecendo as bases das irreps que compõe o produto direto. Sejam $\Phi = \{\phi_1, \phi_2, \dots, \phi_{n_{\mu}}\}$ uma base do espaço vetorial invariante $\mathcal{V}^{n_{\mu}}$ ao qual estão associadas as irreps $\Gamma^{(\mu)}(\mathcal{G})$ e $\Xi = \{\xi_1, \xi_2, \dots, \xi_{n_{\nu}}\}$ uma base do espaço invariante $\mathcal{V}^{n_{\nu}}$ ao qual estão associadas as irreps $\Gamma^{(\nu)}(\mathcal{G})$. Deseja-se encontrar um conjunto $\Psi \doteq \Phi \times \Xi$ de n_{λ} tensores de base $\Psi_m^{(\lambda\tau\lambda)}$ ($m = 1, \dots, n_{\lambda}$), os quais são dados pelas combinações lineares dos *produtos tensoriais* dos vetores das bases Φ e Ξ ,

$$\Psi_m^{(\lambda\tau\lambda)} = \sum_{k=1}^{n_{\mu}} \sum_{\ell=1}^{n_{\nu}} (\mu k, \nu \ell \mid \lambda\tau\lambda m) \phi_k \otimes \xi_{\ell}, \quad (m = 1, \dots, n_{\lambda}) \quad (5.38)$$

onde as quantidades $(\mu k, \nu \ell \mid \lambda\tau\lambda m)$ são os *coeficientes de Clebsch-Gordan*.

Conforme discutido na seção 4.11, os produtos $\phi_k \otimes \xi_{\ell}$ em (5.38) estão contidos na base do espaço tensorial $\mathcal{V}^{n_{\mu}} \otimes \mathcal{V}^{n_{\nu}}$ de dimensão $n_{\mu} n_{\nu}$. O conjunto dos tensores $\Psi_m^{(\lambda\tau\lambda)}$ forma uma base

do espaço tensorial ao qual estão associadas as irreps $\Gamma^{(\lambda)}(\mathcal{G})$. Conforme discutido na seção anterior, sabe-se que o conjunto Ψ somente existe se $\Gamma^{(\lambda)}(\mathcal{G})$ está contido na decomposição de $\Gamma^{(\mu)} \otimes \Gamma^{(\nu)}$, isto é, se $(\mu\nu\lambda) \neq 0$. Por outro lado, se $(\mu\nu\lambda) > 1$, existe esta quantidade de diferentes conjuntos de tensores associados com a multiplicidade de $\Gamma^{(\lambda)}$ em $\Gamma^{(\mu)} \otimes \Gamma^{(\nu)}$. Por esta razão os tensores de base são distinguidos também pelo índice $\tau_\lambda = 1, \dots, (\mu\nu\lambda)$, denominado o *índice de multiplicidade exterior*. O número total de tensores $\Psi_m^{(\lambda\tau_\lambda)}$ deve igualar a cardinalidade do conjunto Ψ , isto é,

$$\sum_{\lambda} (\mu\nu\lambda) n_{\lambda} = n_{\mu} n_{\nu}.$$

Os coeficientes de Clebsch-Gordan (CG) $(\mu k, \nu \ell | \lambda \tau_{\lambda} m)$ compõe uma matriz $n_{\mu} n_{\nu} \times n_{\mu} n_{\nu}$. A relação (5.38) não é nada mais do que uma mudança de base no espaço tensorial, a qual pode ser invertida, resultando nas relações inversas

$$\phi_k \otimes \xi_{\ell} = \sum_{\lambda, \tau_{\lambda}} \sum_{m=1}^{n_{\lambda}} (\lambda \tau_{\lambda} m | \mu k, \nu \ell) \Psi_m^{(\lambda \tau_{\lambda})}, \quad \begin{array}{l} k = 1, \dots, n_{\mu} \\ \ell = 1, \dots, n_{\nu}. \end{array}$$

Para tanto, é necessário que sejam satisfeitas as identidades

$$\begin{aligned} \sum_{k=1}^{n_{\mu}} \sum_{\ell=1}^{n_{\nu}} (\mu k, \nu \ell | \lambda \tau_{\lambda} m) (\lambda' \tau'_{\lambda'} m' | \mu k, \nu \ell) &= \delta_{\lambda \lambda'} \delta_{\tau_{\lambda} \tau'_{\lambda'}} \delta_{m m'}, \\ \sum_{\lambda, \tau_{\lambda}} \sum_{m=1}^{n_{\lambda}} (\lambda \tau_{\lambda} m | \mu k, \nu \ell) (\mu k', \nu \ell' | \lambda \tau_{\lambda} m) &= \delta_{k k'} \delta_{\ell \ell'}. \end{aligned}$$

Como as irreps consideradas são sempre unitárias, as diferentes bases dos espaços vetoriais e tensoriais podem ser tomadas como ortonormais, o que implica que a matriz dos coeficientes CG também é unitária. Neste caso, os seus elementos satisfazem

$$(\lambda \tau_{\lambda} m | \mu k, \nu \ell) = (\mu k, \nu \ell | \lambda \tau_{\lambda} m)^*, \quad (5.39a)$$

$$\sum_{k=1}^{n_{\mu}} \sum_{\ell=1}^{n_{\nu}} (\mu k, \nu \ell | \lambda \tau_{\lambda} m) (\mu k, \nu \ell | \lambda' \tau'_{\lambda'} m')^* = \delta_{\lambda \lambda'} \delta_{\tau_{\lambda} \tau'_{\lambda'}} \delta_{m m'}, \quad (5.39b)$$

$$\sum_{\lambda, \tau_{\lambda}} \sum_{m=1}^{n_{\lambda}} (\mu k, \nu \ell | \lambda \tau_{\lambda} m)^* (\mu k', \nu \ell' | \lambda \tau_{\lambda} m) = \delta_{k k'} \delta_{\ell \ell'} \quad (5.39c)$$

Aplicando agora um elemento $R \in \mathcal{G}$ sobre o tensor $\Psi_m^{(\lambda \tau_{\lambda})}$, resulta de (5.38),

$$R \Psi_m^{(\lambda \tau_{\lambda})} = \sum_{m'=1}^{n_{\lambda}} \Gamma_{m'm}^{(\lambda)}(R) \Psi_{m'}^{(\lambda \tau_{\lambda})} = \sum_{m', k', \ell'} (\mu k', \nu \ell' | \lambda \tau_{\lambda} m') \Gamma_{m'm}^{(\lambda)}(R) \phi_{k'} \otimes \xi_{\ell'}.$$

Por outro lado, aplicando R sobre o lado direito de (5.38), resulta

$$\begin{aligned} R \sum_{k=1}^{n_{\mu}} \sum_{\ell=1}^{n_{\nu}} (\mu k, \nu \ell | \lambda \tau_{\lambda} m) \phi_k \otimes \xi_{\ell} &= \sum_{k=1}^{n_{\mu}} \sum_{\ell=1}^{n_{\nu}} (\mu k, \nu \ell | \lambda \tau_{\lambda} m) (R \phi_k) \otimes (R \xi_{\ell}) \\ &= \sum_{k, \ell} \sum_{k', \ell'} (\mu k, \nu \ell | \lambda \tau_{\lambda} m) \Gamma_{k'k}^{(\mu)}(R) \Gamma_{\ell'\ell}^{(\nu)}(R) \phi_{k'} \otimes \xi_{\ell'}. \end{aligned}$$

Como os tensores de base são LI, resulta

$$\sum_{m'} (\mu k', \nu \ell' | \lambda \tau_{\lambda} m') \Gamma_{m'm}^{(\lambda)}(R) = \sum_{k, \ell} (\mu k, \nu \ell | \lambda \tau_{\lambda} m) \Gamma_{k'k}^{(\mu)}(R) \Gamma_{\ell'\ell}^{(\nu)}(R).$$

As expressões acima obtidas podem ser manipuladas com as propriedades (5.39) dos coeficientes CG. Por exemplo, multiplicando ambos os lados por $(\mu k', \nu \ell' | \lambda' \tau'_{\lambda'} m')^*$ e somando sobre k' e ℓ' , resulta

$$\Gamma_{m'm}^{(\lambda)}(R) \delta_{\lambda \lambda'} \delta_{\tau_{\lambda} \tau'_{\lambda'}} = \sum_{k, \ell, k', \ell'} (\mu k', \nu \ell' | \lambda' \tau'_{\lambda'} m')^* (\mu k, \nu \ell | \lambda \tau_{\lambda} m) \Gamma_{k'k}^{(\mu)}(R) \Gamma_{\ell'\ell}^{(\nu)}(R).$$

Ou, multiplicando ambos os lados por $(\mu i, \nu j \mid \lambda \tau_\lambda m)^*$ e somando sobre λ, τ_λ e m , resulta

$$\Gamma_{ij}^{(\mu)}(R) \Gamma_{k\ell}^{(\nu)}(R) = \sum_{\lambda, \tau_\lambda, m, m'} (\mu j, \nu \ell \mid \lambda \tau_\lambda m)^* (\mu i, \nu k \mid \lambda \tau_\lambda m') \Gamma_{m'm}^{(\lambda)}(R).$$

As equações assim resultantes fornecem relações suficientes para a determinação dos coeficientes de Clebsch-Gordan ou *coeficientes de Wigner* ou *coeficientes de acoplamento vetorial*.

5.8.3 REPRESENTAÇÃO DE UM GRUPO PRODUTO DIRETO

Considera-se agora um grupo que é o produto direto de dois subgrupos que comutam entre si. Sejam $\mathcal{H} = \{H_1 \equiv I, H_2, \dots, H_h\}$ (ordem h) e $\mathcal{G} = \{G_1 \equiv I, G_2, \dots, G_g\}$ (ordem g) estes subgrupos, o grupo produto direto $\mathcal{K} \doteq \mathcal{H} \otimes \mathcal{G}$ (de ordem $k = hg$) possui os elementos

$$\mathcal{K} = \{K_{11} \equiv I, K_{12}, \dots, K_{1g}, K_{21}, \dots, K_{2g}, \dots, K_{h1}, K_{h2}, \dots, K_{hg}\},$$

os quais são obtidos por

$$K_{ij} = H_i G_j = G_j H_i.$$

Sejam $H_i H_m = H_p$ e $G_j G_n = G_q$; então

$$K_{ij} K_{mn} = (H_i G_j) (H_m G_n) = (H_i H_m) (G_j G_n) = H_p G_q = K_{pq}. \quad (5.40)$$

Sejam agora $D^{(\mu)}(\mathcal{H})$ uma representação de \mathcal{H} e $D^{(\nu)}(\mathcal{G})$ uma representação de \mathcal{G} . Então

$$D^{(\mu)}(H_i) D^{(\mu)}(H_m) = D^{(\mu)}(H_p) \quad D^{(\nu)}(G_j) D^{(\nu)}(G_n) = D^{(\nu)}(G_q).$$

Realizando o produto direto destas matrizes, obtém-se

$$\begin{aligned} D^{(\mu)}(H_p) \otimes D^{(\nu)}(G_q) &= \left[D^{(\mu)}(H_i) D^{(\mu)}(H_m) \right] \otimes \left[D^{(\nu)}(G_j) D^{(\nu)}(G_n) \right] \\ &= \left[D^{(\mu)}(H_i) \otimes D^{(\nu)}(G_j) \right] \left[D^{(\mu)}(H_m) \otimes D^{(\nu)}(G_n) \right], \end{aligned}$$

onde a propriedade (5.11a) foi empregada. Definindo novas matrizes pelo produto direto

$$D^{(\mu \times \nu)}(K_{pq}) \doteq D^{(\mu)}(H_p) \otimes D^{(\nu)}(G_q),$$

então

$$D^{(\mu \times \nu)}(K_{pq}) = D^{(\mu \times \nu)}(K_{ij}) D^{(\mu \times \nu)}(K_{mn}). \quad (5.41)$$

Comparando (5.41) com (5.40), conclui-se que as matrizes produto direto assim definidas formam uma representação de \mathcal{K} . Esta representação é identificada por $D^{(\mu \times \nu)}(\mathcal{K})$. Ou seja, o produto direto de representações de dois grupos que comutam é uma representação do grupo produto direto.

Agora, se $D^{(\mu)}(\mathcal{H}) \equiv \Gamma^{(\mu)}(\mathcal{H})$ e $D^{(\nu)}(\mathcal{G}) \equiv \Gamma^{(\nu)}(\mathcal{G})$ são irreps de \mathcal{H} e \mathcal{G} , respectivamente, então $D^{(\mu \times \nu)}(\mathcal{K}) \equiv \Gamma^{(\mu \times \nu)}(\mathcal{K})$ também é irrep de \mathcal{K} . Para verificar isso, recorre-se à condição necessária e suficiente (5.29), a qual fica escrita

$$k = \sum_{K_{ij} \in \mathcal{K}} \left| \chi^{(\mu \times \nu)}(K_{ij}) \right|^2.$$

Pode-se verificar facilmente, de forma semelhante ao que foi feito em (5.13), que $\chi^{(\mu \times \nu)}(K_{ij}) = \chi^{(\mu)}(H_i) \chi^{(\nu)}(G_j)$, ou seja,

$$\sum_{K_{ij} \in \mathcal{K}} \left| \chi^{(\mu \times \nu)}(K_{ij}) \right|^2 = \left[\sum_{H_i \in \mathcal{H}} \left| \chi^{(\mu)}(H_i) \right|^2 \right] \left[\sum_{G_j \in \mathcal{G}} \left| \chi^{(\nu)}(G_j) \right|^2 \right] = hg = k,$$

provando que $\Gamma^{(\mu \times \nu)}(\mathcal{K})$ de fato é irrep de \mathcal{K} .

Finalmente, se N_h e N_g forem respectivamente os números de classes e, portanto, de irreps dos grupos \mathcal{H} e \mathcal{G} , então a condição (5.21) se aplica a ambos. Dada a irrep $\Gamma^{(\mu \times \nu)}(\mathcal{K})$, de dimensão $n_{\mu\nu} = n_\mu n_\nu$, todas as combinações possíveis de μ e ν fornecem a soma a seguir,

$$\sum_{\mu=1}^{N_\mu} \sum_{\nu=1}^{N_\nu} n_{\mu\nu}^2 = \sum_{\mu=1}^{N_\mu} \sum_{\nu=1}^{N_\nu} n_\mu^2 n_\nu^2 = \left(\sum_{\mu=1}^{N_\mu} n_\mu^2 \right) \left(\sum_{\nu=1}^{N_\nu} n_\nu^2 \right) = hg = k.$$

Portanto, a condição (5.21) também se aplica a $\Gamma^{(\mu \times \nu)}(\mathcal{K})$, i. e., $N_{\mu\nu} = N_\mu N_\nu$ é o número total de irreps de \mathcal{K} e os produtos diretos das irreps de \mathcal{H} e \mathcal{G} exaurem todas as irreps de \mathcal{K} .

5.9 APLICAÇÕES FÍSICAS DA TEORIA DE REPRESENTAÇÕES DE GRUPO

Algumas aplicações físicas da teoria de representações de grupos serão discutidas agora. Desde as primeiras contribuições ao assunto, no início do século XX, as principais aplicações são dirigidas para a mecânica quântica, física atômica, molecular e estado sólido e física de partículas e campos. São estas as aplicações a ser discutidas aqui.

Continuando a discussão realizada na seção 5.2.1, o problema fundamental na mecânica quântica está na solução e investigação das propriedades das soluções da equação de Schrodinger

$$\mathcal{H}|\psi(t)\rangle = i\hbar \frac{d}{dt}|\psi(t)\rangle, \quad (5.42)$$

ou $\mathcal{H}|\psi\rangle = E|\psi\rangle$ na forma independente do tempo, sendo $|\psi\rangle$ um vetor *ket* que descreve o estado do sistema e que pertence a um espaço de Hilbert \mathcal{H} , \mathcal{H} um operador linear sobre \mathcal{H} e E a energia total do sistema físico no estado $|\psi\rangle$. A equação acima forma (juntamente com condições de contorno apropriadas) um problema de autovalores,¹⁵ no qual $|\psi\rangle$ é a autofunção de \mathcal{H} e E o seu autovalor. O operador \mathcal{H} , denotado o *Hamiltoniano* do sistema é suposto ser um operador Hermitiano,¹⁶ em cuja situação o autovalor E é real.

Os casos em que a equação de Schrodinger é exatamente solúvel são muito poucos e restritos a sistemas com uma ou duas partículas. Portanto, a grande maioria dos sistemas físicos de interesse, tais como átomos com números atômicos $Z \geq 2$, moléculas e sólidos, não têm solução conhecida. Entretanto, os métodos desenvolvidos neste capítulo e no capítulo 3 podem fornecer informações importantes sobre esses sistemas.

O poder da teoria de representações de grupos aplicada à mecânica quântica está na sua capacidade de fornecer conclusões e informações exatas sobre o sistema físico, partindo somente do conhecimento de suas simetrias, frente às quais o Hamiltoniano permanece invariante. Além disso, ao se estabelecer o grupo de operadores de simetria (o *grupo do Hamiltoniano*), as representações irredutíveis do grupo surgem de forma natural, devido à conexão entre as simetrias do sistema e a degenerescência dos autoestados descritos pela equação de Schrodinger.

Uma primeira evidência já é fornecida pela equação de movimento do valor esperado de um observável. Se O é um observável (possivelmente dependente do tempo) e $\mathcal{O} = \mathcal{O}(t)$ o operador linear associado ao mesmo, então a evolução temporal do valor esperado do observável é dada por

$$\frac{d}{dt}\langle O \rangle = -\frac{i}{\hbar}\langle [O, \mathcal{H}] \rangle + \left\langle \frac{\partial \mathcal{O}}{\partial t} \right\rangle,$$

sendo $[O, \mathcal{H}] = \mathcal{O}\mathcal{H} - \mathcal{H}\mathcal{O}$ o *comutador* de O com \mathcal{H} . Se \mathcal{O} não depender do tempo e o mesmo comutar com o Hamiltoniano, isto é, se $[O, \mathcal{H}] = 0$, então $d\langle O \rangle / dt = 0$ e o observável O é uma *constante de movimento*.

5.9.1 ISOMORFISMO ENTRE TRANSFORMAÇÕES SOBRE SISTEMAS FÍSICOS E TRANSFORMAÇÕES SOBRE ESPAÇOS FUNCIONAIS

A partir deste ponto, ao invés de se manter a descrição da evolução do sistema físico na forma abstrata fornecida pela equação (5.42), será adotada a descrição simplificada provida pela “*representação*” ou “*modo de descrição*” de Schrodinger, segundo a qual o vetor de estado do sistema é descrito pela **função de onda** $\psi = \psi(\mathbf{r}, t)$ sobre o espaço Euclidiano \mathbb{E}^3 (ver exemplo 4.5). Nesta descrição, a equação de Schrodinger (5.42) passa a ser formalmente escrita como

$$\mathcal{H}\psi(\mathbf{r}, t) = i\hbar \frac{\partial}{\partial t}\psi(\mathbf{r}, t), \quad (5.43)$$

sendo que agora o Hamiltoniano é um operador que atua sobre o espaço das funções sobre o \mathbb{E}^3 , *i. e.*, é um funcional na forma $\mathcal{H} = \mathcal{H}\{\mathbf{r}, \nabla; t\}$, por exemplo.

Seja $\mathcal{R} = \{R_i\}$ um grupo abstrato que contenha todas as isometrias de um sistema físico. Os elementos de \mathcal{R} são transformações aplicadas ao sistema, de tal forma que a ação de $R \in \mathcal{R}$

¹⁵De forma mais específica, um problema de *Sturm-Liouville*.

¹⁶Ver definição 4.23.

sobre o sistema físico consiste em modificar as *coordenadas* do mesmo. O termo “coordenadas” aqui possui um significado amplo. Este pode significar de fato as coordenadas das posições das partículas do sistema físico em relação a um determinado sistema de referência em um espaço métrico, mas podem também se referir a coordenadas no espaço de fase, ou a outras propriedades que caracterizam o estado físico do sistema, tais como *spin*, carga elétrica, cor,¹⁷ *etc.*

A dinâmica deste sistema, por sua vez, será determinada por um conjunto de leis (*e. g.*, Leis de Newton, equações de Maxwell, equação de Schroedinger, *etc*) que fornecerão uma descrição quantitativa da evolução do mesmo. Esta descrição será usualmente realizada por intermédio de uma classe de funções que irão depender das coordenadas (no sentido amplo) do sistema. Essas funções serão elementos de um espaço funcional, isto é, de um espaço vetorial (de dimensão finita ou infinita), cujas bases são conjuntos de funções das coordenadas.

Seja agora o grupo $\mathcal{P}_R = \{P_R\}$ formado por operadores lineares que executam transformações sobre as *funções* das coordenadas do sistema, ao invés das coordenadas propriamente ditas. Como o sistema físico permanece indistinguível após ser atuado por uma transformação isométrica $R \in \mathcal{R}$, as leis que regem a sua evolução também devem permanecer invariantes frente a ação de um operador $P_R \in \mathcal{P}_R$ que atua sobre uma função das coordenadas. Para que esta exigência seja satisfeita, é necessário que exista um isomorfismo entre o operador $R \in \mathcal{R}$ e um (e somente um) operador $P_R \in \mathcal{P}_R$. Este isomorfismo é garantido pelo seguinte mapeamento: se \mathbf{r} denota as coordenadas (no sentido amplo) do sistema físico e $f(\mathbf{r})$ é uma função dessas coordenadas, então

$$\begin{aligned} P_R f(R\mathbf{r}) &= f(\mathbf{r}) \\ P_R f(\mathbf{r}) &= f(R^{-1}\mathbf{r}), \end{aligned} \tag{5.44}$$

sendo ambas as expressões equivalentes. Pode-se dizer que o operador P_R muda a forma funcional de $f(\mathbf{r})$ de maneira tal que compensa a transformação nas coordenadas \mathbf{r} executada pelo operador R .

Antes de se mostrar alguns exemplos, será verificado agora que o grupo \mathcal{P}_R é de fato isomórfico a \mathcal{R} . Para tanto, é necessário demonstrar que para $R, S \in \mathcal{R}$, os correspondentes operadores $P_R, P_S \in \mathcal{P}_R$, dados por (5.44), satisfazem

$$P_S P_R = P_{SR}.$$

Procede-se por etapas. Inicialmente, dada uma função das coordenadas $f(\mathbf{r})$, a ação de P_R executa a transformação $f(\mathbf{r}) \rightarrow g(\mathbf{r})$, sendo esta última uma outra função das mesmas coordenadas. Ou seja,

$$P_R f(\mathbf{r}) = f(R^{-1}\mathbf{r}) = g(\mathbf{r}).$$

Assim, $g(\mathbf{r})$ é a nova função que incorpora a ação de R^{-1} sobre as coordenadas na sua forma funcional. Aplicando agora o operador P_S , obtém-se

$$\begin{aligned} P_S [P_R f(\mathbf{r})] &= P_S g(\mathbf{r}) \stackrel{(5.44)}{=} g(S^{-1}\mathbf{r}) = f[R^{-1}(S^{-1}\mathbf{r})] \\ &= f(R^{-1}S^{-1}\mathbf{r}) \xrightarrow[R^{-1}S^{-1}=(SR)^{-1}]{R, S \in \mathcal{R}} f[(SR)^{-1}\mathbf{r}] = P_{SR} f(\mathbf{r}). \end{aligned}$$

Portanto, a transformação P_{SR} correspondente ao elemento $SR \in \mathcal{R}$ é de fato o “produto” das transformações P_S por P_R .

Algumas das transformações isométricas mais comuns e seus respectivos operadores funcionais serão apresentados a seguir.

5.9.1.1 TRANSLAÇÕES ESPACIAIS

Seja um sistema físico cujas coordenadas são descritas como pontos no espaço Euclidiano \mathbb{E}^3 . Uma *translação* no sistema de coordenadas do \mathbb{E}^3 consiste na operação

$$\mathbf{r} \rightarrow \mathbf{r}', \text{ onde } \mathbf{r}' = T_\rho \mathbf{r} \doteq \mathbf{r} + \boldsymbol{\rho},$$

sendo $\mathbf{r} = (x_1, x_2, x_3) \in \mathbb{E}^3$ uma posição no espaço, T_ρ o operador de translação atuando sobre \mathbb{E}^3 e $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3) \in \mathbb{E}^3$ um vetor constante. A isometria do sistema físico frente a ação T irá

¹⁷No sentido da cromodinâmica quântica.

estabelecer os grupos cristalográficos espaciais. Como T pertence a um grupo, sempre existe a transformação inversa $T_\rho^{-1}\mathbf{r} = \mathbf{r} - \boldsymbol{\rho}$.

Seja agora uma função das coordenadas $\psi(\mathbf{r})$. Esta função pode ser, por exemplo, a solução da equação de Schroedinger para uma rede cristalina. Então, como após a transformação $\mathbf{r} \rightarrow \mathbf{r}'$ a relação $\mathbf{r} = \mathbf{r}' - \boldsymbol{\rho}$ é válida, a função $\psi(\mathbf{r})$ passa a ser escrita como

$$\psi(\mathbf{r}) = \psi(\mathbf{r}' - \boldsymbol{\rho}) \xrightarrow{\mathbf{r}' \sim \mathbf{r}} \psi(\mathbf{r} - \boldsymbol{\rho}),$$

sendo que a última forma foi obtida simplesmente eliminando o apóstrofo.

Esta transformação possui duas interpretações. A translação $\mathbf{r} \rightarrow \mathbf{r}'$ ocorreu no sistema de referências, cuja origem foi colocada no ponto $\mathbf{r} = -\boldsymbol{\rho}$ do sistema original. Esta transformação é dita ser do ponto de vista *passivo*. De forma equivalente, a transformação $\psi(\mathbf{r}) \rightarrow \psi(\mathbf{r} - \boldsymbol{\rho})$ consiste em manter o referencial fixo e deslocar a função $\psi(\mathbf{r})$ pela quantidade $\boldsymbol{\rho}$. Esta é uma transformação do ponto de vista *ativo*.¹⁸ Por exemplo, se $\rho_1 > 0$, a ação $T_{\rho_1}\mathbf{r} = Tx_1 = x'_1 = x_1 + \rho_1$ significa que a origem do referencial foi deslocada para a esquerda e colocada no ponto $x_1 = -\rho_1$ do referencial anterior (transformação passiva). De forma equivalente, $\psi(x_1) \rightarrow \psi(x_1 - \rho_1)$ significa que a função foi deslocada em relação à origem para a direita pela quantidade ρ_1 (transformação ativa), pois o valor da função na origem $\psi(0)$ irá ocorrer agora em $x_1 = \rho_1$.

Define-se então um operador linear P_{T_ρ} que atua sobre as funções $\psi(\mathbf{r})$ e que realiza sobre as mesmas a transformação correspondente a T_ρ . Este operador é definido por

$$P_{T_\rho}\psi(\mathbf{r}) = \psi(T^{-1}\mathbf{r}) \equiv \psi'(\mathbf{r}) = \psi(\mathbf{r} - \boldsymbol{\rho}). \quad (5.45)$$

Busca-se agora uma expressão para o operador de translação P_{T_ρ} . Escrevendo $\mathbf{r} = (x, y, z)$, considera-se inicialmente o caso particular onde $\boldsymbol{\rho} = (\rho, 0, 0)$. Então, dada a função $\psi(x, y, z)$,

$$P_{T_\rho}\psi(x, y, z) = \psi(x - \rho, y, z).$$

Desenvolvendo $\psi(x - \rho, y, z)$ em uma série de Taylor em torno de \mathbf{r} , resulta

$$\psi(x - \rho, y, z) = \left(1 - \rho \frac{\partial}{\partial x} + \frac{\rho^2}{2!} \frac{\partial^2}{\partial x^2} + \dots\right) \psi(x, y, z) = \exp\left(-\rho \frac{\partial}{\partial x}\right) \psi(x, y, z),$$

onde o operador $\exp(-\rho\partial/\partial x)$ pode ser interpretado no contexto de uma série de Taylor.

Estendendo essa operação agora para translação arbitrária $\boldsymbol{\rho} = (\rho_x, \rho_y, \rho_z)$, resulta

$$\psi(\mathbf{r} - \boldsymbol{\rho}) = \exp(-\boldsymbol{\rho} \cdot \nabla) \psi(\mathbf{r}) = \exp(-i\boldsymbol{\rho} \cdot \mathbf{p}/\hbar) \psi(\mathbf{r}),$$

onde $\mathbf{p} = -i\hbar\nabla$ é o operador momento linear empregado na descrição de Schroedinger da mecânica quântica. Comparando este resultado com (5.45), identifica-se

$$P_{T_\rho} = \exp\left(-\frac{i}{\hbar}\boldsymbol{\rho} \cdot \mathbf{p}\right) = 1 - \frac{i}{\hbar}\boldsymbol{\rho} \cdot \mathbf{p} + \dots$$

Uma vez que $\boldsymbol{\rho}$ é real e \mathbf{p} é um operador Hermitiano,¹⁹ resulta que P_{T_ρ} é unitário. Neste caso, o operador inverso de P_{T_ρ} satisfaz

$$\psi'(\mathbf{r}) = P_{T_\rho}\psi(\mathbf{r}) \implies P_{T_\rho}^{-1}\psi'(\mathbf{r}) = P_{T_\rho}^\dagger\psi'(\mathbf{r}) = \psi(\mathbf{r}).$$

Uma interpretação mais clara do significado matemático do operador P_{T_ρ} é a seguinte. Supondo que $P_{T_{\delta\mathbf{r}}}$ seja o operador que executa uma translação infinitesimal por $\boldsymbol{\rho} = \delta\mathbf{r}$. Então, adotando a notação $\mathbf{r} = (x_1, x_2, x_3)$, pode-se escrever

$$P_{T_{\delta\mathbf{r}}} = 1 - \frac{i}{\hbar}\mathbf{p} \cdot \delta\mathbf{r} = 1 - \frac{i}{\hbar} \sum_{j=1}^3 p_j \delta x_j \doteq 1 - i \sum_{j=1}^3 I_{x_j} \delta x_j, \text{ onde } I_{x_j} = \frac{p_j}{\hbar}.$$

Este deslocamento arbitrário por $\delta\mathbf{r}$ pode ser escrito em termos de deslocamentos infinitesimais ao longo de cada eixo coordenado como

$$P_{T_{\delta\mathbf{r}}} = \prod_{j=1}^3 P_{T_{\delta x_j}}, \text{ sendo } P_{T_{\delta x_j}} = 1 - iI_{x_j} \delta x_j.$$

¹⁸Transformações ativas ou passivas serão novamente abordadas na seção 6.2.1.

¹⁹Ver definição 4.23.

Portanto, basta estudar o comportamento dos operadores de translação unidimensional $P_{T_{\delta x_j}}$. Duas aplicações consecutivas de $P_{T_{\delta x_j}}$ rendem a translação

$$P_{T_{\delta x_j}}^2 = (1 - iI_{x_j}\delta x_j)^2 = 1 - 2iI_{x_j}\delta x_j,$$

o que equivale a uma translação infinitesimal de magnitude $2\delta x_j$. Após N aplicações sucessivas deste operador, a translação total na direção x_j , $N\delta x_j$, será finita, se

$$\rho_j \doteq \lim_{\substack{\delta x_j \rightarrow 0 \\ N \rightarrow \infty}} N\delta x_j.$$

Então, a translação finita ao longo de x_j será obtida através do processo de limite

$$P_{T_{\rho_j}} = \lim_{\substack{\delta x_j \rightarrow 0 \\ N \rightarrow \infty}} P_{T_{\delta x_j}}^N = \lim_{\substack{\delta x_j \rightarrow 0 \\ N \rightarrow \infty}} (1 - iI_{x_j}\delta x_j)^N = \lim_{N \rightarrow \infty} \left[1 - i \left(\frac{\rho_j}{N} \right) I_{x_j} \right]^N.$$

Dado então o conhecido limite exponencial

$$\lim_{N \rightarrow \infty} \left(1 + \frac{x}{N} \right)^N = e^x,$$

resulta que

$$P_{T_{\rho_j}} = \exp(-i\rho_j I_{x_j}) \implies P_{T_{\rho}} = \prod_{j=1}^3 P_{T_{\rho_j}}.$$

Dessa maneira, a translação executada pelo operador $P_{T_{\rho_j}}$ ocorre devido à existência da quantidade $I_{x_j} = p_j/\hbar$, a qual é denominada o **gerador do grupo de translações**.

Uma outra relação importante é obtida considerando a aplicação do operador comutador

$$[x_i, P_{T_{\delta x_j}}] = x_i P_{T_{\delta x_j}} - P_{T_{\delta x_j}} x_i$$

sobre a função de onda. Cada termo separadamente fornece

$$x_i P_{T_{\delta x_j}} \psi(\mathbf{r}) = x_i \left(1 - \frac{i}{\hbar} \delta x_j p_j \right) \psi(\mathbf{r}), \quad P_{T_{\delta x_j}} x_i \psi(\mathbf{r}) = x_i \left(1 - \frac{i}{\hbar} \delta x_j p_j \right) \psi(\mathbf{r}) - \delta x_j \delta_{ij} \psi(\mathbf{r}),$$

resultando

$$[x_i, P_{T_{\delta x_j}}] \psi(\mathbf{r}) = \delta x_j \delta_{ij} \psi(\mathbf{r}).$$

Mas, do ponto de vista dos operadores,

$$[x_i, P_{T_{\delta x_j}}] = -\frac{i}{\hbar} \delta x_j [x_i, p_j].$$

Portanto, resulta daí a importante relação de comutação

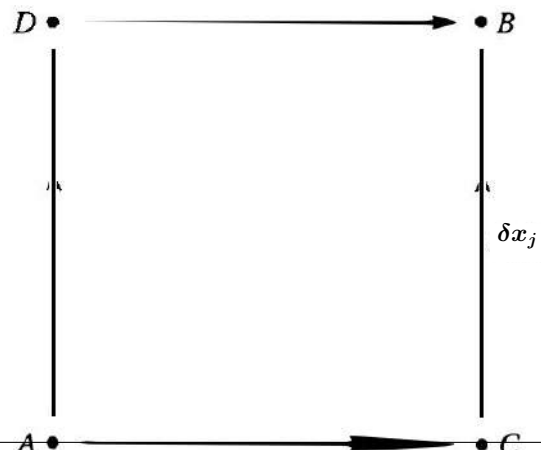
$$[x_i, p_j] = i\hbar \delta_{ij}. \tag{5.46}$$

Por outro lado, a ação do comutador $[P_{T_{\delta x_i}}, P_{T_{\delta x_j}}]$ resulta em

$$[P_{T_{\delta x_i}}, P_{T_{\delta x_j}}] \psi(\mathbf{r}) = -\frac{1}{\hbar^2} \delta x_i \delta x_j [p_i, p_j] \psi(\mathbf{r}).$$

Entretanto, uma propriedade fundamental das translações no espaço é que translações sucessivas, mesmo em diferentes direções, comutam entre si. Isto está ilustrado na figura 5.1. Esta propriedade implica que, necessariamente, $[P_{T_{\delta x_i}}, P_{T_{\delta x_j}}] \psi(\mathbf{r}) = 0$, de onde resulta uma outra relação de comutação essencial,

$$[p_i, p_j] = 0.$$



Isto é, os geradores do grupo de translações comutam. Uma das consequências deste fato é que este grupo é Abelian.

Deseja-se agora investigar a ação do operador P_{T_p} sobre um sistema quantum-mecânico, descrito pela equação de Schroedinger (5.43). Aplicando o operador $i\hbar\partial/\partial t$ sobre $\psi'(\mathbf{r}, t)$,

$$i\hbar\frac{\partial}{\partial t}\psi'(\mathbf{r}, t) = i\hbar\frac{\partial}{\partial t}[P_{T_p}\psi(\mathbf{r}, t)] = i\hbar P_{T_p}\frac{\partial}{\partial t}\psi(\mathbf{r}, t),$$

uma vez que $\partial/\partial t$ e ∇ comutam. Mas, de acordo com (5.43), resulta que

$$i\hbar\frac{\partial}{\partial t}\psi'(\mathbf{r}, t) = P_{T_p}\mathcal{H}\psi(\mathbf{r}, t) = P_{T_p}\mathcal{H}P_{T_p}^\dagger\psi'(\mathbf{r}, t).$$

Portanto, se

$$P_{T_p}\mathcal{H}P_{T_p}^\dagger = \mathcal{H} \text{ ou, de forma equivalente, se } [P_{T_p}, \mathcal{H}] = 0,$$

então

$$i\hbar\frac{\partial}{\partial t}\psi'(\mathbf{r}, t) = \mathcal{H}\psi'(\mathbf{r}, t).$$

Ou seja, a função transformada $\psi'(\mathbf{r}, t)$ também satisfaz a equação de Schroedinger e é uma função de onda do sistema.

Se o operador P_{T_p} comuta com o Hamiltoniano, lembrando da dependência do mesmo com o operador p , segue que $[p, \mathcal{H}] = 0$. Além disso, como P_{T_p} não depende explicitamente do tempo, resulta que *se o sistema físico é invariante frente a todas as translações espaciais, o seu momento linear é uma constante de movimento*, i. e., p é conservado. A mesma lei de conservação é válida para um sistema clássico.

5.9.1.2 ROTAÇÕES

Da mesma forma como ocorre com as translações, uma rotação pode ser executada dos pontos de vista passivo ou ativo, sendo estas também discutidas na seção 6.2.1. Uma rotação no ponto de vista passivo no espaço de configuração consiste na aplicação de um operador $R \in \mathcal{R}$, sendo \mathcal{R} representado por matrizes dos grupos $SO(3)$ ou $SO(2)$.

Como um exemplo simples, considera-se um rotação (passiva) do sistema de coordenadas em torno de um dos eixos coordenados, por exemplo, em torno do eixo x . O operador $R_x(\theta) \in SO(2)$ que executa esta transformação é representado por (seção 6.5.2)

$$R_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \quad R_x^{-1}(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$

juntamente com o operador inverso $R_x^{-1}(\theta)$. Após a aplicação de $R_x(\theta)$, as coordenadas de um ponto no \mathbb{E}^3 , representadas pelo vetor coluna $\mathbf{r} = (x \ y \ z)^T$, passam a ser dadas por

$$\mathbf{r}' = R_x(\theta)\mathbf{r}.$$

O grupo de operadores \mathcal{P}_R , isomórfico a \mathcal{R} , tem seus elementos dados pelo mapeamento (5.44), de tal forma que $P_{R_x(\theta)} \in \mathcal{P}_R$ é o operador que executa a rotação (do ponto de vista ativo) das funções das coordenadas que compensa a ação de $R_x(\theta)$. O exemplo a seguir ilustra a diferença nas ações destes operadores.

Exemplo 5.8 (Orbitais p do átomo de hidrogênio). Sabe-se da solução da equação de Schroedinger para o átomo de hidrogênio que as funções de onda correspondentes aos orbitais p podem ser escritas nas formas $p_x(\mathbf{r}) = x\varphi(r)$, $p_y(\mathbf{r}) = y\varphi(r)$ e $p_z(\mathbf{r}) = z\varphi(r)$, sendo $\varphi(r)$ uma função de $r = |\mathbf{r}|$, isotrópica em \mathbf{r} . Estas funções são tais que $p_x(\mathbf{r})$ está orientada ao longo do eixo x , p_y ao longo de y e p_z ao longo de z , conforme está ilustrado na figura 5.2.

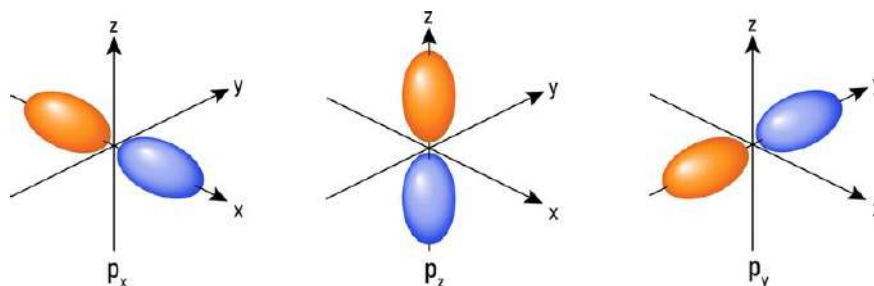


Figura 5.2: Representações gráficas dos orbitais atômicos p_x , p_y e p_z .

Exemplo. Suponha agora que uma rotação de 90° em torno do eixo x (sentido anti-horário) é executada. Do ponto de vista passivo, esta transformação é tal que

$$r \rightarrow r' \implies r \rightarrow r' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ z \\ -y \end{pmatrix} \implies \begin{cases} x' = x \\ y' = z \\ z' = -y. \end{cases}$$

Lembrando que

$$\left[R_x^{-1} \left(\frac{\pi}{2} \right) \right] r = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ -z \\ y \end{pmatrix},$$

de acordo com (5.44), a ação de $P_{R_x(\theta)}$ sobre uma função $f(x, y, z)$ resulta em

$$P_{R_x(\frac{\pi}{2})} f(x, y, z) = P_{R_x(\frac{\pi}{2})} f(r) = f \left[R_x^{-1} \left(\frac{\pi}{2} \right) r \right] = f(x, -z, y).$$

Portanto, se o operador $P_{R_x(\frac{\pi}{2})}$ atuar sobre os orbitais p , os resultados serão

$$\begin{aligned} p'_x &= P_{R_x(\frac{\pi}{2})} p_x = P_{R_x(\frac{\pi}{2})} [x\varphi(r)] = x\varphi(r) = p_x \\ p'_y &= P_{R_x(\frac{\pi}{2})} p_y = P_{R_x(\frac{\pi}{2})} [y\varphi(r)] = -z\varphi(r) = -p_z \\ p'_z &= P_{R_x(\frac{\pi}{2})} p_z = P_{R_x(\frac{\pi}{2})} [z\varphi(r)] = y\varphi(r) = p_y, \end{aligned}$$

uma vez que $r = \sqrt{x^2 + y^2 + z^2}$ é invariante frente a esta rotação. Nota-se na figura 5.2 que os contornos destes orbitais são rotados no sentido horário por 90° em torno do eixo x , compensando a rotação dos eixos coordenados.

A forma explícita do operador $P_{R_x(\theta)}$ pode ser obtida considerando inicialmente a matriz R_1^{-1} para uma rotação infinitesimal $\delta\theta$ sobre as coordenadas $r = (x_1 \ x_2 \ x_3)^T$, em torno do eixo x_1 :

$$R_1^{-1}(\delta\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\delta\theta \\ 0 & \delta\theta & 1 \end{pmatrix} \iff r' = R_1^{-1}(\delta\theta) r \Rightarrow x'_i = \sum_{j=1}^3 (\delta_{ij} - \epsilon_{1ij}\delta\theta) x_j, \quad (i = 1, 2, 3),$$

sendo ϵ_{ijk} o símbolo de Levi-Civita.²⁰ Então, a ação de $P_{R_1(\delta\theta)}$ sobre a função de onda $\psi(r) \equiv \psi(x_1, x_2, x_3) \equiv \psi(\{x_j\})$ resulta em

$$\psi'(\{x_j\}) = P_{R_1(\delta\theta)} \psi(\{x_j\}) = \psi(\{x'_j\}).$$

Desenvolvendo $\psi(\{x'_j\})$ em uma série de Taylor em termos de $\delta\theta$,

$$\begin{aligned} \psi(\{x'_j\}) &= \psi(\{x_j\}) + \frac{d}{d\theta} \psi(\{x'_j\}) \delta\theta = \psi(\{x_j\}) + \left[\sum_{k=1}^3 \frac{\partial}{\partial x_k} \psi(\{x_j\}) \frac{dx'_k}{d\theta} \right] \delta\theta \\ &= \psi(\{x_j\}) - \sum_{k,\ell=1}^3 \epsilon_{1k\ell} x_\ell \frac{\partial \psi}{\partial x_k} \delta\theta = \left(1 - \frac{i}{\hbar} L_1 \delta\theta \right) \psi(\{x_j\}), \end{aligned}$$

²⁰Seção 6.1.2.

sendo

$$L_1 = \sum_{j,k=1}^3 \epsilon_{1jk} x_j p_k = -i\hbar \sum_{j,k=1}^3 \epsilon_{1jk} x_j \frac{\partial}{\partial x_k}$$

a componente na direção x_1 do operador quantum-mecânico associado ao momento angular. Portanto,

$$P_{R_1(\delta\theta)} = 1 - \frac{i}{\hbar} L_1 \delta\theta \doteq 1 - i I_{L_1} \delta\theta, \text{ sendo } I_{L_1} = \frac{L_1}{\hbar}.$$

Uma rotação finita por θ é então obtida via

$$\theta = \lim_{\substack{\delta\theta \rightarrow 0 \\ N \rightarrow \infty}} N \delta\theta$$

e pela aplicação do operador $P_{R_1(\theta)}$, dado por

$$P_{R_1(\theta)} = \lim_{\substack{\delta\theta \rightarrow 0 \\ N \rightarrow \infty}} P_{R_1(\delta\theta)}^N = \lim_{N \rightarrow \infty} \left[1 - i \left(\frac{\theta}{N} \right) I_{L_1} \right]^N = \exp(-i I_{L_1} \theta).$$

Ou seja, $I_{L_1} = L_1/\hbar$ é o **gerador do grupo de rotações (ativas) em torno de x_1** .

Portanto, rotações no espaço de configuração aplicadas a um sistema físico estão relacionadas com o momento angular desse sistema, da mesma forma que translações estão relacionadas com o momento linear. Também de forma equivalente, a *invariância de um sistema físico frente a uma rotação arbitrária no espaço de configuração, em torno de um eixo x_i , implica na conservação de seu momento angular nesta direção, uma vez que o operador L_i irá comutar com o Hamiltoniano*.

Retornando brevemente ao grupo de rotações passivas em torno do eixo x_1 , composto pelas matrizes $R_1(\theta) \in SO(2)$, é possível obter-se também o seu gerador. Para tanto, considera-se a rotação infinitesimal implementada por

$$R_1(\delta\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \delta\theta \\ 0 & -\delta\theta & 1 \end{pmatrix}.$$

Então,

$$r' = R_1(\delta\theta) r \implies x'_i = \sum_{j=1}^3 (\delta_{ij} + \epsilon_{1ij} \delta\theta) x_j, \quad (i = 1, 2, 3),$$

uma vez que

$$R_1(\delta\theta) = R_1^{-1}(-\delta\theta).$$

A partir de $R_1(\delta\theta)$, deseja-se escrever o operador que executa uma rotação arbitrária $R_1(\theta)$ na forma

$$R_1(\theta) = \exp(iM_1\theta) = I_3 + iM_1\theta + \dots,$$

onde I_3 é a matriz identidade e M_1 é uma matriz 3×3 a ser determinada. Realizando o limite $\theta \rightarrow \delta\theta$, observa-se que

$$R_1(\delta\theta) = I_3 + iM_1\delta\theta \implies M_1 = \frac{1}{i\delta\theta} [R_1(\delta\theta) - I_3] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix},$$

a qual pode ser escrita como

$$M_1 = 0 \oplus \sigma_2,$$

onde $0 = 0$ é a matriz 1×1 nula e

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

é uma das matrizes de Pauli. Portanto, as matrizes de Pauli são os geradores do grupo $SO(2)$.

5.9.1.3 DESLOCAMENTO TEMPORAL

Um último exemplo de operador de transformação consiste em realizar um deslocamento temporal sobre a função de onda. Esta transformação é realizada por um operador $P_{T_\tau} \in \mathcal{P}_R$ tal que

$$P_{T_\tau} \psi(\mathbf{r}, t) = \psi'(\mathbf{r}, t) = \psi(\mathbf{r}, t - \tau).$$

Desenvolvendo a função $\psi(\mathbf{r}, t - \tau)$ em uma série de Taylor em torno de t , é fácil concluir que

$$P_{T_\tau} = \exp\left(-\tau \frac{\partial}{\partial t}\right).$$

Porém, de acordo com a equação de Schroedinger (5.43), observa-se que

$$P_{T_\tau} = \exp\left(i\tau \frac{\mathcal{H}}{\hbar}\right).$$

Este resultado mostra também que P_{T_τ} é um operador unitário, uma vez que o Hamiltoniano é Hermitiano.

Se o sistema físico é invariante frente a qualquer deslocamento temporal, a forma de P_{T_τ} mostra que a energia do mesmo é uma constante de movimento e, também, que este operador comuta com o Hamiltoniano,

$$[P_{T_\tau}, \mathcal{H}] = 0, \text{ para todo } \tau.$$

Sistemas que possuem essa simetria têm o seu estado descrito por funções de onda $\psi = \psi(\mathbf{r})$ determinadas pela equação de Schroedinger independente do tempo (5.43)

$$\mathcal{H}\psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad (5.47)$$

onde E é o autovalor de energia e $\psi(\mathbf{r})$ é a autofunção associada a este autovalor. Neste caso, a evolução temporal da função de onda é determinada simplesmente pela ação de P_{T_τ} sobre $\psi(\mathbf{r})$:

$$\Psi(\mathbf{r}, t) = P_{T_t} \psi(\mathbf{r}) = \psi(\mathbf{r}) e^{iEt/\hbar}.$$

5.9.2 O GRUPO DO HAMILTONIANO

Os exemplos anteriores destacaram alguns operadores de transformações que podem ser aplicados à função de onda $\psi(\mathbf{r}, t)$, como solução da equação de Schroedinger (5.43). Observou-se que se essas transformações são isométricas para o sistema físico, os operadores comutam com o Hamiltoniano e seus geradores (que são também observáveis do sistema) são constantes de movimento.

O Hamiltoniano, por si próprio, é um funcional de vários parâmetros do sistema, tais como posição, tempo, momento linear, momento angular, etc. A sua forma funcional específica reflete a simetria do sistema físico cuja dinâmica o mesmo está descrevendo.

Por exemplo, para uma única partícula com massa m movimentando-se sob a energia potencial V , o Hamiltoniano é

$$\mathcal{H} = -\frac{\hbar^2}{2m} \nabla^2 + V.$$

O Laplaciano ∇^2 é invariante frente as transformações ortogonais do sistema de referências, *i. e.*, frente ao grupo $O(3)$ que compreende as rotações próprias ou impróprias, e também invariante frente as translações espaciais. Por sua vez, o potencial V pode apresentar as mesmas simetrias ou outras. Se a partícula se movimenta sob um campo de forças centrais, então $V = V(r)$ e o potencial também é invariante frente o $O(3)$. Em consequência, o Hamiltoniano \mathcal{H} como um todo apresenta a simetria rotacional. Por outro lado, se a partícula encontra-se sob a ação de um potencial periódico, como o de uma rede cristalina, V é invariante frente as translações de algum grupo de isometrias espaciais.

A partir deste ponto, serão considerados sistemas para os quais o tempo é uma constante de movimento e que, portanto, têm sua dinâmica determinada pela equação de Schroedinger independente do tempo (5.47).

Seja então $\psi(\mathbf{r})$ uma função de onda determinada por $\mathcal{H}\psi = E\psi$, onde E é a energia do sistema. Seja agora P_R um operador que atua sobre a função de onda $\psi(\mathbf{r})$ e o qual corresponde a

uma determinada transformação de coordenadas (no sentido amplo) R , através do mapeamento (5.44). Considera-se a transformação

$$\psi'(\mathbf{r}) = P_R \psi(\mathbf{r}).$$

Aplicando o Hamiltoniano na relação acima, obtém-se

$$\mathcal{H}\psi'(\mathbf{r}) = \mathcal{H}P_R\psi(\mathbf{r}).$$

Por outro lado, como sempre existe a transformação inversa P_R^{-1} tal que $P_R^{-1}P_R = P_R P_R^{-1} = 1$, pode-se escrever

$$\mathcal{H}\psi'(\mathbf{r}) = P_R P_R^{-1} \mathcal{H} P_R \psi(\mathbf{r}) = P_R \mathcal{H}' \psi(\mathbf{r}),$$

onde \mathcal{H}' é o Hamiltoniano transformado por $\mathcal{H}' = P_R^{-1} \mathcal{H} P_R$. Entretanto se R, R^{-1} são isometrias do sistema físico, então o Hamiltoniano é invariante frente a P_R, P_R^{-1} , ou seja, $\mathcal{H}' = \mathcal{H}$. Em consequência,

$$P_R \mathcal{H} = \mathcal{H} P_R \implies [P_R, \mathcal{H}] = 0;$$

ou seja, o operador P_R comuta com o Hamiltoniano. Retornando, segue então de (5.47) que

$$\mathcal{H}\psi'(\mathbf{r}) = \mathcal{H}P_R\psi(\mathbf{r}) = P_R \mathcal{H}\psi(\mathbf{r}) = E P_R \psi(\mathbf{r}).$$

Do desenvolvimento acima, seguem então que:

1. Se R é uma isometria do sistema físico, então \mathcal{H} é invariante frente a ação de P_R .
2. Como \mathcal{H} é invariante frente a P_R , segue que $[P_R, \mathcal{H}] = 0$.
3. Como P_R comuta com \mathcal{H} , segue que

$$\psi'(\mathbf{r}) = P_R \psi(\mathbf{r})$$

também é autofunção de \mathcal{H} para o mesmo autovalor de energia E , pois

$$\mathcal{H}\psi' = \mathcal{H}(P_R\psi) = E(P_R\psi) = E\psi'.$$

4. Como $\psi' = P_R\psi$ está associada ao mesmo autovalor E , segue que $\psi(\mathbf{r})$ e $\psi'(\mathbf{r})$ são *autofunções degeneradas*.

Assumindo que as simetrias do sistema físico e, portanto, do Hamiltoniano, são conhecidas, então é em princípio possível determinar-se um conjunto de operadores que executam as transformações correspondentes a essas simetrias. Todos esses operadores comutam com \mathcal{H} e formam o denominado *grupo do Hamiltoniano* ($\mathcal{P}_{\mathcal{H}}$). Este conjunto de operadores de fato forma um grupo frente a operação de concatenação de transformações porque: (i) a transformação trivial $P_R = 1$ é o elemento identidade deste grupo; (ii) para todo operador de transformação existe o elemento inverso que executa a transformação inversa; (iii) o produto de dois operadores quaisquer também é uma transformação que comuta com \mathcal{H} e (iv) a ação de operadores lineares é sempre associativa.

Diversas propriedades importantes do grupo do Hamiltoniano serão agora discutidas.

5.9.2.1 DEGENERESCÊNCIA NORMAL OU ACIDENTAL

Dado um certo sistema físico o qual tem associado o Hamiltoniano \mathcal{H} e sejam $\psi(\mathbf{r})$ e E respectivamente a autofunção e o autovalor de energia determinados por \mathcal{H} através de (5.47). Seja $\mathcal{P}_{\mathcal{H}}$ o grupo do Hamiltoniano \mathcal{H} . Todo $P_R \in \mathcal{P}_{\mathcal{H}}$ satisfaz a relação de comutação $[P_R, \mathcal{H}] = 0$ e gera uma nova autofunção $\psi'(\mathbf{r})$ degenerada a $\psi(\mathbf{r})$ ao mesmo autovalor E através da ação $\psi'(\mathbf{r}) = P_R\psi(\mathbf{r})$.

Pela aplicação de todos os elementos de $\mathcal{P}_{\mathcal{H}}$ a $\psi(\mathbf{r})$ cria-se um conjunto de autofunções degeneradas ao mesmo autovalor E . Em outras palavras, a aplicação de todas as operações de simetria que comutam com o Hamiltoniano gera um conjunto de estados físicos, todos com o mesmo valor de energia. Se este procedimento gerar todas as autofunções degeneradas a E , esta degenerescência é dita *normal* ou *essencial*. Se existir alguma outra função de onda com o mesmo valor de energia, mas que não é gerada pelas operações de simetria do sistema, esta

degenerescência é dita *acidental*, no sentido de que se trata de uma degenerescência sem uma origem evidente a partir das simetrias conhecidas do sistema.

A existência de degenerescências acidentais pode estar ligada a simetrias ocultas no sistema e que não são aparentes no Hamiltoniano. Neste caso, a presença de degenerescências acidentais pode indicar que grupo do Hamiltoniano não está completo. Outra razão para a sua ocorrência está ligada a determinadas combinações dos parâmetros físicos do sistema que levam dois ou mais níveis de energia, usualmente distintos, a se cruzarem nesta particular combinação. A presença de degenerescências acidentais também pode resultar em quantidades físicas conservadas adicionais, relacionadas às mesmas.

Este fato já fica aparente com as autofunções e autovalores do átomo de hidrogênio. Supondo que o grupo de simetria do sistema é o $O(3)$, uma vez que o potencial é central, a degenerescência entre as diferentes autofunções de um mesmo orbital atômico é normal: dada a autofunção p_x , as outras funções podem ser obtidas por rotações de coordenadas. Por outro lado, a degenerescência entre as autofunções de diferentes orbitais com o mesmo autovalor de energia (orbitais $2s$ e $2p$, por exemplo) deve ser considerada acidental. Isto ocorre porque existem simetrias ocultas no Hamiltoniano do átomo de hidrogênio. De fato, a degenerescência das suas funções de onda se torna normal quando um outro operador que comuta com o Hamiltoniano é considerado. Esse outro operador é usualmente o vetor de Runge-Lenz

$$\mathbf{A} = \frac{1}{2} (\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p}) - km \frac{\mathbf{r}}{r},$$

onde k é a constante de unidade na Lei de Coulomb. A comutação do operador de Runge-Lenz com o Hamiltoniano mostra que o grupo do átomo de hidrogênio tem as suas simetrias descritas por um grupo superior ao $O(3)$.

5.9.2.2 REPRESENTAÇÕES DE $\mathcal{P}_{\mathcal{H}}$

Assume-se agora que um dado autovalor $E = E_n$ de \mathcal{H} seja ℓ_n vezes degenerado ($\ell_n = 1, 2, \dots$), excluindo quaisquer degenerescências acidentais. O índice $n = 1, 2, \dots$ passa a ser adotado para se distinguir os distintos autovalores de energia e a quantidade ℓ_n é denominada a **ordem da degenerescência**. Selecionando aquelas autofunções de \mathcal{H} que são também degeneradas ao autovalor E_n , pode-se formar a partir destas o conjunto de autofunções degeneradas dado por

$$\Psi_n = \left\{ \psi_1^{(n)}, \dots, \psi_{\ell_n}^{(n)} \mid \mathcal{H}\psi_\nu^{(n)} = E_n\psi_\nu^{(n)}, (n = 1, 2, \dots; \nu = 1, \dots, \ell_n) \right\}.$$

A partir deste ponto, será empregada a notação $\psi_\nu^{(n)}(\mathbf{r})$ para identificar a ν -ésima autofunção pertencente ao conjunto Ψ_n , formado pelos autovetores degenerados à energia E_n .

As funções em Ψ_n formam um subespaço do espaço de Hilbert \mathcal{H} ao qual as autofunções de \mathcal{H} pertencem²¹ e, portanto, sempre é possível determinar um conjunto de funções que sejam tanto LI quanto ortonormais através de combinações lineares de funções em Ψ_n , via algum processo de ortogonalização como o de Gram-Schmidt. Ou seja, a partir das autofunções em Ψ_n sempre é possível se formar uma base em um subespaço, de dimensão menor ou igual a ℓ_n , do espaço de Hilbert \mathcal{H} .

A importância que os subespaços formados pelas bases em Ψ_n possuem está no fato de que estes são *subespaços invariantes de \mathcal{H} com relação ao grupo $\mathcal{P}_{\mathcal{H}}$* .²² Estes subespaços devem ser invariantes porque a aplicação de qualquer operador em $\mathcal{P}_{\mathcal{H}}$ em uma dada autofunção $\psi_\nu^{(n)} \in \Psi_n$ deve, necessariamente, gerar outra autofunção $\psi_\nu'^{(n)}$ associada ao mesmo autovalor E_n e pertencente ao conjunto Ψ_n . Portanto, o espaço de Hilbert \mathcal{H} ao qual todas as soluções de (5.47) pertencem pode ser contruído a partir da soma direta

$$\mathcal{H} = \bigoplus_{n=1,2,\dots} \mathcal{H}_n = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \dots,$$

sendo \mathcal{H}_n o subespaço formado pela base em Ψ_n .

Será assumido então que as autofunções $\psi_\nu^{(n)}(\mathbf{r}) \in \Psi_n$ (com $|\Psi_n| = \ell_n$) formam uma base ortonormal do subespaço \mathcal{H}_n . Considera-se agora um operador $P_R \in \mathcal{P}_{\mathcal{H}}$, o qual corresponde à transformação de coordenadas R via o mapeamento (5.44). É sempre possível construir uma

²¹Ver teorema 4.2

²²Ver definição 4.16

representação para $P_R \leftrightarrow R$. De acordo com a discussão realizada na seção 5.2.2, esta representação é obtida a partir da aplicação de P_R sobre $\psi_\nu^{(n)}(\mathbf{r})$, da qual resulta a função $\psi'(\mathbf{r})$ que descreve um outro estado físico do sistema com a mesma energia. A função $\psi'(\mathbf{r})$ pode sempre ser decomposta em termos dos vetores da base Ψ_n , resultando, dessa maneira,

$$\psi'(\mathbf{r}) = P_R \psi_\nu^{(n)} = \sum_{\mu=1}^{\ell_n} \psi_\mu^{(n)} \langle \psi_\mu^{(n)}, P_R \psi_\nu^{(n)} \rangle = \sum_{\mu=1}^{\ell_n} \psi_\mu^{(n)} \Gamma_{\mu\nu}^{(n)}(R), \quad (5.48a)$$

sendo

$$\Gamma_{\mu\nu}^{(n)}(R) \doteq \langle \psi_\mu^{(n)}, P_R \psi_\nu^{(n)} \rangle \rightsquigarrow \Gamma^{(n)}(R) \doteq \left[\Gamma_{\mu\nu}^{(n)}(R) \right]_{(\ell_n \times \ell_n)} \quad (5.48b)$$

a representação de P_R na base Ψ_n .

O conjunto

$$\Gamma^{(n)}(\mathcal{H}) = \left\{ \Gamma^{(n)}(R), \forall P_R \in \mathcal{P}_{\mathcal{H}} \right\} \quad (5.48c)$$

forma, portanto, uma *representação de dimensão ℓ_n do grupo do Hamiltoniano \mathcal{H}* . Estas representações são necessariamente *irredutíveis*, uma vez que a aplicação de qualquer $P_R \in \mathcal{P}_{\mathcal{H}}$ sobre qualquer $\psi_\nu^{(n)}$ sempre vai resultar em um vetor contido em \mathcal{H}_n (por este ser invariante); pela mesma razão, nunca será possível representar $P_R \psi_\nu^{(n)}$ pela soma direta de matrizes menores.

Uma demonstração de que $\Gamma^{(n)}(\mathcal{H})$ é irredutível segue da seguinte maneira. Supõe-se que a representação $\Gamma^{(n)}(\mathcal{H})$, obtida a partir do conjunto Ψ_n seja redutível. Então \mathcal{H}_n é dado pela soma direta de pelo menos dois subespaços menores. Existe então uma matriz S unitária que executa uma transformação de similaridade, correspondente a uma mudança de base $\Psi_n \rightarrow \Phi_n$, que irá colocar cada matriz em $\Gamma^{(n)}(\mathcal{H})$ na forma bloco-diagonal. Isto implica que agora o resultado da aplicação de $P_R \in \mathcal{P}_{\mathcal{H}}$ a qualquer nova autofunção $\varphi^{(n)} \in \Phi_n$ corresponderá a uma combinação linear que contém somente autofunções de seu subespaço e nenhuma dos complementares em \mathcal{H}_n . Isto significa que $P_R \varphi^{(n)}$ é degenerado ao mesmo resultado obtido a partir de qualquer outro subespaço complementar e, portanto, a degenerescência em E_n é acidental, contrariando a hipótese inicial de que a degenerescência é normal. Portanto, a representação $\Gamma^{(n)}(\mathcal{H})$ não pode ser redutível; ela deve ser necessariamente irredutível.

Pode-se repetir aqui a demonstração de que $\Gamma^{(n)}(\mathcal{H})$ de fato forma uma representação de $\mathcal{P}_{\mathcal{H}}$. Dadas as transformações $R, S \in \mathcal{R}$, com os correspondentes operadores $P_S, P_R \in \mathcal{P}_{\mathcal{H}}$, então $SR \leftrightarrow P_{SR} = P_S P_R$. Dado agora qualquer $\psi_\nu^{(n)} \in \Psi_n$, a representação de P_{SR} na base Ψ_n é então

$$\begin{aligned} P_{SR} \psi_\nu^{(n)} &= P_S P_R \psi_\nu^{(n)} = P_S \sum_{\kappa=1}^{\ell_n} \psi_\kappa^{(n)} \Gamma_{\kappa\nu}^{(n)}(R) \\ &= \sum_{\kappa=1}^{\ell_n} \left(P_S \psi_\kappa^{(n)} \right) \Gamma_{\kappa\nu}^{(n)}(R) = \sum_{\kappa, \lambda=1}^{\ell_n} \psi_\lambda^{(n)} \Gamma_{\lambda\kappa}^{(n)}(S) \Gamma_{\kappa\nu}^{(n)}(R) \\ &= \sum_{\lambda=1}^{\ell_n} \psi_\lambda^{(n)} \left[\Gamma^{(n)}(S) \Gamma^{(n)}(R) \right]_{\lambda\nu}. \end{aligned}$$

Ou seja,

$$P_{SR} \psi_\nu^{(n)} = \sum_{\lambda=1}^{\ell_n} \psi_\lambda^{(n)} \Gamma_{\lambda\nu}^{(n)}(SR) \implies \Gamma^{(n)}(SR) = \Gamma^{(n)}(S) \Gamma^{(n)}(R),$$

mostrando que $\Gamma^{(n)}(\mathcal{H})$ é de fato uma representação de $\mathcal{P}_{\mathcal{H}}$.

Portanto, conclui-se que *o conjunto das ℓ_n autofunções $\psi_\nu^{(n)}(\mathbf{r})$ que descrevem estados degenerados de energia E_n consistem nas funções de base que formam uma representação irredutível ℓ_n -dimensional $\Gamma^{(n)}(\mathcal{H})$ do grupo do Hamiltoniano.*

Pode-se mostrar facilmente que esta representação é também *unitária*, se a base Ψ_n for formada por autofunções ortonormais e se os elementos de $\mathcal{P}_{\mathcal{H}}$ forem operadores unitários com um parâmetro contínuo. Conforme mencionado nos exemplos apresentados no início desta seção, os operadores que formam o grupo $\mathcal{P}_{\mathcal{H}}$ atuam sobre as funções de onda e como estas, por sua vez, descrevem estados físicos (mensuráveis) do sistema, é assumido então que os membros de $\mathcal{P}_{\mathcal{H}}$ sejam expressos em termos de operadores Hermitianos (como os momentos linear ou angular), pois somente neste caso pode-se garantir que o valor esperado do observável seja real.²³ Esses

²³Ver discussão na seção 5.2.1.

operadores Hermitianos servirão de geradores infinitesimais dos grupos de transformação e se esses grupos possuírem somente um parâmetro livre, é sempre possível escrever os elementos de $\mathcal{P}_{\mathcal{H}}$ como operadores unitários.

Demonstração. Emprega-se a definição usual do produto interno de dois vetores de um espaço de Hilbert \mathcal{H} complexo, sendo que estes vetores são funções das coordenadas do espaço Euclidiano \mathbb{E}^3 . Neste caso, dados $\psi, \varphi \in \mathcal{H}$, o seu produto interno é

$$\langle \psi, \varphi \rangle = \int d^3r \psi^* \varphi.$$

Dados agora os autovetores $\psi_\mu^{(n)}, \psi_\nu^{(n)} \in \Psi_n$, como a base é ortonormal, resulta que

$$\langle \psi_\mu^{(n)}, \psi_\nu^{(n)} \rangle = \int d^3r \psi_\mu^{(n)*} \psi_\nu^{(n)} = \delta_{\mu\nu}.$$

Como o Hamiltoniano é Hermitiano, um operador $P_R \in \mathcal{P}_{\mathcal{H}}$ sempre pode ser construído de forma que este seja unitário. Neste caso, de acordo com o teorema 4.22,

$$\langle \psi_\mu^{(n)}, \psi_\nu^{(n)} \rangle = \langle P_R \psi_\mu^{(n)}, P_R \psi_\nu^{(n)} \rangle = \delta_{\mu\nu}.$$

Este resultado pode ser compreendido interpretando a ação de P_R como uma “rotação genérica” das coordenadas que não altera o produto interno. Então,

$$\begin{aligned} \delta_{\mu\nu} &= \langle P_R \psi_\mu^{(n)}, P_R \psi_\nu^{(n)} \rangle = \left\langle \sum_{\kappa=1}^{\ell_n} \psi_\kappa^{(n)} \Gamma_{\kappa\mu}^{(n)}(R), \sum_{\lambda=1}^{\ell_n} \psi_\lambda^{(n)} \Gamma_{\lambda\nu}^{(n)}(R) \right\rangle \\ &= \sum_{\kappa, \lambda=1}^{\ell_n} \langle \psi_\kappa^{(n)}, \psi_\lambda^{(n)} \rangle \Gamma_{\kappa\mu}^{(n)*}(R) \Gamma_{\lambda\nu}^{(n)}(R) = \sum_{\kappa=1}^{\ell_n} \Gamma_{\kappa\mu}^{(n)*}(R) \Gamma_{\kappa\nu}^{(n)}(R) \\ &= \sum_{\kappa=1}^{\ell_n} \Gamma_{\mu\kappa}^{(n)\dagger}(R) \Gamma_{\kappa\nu}^{(n)}(R) = \left[\Gamma^{(n)\dagger}(R) \Gamma^{(n)}(R) \right]_{\mu\nu}. \end{aligned}$$

Ou seja,

$$\Gamma^{(n)\dagger}(R) \Gamma^{(n)}(R) = I_{\ell_n}$$

e a representação é unitária. □

A relação entre a ortonormalidade das funções de base e a unitariedade da representação $\Gamma^{(n)}(\mathcal{H})$ discutida acima também pode ser considerada no sentido inverso. Supondo que Ψ_n seja formado por autofunções linearmente independentes, será mostrado agora que se os elementos do grupo $\mathcal{P}_{\mathcal{H}}$ forem operadores unitários então as autofunções serão necessariamente ortogonais entre si.

Demonstração. Sejam as autofunções $\psi_\nu^{(n)} \in \Psi_n$ e $\psi_\mu^{(m)} \in \Psi_m$, associadas respectivamente aos autovalores E_n e E_m . Seja também $P_R \in \mathcal{P}_{\mathcal{H}}$ um operador unitário. O produto interno destas funções satisfaz, portanto, a identidade

$$\langle \psi_\mu^{(m)}, \psi_\nu^{(n)} \rangle = \langle P_R \psi_\mu^{(m)}, P_R \psi_\nu^{(n)} \rangle.$$

Introduzindo a representação para P_R dada por (5.48), resulta

$$\begin{aligned} \langle \psi_\mu^{(m)}, \psi_\nu^{(n)} \rangle &= \left\langle \sum_{\kappa=1}^{\ell_m} \psi_\kappa^{(m)} \Gamma_{\kappa\mu}^{(m)}(R), \sum_{\lambda=1}^{\ell_n} \psi_\lambda^{(n)} \Gamma_{\lambda\nu}^{(n)}(R) \right\rangle \\ &= \sum_{\kappa=1}^{\ell_m} \sum_{\lambda=1}^{\ell_n} \Gamma_{\kappa\mu}^{(m)*}(R) \Gamma_{\lambda\nu}^{(n)}(R) \langle \psi_\kappa^{(m)}, \psi_\lambda^{(n)} \rangle. \end{aligned}$$

Observa-se que o lado esquerdo da identidade acima não depende de P_R . Então, o mesmo produto interno à esquerda ocorrerá para todos os elementos de $\mathcal{P}_{\mathcal{H}}$. Somando todas as equações correspondentes a todos os elementos do grupo, resulta

$$g \langle \psi_\mu^{(m)}, \psi_\nu^{(n)} \rangle = \sum_{\kappa=1}^{\ell_m} \sum_{\lambda=1}^{\ell_n} \sum_{P_R \in \mathcal{P}_{\mathcal{H}}} \Gamma_{\kappa\mu}^{(m)*}(R) \Gamma_{\lambda\nu}^{(n)}(R) \langle \psi_\kappa^{(m)}, \psi_\lambda^{(n)} \rangle,$$

onde $g = |\mathcal{P}_{\mathcal{H}}|$ será assumido finito, por enquanto.

Segundo o teorema 5.1, as representações de P_R em \mathcal{H}_n e \mathcal{H}_m ou são unitárias ou podem se tornar unitárias frente a transformações de similaridade, que correspondem a mudanças de base em Ψ_n e Ψ_m . Portanto, de acordo com o grande teorema da ortogonalidade 5.2, resulta então

$$\langle \psi_{\mu}^{(m)}, \psi_{\nu}^{(n)} \rangle = \frac{1}{\ell_n} \delta_{nm} \delta_{\mu\nu} \sum_{\kappa=1}^{\ell_m} \sum_{\lambda=1}^{\ell_n} \delta_{\kappa\lambda} \langle \psi_{\kappa}^{(m)}, \psi_{\lambda}^{(n)} \rangle. \quad (5.49a)$$

O resultado acima mostra que se $n \neq m$ e/ou $\mu \neq \nu$,

$$\langle \psi_{\mu}^{(m)}, \psi_{\nu}^{(n)} \rangle = 0, \quad \begin{pmatrix} n \neq m \\ \mu \neq \nu \end{pmatrix}. \quad (5.49b)$$

Ou seja, as propriedades do grupo $\mathcal{P}_{\mathcal{H}}$ mostram que as autofunções $\psi_{\mu}^{(m)}$ e $\psi_{\nu}^{(n)}$ são necessariamente ortogonais quanto estas estão associadas a distintos autovalores de energia ou são distintas autofunções do mesmo autovalor E_n . Por outro lado, se $m = n$ e $\mu = \nu$,

$$\langle \psi_{\nu}^{(n)}, \psi_{\nu}^{(n)} \rangle = \frac{1}{\ell_n} \sum_{\kappa, \lambda=1}^{\ell_n} \delta_{\kappa\lambda} \langle \psi_{\kappa}^{(n)}, \psi_{\lambda}^{(n)} \rangle = \frac{1}{\ell_n} \sum_{\kappa=1}^{\ell_n} \langle \psi_{\kappa}^{(n)}, \psi_{\kappa}^{(n)} \rangle.$$

Ou seja, obtém-se a seguinte relação entre as normas das autofunções:

$$\left\| \psi_{\nu}^{(n)} \right\|^2 = \frac{1}{\ell_n} \sum_{\kappa=1}^{\ell_n} \left\| \psi_{\kappa}^{(n)} \right\|^2, \quad \forall \nu.$$

Mas, somando ambos os lados sobre o índice $\nu = 1, \dots, \ell_n$, conclui-se que, necessariamente,

$$\left\| \psi_{\nu}^{(n)} \right\| = \left\| \psi_{\kappa}^{(n)} \right\|, \quad (1 \leq \nu, \kappa \leq \ell_n). \quad (5.49c)$$

Ou seja, todas as autofunções em Ψ_n possuem a mesma norma. \square

Embora as conclusões acima tenham sido obtidas supondo que $\mathcal{P}_{\mathcal{H}}$ é finito, as mesmas valem também para grupos do Hamiltoniano contínuos e compactos. Este resultado é importante porque demonstra a ortogonalidade das autofunções, mesmo que estas sejam degeneradas na energia. Este resultado foi obtido de uma forma genérica, empregando somente as propriedades de simetria do sistema físico. A bem conhecida ortogonalidade das autofunções do átomo de Hidrogênio, tanto dentro de um orbital quanto entre diferentes orbitais, é um exemplo deste resultado.

5.9.2.3 TEORIA DE GRUPOS E “BONS” NÚMEROS QUÂNTICOS

Em mecânica quântica, um “bom” número quântico é aquele número/índice (ou conjunto de números/índices) que identificam os diferentes autovalores de observáveis que são constantes de movimento. Uma condição necessária e suficiente para tanto ocorre se o operador associado ao observável comutar com o Hamiltoniano, pois neste caso o valor esperado do observável será constante.

A teoria de representações de grupos oferece uma relação direta entre as irreps do grupo do Hamiltoniano e bons números quânticos. De acordo com a discussão realizada na seção anterior, a representação de qualquer $P_R \in \mathcal{P}_{\mathcal{H}}$ é obtida a partir de uma dada autofunção $\psi_{\nu}^{(n)} \in \Psi_n$ via as expressões (5.48). Estas expressões mostram que $\psi_{\nu}^{(n)}$ está sempre relacionada com a ν -ésima coluna de qualquer matriz do conjunto $\Gamma^{(n)}(\mathcal{H})$, que forma a n -ésima irrep do grupo do Hamiltoniano. Estabelece-se então a correspondência $\psi_{\nu}^{(n)} \rightsquigarrow \Gamma_{\mu\nu}^{(n)}(\mathcal{H})$ e diz-se que $\psi_{\nu}^{(n)}$ pertence à (ou que se transforma de acordo com a) ν -ésima coluna da n -ésima representação irredutível de $\mathcal{P}_{\mathcal{H}}$.

Caso fosse realizada uma mudança de base $\Psi_n \rightarrow \Psi'_n$ em \mathcal{H}_n através da relação²⁴

$$\psi_{\nu}^{(n)'} = \sum_{\mu=1}^{\ell_n} \psi_{\mu}^{(n)} A_{\mu\nu} \iff \psi_{\nu}^{(n)} = \sum_{\mu=1}^{\ell_n} \psi_{\mu}^{(n)'} A_{\mu\nu}^{-1},$$

²⁴Ver seção 4.3.1.

onde $\psi_\nu^{(n)} \in \Psi_n$, $\psi_\mu^{(n)'} \in \Psi_n'$ e $A = [A_{\mu\nu}] : (\ell_n \times \ell_n)$ uma matriz não singular, então, para todo $P_R \in \mathcal{P}_{\mathcal{H}}$,

$$\begin{aligned} P_R \psi_\nu^{(n)'} &= \sum_{\mu=1}^{\ell_n} P_R \psi_\mu^{(n)} A_{\mu\nu} \stackrel{(5.48)}{=} \sum_{\mu,\kappa=1}^{\ell_n} \psi_\kappa^{(n)} \Gamma_{\kappa\mu}^{(n)}(R) A_{\mu\nu} \\ &= \sum_{\mu,\kappa,\lambda=1}^{\ell_n} \psi_\lambda^{(n)'} A_{\lambda\kappa}^{-1} \Gamma_{\kappa\mu}^{(n)}(R) A_{\mu\nu} = \sum_{\lambda=1}^{\ell_n} \psi_\lambda^{(n)'} \left[A^{-1} \Gamma^{(n)}(R) A \right]_{\lambda\nu} \\ &= \sum_{\lambda=1}^{\ell_n} \psi_\lambda^{(n)'} \Gamma_{\lambda\nu}^{(n)'}(R), \end{aligned}$$

onde

$$\Gamma^{(n)'}(R) \doteq A^{-1} \Gamma^{(n)}(R) A$$

é a representação de P_R na base Ψ_n' .

Contudo, esta relação nada mais é senão a transformação de similaridade (5.5); ou seja, $\Gamma^{(n)'}(R)$ é equivalente a $\Gamma^{(n)}(R)$. Além disso, $\psi_\nu^{(n)'}(r)$ também se transforma de acordo com a ν -ésima coluna de $\Gamma^{(n)'}(R)$. Portanto, a menos de uma transformação de similaridade, *existe uma única representação irredutível do grupo do Hamiltoniano correspondente a cada autovalor de energia*.

Um conjunto de autofunções pode sempre ser classificado ou identificado de uma única maneira, de acordo com a representação irredutível à qual pertence. Desta forma, a teoria de representações de grupos fornece os “bons números quânticos” para a dinâmica do sistema físico através dos índices relacionados ao autovalor de energia (n) e à coluna da irrep $\Gamma^{(n)}(\mathcal{H})$. A degenerescência (normal) do autovalor é simplesmente a dimensionalidade da representação. Caso seja possível calcular as dimensionalidades de todas as irreps do grupo do Hamiltoniano, as ordens das degenerescências (normais) serão automaticamente obtidas.

Se ocorrerem degenerescências acidentais associadas a um dado autovalor E_n , as distintas bases dos subespaços invariantes correspondentes podem ser facilmente distinguidas, bastando para isso introduzir um índice adicional. Dessa forma, a base

$$\Psi_{pn} \doteq \left\{ \psi_{p\nu}^{(n)}(r) \mid \mathcal{H} \psi_{p\nu}^{(n)} = E_{pn} \psi_{p\nu}^{(n)}, (n = 1, 2, \dots; p = 1, 2, \dots; \nu = 1, \dots, \ell_n; E_{1n} = E_{2n} = \dots) \right\}$$

é formada por autofunções tais que uma dada $\psi_{p\nu}^{(n)} \in \Psi_{pn}$ é a autofunção que se transforma de acordo com a ν -ésima coluna da n -ésima irrep de $\mathcal{P}_{\mathcal{H}}$ que ocorre pela p -ésima vez na distribuição de níveis de energia do sistema físico. Nesta situação, é adequado identificar os autovalores por E_{pn} e os bons números quânticos serão portanto $\{p, n, \nu\}$.

Esta discussão também implica que uma perturbação aplicada ao sistema físico somente poderá quebrar degenerescências se, e somente se, a sua inclusão no Hamiltoniano reduzir o grupo de simetria do sistema, alterando assim as suas representações irredutíveis.

Exemplo 5.9. Um exemplo simples, envolvendo um grupo $\mathcal{P}_{\mathcal{H}}$ finito, consiste em um elétron movendo-se sob o potencial de três prótons localizados nos vértices de um triângulo equilátero.

Neste caso, o sistema físico claramente possui as isometrias do triângulo, discutidas pela primeira vez nos exercícios 3.10 e 3.11. Estas isometrias formam o grupo \mathcal{C}_{3v} , isomórfico ao S_3 e que possui três representações irredutíveis (exercício 5.1), de dimensões 1, 1 e 2, apresentadas na tabela 5.1.

Portanto, este sistema possui somente 3 (três) autoestados de energia, sendo que os dois primeiros são não degenerados e o terceiro é duplamente degenerado. De acordo com (5.48), a autofunção do estado E_1 ($\psi_1^{(1)}(r)$) permanece invariante frente a todas as operações de simetria do grupo, uma vez que estas são operacionalizadas por $\Gamma^{(1)}(S_3)$. Por outro lado, a autofunção do estado E_2 ($\psi_2^{(2)}(r)$), em ações são descritas por $\Gamma^{(2)}(S_3)$, permanece invariante frente as rotações próprias, operadas por $\{I, \pi_4, \pi_5\}_{S_3} \equiv \{E, C_3^2, C_3\}_{\mathcal{C}_{3v}}$, mas muda de sinal frente as rotações impróprias $\{\pi_2, \pi_3, \pi_6\}_{S_3} \equiv \{\sigma_{v16}, \sigma_{v35}, \sigma_{v24}\}_{\mathcal{C}_{3v}}$. Finalmente, as duas autofunções do estado E_3 são transformadas de acordo com $\Gamma^{(3)}(S_3)$:

$$P_{\pi_i} \psi_\nu^{(3)} = \sum_{\mu=1}^2 \psi_\mu^{(3)} \Gamma_{\mu\nu}^{(3)}(\pi_i), \quad \left(\begin{array}{l} i = 1, \dots, 6 \\ \nu = 1, 2 \end{array} \right);$$

ou seja, $\psi_\nu^{(3)}$ transforma-se de acordo com a ν -ésima coluna de $\Gamma^{(3)}(S_3)$. Por exemplo, uma reflexão em torno do plano 3 – 5 do triângulo (figura 3.5) irá realizar as transformações

$$\psi_{1,2}^{(3)} \xrightarrow{P_{\sigma_{v35}}} \psi_{1,2}^{(3)'} : \psi_\nu^{(3)'} = P_{\sigma_{v35}} \psi_\nu^{(3)} = \sum_{\mu=1}^2 \psi_\mu^{(3)} \Gamma_{\mu\nu}^{(3)}(\sigma_{v35}), (\nu = 1, 2),$$

$$\begin{pmatrix} \psi_1^{(3)'} \\ \psi_2^{(3)'} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \psi_1^{(3)} & \psi_2^{(3)} \end{pmatrix} \begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \psi_1^{(3)} - \sqrt{3}\psi_2^{(3)} \\ -\sqrt{3}\psi_1^{(3)} - \psi_2^{(3)} \end{pmatrix}.$$

5.9.3 GRUPOS ABELIANOS E O TEOREMA DE BLOCH

Algumas conclusões importantes já podem ser obtidas para sistemas físicos que possuem propriedades de simetria que formam grupos Abelianos. Um exemplo de sistema com essa propriedade é uma rede cristalina, a qual é isométrica frente a algum grupo cristalográfico espacial.²⁵

Em um grupo Abeliano, cada elemento forma uma classe por si próprio²⁶ e, por consequência, o número de representações irredutíveis é igual ao número de elementos do grupo. Contudo, como (5.21) deve sempre ser satisfeita, neste caso

$$\sum_{n=1}^g \ell_n^2 = g, \text{ sendo } g = |\mathcal{P}_{\mathcal{H}}|,$$

e a única possibilidade é $\ell_n = 1$, para $n = 1, 2, \dots, g$.

Portanto, um grupo Abeliano de ordem g tem g representações irredutíveis unidimensionais e um sistema físico cujas isometrias pertencem a esse grupo possui o mesmo número de autoestados de energia. Um sistema cujo grupo do Hamiltoniano é Abeliano tem seus autoestados descritos por autofunções todas não degeneradas.

5.9.3.1 GRUPOS CÍCLICOS

Se além de Abeliano, o grupo for também cíclico, então $\mathcal{P}_{\mathcal{H}} = \langle P_A : P_A^g = I \rangle = \{P_A, P_A^2, \dots, P_A^g = I\}$, sendo P_A o gerador do grupo.

Denotando por $\Gamma(A) = r$ o valor que representa P_A em uma certa irrep de $\mathcal{P}_{\mathcal{H}}$, a discussão realizada ao longo da seção 5.6 mostra que $\Gamma(P_A^m) = [\Gamma(P_A)]^m = r^m$ ($m = 2, \dots, g$). Em particular, como $\Gamma(P_A^g) = \Gamma(I) = 1$, resulta que a quantidade r satisfaz $r^g = 1$. Ou seja, as g irreps de P_A são dadas pelas g raízes da unidade. Assim, pode-se escolher

$$r^g = 1 \implies \Gamma^{(n)}(P_A) = r_n = e^{2\pi(n-1)i/g}, (n = 1, \dots, g),$$

sendo que esta escolha foi realizada para que $\Gamma^{(1)}(P_A) = 1$ seja a irrep trivial.

As irreps de P_A^2 são, por conseguinte, obtidas a partir de $\Gamma^{(n)}(P_A^2) = [\Gamma^{(n)}(P_A)]^2 = r_n^2 = e^{4\pi(n-1)i/g}$. Assim, para qualquer elemento de $\mathcal{P}_{\mathcal{H}}$ pode-se escrever

$$\Gamma^{(n)}(P_A^m) = e^{2\pi m(n-1)i/g}, (n, m = 1, \dots, g). \quad (5.50)$$

Nota-se que com as escolhas aqui realizadas, a irrep trivial corresponde a $\Gamma^{(1)}(P_A^m) = 1$, ($m = 1, \dots, g$). A tabela 5.4 lista algumas das irreps de $\mathcal{P}_{\mathcal{H}}$. A mesma tabela também serve para listar os caracteres das representações.

5.9.3.2 O TEOREMA DE BLOCH

O teorema de Bloch possui grande importância da dinâmica de uma rede cristalina e pode ser deduzido facilmente considerando a simetria translacional inerente a estes sistemas. Este teorema será deduzido agora para uma rede unidimensional, mas o mesmo pode ser estendido para outras dimensões.

Conforme discutido na seção 3.4.3, em uma dimensão o grupo cíclico de ordem N é o grupo de simetria do Hamiltoniano para o potencial periódico de uma rede cristalina com N períodos

²⁵Seção 3.4.3.

²⁶Ver página 101.

Tabela 5.4: Tabela das representações irredutíveis do grupo cíclico de ordem g $\langle P_A : P_A^g = I^\alpha \rangle$, sendo $\omega = e^{2\pi i/g}$.

	I	P_A	P_A^2	P_A^3	\dots	P_A^{g-1}
$\Gamma^{(1)}$	1	1	1	1	\dots	1
$\Gamma^{(2)}$	1	ω	ω^2	ω^3	\dots	$\omega^{(g-1)}$
$\Gamma^{(3)}$	1	ω^2	ω^4	ω^6	\dots	$\omega^{2(g-1)}$
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
$\Gamma^{(g)}$	1	$\omega^{(g-1)}$	$\omega^{2(g-1)}$	$\omega^{3(g-1)}$	\dots	$\omega^{(g-1)^2}$

em um anel circular ou em uma rede linear com condições de contorno periódicas, sendo a o parâmetro de rede. Neste caso, um elemento do grupo translação \mathcal{T} é o operador T_1 , cuja ação sobre a coordenada x ao longo da rede é

$$T_1 x = x' = x + a,$$

ou seja, transladar a origem para a esquerda por uma distância igual ao parâmetro de rede. O operador inverso é $T_1^{-1} x = x - a$ e o operador que executa a translação por $m \leq N$ períodos de rede é simplesmente $T_m x = T_1^m x = x + ma$, onde, devido às condições de contorno, $T_N x = T_1^N x = x + Na = x = Ex$. Assim, observa-se que T_1 é o gerador do grupo de translação, o qual pode ser descrito por $\mathcal{T} = \langle T_1 : T_1^N = E \rangle$.

De acordo com o mapeamento (5.44), o operador P_{T_1} que efetua a transformação correspondente na função de onda é dado por

$$P_{T_1} \psi(x) = \psi(T_1^{-1} x) = \psi(x - a).$$

Da mesma forma, o operador $P_{T_m} \leftrightarrow T_m$ tem sua ação dada por $P_{T_m} \psi(x) = \psi(T_m^{-1} x) = \psi(x - ma)$.

Conforme a discussão realizada a respeito das propriedades do grupo do Hamiltoniano, dada a equação de Schroedinger (5.47) para um potencial periódico, o número total de autovalores de energia possíveis é igual ao número de representações irredutíveis do grupo $\mathcal{P}_{\mathcal{U}}$; ou seja, $E = E_n$, com $1 \leq n \leq N$. Além disso, como o grupo \mathcal{T} é cíclico, todos os autovalores são não degenerados, pois as irreps são unidimensionais. Finalmente, de acordo com (5.48), a ação de qualquer operador em \mathcal{T} implica na transformação de $\psi(x)$ de acordo com alguma representação irredutível de $\mathcal{P}_{\mathcal{U}}$.

Assim, dada autofunção $\psi^{(n)}(x)$ associada ao autovalor E_n , esta deve se transformar sob a ação de P_{T_1} de acordo com a n -ésima irrep de $\mathcal{P}_{\mathcal{U}}$. De acordo com (5.50), resulta então

$$\psi^{(n)}(x - a) = P_{T_1} \psi^{(n)}(x) = \Gamma^{(n)}(T_1) \psi^{(n)}(x) = e^{2\pi(n-1)i/N} \psi^{(n)}(x) = \omega^{n-1} \psi^{(n)}(x),$$

onde o parâmetro ω está na tabela 5.4. Como as irreps na tabela são puramente fases unimodulares, a seguinte propriedade é imediata:

$$\left| \psi^{(n)}(x - a) \right|^2 = \left| \psi^{(n)}(x) \right|^2.$$

Portanto, a forma mais geral para $\psi^{(n)}(x)$ é

$$\psi^{(n)}(x) = e^{-i\phi_n(x)} u_n(x),$$

onde $\phi_n(x)$ é uma função real denominada função de fase e $u_n(x)$ apresenta a periodicidade da rede: $u_n(x - a) = u_n(x)$.

A forma explícita de $\phi_n(x)$ pode ser determinada. Aplicando uma translação qualquer,

$$P_{T_m} \psi^{(n)}(x) \stackrel{(5.50)}{=} \omega^{m(n-1)} \psi^{(n)}(x) = \omega^{m(n-1)} e^{-i\phi_n(x)} u_n(x).$$

Por outro lado,

$$P_{T_m} \psi^{(n)}(x) = \psi^{(n)}(x - ma) = e^{-i\phi_n(x-ma)} u_n(x).$$

Igualando ambos os resultados, conclui-se que a função de fase $\phi_n(x)$ deve satisfazer

$$\phi_n(x - ma) = \phi_n(x) - \frac{2\pi m(n-1)}{N}.$$

Desenvolvendo $\phi_n(x + ma)$ em uma série de Taylor em torno de x ,

$$\phi_n(x - ma) = \sum_{r=0}^{\infty} \frac{(-m)^r a^r}{r!} \phi_n^{(r)}(x) = \phi_n(x) - ma\phi_n'(x) + \frac{1}{2}m^2 a^2 \phi_n''(x) + \dots,$$

conclui-se que a identidade somente pode ser satisfeita se $\phi_n''(x) = 0$, ou seja, se esta for linear em x . Obtém-se também que

$$\phi_n'(x) = \frac{2\pi(n-1)}{L} \Rightarrow \phi_n(x) = \frac{2\pi(n-1)}{L}x + \alpha,$$

onde $L = Na$ é a extensão da rede periódica e α é uma constante.

Portanto, definindo a nova quantidade $k_n = 2\pi(n-1)/L$, denominada o **número de onda**, a forma geral da autofunção é

$$\psi_n(x) = e^{-ik_n x} u_n(x).$$

Este resultado é o teorema de Bloch (unidimensional) e esta forma para $\psi_n(x)$ é denominada **função de Bloch**.

A dedução recém concluída pode ser estendida para uma rede cristalina real (em 3D). Fazendo isso, pode-se mostrar que o número de onda é generalizado para o **vetor de onda**

$$\mathbf{k}_{n_1, n_2, n_3} = 2\pi \sum_{r=1}^3 \frac{n_r - 1}{N_r} \mathbf{b}_r,$$

onde $\{\mathbf{b}_r\}$ são os vetores translacionais da rede recíproca. Com isso, a função de Bloch fica escrita

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{k} \cdot \mathbf{r}} u_{\mathbf{k}}(\mathbf{r}).$$

De acordo com a discussão realizada na seção 5.9.2.3, o vetor de onda \mathbf{k} é um bom número quântico para a função de onda de um elétron em um cristal, desde que a simetria translacional não seja quebrada por alguma perturbação. Neste caso, o vetor de onda aparece relacionado com o espalhamento do elétron e $\hbar\mathbf{k}$ é o momento de um quantum de vibração da rede.

5.9.4 FUNÇÕES DE BASE PARA REPRESENTAÇÕES IRREDUTÍVEIS

Os exemplos apresentados na seção anterior já dão uma indicação da utilidade da teoria de representação para obter-se informações importantes a respeito da dinâmica do sistema quando o seu grupo de simetria for Abelian.

Serão agora desenvolvidos métodos mais gerais, que irão também fornecer informações importantes quando os estados de energia são degenerados, em cuja situação as representações irredutíveis possuem dimensões maiores que um. Nesta situação, conforme salientado na discussão a respeito de bons números quânticos (seção 5.9.2.3), as funções de base dos subespaços vetoriais necessitam de, no mínimo, dois índices para a sua identificação, um índice para indicar a representação irredutível à qual pertencem e outro para indicar a coluna (ou linha) dentro da representação.

Retomando então a nomenclatura introduzida na seção 5.9.2.3, $\psi_{\nu}^{(n)}(\mathbf{r})$ é a função de base pertencente à ν -ésima coluna da n -ésima irrep. As outras autofunções $\psi_{\mu}^{(n)}(\mathbf{r})$ necessárias para completar a base da representação são denominadas as **parceiras** de $\psi_{\nu}^{(n)}(\mathbf{r})$. Retornando a (5.48), a ação de qualquer elemento P_R do grupo do Hamiltoniano $\mathcal{P}_{\mathcal{H}}$ sobre $\psi_{\nu}^{(n)}(\mathbf{r})$ pode ser expresso com o auxílio das parceiras como

$$P_R \psi_{\nu}^{(n)} = \sum_{\mu=1}^{\ell_n} \psi_{\mu}^{(n)} \Gamma_{\mu\nu}^{(n)}(R),$$

sendo ℓ_n a ordem da degenerescência do n -ésimo estado de energia. Multiplicando agora ambos os lados por $\Gamma_{\mu'\nu'}^{(m)*}(R)$ e somando sobre os elementos do grupo, o teorema da ortogonalidade 5.2 fornece

$$\sum_{P_R \in \mathcal{P}_{\mathcal{H}}} \Gamma_{\mu'\nu'}^{(m)*}(R) P_R \psi_{\nu}^{(n)} = \sum_{\mu=1}^{\ell_n} \sum_{P_R \in \mathcal{P}_{\mathcal{H}}} \psi_{\mu}^{(n)} \Gamma_{\mu\nu}^{(n)}(R) \Gamma_{\mu'\nu'}^{(m)*}(R)$$

$$\begin{aligned} &= \frac{g}{\ell_n} \sum_{\mu=1}^{\ell_n} \psi_{\mu}^{(n)} \delta_{mn} \delta_{\mu\mu'} \delta_{\nu\nu'} \\ &= \frac{g}{\ell_n} \delta_{mn} \delta_{\nu\nu'} \psi_{\mu'}^{(n)}, \end{aligned}$$

lembrando que $g = |\mathcal{P}_{\mathcal{H}}|$.

O resultado acima permite definir o **operador de transferência**

$$\mathfrak{p}_{\mu\nu}^{(n)} \doteq \frac{\ell_n}{g} \sum_{P_R \in \mathcal{P}_{\mathcal{H}}} \Gamma_{\mu\nu}^{(n)*}(R) P_R, \quad (5.51a)$$

com o qual pode-se escrever o resultado como

$$\mathfrak{p}_{\mu\kappa}^{(k)} \psi_{\nu}^{(n)} = \delta_{nk} \delta_{\nu\kappa} \psi_{\mu}^{(k)}. \quad (5.51b)$$

Ou seja, a atuação do operador $\mathfrak{p}_{\mu\kappa}^{(m)}$ sobre uma função de base gera um resultado nulo, a não ser que esta pertença à κ -ésima coluna da k -ésima representação irredutível. Neste caso, se $n = k$ e $\nu = \kappa$, resulta que

$$\mathfrak{p}_{\mu\nu}^{(n)} \psi_{\nu}^{(n)} = \psi_{\mu}^{(n)},$$

ou seja, ao se aplicar (5.51a) na função de base $\psi_{\nu}^{(n)}$, obteve-se a sua parceira $\psi_{\mu}^{(n)}$. Esta propriedade do operador $\mathfrak{p}_{\mu\nu}^{(n)}$ fornece um método para se obter todas funções de base de uma dada irrep conhecendo somente uma delas.

Retornando ao último resultado acima, se $\mu = \nu$, resulta que

$$\mathfrak{p}_{\nu\nu}^{(n)} \psi_{\nu}^{(n)} = \psi_{\nu}^{(n)}.$$

Ou seja, $\psi_{\nu}^{(n)}$ é uma autofunção de $\mathfrak{p}_{\nu\nu}^{(n)}$ com autovalor unitário. Esta propriedade permite identificar de forma única os índices de qualquer função de base e até mesmo se uma dada autofunção associada a E_n é de fato uma função de base. Além disso, como $\mathfrak{p}_{\nu\nu}^{(n)}$ é um operador linear, qualquer combinação de funções pertencentes à ν -ésima coluna de $\Gamma^{(n)}$, mas oriundas de diferentes escolhas de bases (tal como $a\psi_{\nu}^{(n)} + b\psi_{\nu}^{\prime(n)}$), também pertencerá às mesmas coluna e representação. Finalmente, como o autovalor é unitário, o operador é tal que $\mathfrak{p}_{\nu\nu}^{(n)} \mathfrak{p}_{\nu\nu}^{(n)} = \mathfrak{p}_{\nu\nu}^{(n)}$. Operadores que possuem esta propriedade são denominados *idempotentes*.

Com base nas propriedades do operador de transferência, o teorema a seguir pode ser formulado.

Teorema 5.8. *Se $\Gamma^{(1)}, \Gamma^{(2)}, \dots, \Gamma^{(N)}$ são as representações irredutíveis de um grupo de operadores P_R , então qualquer função F no espaço operado por P_R pode ser decomposto em uma soma na forma*

$$F = \sum_{n=1}^N \sum_{\nu=1}^{\ell_n} f_{\nu}^{(n)},$$

onde $f_{\nu}^{(n)}$ pertence à ν -ésima coluna da n -ésima representação irredutível do grupo.

Demonstração. Considera-se todas as funções $F, F_2', F_3', \dots, F_g'$ obtidas pelas aplicações de todos os operadores P_R sobre F . Primeiro descarta-se todas as funções que são LD das outras e ortonormaliza-se as demais (e. g. via o processo de Gram-Schmidt). Denotando o conjunto resultante por $\Phi = \{F, F_2, \dots, F_m\}$ ($m \leq g$), estas funções formam a base para uma representação unitária do grupo, identificada por $\hat{D}(\mathcal{H})$, de tal forma que

$$P_R F_k = \sum_{r=1}^m F_r \hat{D}_{rk}(R), \quad (k = 1, \dots, m).$$

Agora, há duas possibilidades: ou a representação \hat{D} é irredutível ou não. Se for irredutível, então F_k pertence à sua k -ésima coluna e o teorema está provado. Se não o for, então existe

uma matriz não singular S que gera uma transformação de similaridade que irá colocar todas as matrizes em \hat{D} na forma bloco-diagonal. Neste caso, as novas funções

$$F_\ell'' = \sum_{k=1}^m F_k S_{k\ell}, \quad (\ell = 1, \dots, m)$$

podem ser escritas como combinações lineares de funções do tipo $f_\nu^{(n)}$, as quais são as funções de base das irreps. Usando então a matriz inversa S^{-1} , pode-se finalmente escrever as funções em Φ (em particular F) em termos das funções $f_\nu^{(n)}$. \square

Com base neste teorema, é possível empregar o operador de transferência (5.51a) para a determinação das funções de base. Como foi constatado que

$$\mathfrak{p}_{\kappa\kappa}^{(k)} f_\nu^{(n)} = \delta_{nk} \delta_{\nu\kappa} f_\kappa^{(k)}, \quad (5.52a)$$

a aplicação de $\mathfrak{p}_{\kappa\kappa}^{(k)}$ sobre uma função qualquer F resulta

$$\mathfrak{p}_{\kappa\kappa}^{(k)} F = \mathfrak{p}_{\kappa\kappa}^{(k)} \left(\sum_{n=1}^g \sum_{\nu=1}^{\ell_n} f_\nu^{(n)} \right) = \sum_{n=1}^g \sum_{\nu=1}^{\ell_n} \mathfrak{p}_{\kappa\kappa}^{(k)} f_\nu^{(n)} \implies \mathfrak{p}_{\kappa\kappa}^{(k)} F = f_\kappa^{(k)}. \quad (5.52b)$$

Assim, $\mathfrak{p}_{\kappa\kappa}^{(k)}$ atua sobre qualquer função F projetando sua componente ao longo da κ -ésima coluna da irrep $\Gamma^{(k)}$. Por esta razão, $\mathfrak{p}_{\kappa\kappa}^{(k)}$ é denominado **operador de projeção**.

É importante mencionar que a função $f_\nu^{(n)}(r)$ projetada pelo operador $\mathfrak{p}_{\nu\nu}^{(n)}$ aplicado a uma função arbitrária $F(r)$ está diretamente relacionada à função de base $\psi_\nu^{(n)}(r)$, pertencente à ν -ésima coluna da irrep $\Gamma^{(n)}$, por $f_\nu^{(n)}(r) = c_\nu^{(n)} \psi_\nu^{(n)}(r)$, sendo $c_\nu^{(n)} \in \mathbb{C}$ uma constante a ser determinada, e. g., pela normalização da autofunção.

Com os recursos deduzidos acima, é possível obter-se as funções de base de qualquer representação irreduzível (por exemplo, $\Gamma^{(n)}$). Partindo de uma função F arbitrária, projeta-se inicialmente a mesma ao longo de uma coluna da irrep, obtendo assim $f_\nu^{(n)}$ (por exemplo). Esta função é então normalizada, obtendo assim a função de base $\psi_\nu^{(n)}$ adequada. A partir daí, o uso sistemático do operador de transferência $\mathfrak{p}_{\mu\nu}^{(n)}$ irá gerar todas as suas parceiras, uma vez que $\mathfrak{p}_{\mu\nu}^{(n)} \psi_\nu^{(n)} = \psi_\mu^{(n)}$. O físico americano John Hasbrouck Van Vleck (1899–1980) denominou este procedimento de *máquina geradora das funções de base*.

Os resultados recém obtidos requerem o conhecimento total das representações irreduzíveis do grupo do Hamiltoniano. Em algumas situações, este conhecimento é impossível ou muito difícil de ser obtido. Em comparação, a dedução da tabela de caracteres é mais fácil de ser obtida, se for aplicado um procedimento semelhante ao discutido na seção 5.6.6. Nesta situação, ainda é possível obter-se informações a respeito das funções de base, conhecendo os caracteres da representação. Com este intuito, retorna-se a (5.51a), coloca-se $\mu = \nu$ e soma-se sobre ν para obter

$$\begin{aligned} \mathfrak{p}^{(n)} &\doteq \sum_{\nu=1}^{\ell_n} \mathfrak{p}_{\nu\nu}^{(n)} = \frac{\ell_n}{g} \sum_{P_R \in \mathcal{P}_\chi} \left(\sum_{\nu=1}^{\ell_n} \Gamma_{\nu\nu}^{(n)*}(R) \right) P_R, \\ \mathfrak{p}^{(n)} &= \frac{\ell_n}{g} \sum_{P_R \in \mathcal{P}_\chi} \chi^{(n)*}(R) P_R, \end{aligned} \quad (5.53)$$

sendo $\chi^{(n)}(R)$ o caractere de P_R na n -ésima representação irreduzível.

A ação de $\mathfrak{p}^{(n)}$ sobre uma função qualquer das coordenadas pode ser vista a partir de (5.52b),

$$\mathfrak{p}_{\nu\nu}^{(n)} F = f_\nu^{(n)} \implies \sum_{\nu} \mathfrak{p}_{\nu\nu}^{(n)} F = \sum_{\nu} f_\nu^{(n)} \implies \mathfrak{p}^{(n)} F = f^{(n)},$$

isto é, a partir de uma função arbitrária, $\mathfrak{p}^{(n)}$ projeta a parte pertencendo à n -ésima irrep, a qual consiste na soma das funções de base da mesma. Este resultado, da mesma forma que o caractere, não é afetado por transformações de similaridade que alteram as funções de base.

Por sua vez, a ação de $\mathfrak{p}^{(n)}$ sobre a função $f^{(n)}$ pode ser visualizada a partir de (5.52a),

$$\mathfrak{p}_{\kappa\kappa}^{(n)} f_{\nu}^{(n)} = \delta_{\nu\kappa} f_{\kappa}^{(n)} \implies \sum_{\kappa} \mathfrak{p}_{\kappa\kappa}^{(n)} f_{\nu}^{(n)} = \sum_{\kappa} \delta_{\nu\kappa} f_{\kappa}^{(n)} \implies \mathfrak{p}^{(n)} f_{\nu}^{(n)} = f_{\nu}^{(n)}.$$

Então,

$$\mathfrak{p}^{(n)} \sum_{\nu} f_{\nu}^{(n)} = \sum_{\nu} f_{\nu}^{(n)} \implies \mathfrak{p}^{(n)} f^{(n)} = f^{(n)}.$$

Finalmente, retornando à ação genérica do operador de transferência sobre uma função de base, dada por (5.51b), fazendo $\kappa = \mu$ e somando sobre μ e k , resulta

$$\sum_k \mathfrak{p}^{(k)} \psi_{\nu}^{(n)} = \psi_{\nu}^{(n)}.$$

Ou seja,

$$\sum_k \mathfrak{p}^{(k)} = P_I,$$

sendo P_I a identidade do grupo do Hamiltoniano.

Exemplo 5.10 (Operador de projeção do grupo C_{1h}). O grupo C_{1h} contém apenas duas operações (eq. 3.7): a identidade E e a reflexão σ_h , a qual realiza a transformação $\sigma_h x = -x$. Os correspondentes operadores que atuam no espaço funcional são P_E e P_{σ_h} , cujas ações sobre uma função $F(x)$, de acordo com (5.44), são $P_E F(x) = F(x)$ e $P_{\sigma_h} F(x) = F(-x)$. A partir agora da tabela de caracteres do grupo (tabela 5.3), os operadores de projeção do grupo são dados por (5.53):

$$\mathfrak{p}^{(1)} = \frac{1}{2} (P_E + P_{\sigma_h}), \quad \mathfrak{p}^{(2)} = \frac{1}{2} (P_E - P_{\sigma_h}).$$

Estes operadores, ao atuarem sobre uma função arbitrária $F(x)$, resultam em

$$\begin{aligned} \mathfrak{p}^{(1)} F(x) &= \frac{1}{2} [P_E F(x) + P_{\sigma_h} F(x)] = \frac{1}{2} [F(x) + F(x)], \\ \mathfrak{p}^{(2)} F(x) &= \frac{1}{2} [P_E F(x) - P_{\sigma_h} F(x)] = \frac{1}{2} [F(x) - F(-x)]. \end{aligned}$$

Claramente, as funções resultantes são, respectivamente, par e ímpar frente a uma reflexão, como seria esperado para estas pertecerem às representações $\Gamma^{(1)}$ e $\Gamma^{(2)}$, respectivamente.

Exemplo 5.11 (Funções de base simetrizadas do C_{3v}). Seja uma molécula formada por três núcleos atômicos puntiformes situados nos vértices de um triângulo equilátero (conforme ilustrado na figura 5.3) e por um elétron. Supondo que as funções de onda do elétron ligado separadamente a cada núcleo, denotadas por $\psi_j(r)$ ($j = a, b, c$), correspondam ao estado fundamental de um átomo monoelétrônico, deseja-se construir as funções de base da molécula a partir das autofunções atômicas empregando as suas propriedades de simetria, as quais são as isometrias de um triângulo equilátero.

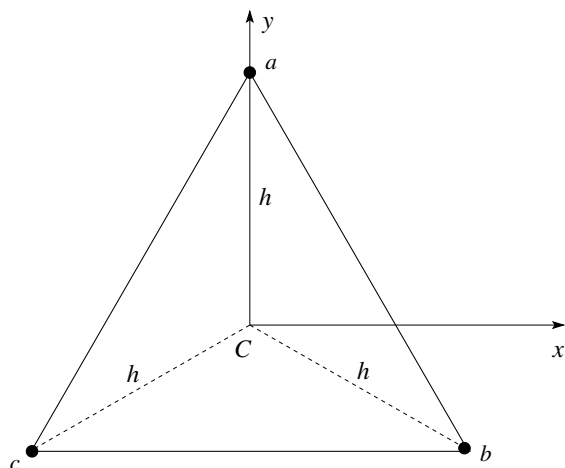


Figura 5.3: Molécula formada por três núcleos atômicos nos vértices de um triângulo equilátero e um elétron.

Seja $\mathcal{R} = \{R_j, j = 1, \dots, 6\}$ o grupo de transformações passivas aplicado às coordenadas das autofunções. Como este grupo irá realizar transformações somente no plano xy , deixando o eixo z invariante, existe um isomorfismo natural entre \mathcal{R} e o grupo \mathcal{C}_{3v} , o qual foi discutido no exercício 3.10. Por exemplo, sendo R_1 o operador que realiza uma rotação do referencial em torno do eixo z no sentido *anti-horário* por um ângulo de 120° , observa-se que $R_1 \cong C_3^2$. Para verificar isso, se $r = (xy)^T$ é a matriz coluna que contém as coordenadas de um ponto no sistema, o operador R_1 pode ser representado pela matriz de rotação $R(2\pi/3) \in SO(2)$,²⁷ onde

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

Ou seja, chamando $\alpha = \sqrt{3}/2$ e $\beta = 1/2$,

$$r' = R_1 r \rightsquigarrow r' = R\left(\frac{2}{3}\pi\right) r = \Gamma^{(3)}(\pi_4) r = \begin{pmatrix} -\beta & \alpha \\ -\alpha & -\beta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -\beta x + \alpha y \\ -\alpha x - \beta y \end{pmatrix}.$$

Dessa maneira, o isomorfismo $\mathcal{R} \cong \mathcal{C}_{3v}$ permite estabelecer

$$R_1 \cong C_3^2, \quad R_2 \cong C_3, \quad R_3 \cong \sigma_{v16}, \quad R_4 \cong \sigma_{v24}, \quad R_5 \cong \sigma_{v35}.$$

Seja agora \mathcal{P}_R o grupo de transformações ativas aplicado às autofunções atômicas $\psi_j(r)$. Para determinar a ação dos operadores de \mathcal{P}_R de acordo com o mapeamento (5.44), observa-se que $\psi_j(x, y, z) = \psi(r_j)$ ($j = a, b, c$), sendo $r_a = \sqrt{x^2 + (y-h)^2 + z^2}$, $r_b = \sqrt{(x-\alpha h)^2 + (y+\beta h)^2 + z^2}$ e $r_c = \sqrt{(x+\alpha h)^2 + (y+\beta h)^2 + z^2}$. Como as transformações mantêm a coordenada z invariante, observa-se que

$$R_1^{-1} r = C_3 r \rightsquigarrow r' = \Gamma^{(3)}(\pi_5) r = \begin{pmatrix} -\beta & -\alpha \\ \alpha & -\beta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \implies \begin{cases} x' = -\beta x - \alpha y \\ y' = \alpha x - \beta y. \end{cases}$$

Então,

$$\begin{aligned} r_a &\rightarrow r'_a = \sqrt{x'^2 + (y' - h)^2 + z^2} = \sqrt{(-\beta x - \alpha y)^2 + (\alpha x - \beta y - h)^2 + z^2} \\ &= \sqrt{(x - \alpha h)^2 + (y + \beta h)^2 + z^2} = r_b, \\ r_b &\rightarrow r'_b = \sqrt{(x + \alpha h)^2 + (y + \beta h)^2 + z^2} = r_c, \\ r_c &\rightarrow r'_c = \sqrt{x^2 + (y - h)^2 + z^2} = r_a. \end{aligned}$$

Ou seja, $P_{R_1} \psi_a = \psi_b$, $P_{R_1} \psi_b = \psi_c$ e $P_{R_1} \psi_c = \psi_a$ e a ação do operador P_{R_1} consiste em rodar a molécula por um ângulo de 120° no sentido *horário*. Da mesma forma,

$$\begin{aligned} R_4^{-1} r = \sigma_{v24} r \rightsquigarrow r' = \Gamma^{(3)}(\pi_6) r = \begin{pmatrix} \beta & \alpha \\ \alpha & -\beta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &\implies \begin{cases} x' = \beta x + \alpha y \\ y' = \alpha x - \beta y \end{cases} \\ r'_a = \sqrt{(x - \alpha h)^2 + (y + \beta h)^2 + z^2} &= r_b \\ r'_b = \sqrt{x^2 + (y - h)^2 + z^2} &= r_a \\ r'_c = \sqrt{(x + \alpha h)^2 + (y + \beta h)^2 + z^2} &= r_c. \end{aligned}$$

Portanto,

$$\begin{array}{lll} P_{R_2} \psi_a = \psi_c & P_{R_2} \psi_b = \psi_a & P_{R_2} \psi_c = \psi_b \\ P_{R_3} \psi_a = \psi_a & P_{R_3} \psi_b = \psi_c & P_{R_3} \psi_c = \psi_b \\ P_{R_4} \psi_a = \psi_b & P_{R_4} \psi_b = \psi_a & P_{R_4} \psi_c = \psi_c \\ P_{R_5} \psi_a = \psi_c & P_{R_5} \psi_b = \psi_b & P_{R_5} \psi_c = \psi_a. \end{array}$$

²⁷Ver seção 6.5.2.

A partir de qualquer uma das autofunções atômicas é possível agora obter as funções de base simetrizadas para esta molécula. Para tanto, toma-se por exemplo a autofunção $\psi_a(\mathbf{r})$ e aplica-se sobre a mesma o operador de projeção (5.52a) $\mathcal{P}_{\mu\mu}^{(n)} = \frac{\ell_n}{6} \sum_{R_i} \Gamma_{\mu\mu}^{(n)*}(R_i) P_{R_i}$. De acordo com o teorema 5.8, este procedimento irá fornecer as funções de base do n -ésimo autoestado de energia da molécula. Como a molécula possui a simetria do triângulo equilátero, conclui-se que: (i) o número de autoestados de energia da molécula é igual ao número de irreps do grupo \mathcal{C}_{3v} e (ii) a degenerescência de cada autoestado n ($= 1, 2, 3$) é igual à dimensão da n -ésima irrep do \mathcal{C}_{3v} .

Estado $n = 1$. Para este estado, $\ell_1 = 1$, correspondendo ao estado fundamental, o qual não é degenerado. O operador de projeção é $\mathcal{P}^{(1)} = \frac{1}{6} \sum_{R_i} \Gamma^{(1)*}(R_i) P_{R_i} = \frac{1}{6} \sum_{R_i} P_{R_i}$. Portanto,

$$\begin{aligned} \mathcal{P}^{(1)}\psi_a &= \frac{1}{6} \sum_{R_i} P_{R_i}\psi_a = \frac{1}{6} (\psi_a + P_{R_1}\psi_a + P_{R_2}\psi_a + P_{R_3}\psi_a + P_{R_4}\psi_a + P_{R_5}\psi_a) \\ &= \frac{1}{6} (\psi_a + \psi_b + \psi_c + \psi_a + \psi_b + \psi_c) = \frac{1}{3} (\psi_a + \psi_b + \psi_c). \end{aligned}$$

Ou seja, calculando a constante de normalização,

$$\psi_{\text{mol}}^{(1)} = \frac{1}{\sqrt{3}} (\psi_a + \psi_b + \psi_c).$$

Estado $n = 2$. Para este estado, $\ell_2 = 1$ e $\mathcal{P}^{(2)} = \frac{1}{6} \sum_{R_i} \Gamma^{(2)*}(R_i) P_{R_i}$. Então,

$$\mathcal{P}^{(2)}\psi_a = \frac{1}{6} \sum_{R_i} \Gamma^{(2)*}(R_i) P_{R_i}\psi_a = \frac{1}{6} (\psi_a + \psi_b + \psi_c - \psi_a - \psi_b - \psi_c) = 0.$$

Ou seja, não há projeção sobre o subespaço $n = 2$. As autofunções atômicas não formam uma função molecular neste caso. Isto pode ser verificado aplicando $\mathcal{P}^{(2)}$ a ψ_b e ψ_c .

Estado $n = 3$. Agora, $\ell_3 = 2$; o autoestado é duplamente degenerado e

$$\mathcal{P}_{\mu\mu}^{(3)}\psi_a = \frac{1}{3} \sum_{R_i \in \mathcal{R}} \Gamma_{\mu\mu}^{(3)*}(R_i) P_{R_i}\psi_a, \quad (\mu = 1, 2).$$

- Para $\mu = 1$:

$$\mathcal{P}_{11}^{(3)}\psi_a = \frac{1}{3} \sum_{R_i \in \mathcal{R}} \Gamma_{11}^{(3)*}(R_i) P_{R_i}\psi_a = \frac{1}{3} \left(\psi_a - \frac{1}{2}\psi_b - \frac{1}{2}\psi_c - \psi_a + \frac{1}{2}\psi_b + \frac{1}{2}\psi_c \right) = 0.$$

Ou seja, não há projeção de ψ_a sobre o subespaço $n = 3$.

- Para $\mu = 2$:

$$\begin{aligned} \mathcal{P}_{22}^{(3)}\psi_a &= \frac{1}{3} \sum_{R_i \in \mathcal{R}} \Gamma_{22}^{(3)*}(R_i) P_{R_i}\psi_a \\ &= \frac{1}{3} \left(\psi_a - \frac{1}{2}\psi_b - \frac{1}{2}\psi_c + \psi_a - \frac{1}{2}\psi_b - \frac{1}{2}\psi_c \right) = \frac{1}{3} (2\psi_a - \psi_b - \psi_c). \end{aligned}$$

Ou seja,

$$\psi_{\text{mol},2}^{(3)} = \frac{1}{\sqrt{6}} (2\psi_a - \psi_b - \psi_c).$$

A função $\psi_{\text{mol},2}^{(3)}$ é uma das funções da base do subespaço $n = 3$, o qual tem dimensão 2. A sua parceira é obtida a partir do operador de transferência (5.51a),

$$\begin{aligned} \psi_1^{(3)} &= \mathcal{P}_{12}^{(3)}\psi_{\text{mol},2}^{(3)} = \frac{1}{3} \sum_{R_i \in \mathcal{R}} \Gamma_{12}^{(3)*}(R_i) P_{R_i}\psi_{\text{mol},2}^{(3)}, \\ 2\sqrt{3}\psi_1^{(3)} &= P_{R_1}\psi_{\text{mol},2}^{(3)} - P_{R_2}\psi_{\text{mol},2}^{(3)} + 0 \cdot P_{R_3}\psi_{\text{mol},2}^{(3)} + P_{R_4}\psi_{\text{mol},2}^{(3)} - P_{R_5}\psi_{\text{mol},2}^{(3)}, \\ \sqrt{2}\psi_1^{(3)} &= \psi_b - \psi_c, \end{aligned}$$

a qual fica, após a normalização,

$$\psi_{\text{mol},1}^{(3)} = \frac{1}{\sqrt{3}} (\psi_b - \psi_c).$$

5.9.5 PERTURBAÇÕES, REGRAS DE SELEÇÃO E SIMETRIA

Nesta seção será realizada uma breve discussão sobre como o formalismo da teoria de grupos pode auxiliar na obtenção de resultados aproximados envolvendo sistemas quânticos. Em particular, será abordado o problema do tratamento de perturbações que são aplicadas a um sistema inicialmente em um estado estacionário, levando o mesmo a transições para diferentes estados. Um dos objetivos do uso da teoria de grupos neste caso consiste na obtenção das regras de seleção que definem quais transições são permitidas ou quais são proibidas.

5.9.5.1 PERTURBAÇÕES SOBRE AUTOESTADOS

Em geral, a solução da equação de Schroedinger completa (5.43) não existe; de fato, na maior parte das situações, mesmo a forma estacionária (5.47) não é exatamente solúvel. Contudo, para simplificar a discussão, será suposto que as soluções da equação independente do tempo sejam conhecidas.

Assume-se que o sistema físico esteja inicialmente em um estado estacionário, determinado pela solução de (5.47), quando o Hamiltoniano é formalmente escrito como $\mathcal{H} = \mathcal{H}_0$. Aplica-se então uma perturbação sobre o sistema de tal forma que o Hamiltoniano é alterado por $\mathcal{H}_0 \rightarrow \mathcal{H}$, onde \mathcal{H} inclui o efeito da perturbação. Se esse efeito puder ser considerado suficientemente “pequeno” (sob algum critério), então a perturbação pode ser adicionada linearmente ao Hamiltoniano original e pode-se então escrever

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}',$$

onde \mathcal{H}' é o termo que inclui a perturbação.

Seja \mathcal{P}_0 o grupo das transformações de simetria de \mathcal{H}_0 . Em geral, existe um ou mais operadores $P_R \in \mathcal{P}_0$ para os quais $[P_R, \mathcal{H}'] \neq 0$, i. e., existem operações para as quais \mathcal{H}_0 é invariante, mas \mathcal{H}' não o é. Isto se deve ao fato de que a ação da perturbação remove uma ou mais simetrias do sistema físico. Supõe-se que, neste caso, o sistema atuado pela perturbação apresente um conjunto reduzido de isometrias, o qual é um subconjunto das isometrias originais. Por isso, assume-se que ainda existe um grupo de transformações de simetria de \mathcal{H}' , denotado por \mathcal{P}' e tal que todos os operadores $P'_R \in \mathcal{P}'$ satisfaçam $[P'_R, \mathcal{H}'] = 0$, i. e., deixam \mathcal{H}' invariante. Supõe-se também que $\mathcal{P}' \subset \mathcal{P}_0$, i. e., é um subgrupo das operações de simetria de \mathcal{H}_0 . Com isso, o Hamiltoniano completo \mathcal{H} permanece invariante apenas sob as transformações de simetria comuns a \mathcal{H}_0 e \mathcal{H}' , ou seja, o grupo de simetrias de \mathcal{H} passa a ser \mathcal{P}' .

Retornando ao Hamiltoniano não perturbado \mathcal{H}_0 , uma vez que suas propriedades foram supostas conhecidas, sabe-se quantos autovalores de energia E_n existem e as ordens de suas degenerescências. Conhecem-se também os conjuntos de autofunções $\Psi_n = \{\psi_1^{\{n\}}, \dots, \psi_{\ell_n}^{\{n\}}\}$, as quais formam as bases ortonormais dos subespaços \mathcal{H}_n e a partir das quais as representações irreduzíveis $\Gamma^{(n)}(\mathcal{P}_0)$ são deduzidas.

Aplicando então a perturbação \mathcal{H}' sobre o sistema, o grupo de simetrias é reduzido para \mathcal{P}' . Como \mathcal{P}' é subgrupo de \mathcal{P}_0 , as funções de base em Ψ_n ainda geram uma representação de dimensão ℓ_n para \mathcal{P}' , mas esta representação será, em geral, redutível. Essa representação pode ser reduzida à forma bloco-diagonal, quando então surgirão novos subconjuntos de funções de base, expressos a partir de Ψ_n , tais que uma dada função de um determinado subconjunto será expressa apenas em termos de suas parceiras no subconjunto sob a ação de qualquer $P'_R \in \mathcal{P}'$. Como as representações irreduzíveis de \mathcal{P}_0 foram determinadas a partir das autofunções degeneradas (normais) aos autovalores de \mathcal{H}_0 , os novos subconjuntos serão compostos por funções degeneradas a novos autovalores, exceto no caso de degenerescências acidentais; em outras palavras, os autovalores de energia originais $\{E_n^{(0)}\}$ dividem-se em novos níveis devido à redução na simetria do sistema físico.

5.9.5.2 O TEOREMA DOS ELEMENTOS DE MATRIZ E REGRAS DE SELEÇÃO

Uma importante aplicação do tratamento perturbativo da mecânica quântica está na determinação das probabilidades temporais (ou taxas temporais) de transições entre dois estados estacionários de energia sob a ação de uma perturbação dependente (de forma “lenta”) do tempo. É assumido que antes da perturbação incidir sobre o sistema, este estava em um estado estacionário bem definido, descrito pela autofunção $\psi_\nu^{(n)}(\mathbf{r})$, associada ao autovalor E_n ,²⁸ os quais são

²⁸Ou $\psi_{p\nu}^{(n)}(\mathbf{r})$ e E_{pn} , caso existam degenerescências acidentais.

soluções do Hamiltoniano \mathcal{H}_0 .

Uma perturbação dependente do tempo é então aplicada ao sistema em um dado instante $t = t_0$. Um tipo comum de perturbação deste tipo consiste na incidência de radiação eletromagnética sobre átomos, moléculas ou sólidos. Essa perturbação não somente reduz a simetria do sistema, como também promove transições entre diferentes estados, de tal maneira que, se no instante $t = t_f > t_0$ a perturbação é removida e o sistema torna a ser regido por \mathcal{H}_0 , existe uma probabilidade não nula de o mesmo se encontrar em um outro autoestado $\psi_\mu^{(m)}$, distinto do inicial.

Empregando a teoria de perturbações dependentes do tempo na mecânica quântica, mostra-se que a probabilidade de transição por unidade de tempo na presença de uma perturbação pode ser expressa genericamente pela *regra de ouro de Fermi*

$$w = \frac{2\pi}{\hbar} \rho(m) |H'_{m\mu, n\nu}|^2, \text{ sendo } H'_{m\mu, n\nu} = \langle \psi_\mu^{(m)}, \mathcal{H}' \psi_\nu^{(n)} \rangle.$$

Na expressão acima, $\rho(m)$ é a densidade de estados finais e $H'_{m\mu, n\nu}$ é o elemento da matriz do Hamiltoniano perturbado \mathcal{H}' que conecta o estado inicial $n\nu$ com o estado final $m\mu$. A partir desta expressão fica evidente que a transição $n\nu \rightarrow m\mu$ somente poderá ocorrer se $H'_{m\mu, n\nu} \neq 0$. Transições que satisfazem esta condição são denominadas *permitidas*, ao passo que aquelas para as quais $H'_{m\mu, n\nu} = 0$ são ditas *proibidas*.

A teoria de representações de grupos pode fornecer informações acerca de quais transições são permitidas ou não. Para tanto, considera-se a ação da perturbação \mathcal{H}' sobre a autofunção do estado inicial $\psi_\nu^{(n)}$. Como $\mathcal{H}' \psi_\nu^{(n)}$ continua sendo uma função pertencente ao espaço de Hilbert \mathcal{H} completo, o teorema 5.8 mostrou que a mesma pode ser decomposta em termos do conjunto completo de autofunções de \mathcal{H}_0 por

$$\mathcal{H}' \psi_\nu^{(n)} = \sum_m \sum_\mu \psi_\mu^{(m)} c(n, \nu; m, \mu),$$

sendo $\{c(n, \nu; m, \mu)\} \in \mathbb{C}$ constantes a ser determinadas. Como as autofunções \mathcal{H}_0 são ortonormais, estas constantes são formalmente dadas por

$$c(n, \nu; m, \mu) = \langle \psi_\mu^{(m)}, \mathcal{H}' \psi_\nu^{(n)} \rangle = H'_{m\mu, n\nu}.$$

Ou seja,

$$\mathcal{H}' \psi_\nu^{(n)} = \sum_m \sum_\mu \psi_\mu^{(m)} H'_{k\kappa, m\mu}.$$

O que mostra que a componente de $\mathcal{H}' \psi_\nu^{(n)}$ ao longo da função de base $\psi_\kappa^{(k)}$ é justamente determinada pela matriz de transição $H'_{k\kappa, n\nu}$.

Na linguagem da teoria de grupos, o resultado acima mostra que a transição $n\nu \rightarrow m\mu$ somente é permitida se o vetor $\mathcal{H}' \psi_\nu^{(n)}$ possuir uma componente ao longo do estado final. Isto demonstra o teorema a seguir.

Teorema 5.9 (Elemento da matriz de transição). *Se a função $\mathcal{H}' \psi_\nu^{(n)}$ não contiver uma componente que se transforma de acordo com a μ -ésima coluna da representação irredutível $\Gamma^{(m)}$ (\mathcal{P}_0), o elemento de matriz*

$$H'_{m\mu, n\nu} = \langle \psi_\mu^{(m)}, \mathcal{H}' \psi_\nu^{(n)} \rangle$$

deve ser nulo.

É necessário ressaltar aqui que ainda assim pode ocorrer que transições em princípio permitidas pelo teorema acima resultem com $H'_{m\mu, n\nu} = 0$, mas neste caso isto ocorre devido a circunstâncias particulares.

Para determinar se uma transição é permitida ou proibida de acordo com este teorema, o seguinte procedimento genérico é adotado. Lembrando que, por hipótese, $\mathcal{P}' \subset \mathcal{P}_0$, a função $\psi' = \mathcal{H}' \psi_\nu^{(n)}$ pode ser usada para gerar uma representação $D_{\mathcal{H}' \psi_\nu^{(n)}}$ do grupo de simetria \mathcal{P}_0 . Contudo, essa representação em geral será redutível, uma vez que ψ' não é autofunção de \mathcal{H}_0 , exceto nos (raros) casos em que $[\mathcal{H}', \mathcal{H}_0] = 0$.

A perturbação \mathcal{H}' , de forma isolada, é suposta ser uma função das coordenadas ou um funcional das coordenadas. Neste caso, pode-se usar o mesmo para gerar uma representação D' de \mathcal{P}_0 , também redutível em princípio. Além disso, a autofunção sobre a qual a perturbação atua

pertence à representação irredutível $\Gamma^{(n)}(\mathcal{P}_0)$. Então, para se construir de forma genérica a representação à qual $\mathcal{H}'\psi_\nu^{(n)}$ pertence, sem possuir maiores detalhes acerca da forma de atuação da perturbação, pode-se empregar a técnica de construção de produtos diretos de representações, discutida na seção 5.8.

A função $\psi' = \mathcal{H}'\psi_\nu^{(n)}$, considerada como o “produto” da “função” \mathcal{H}' com a função $\psi_\nu^{(n)}$, pertence à representação

$$D_{\mathcal{H}'\psi_\nu^{(n)}} = D' \otimes \Gamma^{(n)}(\mathcal{P}_0).$$

Esta representação de produto direto, por sua vez, pode ser decomposta nas irreps de \mathcal{P}_0 , de acordo com (5.37a),

$$D_{\mathcal{H}'\psi_\nu^{(n)}} = D' \otimes \Gamma^{(n)}(\mathcal{P}_0) = \sum_n a_n \Gamma^{(n)}(\mathcal{P}_0).$$

Portanto, fica claro que se o produto direto $D' \otimes \Gamma^{(n)}(\mathcal{P}_0)$ não contiver nenhum componente na representação $\Gamma^{(m)}$ (i. e., $a_m = 0$), então o elemento de matriz $H'_{m\mu, n\nu}$ deve ser nulo.

Deve ser observado que esta é uma condição *fraca*; ou seja, para determinar se uma transição é proibida, basta verificar se $a_m = 0$. Contudo, o resultado $a_m \neq 0$ não é suficiente para determinar se a transição é permitida; para tanto, é necessário verificar se $\mathcal{H}'\psi_\nu^{(n)}$ contém alguma parte que se transforma de acordo com a μ -ésima coluna de $\Gamma^{(m)}$.

Um outro procedimento semelhante ao acima e que irá fornecer a mesma conclusão é devido à seguinte propriedade da decomposição de produtos diretos de representações:

Propriedade 5.3. Dadas $\Gamma^{(\alpha)}$ e $\Gamma^{(\beta)}$ duas representações irredutíveis não equivalentes de um mesmo grupo \mathcal{G} :

- (i) O produto direto $\Gamma^{(\alpha)*} \otimes \Gamma^{(\beta)}$ não contém a representação identidade.
- (ii) Os produtos diretos $\Gamma^{(\alpha)*} \otimes \Gamma^{(\alpha)}$ ou $\Gamma^{(\beta)*} \otimes \Gamma^{(\beta)}$ contêm a representação identidade somente uma vez.

Com base nesta propriedade, o procedimento alternativo consiste em realizar o produto direto de $\Gamma^{(m)}$ por $D_{\mathcal{H}'\psi_\nu^{(n)}}$ e realizar a decomposição

$$\Gamma^{(m)*}(\mathcal{P}_0) \otimes D' \otimes \Gamma^{(n)}(\mathcal{P}_0) = \sum_n b_n \Gamma^{(n)}(\mathcal{P}_0).$$

Se $b_1 = 0$, i. e. o produto direto não contém a representação identidade, então $\Gamma^{(m)}$ e $D_{\mathcal{H}'\psi_\nu^{(n)}}$ são não equivalentes e o elemento de matriz é nulo.

Em resumo, a teoria de representações de grupos determina, com base nas simetrias do sistema físico, se uma determinada transição é proibida. Se a transição for permitida, cálculos adicionais se fazem necessários para determinar a probabilidade da transição.

Exemplo 5.12 (Transições eletrônicas induzidas por ondas eletromagnéticas). O Hamiltoniano clássico de um elétron confinado por um potencial $V(\mathbf{r})$, sobre o qual se incide um campo de radiação eletromagnética é

$$H = \frac{1}{2m_e} [\mathbf{p} + e\mathbf{A}(\mathbf{r}, t)]^2 + V(\mathbf{r}) - e\Phi(\mathbf{r}, t),$$

sendo \mathbf{p} o momento conjugado, e a carga elementar e $\Phi(\mathbf{r}, t)$ e $\mathbf{A}(\mathbf{r}, t)$ os potenciais escalar e vetor do campo eletromagnético, respectivamente. Para a obtenção do correspondente operador \mathcal{H} , a seguinte generalização do comutador (5.46) é empregada,

$$[f(\mathbf{r}), p_j] = i\hbar \frac{\partial f}{\partial x_j}, \quad (j = 1, 2, 3).$$

Com isso, obtém-se

$$\mathcal{H} = \frac{p^2}{2m_e} + \frac{e}{m_e} \mathbf{A} \cdot \mathbf{p} - \frac{ie\hbar}{2m_e} \nabla \cdot \mathbf{A} + \frac{e^2 A^2}{2m_e} + V(\mathbf{r}) - e\Phi.$$

Adotando o calibre de Coulomb, resulta que $\nabla \cdot \mathbf{A} = \Phi = 0$. Assumindo também que a perturbação introduzida pela radiação é de pequena amplitude (i. e., $e|\mathbf{A} \cdot \mathbf{p}| \ll m_e |V|$), resulta que o operador Hamiltoniano pode ser escrito como

$$\mathcal{H} \simeq \mathcal{H}_0 + \mathcal{H}', \quad \text{sendo}$$

$$\mathcal{H}_0 = \frac{p^2}{2m_e} + V(\mathbf{r}), \quad \mathcal{H}' = \frac{e}{m_e} \mathbf{A} \cdot \mathbf{p}.$$

Exemplo. Com isto, resulta que o elemento da matriz de transição é dado por

$$H'_{m\mu, n\nu} = \frac{e}{m_e} \sum_{j=1}^3 \left\langle \psi_\mu^{(m)}, A_j p_j \psi_\nu^{(n)} \right\rangle.$$

Exemplo 5.13 (Transições do tipo dipolo elétrico). O elemento de matriz obtido na seção anterior pode ser aproximado para certas situações. Uma vez que a perturbação incidente é uma onda eletromagnética, o potencial vetor pode ser escrito como

$$A_j(\mathbf{r}, t) = A_{0j} e^{i\mathbf{k}\cdot\mathbf{r} - \omega t},$$

onde \mathbf{k} e ω são respectivamente o vetor de onda e a sua frequência (angular). Pode ser mostrado neste caso que a probabilidade de transição do autoestado E_n ao autoestado E_m sob a ação da onda será dada por

$$w = \frac{2\pi e}{\hbar m_e} \rho(m) \sum_{j=1}^3 \left| \left\langle \psi_\mu^{(m)}, A_{0j} e^{i\mathbf{k}\cdot\mathbf{r}} p_j \psi_\nu^{(n)} \right\rangle \right|^2 \delta(E_m - E_n - \hbar\omega).$$

Em muitos casos de interesse, o comprimento da onda (dado por $\lambda = 2\pi/k$) é muito maior que a dimensão típica da função de onda. Nesta situação, pode-se realizar a aproximação $\mathbf{k}\cdot\mathbf{r} \ll 1 \rightsquigarrow \exp(i\mathbf{k}\cdot\mathbf{r}) \approx 1$. Neste caso,

$$w \approx \frac{2\pi e}{\hbar m_e} \rho(m) \sum_{j=1}^3 |A_{0j}|^2 \left| (p_j)_{\mu\nu}^{mn} \right|^2 \delta(E_m - E_n - \hbar\omega), \text{ onde } (p_j)_{\mu\nu}^{mn} = \left\langle \psi_\mu^{(m)}, p_j \psi_\nu^{(n)} \right\rangle.$$

Mas, uma relação básica para a dinâmica de um elemento de matriz $F_{k\ell} = \langle \psi_k, F \psi_\ell \rangle$ é

$$\frac{dF_{k\ell}}{dt} = \frac{\partial F_{k\ell}}{\partial t} + i\hbar [\mathcal{H}, F]_{k\ell}.$$

Como o operador p_j não depende explicitamente do tempo, resulta então que

$$(p_j)_{\mu\nu}^{mn} = \frac{i}{\hbar} m_e (E_m - E_n) (x_j)_{\mu\nu}^{mn} = \frac{i}{\hbar} m_e (E_m - E_n) \left\langle \psi_\mu^{(m)}, x_j \psi_\nu^{(n)} \right\rangle.$$

Ou seja,

$$w \approx \frac{2\pi m_e}{\hbar^3} \rho(m) (E_m - E_n)^2 \sum_{j=1}^3 |A_{0j}|^2 \left| \left\langle \psi_\mu^{(m)}, e x_j \psi_\nu^{(n)} \right\rangle \right|^2 \delta(E_m - E_n - \hbar\omega).$$

O aparecimento da componente do momento de dipolo elétrico $\mathbf{p}_e = e\mathbf{r}$ ($p_{ej} = e x_j$) mostra que a perturbação introduzida pela onda deforma a nuvem eletrônica do sistema de maneira a induzir a formação de um dipolo elétrico. Por esta razão, as transições permitidas pelo elemento de matriz acima são denominadas *transições do tipo dipolo elétrico*.

O exemplo a seguir apresenta uma aplicação particular da metodologia delineada nesta seção para determinar as regras de seleção para transições.

Exemplo 5.14 (Regras de seleções para transições do tipo dipolar em uma molécula triangular). Considere um sistema formado por três átomos idênticos ligados quimicamente de tal forma que no estado estacionário os mesmos se encontrem nos vértices de um triângulo equilátero. Este sistema possui a evidente isometria do grupo pontual \mathcal{C}_{3v} , obtido no exercício 3.10. Por sua vez, as suas representações irredutíveis foram obtidas no exercício 5.1 (pelo isomorfismo $S_3 \leftrightarrow \mathcal{C}_{3v}$) e a tabela de caracteres deste grupo foi obtida no exemplo 5.3. A tabela 5.5 é repetida aqui, explicitamente em termos das classes do \mathcal{C}_{3v} .

Exemplo. Uma perturbação (de pequena amplitude) que incide sobre esta molécula irá excitar oscilações harmônicas na mesma, de tal forma que a simetria triangular é quebrada. Em outras palavras, a perturbação irá provocar transições entre estados vibracionais da molécula. Se as oscilações induzidas pela perturbação são do tipo dipolo elétrico, então a probabilidade de transição depende do elemento de matriz $\langle \psi_\mu^{(m)}, x_j \psi_\nu^{(n)} \rangle$ ($j = 1, 2, 3$). Neste caso, para determinar as regras de seleções de dipolo elétrico para este sistema, é necessário primeiro determinar as

Tabela 5.5: Tabela de caracteres do grupo C_{3v} .

Caracteres	Classes	C_E	$C_{\sigma_{v16}}$	C_{C_3}
		$\{E\}$	$\{\sigma_{v16}, \sigma_{v24}, \sigma_{v35}\}$	$\{C_3, C_3^2\}$
$\chi^{(1)}$		+1	+1	+1
$\chi^{(2)}$		+1	-1	+1
$\chi^{(3)}$		+2	0	-1

propriedades de transformação do operador posição $\mathbf{r} = (x, y, z)$. Para tanto, será aplicado o procedimento delineado nesta seção.

Assumindo que a molécula se situa no plano $x - y$, o eixo z está orientado ao longo do eixo de rotação do triângulo. Supondo que as oscilações ocorrem somente no plano da molécula, a coordenada z permanece invariante frente a todas as operações de transformação do C_{3v} . Por outro lado, as coordenadas (x, y) podem servir como funções de base para gerar a representação irreduzível de dimensão 2, i. e., $\Gamma^{(3)}(C_{3v})$. Portanto, o operador \mathbf{r} transforma-se, neste sistema de coordenadas, de acordo com

$$D'(\mathbf{r}) = \Gamma^{(1)} \oplus \Gamma^{(3)},$$

onde a representação identidade $\Gamma^{(1)}$ ocorre devido à invariância do sistema na direção z .

Como é necessário realizar produtos diretos de representações irreduzíveis do grupo \mathcal{P}_0 , isomórfico a C_{3v} , é útil verificar quais são as possíveis decomposições em irreps que podem ser obtidas. Usando resultados obtidos nas seções 5.6.5 e 5.8, obtém-se a tabela abaixo, a qual apresenta os caracteres dos produtos diretos das irreps e suas decomposições.

\otimes	C_E	$C_{\sigma_{v16}}$	C_{C_3}	$\Gamma \otimes \Gamma = \bigoplus a\Gamma$
$\Gamma^{(1)} \otimes \Gamma^{(1)}$	1	1	1	$\Gamma^{(1)}$
$\Gamma^{(1)} \otimes \Gamma^{(2)}$	1	-1	1	$\Gamma^{(2)}$
$\Gamma^{(1)} \otimes \Gamma^{(3)}$	2	0	-1	$\Gamma^{(3)}$
$\Gamma^{(2)} \otimes \Gamma^{(2)}$	1	1	1	$\Gamma^{(1)}$
$\Gamma^{(2)} \otimes \Gamma^{(3)}$	2	0	-1	$\Gamma^{(3)}$
$\Gamma^{(3)} \otimes \Gamma^{(3)}$	4	0	1	$\Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}$

Observa-se como as possibilidades mencionadas na propriedade 5.3 são de fato satisfeitas.

Com base nesta tabela, pode-se começar a construir as diferentes possibilidades da representação $D_{\mathcal{H}'\psi_v^{(n)}} = D'(\mathbf{r}) \otimes \Gamma^{(n)}$. Essas possibilidades são apresentadas na tabela abaixo, juntamente com as decomposições dos resultados, os quais foram obtidos empregando as propriedades 5.2:

$D'(\mathbf{r})$	\otimes	$\Gamma^{(n)}$	=	$\Gamma^{(1)} \otimes \Gamma^{(1)} \oplus \Gamma^{(3)} \otimes \Gamma^{(1)}$	=	$\Gamma^{(1)} \oplus \Gamma^{(3)}$
$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(1)}$	=	$\Gamma^{(1)} \otimes \Gamma^{(1)} \oplus \Gamma^{(3)} \otimes \Gamma^{(1)}$	=	$\Gamma^{(1)} \oplus \Gamma^{(3)}$
$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(2)}$	=	$\Gamma^{(1)} \otimes \Gamma^{(2)} \oplus \Gamma^{(3)} \otimes \Gamma^{(2)}$	=	$\Gamma^{(2)} \oplus \Gamma^{(3)}$
$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(3)}$	=	$\Gamma^{(1)} \otimes \Gamma^{(3)} \oplus \Gamma^{(3)} \otimes \Gamma^{(3)}$	=	$\Gamma^{(3)} \oplus \Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}$

Finalmente, verifica-se agora se a decomposição de $\Gamma^{(m)*} \otimes D'(\mathbf{r}) \otimes \Gamma^{(n)}$ contém a representação identidade. Em caso contrário, a transição é proibida. Isto será realizado considerando todas as possibilidades em que $m \neq n$. A tabela abaixo resume os resultados:

$\Gamma^{(m)}$	\otimes	$D'(\mathbf{r})$	\otimes	$\Gamma^{(n)}$	=	Transição
$\Gamma^{(2)}$	\otimes	$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(1)}$	=	$\Gamma^{(2)} \oplus \Gamma^{(3)}$ 1 \rightarrow 2 Proibida
$\Gamma^{(1)}$	\otimes	$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(2)}$	=	$\Gamma^{(2)} \oplus \Gamma^{(3)}$ 2 \rightarrow 1 Proibida
$\Gamma^{(3)}$	\otimes	$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(2)}$	=	$\Gamma^{(3)} \oplus \Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}$ 2 \rightarrow 3 Permitida
$\Gamma^{(2)}$	\otimes	$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(3)}$	=	$\Gamma^{(3)} \oplus \Gamma^{(2)} \oplus \Gamma^{(1)} \oplus \Gamma^{(3)}$ 3 \rightarrow 2 Permitida
$\Gamma^{(3)}$	\otimes	$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(1)}$	=	$\Gamma^{(3)} \oplus \Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}$ 1 \rightarrow 3 Permitida
$\Gamma^{(1)}$	\otimes	$(\Gamma^{(1)} \oplus \Gamma^{(3)})$	\otimes	$\Gamma^{(3)}$	=	$\Gamma^{(3)} \oplus \Gamma^{(1)} \oplus \Gamma^{(2)} \oplus \Gamma^{(3)}$ 3 \rightarrow 1 Permitida

Nota-se que somente verificou-se a condição *fraca*, isto é, somente verificou-se quais transições são proibidas. Para confirmar se as transições listadas como permitidas acima realmente podem ocorrer, é necessário ainda verificar se $D'(\mathbf{r}) \otimes \Gamma^{(n)}$ possui algum componente que se transforma conforme a μ -ésima coluna de $\Gamma^{(m)}$.

5.10 GRUPOS E ÁLGEBRAS DE LIE E SUAS REPRESENTAÇÕES

Até este ponto, foram discutidos de forma predominante os grupos finitos, particularmente os grupos cíclicos (seção 3.2.2), os grupos de permutações (seção 3.2.3) e os grupos de simetrias pontuais (seção 3.4.1) de uma estrutura cristalina regular.

Em se tratando de grupos de simetria, existem diversos exemplos de sistemas físicos que permanecem invariantes frente a um conjunto infinito de transformações. Um exemplo corriqueiro consiste em uma partícula submetida a um campo de força central, como ocorre com o elétron em um átomo de Hidrogênio. O Hamiltoniano deste sistema é invariante a todas as rotações no espaço Euclidiano \mathbb{E}^3 (exemplo 4.5), as quais compõe o grupo contínuo $SO(3)$. Portanto, muitos sistemas físicos demandam o estudo de grupos infinitos (discretos ou contínuos) e suas representações.

Grupos infinitos foram definidos e brevemente discutidos na seção 3.1.1. Nesta seção estes serão discutidos em mais detalhes, com o interesse focado em uma classe particular de grupos contínuos denominados *grupos de Lie*, os quais foram estudados pela primeira vez pelo matemático norueguês Marius Sophus Lie (1842 – 1899). Devido à grande importância que a teoria de representações de grupos possui para grupos contínuos em geral e para os grupos de Lie em particular, este assunto será discutido nesta seção, ao invés do capítulo 3.

5.10.1 DEFINIÇÕES INICIAIS

Parte-se de um grupo finito $\mathcal{G} = \{G; *\}$ de ordem g . Denominando g_i ($i = 1, \dots, g$) um elemento qualquer do grupo, o conjunto $G = \{g_i\}$ destes pode ser interpretado de forma abstrata como um conjunto de “pontos”, sendo g_i um “ponto” em particular de G . Este conjunto de pontos será denominado como a **variedade do grupo**.

Para qualquer conjunto discreto, sempre é possível realizar-se uma associação de cada elemento do mesmo com um número do conjunto \mathbb{N} . Assim, ao ponto rotulado por $a \in \mathbb{N}$, pode-se associar o elemento $g_a \in G$, em cuja situação o índice a pode assumir os valores $1 \leq a \leq g$. No produto $g_a * g_b = g_c$, se o índice a for mantido fixo e o índice b variar de forma que o elemento g_b varre G , o resultado g_c também irá varrer o grupo, de acordo com a linha de g_a na tabela de multiplicações de \mathcal{G} . Pode-se afirmar então que a tabela de multiplicações define uma função

$$c = \phi(a, b)$$

entre os índices $a, b, c \in \mathbb{N}$.

Quando o número de elementos do grupo tende ao infinito, este pode ser tornar um grupo discreto ou contínuo. No segundo caso, que é o de interesse aqui, os índices a, b e c passam a variar de forma contínua e o elemento $g_a \in \mathcal{G}$ passa a ser identificado por uma função real $g(a)$. Em geral, mais de um parâmetro contínuo pode ser necessário para identificar um elemento de \mathcal{G} . Por exemplo, a translação

$$x' = ax + b$$

forma um grupo de transformação com 2 parâmetros: a e b , os quais podem variar continuamente nos intervalos $-\infty < a, b < \infty$. Esta translação pode ser descrita em termos do operador de transformação $T(a, b)$, o qual executa a ação $T(a, b)x = ax + b$. O número de parâmetros contínuos necessários para definir um elemento do grupo serve para categorizar o mesmo; neste ponto, repete-se a definição 3.5 de uma maneira conveniente para esta seção.

Definição 5.12 (Grupo contínuo de r parâmetros). Um grupo é denominado *grupo contínuo de r parâmetros* quando todos os seus elementos dependem de um conjunto de parâmetros contínuos reais $\{a_\lambda\}$ para $\lambda = 1, \dots, r$. Um determinado elemento deste grupo é denotado por

$$g(a_1, \dots, a_r) \doteq g(a).$$

Se o número de parâmetros contínuos necessários para identificar os elementos do grupo é finito, então este é denominado um *grupo contínuo finito* e o parâmetro r é dito ser a *ordem* ou a *dimensão do grupo contínuo*. O conjunto de parâmetros $\{a_\lambda\}$ forma um espaço métrico Euclidiano²⁹ \mathcal{S}_r denominado o *espaço de parâmetros*. Se o domínio de variação dos parâmetros é finito, a variedade do grupo é dita ser *fechada*.

²⁹Ver exemplo 4.2.

Uma definição mais rigorosa para um grupo contínuo de r parâmetros envolve a elaboração e imposição de conceitos tais como “proximidade” ou continuidade entre os elementos da variedade do grupo. Relacionada com essa imposição, surge a necessidade de definição da vizinhança de um determinado elemento da variedade. Essa formalização, que não será aqui realizada, leva à definição de um **grupo topológico**, como sendo aquele cuja variedade do grupo forma um espaço topológico (definição 4.39).

De uma forma coloquial, a continuidade imposta à variedade do grupo é expressa em termos da métrica ou função distância do espaço de parâmetros. Dois elementos $g(a)$ e $g(a')$ de \mathcal{G} estão “próximos” entre si se existe um número $\epsilon \in \mathbb{R}$ para o qual

$$d(a, a') = \|a - a'\| \doteq \sqrt{\sum_{i=1}^r (a_i - a'_i)^2} < \epsilon.$$

As outras imposições realizadas sobre um grupo contínuo de r parâmetros são os axiomas de grupo (definição 3.1). Dados $g(a), g(b) \in \mathcal{G}$:

- Deve existir um conjunto de parâmetros $a^0 \in \mathcal{S}_r$ tal que para todo $a \in \mathcal{S}_r$

$$g(a^0)g(a) = g(a)g(a^0) = g(a),$$

sendo que $g(a^0) \in \mathcal{G}$ é o elemento identidade do grupo.

- Para todo $a \in \mathcal{S}_r$ deve existir um conjunto de parâmetros $\bar{a} \in \mathcal{S}_r$ tal que

$$g(\bar{a})g(a) = g(a)g(\bar{a}) = g(a^0).$$

O elemento $g(\bar{a}) \equiv g^{-1}(a)$ é o elemento inverso de $g(a)$.

- Para todo $a, b \in \mathcal{S}_r$ deve existir um conjunto de parâmetros $c \in \mathcal{S}_r$ tal que

$$g(a)g(b) = g(c),$$

sendo que $g(c) \in \mathcal{G}$.

Esta relação de clausura implica na existência de um conjunto de funções contínuas $\{\phi_k(a; b)\}$ dos parâmetros a e b tais que

$$c = \phi(a; b), \text{ ou seja, } c_k = \phi_k(a_1, \dots, a_r; b_1, \dots, b_r) \quad (k = 1, \dots, r).$$

Impõe-se agora condições adicionais à definição de um grupo contínuo. As funções contínuas $\{\phi_k(a; b)\}$ do espaço de parâmetros são exigidas serem *funções suaves*, i. e., devem possuir derivadas em todas as ordens com respeito a ambos os conjuntos de parâmetros a e b . Além disso, exige-se que o parâmetro \bar{a} seja uma função analítica³⁰ do parâmetro a . Estas exigências adicionais levam à definição de um *grupo de Lie de r parâmetros*.

Definição 5.13 (Grupo de Lie de r parâmetros). Seja \mathcal{G} um grupo contínuo de r parâmetros. Sejam $g(a), g(b) \in \mathcal{G}$ determinados pelos parâmetros $a \doteq (a_1, \dots, a_r)$ e $b \doteq (b_1, \dots, b_r)$. Seja $g(c) \in \mathcal{G}$ determinado por $g(c) = g(a)g(b)$ de tal forma que $c \doteq (c_1, \dots, c_r)$ é obtido a partir da função contínua $c = \phi(a; b)$, onde $\phi \doteq (\phi_1, \dots, \phi_k)$ contém um conjunto de funções contínuas tais que $c_k = \phi_k(a; b)$ ($k = 1, \dots, r$). Sejam também $g(a^0), g(\bar{a}) \in \mathcal{G}$ tais que $g(a^0)$ é o elemento identidade de \mathcal{G} e $g(\bar{a})$ o elemento inverso de $g(a)$. Este grupo é denominado um *grupo de Lie de r parâmetros* se:

1. As funções $\{\phi_k(a_1, \dots, a_r; b_1, \dots, b_r)\}$ ($k = 1, \dots, r$) possuem derivadas em todas as ordens em relação aos parâmetros $\{a_k\}$ e $\{b_k\}$.
2. O conjunto de parâmetros $\bar{a} \doteq (\bar{a}_1, \dots, \bar{a}_r)$ que determina o elemento inverso $g(\bar{a})g(a) = g(a)g(\bar{a}) = g(a^0)$ é uma função analítica do conjunto de parâmetros a , i. e., a função $\bar{a} = \bar{a}(a_1, \dots, a_r)$ pode ser desenvolvida em uma série de potências dos parâmetros a_1, \dots, a_r .

³⁰Ver seções 2.4.5 e 2.8.3.

Como a função $\phi = \phi(a; b)$ é suave, a variedade do grupo é diferenciável.³¹ A definição acima de um grupo de Lie será agora adaptada ao tipo de grupo de interesse para a física: grupos de operadores de transformação, ou, simplesmente, grupos de transformação. Assim, uma outra definição pode ser apresentada, especificamente para grupos de transformação.

Definição 5.14 (Grupo de Lie de transformações de r parâmetros). Seja \mathcal{V}^n um espaço vetorial de dimensão n e $\Psi = \{\psi_1, \dots, \psi_n\} \subset \mathcal{V}^n$ uma base de \mathcal{V}^n tal que todo vetor $x \in \mathcal{V}^n$ pode ser decomposto de forma unívoca com componentes $\{x_i\}$ ($i = 1, \dots, n$) com respeito à base Ψ . Seja $\mathcal{G} = \{f(x; a)\}$ um conjunto de operadores lineares $f: \mathcal{V} \mapsto \mathcal{V}$, sendo que $x \doteq (x_1, \dots, x_n)$ são as componentes do vetor x e $a \doteq (a_1, \dots, a_r)$ um conjunto de r parâmetros, tais que existe $x' \in \mathcal{V}^n$ cujas coordenadas $\{x'_i\}$ correspondem às imagens dos mapeamentos

$$x'_i = f_i(x_1, \dots, x_n; a_1, \dots, a_r), \quad (i = 1, \dots, n); \quad (5.54a)$$

ou, simbolicamente,

$$x' = f(x; a). \quad (5.54b)$$

Se as funções $\{f(x; a)\}$ são funções analíticas dos parâmetros a e satisfazem as condições da definição 5.13, então o grupo \mathcal{G} é denominado um grupo de Lie de transformações de r parâmetros.

A exigência de cumprimento por parte do conjunto $\mathcal{G} = \{f(x; a)\}$ dos axiomas para a formação de um grupo de Lie possui as seguintes implicações:

- *Clausura:* dado um vetor $x \in \mathcal{V}^n$, devem existir $x', x'' \in \mathcal{V}^n$ que resultam de duas transformações sucessivas, obtidas respectivamente a partir dos conjuntos de r parâmetros a e b da seguinte maneira:

$$x_i \longrightarrow x'_i = f_i(x_1, \dots, x_n; a_1, \dots, a_r) \longrightarrow x''_i = f_i(x'_1, \dots, x'_n; b_1, \dots, b_r) \quad (i = 1, \dots, n).$$

Para que a condição de clausura seja satisfeita, a composição de transformações acima deve corresponder a uma única transformação, ou seja, deve existir um conjunto de parâmetros $c \doteq (c_1, \dots, c_r)$ tais que

$$x'' = f(f(x; a); b) = f(x; c), \quad (5.55a)$$

e os quais são uma função analítica

$$c = \phi(a; b), \text{ ou seja, } c_k = \phi_k(a_1, \dots, a_r; b_1, \dots, b_r) \quad (k = 1, \dots, r) \quad (5.55b)$$

dos parâmetros a e b . Esta condição também pode ser escrita como

$$f(f(x; a); b) = f(x; \phi(a; b)). \quad (5.55c)$$

- *Elemento identidade:* deve existir um conjunto de parâmetros a^0 tais que

$$x' = f(x; a^0) = x.$$

Em muitas situações, é conveniente escolher o grupo de transformações de tal forma que o elemento identidade é obtido como a imagem da origem do espaço de parâmetros, *i. e.*, $a^0 \equiv 0 = (0, \dots, 0) \Rightarrow x = f(x; 0)$.

- *Elemento inverso:* dada a transformação $x' = f(x; a)$, deve existir um conjunto de parâmetros \bar{a} tais que

$$x'' = f(x'; \bar{a}) = f(f(x; a); \bar{a}) = x.$$

Existem duas importantes condições impostas como consequência:

1. As relações (5.54) devem ser inversíveis, *i. e.*, deve ser possível exprimir os $\{x_i\}$ em termos dos $\{x'_i\}$. A condição para tanto é que o Jacobiano da transformação não pode ser nulo,

$$J = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{vmatrix} \neq 0.$$

³¹Ver definição 4.42.

2. Os parâmetros $\{\bar{a}_k\}$ devem ser funções analíticas dos $\{a_k\}$, pois

$$x = f(f(x; a); \bar{a}) = f(x; a^0)$$

implica a partir de (5.55) que

$$a^0 = \phi(a; \bar{a}) \implies \bar{a} = \psi(a; a^0),$$

uma vez que os parâmetros a^0 são fixos (únicos).

Os grupos de Lie de transformações definidos em 5.14 pressupõe a existência de um conjunto de funções $\{f(x; a)\} = \{f_1(x; a), \dots, f_n(x; a)\}$, as quais são funções suaves não somente dos parâmetros $\{a_i\}$, mas também dos componentes $\{x_i\}$. Por isso, estes grupos são exemplos de **grupos de Lie matriciais**, os quais descrevem grupos de transformações realizadas na forma de matrizes que atuam sobre os componentes de um vetor do espaço vetorial. Alguns exemplos simples ajudam a clarificar as definições apresentadas acima.

Exemplo 5.15 (Transformações de alongamento unidimensionais). Dados x, x' e $a \neq 0$ tais que

$$x' = ax.$$

Neste caso, $\mathcal{G} = \{a\} \equiv \mathbb{R} \setminus \{0\}$ é um grupo de Lie de transformações de 1 parâmetro, no qual:

- *Clausura:* $\exists a, b, c \in \mathcal{G}$ tais que $c = ab$.
- *Elemento identidade:* $a = 1$.
- *Elemento inverso:* $\forall a \in \mathcal{G}$ tal que $x' = ax$, $\exists \bar{a} = 1/a$ tal que

$$x'' = \frac{1}{a}x' = \frac{1}{a}(ax) = x.$$

Exemplo 5.16 (Transformações de alongamento/deslocamento unidimensionais). Dados x, x' e os parâmetros reais $a_1 \neq 0$ e a_2 tais que

$$x' = f(x; a_1, a_2) \doteq T(a_1, a_2)x = a_1x + a_2.$$

Neste caso, $\mathcal{G} = \{T(a_1, a_2)\}$ forma um grupo de Lie de transformações de 2 parâmetros, no qual:

- *Clausura:* $\exists a_1, b_1, c_1, a_2, b_2, c_2$ tais que dadas as transformações sucessivas

$$x' = f(x; a_1, a_2) = a_1x + a_2 \longrightarrow x'' = f(x'; b_1, b_2) = b_1x' + b_2 = b_1(a_1x + a_2) + b_2,$$

a condição de clausura implica em existir

$$x'' = f(x; c_1, c_2) = c_1x + c_2.$$

De (5.55) resulta que

$$f(f(x; a_1, a_2); b_1, b_2) = f(x; c_1, c_2) \Rightarrow c_1 = a_1b_1, \quad c_2 = a_2b_1 + b_2.$$

- *Elemento identidade:* $a^0 = (1, 0)$, i. e., $a_1 = 1$ e $a_2 = 0$, pois

$$x = f(x; a^0) = 1 \cdot x + 0 = x.$$

- *Elemento inverso:* $\bar{a} = (\bar{a}_1, \bar{a}_2)$, com $\bar{a}_1 = 1/a_1$ e $\bar{a}_2 = -a_2/a_1$, pois

$$x' = a_1x + a_2 \implies x'' = f(x'; \bar{a}) = \bar{a}_1x' + \bar{a}_2 = \frac{1}{a_1}(a_1x + a_2) - \frac{a_2}{a_1} = x.$$

Verifica-se que este grupo não é Abelian para $a_1 \neq 1$, pois se

$$T_1(a_1, a_2)x = f(x; a_1, a_2) \quad \text{e} \quad T_2(b_1, b_2)x = f(x; b_1, b_2)$$

então

$$\begin{aligned} x'' &= T_1T_2x = f(f(x; b_1, b_2); a_1, a_2) = a_1(b_1x + b_2) + a_2 \\ x''' &= T_2T_1x = f(f(x; a_1, a_2); b_1, b_2) = b_1(a_1x + a_2) + b_2, \end{aligned}$$

ou seja, $x'' \neq x'''$.

Exemplo 5.17 (Grupo geral linear de n dimensões, $GL(n, \mathbb{K})$). Como sempre, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Dado o espaço vetorial \mathcal{V}^n cujos vetores têm componentes representados coluna $x = (x_1 \cdots x_n)^T$, então para toda matriz $A \in GL(n, \mathbb{K})$ tal que $\det A \neq 0$,

$$x' = Ax \Rightarrow x'_i = \sum_{j=1}^n a_{ij}x_j \quad (i = 1, \dots, n).$$

- *Clausura:* Existem $A, B, C \in GL(n, \mathbb{K})$ tais que

$$C = AB.$$

- *Elemento identidade:* $A = I_n$.

- *Elemento inverso:* A^{-1} .

O grupo $GL(n, \mathbb{K})$ é um grupo de Lie de transformações de n^2 parâmetros.

Exemplo 5.18 (Grupo unimodular de n dimensões, $SL(n, \mathbb{K})$). Consiste no grupo $GL(n, \mathbb{K})$ com a restrição adicional que para todo $A \in SL(n, \mathbb{K})$, $\det A = +1$. Trata-se de um grupo de Lie de $n^2 - 1$ parâmetros.

Exemplo 5.19 (Grupo ortogonal de n dimensões, $O(n)$). Consiste no subgrupo $GL(n, \mathbb{R})$ composto por matrizes reais, com a condição adicional de que a norma de \mathcal{V}^n , dada por $\sum_{i=1}^n x_i^2$, permaneça invariante. De acordo com a discussão realizada na seção 6.2.1, isto implica que as matrizes são ortogonais, *i. e.*, toda matriz $A \in O(n)$ é tal que $AA^T = A^T A = I_n$. Este grupo, portanto, é composto por todas as rotações (próprias ou impróprias) em \mathcal{V}^n .

A condição de ortogonalidade implica que $\det A = \pm 1$ e que

$$\sum_{k=1}^n a_{ik}a_{jk} = \sum_{k=1}^n a_{ki}a_{kj} = \delta_{ij},$$

o que resulta em n condições do tipo $\sum_k a_{ik}^2 = 1$ e $n(n-1)/2$ condições do tipo $\sum_k a_{ik}a_{jk} = 0$ para $i \neq j$. Portanto, o número de parâmetros independentes é $n^2 - [n + n(n-1)/2] = n(n-1)/2$. Ou seja, o grupo $O(n)$ é um grupo de transformações de $n(n-1)/2$ parâmetros.

Exemplo 5.20 (Grupo ortogonal especial de n dimensões, $SO(n)$). Trata-se do grupo $O(n)$ com a condição adicional de que todas as matrizes $A \in SO(n)$ são tais que $\det A = +1$, correspondendo a somente rotações próprias em \mathcal{V}^n . O número de parâmetros do grupo também é $n(n-1)/2$.

Exemplo 5.21 (Grupo unitário de n dimensões, $U(n)$). Consiste no subgrupo $GL(n, \mathbb{C})$ composto por matrizes complexas que atuam sobre um espaço vetorial complexo de dimensão n , com a restrição de que as matrizes sejam unitárias, *i. e.*, todas $A \in U(n)$ são tais que

$$AA^\dagger = A^\dagger A = I_n.$$

Esta condição implica em que $|\det A| = +1$ e que

$$\sum_{k=1}^n a_{ik}a_{jk}^* = \sum_{k=1}^n a_{ki}^*a_{kj} = \delta_{ij}.$$

Como os elementos $\{a_{ij}\}$ são complexos, $a_{ij} = \text{Re } a_{ij} + i \text{Im } a_{ij}$, o que significa que a matriz A possui $2n^2$ parâmetros reais. Contudo, a condição de unitariedade impõe n condições do tipo $\sum_k |a_{ik}|^2 = 1$ e $2\binom{n}{2} = n(n-1)$ condições do tipo $\sum_k a_{ik}a_{jk}^* = \sum_k a_{ki}^*a_{kj} = 0$ ($i \neq j$), resultando em um total de $2n^2 - [n + n(n-1)] = n^2$ parâmetros reais independentes (contando separadamente as partes real e imaginária de a_{ij}). Ou seja, o grupo $U(n)$ é um grupo de transformações de n^2 parâmetros.

Além disso, a condição de unitariedade também implica que $|a_{ij}|^2 \leq 1$ ($\forall i, j = 1, \dots, n$). Portanto, o parâmetros reais devem variar sobre um intervalo finito, implicando em que o grupo é fechado. Em consequência, todos os subgrupos de $U(n)$, tais como o $O(n)$, são também grupos de Lie fechados.

Quando o $O(n)$ está representando um grupo de operadores unitários, as suas matrizes são tais que a norma e o produto interno de vetores no espaço complexo são preservados.³²

³²Teorema 4.22.

Exemplo 5.22 (Grupo especial unitário de n dimensões, $SU(n)$). Trata-se do grupo $U(n)$ com a restrição adicional de que $\det A = +1$, tratando-se, assim, de um grupo de $n^2 - 1$ parâmetros reais.

5.10.2 TRANSFORMAÇÕES INFINITESIMAIS DO GRUPO

Em (5.55a-c) mostrou-se que os parâmetros $\{c_i\}$ descrevem uma concatenação de transformações em termos dos parâmetros a e b . Como sempre deve existir um elemento inverso para cada transformação, a relação $c = \phi(a; b)$ deve ser inversível, i. e., deve ser sempre possível escrever $a = \psi(b; c)$ ou $b = \zeta(a; c)$. Para que isso ocorra, os Jacobianos $|\partial\phi_k/\partial a_\ell|$ e $|\partial\phi_k/\partial b_\ell|$ não podem ser nulos.

A transformação $x' = f(x; a)$ toma todos os pontos do espaço \mathcal{V}^n de suas “posições” iniciais $x = (x_1, \dots, x_n)$ e os leva às suas posições finais em x' . Como as funções $\{f_i(x; a)\}$ são analíticas nos parâmetros $\{a_k\}$, pode-se definir o elemento identidade do grupo de transformações como o parâmetro $a^0 = (0, \dots, 0)$ e se considerar variações infinitesimais dos parâmetros, o que gera uma **transformação infinitesimal**.

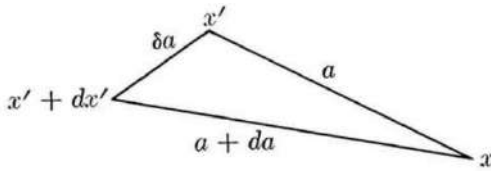


Figura 5.4: Uma variação infinitesimal $a \rightarrow a + da$ que leva à transformação infinitesimal $x \rightarrow x' \rightarrow x' + dx'$.

A figura 5.4 ilustra o conceito de uma transformação infinitesimal para um grupo de transformações de um parâmetro $\mathcal{G} = \{f(x; a)\}$ que atua em um espaço de dimensão 1. Uma transformação inicial com o parâmetro a leva x a $x' = f(x; a)$. O parâmetro $a + da$, nas vizinhanças de a irá levar o mesmo vetor x ao vetor $x'' = x' + dx' = f(x; a + da)$. Por outro lado, como $f(x; a)$ é analítica, deve existir um parâmetro δa próximo a $a^0 = 0$ que leva o vetor x' ao vetor $x' + dx' = f(x'; \delta a)$.

Assim, existem dois caminhos alternativos que levam x para $x' + dx'$:

$$x' + dx' = f(x; a + da), \text{ ou} \\ x' = f(x; a) \rightarrow x' + dx' = f(x'; \delta a).$$

Desenvolvendo a última expressão em uma série de Taylor sobre $a = a^0$ e mantendo somente o primeiro termo, resulta

$$x' + dx' = \underbrace{f(x'; 0)}_{=x'} + \left. \frac{\partial f(x'; a)}{\partial a} \right|_{a=0} \delta a \implies dx' = \left. \frac{\partial f(x'; a)}{\partial a} \right|_{a=0} \delta a \doteq u(x') \delta a.$$

Por outro lado,

$$x \rightarrow x' = f(x; a) \rightarrow x' + dx' = f(f(x; a); \delta a) = f(x; a + da),$$

implica, a partir de (5.55), que

$$a + da = \phi(a; \delta a),$$

de forma que

$$da = \left. \frac{\partial\phi(a; b)}{\partial b} \right|_{b=0} \delta a \iff \delta a = \psi(a) da.$$

Substituindo este último resultado no anterior, resulta

$$dx' = u(x') \psi(a) da \implies \frac{dx'}{u(x')} = \psi(a) da.$$

Chamando

$$\int \frac{dx'}{u(x')} = y(x') \text{ e } \int_0^a \psi(a') da' = t,$$

a integração do resultado anterior de $a' = 0$ a $a' = a$ resulta em

$$y' - y = t, \text{ sendo } y' = y(a) \text{ e } y = y(0).$$

Ou seja, um grupo de transformações de 1 parâmetro é sempre isomorfo ao grupo de translações (exemplo 5.16, com $a_1 = 1$), o qual é Abelian.

Portanto, independente da transformação específica que está sendo considerada, um grupo de Lie de um parâmetro será sempre um grupo Abeliano, com as propriedades

$$f(t_1)f(t_2) = f(t_1 + t_2), \quad f^{-1}(t) = f(-t),$$

as quais são diretamente obtidas do exemplo 5.16, sendo t denominado o *parâmetro canônico* do grupo.

Considera-se agora o caso geral dado em (5.54) e a dupla sequência de transformações infinitesimais $x_i \xrightarrow{a} x'_i \xrightarrow{\delta a} x'_i + dx'_i$ e $x_i \xrightarrow{a+da} x'_i + dx'_i$ para as n componentes x_i e com os r parâmetros (a_1, \dots, a_r) . Para o primeiro caminho,

$$\begin{aligned} x'_i &= f_i(x_1, \dots, x_n; a_1, \dots, a_r), & (i = 1, \dots, n), \\ x'_i + dx'_i &= f_i(x'_1, \dots, x'_n; \delta a_1, \dots, \delta a_r). \end{aligned} \quad (5.56a)$$

Desenvolvendo $f_i(x; \delta a)$ em torno de $\delta a = 0$ (a identidade) e mantendo somente os dois primeiros termos da série, resulta que

$$dx'_i = \sum_{k=1}^r u_{ik}(x') \delta a_k, \quad \text{sendo } u_{ik}(x) \doteq \left. \frac{\partial}{\partial a_k} f_i(x; a) \right|_{a=0}. \quad (5.56b)$$

Por outro lado, o segundo caminho implica que devem existir funções ϕ_ℓ ($\ell = 1, \dots, r$) tais que

$$a_\ell + da_\ell = \phi_\ell(a_1, \dots, a_r; \delta a_1, \dots, \delta a_r).$$

Esta condição vem da relação de clausura do grupo e de (5.55b). Desenvolvendo agora ϕ_ℓ em torno de $\delta a = 0$ e mantendo somente os dois primeiros termos, resulta

$$da_\ell = \sum_{m=1}^r \Theta_{\ell m}(a) \delta a_m, \quad \text{sendo } \Theta_{\ell m}(a) \doteq \left. \frac{\partial}{\partial b_m} \phi_\ell(a; b) \right|_{b=0}. \quad (5.56c)$$

Se for determinado que $a = 0$, isto implica que o primeiro caminho não foi realizado, o que significa que $da_\ell = \delta a_\ell$. Portanto, os elementos da matriz Θ devem satisfazer a condição inicial $\Theta_{\ell m}(0) = \delta_{\ell m}$.

As relações (5.56c) podem ser invertidas, de forma que os δa 's são expressos em termos dos da 's, se existir uma matriz $\Psi(a_1, \dots, a_r) = [\psi_{k\ell}(a)]$ tal que

$$\Psi\Theta = I_r, \quad \text{com } \psi_{k\ell}(a = 0) = \delta_{k\ell}. \quad (5.56d)$$

Neste caso,

$$\delta a_k = \sum_{\ell=1}^r \psi_{k\ell}(a) da_\ell, \quad (k = 1, \dots, r).$$

Substituindo então este resultado em (5.56b), resulta

$$dx'_i = \sum_{k,\ell=1}^r u_{ik}(x') \psi_{k\ell}(a) da_\ell,$$

a qual expressa a diferencial da coordenada x'_i em termos dos da 's. Como $x'_i = x'_i(x; a)$, isto implica que

$$\frac{\partial x'_i}{\partial a_\ell} = \sum_{k=1}^r u_{ik}(x') \psi_{k\ell}(a), \quad (i=1, \dots, n; \ell=1, \dots, r). \quad (5.56e)$$

Este resultado constitui um sistema de equações para as coordenadas x'_i em função dos parâmetros a , com a condição inicial $x'_i(a = 0) = x_i$ ($i = 1, \dots, n$).

As equações (5.56e) levam a um resultado fundamental para o estudo dos grupos de Lie. Dada uma função qualquer das coordenadas do sistema, $F = F(x_1, \dots, x_n)$, se forem aplicadas as transformações (5.54) às mesmas, a função será transformada de acordo com $F(x) \xrightarrow{x \rightarrow x'} F' = F(x')$. Realizando a transformação infinitesimal (5.56a) resulta então

$$F \rightarrow F' = F + dF = F(x + dx) \implies dF = \sum_{i=1}^n \frac{\partial F}{\partial x_i} dx_i.$$

Inserindo então (5.56b), resulta

$$dF = \sum_{i=1}^n \frac{\partial F}{\partial x_i} \sum_{\ell=1}^r u_{i\ell}(x) \delta a_\ell = \sum_{\ell=1}^r \delta a_\ell X_\ell F, \text{ sendo}$$

$$X_\ell \doteq \sum_{i=1}^n u_{i\ell}(x) \frac{\partial}{\partial x_i}. \quad (5.57)$$

Ou seja, sob a transformação infinitesimal,

$$F' = \left[1 + \sum_{\ell=1}^r \delta a_\ell X_\ell \right] F(x).$$

Os operadores diferenciais X_ℓ são denominados os **geradores infinitesimais** do grupo. Por exemplo, se for escolhido $F(x) = x_i$, resulta que

$$x'_i = \left[1 + \sum_{\ell=1}^r \delta a_\ell X_\ell \right] x_i = x_i + \sum_{\ell=1}^r u_{i\ell}(x) \delta a_\ell,$$

o que reproduz o resultado (5.56b).

A transformação $x_i \rightarrow x'_i = x_i + \sum_{\ell} u_{i\ell} \delta a_\ell$ foi levada a cabo por meio de uma variação infinitesimal dos parâmetros a . Se somente for variado o parâmetro a_k ($1 \leq k \leq r$), então, obviamente, $x'_i = x_i + u_{ik} \delta a_k$. Se δa_k não for infinitesimal, mas suficientemente pequeno, então

$$x'_i \approx x_i + u_{ik} \Delta a_k = (1 + \Delta a_k X_k) x_i,$$

onde $|\Delta a_k| \ll 1$. Após um número grande N de passos finitos Δa_k , o parâmetro irá valer $a_k = N \Delta a_k$ e a coordenada x_i sofre uma variação que pode ser aproximada pela sequência de passos finitos

$$x_i^{(N \times)} \approx (1 + \Delta a_k X_k)^N x_i = \left(1 + \frac{1}{N} a_k X_k \right)^N x_i.$$

Se for realizado agora o limite $\Delta a_k \rightarrow 0$, porém simultaneamente com $N \rightarrow \infty$, de tal forma que $a_k = N \Delta a_k$ resulta em um valor finito, emprega-se a identidade exponencial

$$\lim_{N \rightarrow \infty} \left(1 + \frac{1}{N} x \right)^N = e^x$$

para se obter a expressão formal

$$x_i(0; \dots, a_k, \dots) = \exp(a_k X_k) x_i.$$

Realizando variações arbitrárias em todos os parâmetros, as mesmas são todas realizadas de forma independente entre si, o que permite escrever então a expressão final

$$x_i(a_1, \dots, a_r) = P(x; a) x_i, \text{ sendo } P(x; a) \doteq \exp \left(\sum_{k=1}^r a_k X_k \right) \quad (5.58)$$

o operador que executa a transformação finita. Este resultado mostra como o conjunto de geradores $\{X_k\}$ realiza as transformações do grupo.

Antes de prosseguir com a determinação das propriedades restantes do grupo de Lie, serão apresentados alguns exemplos de determinação dos geradores infinitesimais.

Exemplo 5.23 (Geradores do grupo de alongamento/deslocamento unidimensional). Retornando ao grupo de 2 parâmetros discutido no exemplo 5.16, se $x' = f(x; a, b) = ax + b$, a identidade do grupo é $a = 1$ e $b = 0$, de onde resulta que a transformação infinitesimal será

$$x' = (1 + \delta a) x + \delta b = x + x \delta a + \delta b.$$

Então, de (5.56b),

$$u_a = \left. \frac{\partial}{\partial a} f(x; a, b) \right|_{a=0} = x, \quad u_b = \left. \frac{\partial}{\partial b} f(x; a, b) \right|_{b=0} = 1,$$

e então de (5.57) resultam os geradores do grupo

$$X_a = u_a(x) \frac{\partial}{\partial x} = x \frac{\partial}{\partial x}, \quad X_b = u_b(x) \frac{\partial}{\partial x} = \frac{\partial}{\partial x}.$$

Nota-se que o comutador $[X_a, X_b] \doteq X_a X_b - X_b X_a$ resulta

$$[X_a, X_b] = x \frac{\partial}{\partial x} \frac{\partial}{\partial x} - \frac{\partial}{\partial x} \left(x \frac{\partial}{\partial x} \right) = -\frac{\partial}{\partial x} = -X_b;$$

isto é, o comutador é uma função dos geradores do grupo.

Exemplo 5.24 (Geradores do grupo geral linear de 2 dimensões, $GL(2, \mathbb{R})$). A partir do exemplo 5.17, se $x = (x_1 \ x_2)^T$, então

$$x' = f(x_1, x_2; a) = Ax \implies \begin{cases} x'_1 = f_1(x_1, x_2; a_1, a_2) = a_1 x_1 + a_2 x_2 \\ x'_2 = f_2(x_1, x_2; a_3, a_4) = a_3 x_1 + a_4 x_2, \end{cases}$$

constituindo um grupo de Lie de 4 parâmetros. Agora, a identidade do grupo é $A = I_2$ e, portanto,

$$\begin{aligned} x'_1 &= (1 + \delta a_1) x_1 + \delta a_2 x_2 \\ x'_2 &= \delta a_3 x_1 + (1 + \delta a_4) x_2. \end{aligned}$$

Então, de (5.56b) e (5.57) resultam

$$\begin{aligned} u_{11}(x_1, x_2) &= x_1, & u_{12}(x_1, x_2) &= x_2, & u_{13}(x_1, x_2) &= u_{14}(x_1, x_2) = 0 \\ u_{23}(x_1, x_2) &= x_1, & u_{24}(x_1, x_2) &= x_2, & u_{21}(x_1, x_2) &= u_{22}(x_1, x_2) = 0, \end{aligned}$$

resultando os geradores

$$X_1 = x_1 \frac{\partial}{\partial x_1}, \quad X_2 = x_2 \frac{\partial}{\partial x_1}, \quad X_3 = x_1 \frac{\partial}{\partial x_2}, \quad X_4 = x_2 \frac{\partial}{\partial x_2}.$$

Os comutadores dos geradores ficam agora

$$\begin{aligned} [X_1, X_2] &= -X_2, & [X_1, X_3] &= X_3, & [X_1, X_4] &= 0, \\ [X_2, X_3] &= X_4 - X_1, & [X_2, X_4] &= -X_2, & [X_3, X_4] &= X_3. \end{aligned}$$

Novamente, os comutadores dos geradores sempre são expressos como combinações lineares dos mesmos.

Exemplo 5.25 (Gerador do grupo $SO(2)$ e rotações finitas). Do exemplo 5.20, o $SO(2)$ é um grupo de 1 parâmetro, sendo que as transformações são matrizes ortogonais $RR^T = I_2$ unimodulares ($\det R = +1$) que atuam em um espaço de dimensão 2. Ou seja, se $x = (x_1 \ x_2)^T$ e denotando por ϕ o parâmetro do grupo, então

$$x' = f(x_1, x_2; \phi) = R(\phi)x,$$

sendo que a identidade do grupo é $R(a = 0) = I_3$. Se uma transformação infinitesimal for escrita como a soma matricial

$$x' = R(\delta\phi)x = [I_2 + B(\delta\phi)]x,$$

sendo $\delta\phi$ o infinitésimo do parâmetro do grupo, como R é ortogonal, resulta que, mantendo somente os termos em primeira ordem no infinitésimo,

$$RR^T = (I_2 + B)(I_2 + \tilde{B}) = I_2 + \tilde{B} + B = I_2 \implies \tilde{B} + B = 0;$$

ou seja, $B = -\tilde{B}$; a matriz B é antissimétrica. Neste caso, pode-se escrever

$$B = \begin{pmatrix} 0 & \delta\phi \\ -\delta\phi & 0 \end{pmatrix} \implies \begin{cases} x'_1 = x_1 + x_2 \delta\phi \\ x'_2 = x_2 - x_1 \delta\phi. \end{cases}$$

Então, de (5.56b) e (5.57) obtém-se

$$X = x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2},$$

o qual é o gerador do $SO(2)$.

De acordo com a discussão realizada acima, o operador de rotação, para um ângulo finito ϕ , pode ser escrito formalmente como

$$R(\phi) = e^{\phi X}.$$

A partir da transformação infinitesimal é possível obter-se a representação de uma matriz de rotação $R(\phi)$ por um ângulo finito ϕ . Retornando a

$$x'(\delta\phi) = R(\delta\phi)x \implies R(\delta\phi) = I_2 + \delta\phi X, \text{ sendo } X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Então, de acordo com a argumentação acima, realizando um número arbitrariamente grande N de rotações infinitesimais consecutivas, resulta a rotação pelo ângulo finito $\phi = N\delta\phi$. Neste caso,

$$x'(\phi) = \lim_{N \rightarrow \infty} R^N(\delta\phi)x = R(\phi)x.$$

Mas,

$$R(\phi) \stackrel{N \gg 1}{\approx} R^N(\delta\phi) = \left(I_2 + \frac{\phi}{N} X \right)^N = \sum_{n=0}^N \binom{N}{n} \left(\frac{\phi}{N} X \right)^n.$$

Para $N \gg 1$, pode-se usar a expansão assintótica³³

$$\binom{N}{n} \frac{1}{N^n} \stackrel{N \gg 1}{\approx} \frac{1}{n!} \left[1 + \frac{(1-n)n}{2N} + \mathcal{O}\left(\frac{1}{N^2}\right) \right].$$

Portanto, mantendo somente o termo mais significativo,

$$R(\phi) = \sum_{n=0}^{\infty} \frac{\phi^n}{n!} X^n = I_2 + \phi X + \frac{\phi^2}{2!} X^2 + \frac{\phi^3}{3!} X^3 + \dots = e^{\phi X}.$$

Observando que para $k = 0, 1, 2, \dots$,

$$X^{2k} = (-1)^k I_2, \quad X^{2k+1} = (-1)^k X,$$

resulta

$$\begin{aligned} R(\phi) &= \left(1 - \frac{\phi^2}{2!} + \dots \right) I_2 + \left(\phi - \frac{\phi^3}{3!} + \dots \right) X = \cos \phi \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin \phi \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \\ R(\phi) &= \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}, \end{aligned}$$

a qual é a matriz que executa uma rotação dos eixos coordenados de um referencial bidimensional em torno de um eixo de rotação perpendicular ao plano do referencial.³⁴

De fato, o resultado $R(\phi) = e^{\phi X}$ obtido no exemplo anterior é sustentado também por teoremas válidos para grupos matriciais. Dada uma matriz $n \times n$ complexa A , a exponencial da mesma, a qual é desenvolvida na forma de uma série de potências

$$e^A = \sum_{m=0}^{\infty} \frac{A^m}{m!} = I_n + A + \frac{1}{2!} A^2 + \dots,$$

sempre converge. Além disso, se \mathcal{A} é um subgrupo do $GL(n, \mathbb{C})$ e é também um grupo de um parâmetro t , então sempre existe uma única matriz $n \times n$ complexa X tal que

$$A(t) = e^{tX}, \text{ sendo } A(t) \in \mathcal{A}.$$

³³Referência: <http://functions.wolfram.com/06.03.06.0003.01>.

³⁴Ver discussão na seção 6.2.1.

Exemplo 5.26 (Geradores do grupo $SO(3)$). Novamente do exemplo 5.20, o $SO(3)$ é um grupo de 3 parâmetros, sendo que as transformações são matrizes ortogonais $A\tilde{A} = I_3$ unimodulares ($\det A = +1$) que atuam em um espaço de dimensão 3. Ou seja, se $x = (x_1 \ x_2 \ x_3)^T$, então

$$x' = f(x_1, x_2, x_3; a_1, a_2, a_3) = A(a_1, a_2, a_3)x,$$

sendo que a identidade do grupo é $A(a=0) = I_3$. Se uma transformação infinitesimal for escrita como a soma matricial

$$x' = A(\epsilon_1, \epsilon_2, \epsilon_3)x = [I_3 + B(\epsilon_1, \epsilon_2, \epsilon_3)]x,$$

sendo os ϵ_i 's infinitésimos dos parâmetros do grupo, como A é ortogonal, resulta que, mantendo somente os termos em primeira ordem nos infinitésimos,

$$A\tilde{A} = (I_3 + B)(I_3 + \tilde{B}) = I_3 + \tilde{B} + B = I_3 \implies \tilde{B} + B = 0;$$

ou seja, $B = -\tilde{B}$; a matriz B é antissimétrica. Neste caso, pode-se escrever

$$B = \begin{pmatrix} 0 & \epsilon_3 & -\epsilon_2 \\ -\epsilon_3 & 0 & \epsilon_1 \\ \epsilon_2 & -\epsilon_1 & 0 \end{pmatrix} \implies f_i(x; \epsilon) = x_i + \sum_{j,k=1}^3 \epsilon_{ijk} \epsilon_k x_j, \quad (i = 1, 2, 3),$$

sendo ϵ_{ijk} o símbolo de Levi-Civita.³⁵

Então, de (5.56b) e (5.57) obtém-se

$$u_{i\ell}(x) = \left. \frac{\partial}{\partial \epsilon_\ell} f_i(x; \epsilon) \right|_{\epsilon=0} = \sum_{j,k=1}^3 \epsilon_{ijk} \delta_{k\ell} x_j = \sum_{j=1}^3 \epsilon_{ij\ell} x_j.$$

Com isso, os geradores do grupo ficam

$$X_i = \sum_{j=1}^3 u_{ji}(x) \frac{\partial}{\partial x_j} = \sum_{j,k=1}^3 \epsilon_{ijk} x_k \frac{\partial}{\partial x_j}, \quad (i = 1, 2, 3).$$

Os comutadores do grupo são:

$$\begin{aligned} [X_i, X_j] &= X_i X_j - X_j X_i, \quad (i \neq j = 1, 2, 3), \\ &= \sum_{p,q,r=1}^3 \epsilon_{pqi} \left(\epsilon_{rpk} x_q \frac{\partial}{\partial x_r} - \epsilon_{qrj} x_r \frac{\partial}{\partial x_p} \right), \\ [X_i, X_j] &= -x_i \frac{\partial}{\partial x_j} + x_j \frac{\partial}{\partial x_i}, \end{aligned}$$

onde foi empregada a identidade (6.3g). A mesma identidade mostra que este resultado pode ser colocado na forma

$$[X_i, X_j] = \sum_{k=1}^3 \epsilon_{ijk} X_k, \quad (i, j = 1, 2, 3).$$

Recordando da discussão realizada na seção 5.9.1.2, se for introduzido o operador de momento angular na mecânica quântica,

$$L_i = \sum_{j,k=1}^3 \epsilon_{ijk} x_j p_k \xrightarrow{p=-i\hbar\nabla} i\hbar \sum_{j,k=1}^3 \epsilon_{ijk} x_k \frac{\partial}{\partial x_j}, \quad (i = 1, 2, 3),$$

observa-se que o gerador infinitesimal do $SO(3)$ é próprio operador L ,

$$X_i = \sum_{j,k=1}^3 \epsilon_{ijk} x_k \frac{\partial}{\partial x_j} = -\frac{i}{\hbar} L_i, \quad (i = 1, 2, 3).$$

³⁵Ver seção 6.1.2.

Obtém-se assim as conhecidas relações de comutação do momento angular

$$[L_i, L_j] = i\hbar \sum_{k=1}^3 \epsilon_{ijk} L_k, \quad (i, j = 1, 2, 3) \Rightarrow \begin{cases} [L_1, L_2] = i\hbar L_3 \\ [L_3, L_1] = i\hbar L_2 \\ [L_2, L_3] = i\hbar L_1. \end{cases}$$

Finalmente, colocando os geradores do $SO(3)$, escritos em termos do momento angular, no operador de transformação finita (5.58), pode-se escrever o mesmo como

$$R(\mathbf{a}) = e^{-i\mathbf{a}\cdot\mathbf{L}/\hbar},$$

onde o operador $R(\mathbf{a})$ é interpretado como o operador que executa uma rotação arbitrária finita no espaço \mathbb{E}^3 . Neste caso, o vetor \mathbf{a} é um vetor axial associado a uma rotação em torno de algum eixo que passa pela origem.

5.10.3 CONSTANTES DE ESTRUTURA

Os exemplos anteriores mostraram que, nos grupos em questão, os comutadores dos geradores infinitesimais sempre podem ser escritos em termos de combinações lineares dos próprios geradores. Será mostrado agora que esta propriedade é universal para grupos de Lie.

Nas transformações infinitesimais (5.56a), supõe-se que os r parâmetros a variam de forma independente entre si. Em consequência, na variação infinitesimal da coordenada x_i em (5.56b), $dx_i = 0$ se e somente se todas as variações δa_k forem simultaneamente nulas. Isto implica em uma relação de independência linear entre as funções $u_{ik}(x)$ ($k = 1, \dots, r$).

Retorna-se então a (5.56e), a qual será escrita como

$$\frac{\partial x_i}{\partial a_\lambda} = u_{i\kappa}(x) \psi_{\kappa\lambda}(a), \quad \begin{matrix} i=1, \dots, n \\ \lambda, \kappa=1, \dots, r. \end{matrix} \quad (5.59)$$

Para simplificar a notação, já na expressão acima adotou-se, em primeiro lugar, o uso de índices gregos para denotar os r distintos parâmetros do grupo, mantendo os índices latinos para denotar os componentes do vetor. Em segundo lugar, adotou-se a *convenção de somas implícitas de índices repetidos*,³⁶ segundo a qual existe na expressão acima uma soma oculta sobre o índice $\kappa = 1, \dots, r$, uma vez que este índice está repetido.

Como a transformação (5.54) é por hipótese uma função analítica de a , então

$$\frac{\partial^2 x_i}{\partial a_\lambda \partial a_\mu} = \frac{\partial^2 x_i}{\partial a_\mu \partial a_\lambda};$$

ou seja,

$$\begin{aligned} \frac{\partial}{\partial a_\lambda} [u_{i\kappa}(x) \psi_{\kappa\mu}(a)] - \frac{\partial}{\partial a_\mu} [u_{i\kappa}(x) \psi_{\kappa\lambda}(a)] &= 0, \\ u_{i\kappa} \left(\frac{\partial \psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial \psi_{\kappa\lambda}}{\partial a_\mu} \right) + \psi_{\kappa\mu} \frac{\partial u_{i\kappa}}{\partial a_\lambda} - \psi_{\kappa\lambda} \frac{\partial u_{i\kappa}}{\partial a_\mu} &= 0, \end{aligned}$$

onde foi levado em conta que as funções $u_{i\kappa}$ são funções dos componentes x_i , os quais, de acordo com (5.59), são funções dos parâmetros a . Portanto, se $u_{i\kappa} = u_{i\kappa}(x(a))$,

$$\frac{\partial u_{i\kappa}}{\partial a_\lambda} = \frac{\partial u_{i\kappa}}{\partial x_j} \frac{\partial x_j}{\partial a_\lambda} = \frac{\partial u_{i\kappa}}{\partial x_j} u_{j\nu} \psi_{\nu\lambda}. \quad (5.60)$$

Portanto,

$$\begin{aligned} u_{i\kappa} \left(\frac{\partial \psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial \psi_{\kappa\lambda}}{\partial a_\mu} \right) + u_{j\nu} \frac{\partial u_{i\kappa}}{\partial x_j} \psi_{\kappa\mu} \psi_{\nu\lambda} - \underbrace{u_{j\nu} \frac{\partial u_{i\kappa}}{\partial x_j} \psi_{\kappa\lambda} \psi_{\nu\mu}}_{\kappa \leftrightarrow \nu} &= 0, \\ u_{i\kappa} \left(\frac{\partial \psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial \psi_{\kappa\lambda}}{\partial a_\mu} \right) + \left(u_{j\nu} \frac{\partial u_{i\kappa}}{\partial x_j} - u_{j\kappa} \frac{\partial u_{i\nu}}{\partial x_j} \right) \psi_{\kappa\mu} \psi_{\nu\lambda} &= 0. \end{aligned}$$

³⁶Ver a convenção completa na seção 6.1.1.

Lembrando agora da relação (5.56d), $\psi_{\kappa\mu}(a)\Theta_{\mu\lambda}(a) = \delta_{\kappa\lambda}$, chega-se ao resultado

$$\begin{aligned} \left(\frac{\partial\psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial\psi_{\kappa\lambda}}{\partial a_\mu}\right)u_{i\kappa}\Theta_{\mu\tau}\Theta_{\lambda\sigma} + \left(u_{j\nu}\frac{\partial u_{i\kappa}}{\partial x_j} - u_{j\kappa}\frac{\partial u_{i\nu}}{\partial x_j}\right)\psi_{\kappa\mu}\Theta_{\mu\tau}\psi_{\nu\lambda}\Theta_{\lambda\sigma} &= 0, \\ \left(\frac{\partial\psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial\psi_{\kappa\lambda}}{\partial a_\mu}\right)u_{i\kappa}\Theta_{\mu\tau}\Theta_{\lambda\sigma} + \left(u_{j\nu}\frac{\partial u_{i\kappa}}{\partial x_j} - u_{j\kappa}\frac{\partial u_{i\nu}}{\partial x_j}\right)\delta_{\kappa\tau}\delta_{\nu\sigma} &= 0, \\ u_{j\tau}\frac{\partial u_{i\sigma}}{\partial x_j} - u_{j\sigma}\frac{\partial u_{i\tau}}{\partial x_j} &= \left(\frac{\partial\psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial\psi_{\kappa\lambda}}{\partial a_\mu}\right)\Theta_{\mu\tau}\Theta_{\lambda\sigma}u_{i\kappa}. \end{aligned}$$

Definindo a quantidade

$$c_{\tau\sigma}^\kappa(a) \doteq \left(\frac{\partial\psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial\psi_{\kappa\lambda}}{\partial a_\mu}\right)\Theta_{\mu\tau}\Theta_{\lambda\sigma}, \quad (5.61)$$

escreve-se

$$u_{j\tau}\frac{\partial u_{i\sigma}}{\partial x_j} - u_{j\sigma}\frac{\partial u_{i\tau}}{\partial x_j} = c_{\tau\sigma}^\kappa u_{i\kappa}. \quad (5.62)$$

Derivando agora ambos os lados em relação a a_ρ ,

$$\frac{\partial}{\partial a_\rho} \left(u_{j\tau}\frac{\partial u_{i\sigma}}{\partial x_j}\right) - \frac{\partial}{\partial a_\rho} \left(u_{j\sigma}\frac{\partial u_{i\tau}}{\partial x_j}\right) = \frac{\partial c_{\tau\sigma}^\kappa}{\partial a_\rho} u_{i\kappa} + c_{\tau\sigma}^\kappa \frac{\partial}{\partial a_\rho} u_{i\kappa}.$$

Nos termos que contêm derivadas das funções $u(x)$ e suas derivadas em x , realiza-se a interpretação $\partial/\partial a_\rho = (\partial x_k/\partial a_\rho)(\partial/\partial x_k)$, conforme foi realizado em (5.60). Portanto,

$$\frac{\partial x_k}{\partial a_\rho} \frac{\partial}{\partial x_k} \left(u_{j\tau}\frac{\partial u_{i\sigma}}{\partial x_j}\right) - \frac{\partial x_k}{\partial a_\rho} \frac{\partial}{\partial x_k} \left(u_{j\sigma}\frac{\partial u_{i\tau}}{\partial x_j}\right) = \frac{\partial c_{\tau\sigma}^\kappa}{\partial a_\rho} u_{i\kappa} + c_{\tau\sigma}^\kappa \frac{\partial x_k}{\partial a_\rho} \frac{\partial u_{i\kappa}}{\partial x_k}.$$

Por outro lado, retornando a (5.62) e aplicando diretamente o último operador, resulta

$$\frac{\partial x_k}{\partial a_\rho} \frac{\partial}{\partial x_k} \left(u_{j\tau}\frac{\partial u_{i\sigma}}{\partial x_j}\right) - \frac{\partial x_k}{\partial a_\rho} \frac{\partial}{\partial x_k} \left(u_{j\sigma}\frac{\partial u_{i\tau}}{\partial x_j}\right) = c_{\tau\sigma}^\kappa \frac{\partial x_k}{\partial a_\rho} \frac{\partial u_{i\kappa}}{\partial x_k}.$$

Comparando as duas últimas expressões, conclui-se que

$$\frac{\partial c_{\tau\sigma}^\kappa}{\partial a_\rho} u_{i\kappa} = 0, \quad (i=1, \dots, n)_{(\tau, \sigma, \rho=1, \dots, r)}$$

Como já foi estabelecido que as funções $u_{i\kappa}(x)$ ($\kappa = 1, \dots, r$) são linearmente independentes, este resultado implica que

$$\frac{\partial c_{\tau\sigma}^\kappa}{\partial a_\rho} = 0 \implies c_{\tau\sigma}^\kappa(a) = \text{cte.}, \quad (\kappa, \tau, \sigma = 1, \dots, r).$$

Estes coeficientes são as **constantes de estrutura do grupo de Lie**.

Retornando à definição (5.61), resulta que os elementos da matriz Ψ satisfazem

$$\frac{\partial\psi_{\kappa\mu}}{\partial a_\lambda} - \frac{\partial\psi_{\kappa\lambda}}{\partial a_\mu} = c_{\tau\sigma}^\kappa \psi_{\tau\mu} \psi_{\sigma\lambda}, \quad (5.63)$$

onde foi empregado novamente (5.56d).

Consideram-se agora os geradores infinitesimais do grupo, definidos por (5.57). Calculando o comutador de dois geradores, obtém-se

$$\begin{aligned} [X_\rho, X_\sigma] &= X_\rho X_\sigma - X_\sigma X_\rho \\ &= u_{i\rho} \frac{\partial}{\partial x_i} u_{j\sigma} \frac{\partial}{\partial x_j} - u_{j\sigma} \frac{\partial}{\partial x_j} u_{i\rho} \frac{\partial}{\partial x_i} \\ &= \left(u_{i\rho} \frac{\partial u_{j\sigma}}{\partial x_i} - u_{i\sigma} \frac{\partial u_{j\rho}}{\partial x_i}\right) \frac{\partial}{\partial x_j}. \end{aligned}$$

Mas, de acordo com (5.62),

$$[X_\rho, X_\sigma] = c_{\rho\sigma}^\kappa u_{j\kappa} \frac{\partial}{\partial x_j} = c_{\rho\sigma}^\kappa X_\kappa. \quad (5.64a)$$

Ou seja, de fato o comutador de dois geradores infinitesimais sempre pode ser escrito como uma combinação linear dos mesmos geradores, sendo que os coeficientes da combinação são as constantes de estrutura do grupo.

Outras propriedades das constantes de estrutura são facilmente obtidas. Claramente, se em (5.62) forem trocados os índices $\tau \leftrightarrow \sigma$,

$$c_{\sigma\tau}^k u_{i\kappa} = u_{j\sigma} \frac{\partial u_{i\tau}}{\partial x_j} - u_{j\tau} \frac{\partial u_{i\sigma}}{\partial x_j} = -c_{\tau\sigma}^k u_{i\kappa} \implies (c_{\sigma\tau}^k + c_{\tau\sigma}^k) u_{i\kappa} = 0 \implies c_{\sigma\tau}^k = -c_{\tau\sigma}^k. \quad (5.64b)$$

Seja agora a identidade de Jacobi de três objetos, definida na seção 3.7.3, é fácil verificar que quaisquer três geradores do grupo satisfazem

$$[X_\rho, [X_\sigma, X_\tau]] + [X_\sigma, [X_\tau, X_\rho]] + [X_\tau, [X_\rho, X_\sigma]] = 0.$$

Introduzindo a relação (5.64a), conclui-se então que

$$(c_{\rho\sigma}^\mu c_{\mu\tau}^\nu + c_{\sigma\tau}^\mu c_{\mu\rho}^\nu + c_{\tau\rho}^\mu c_{\mu\sigma}^\nu) X_\nu = 0 \implies c_{\rho\sigma}^\mu c_{\mu\tau}^\nu + c_{\sigma\tau}^\mu c_{\mu\rho}^\nu + c_{\tau\rho}^\mu c_{\mu\sigma}^\nu = 0, \quad (\rho, \sigma, \nu, \tau = 1, \dots, r). \quad (5.64c)$$

Resumindo, partindo da definição 5.14 de um grupo de Lie de transformações, chega-se ao sistema de equações (5.56e), o qual descreve a variação das coordenadas de um vetor do espaço vetorial com os parâmetros do grupo. Como as transformações são analíticas nos parâmetros, chegou-se então a (5.62) e (5.63) e, finalmente, às propriedades (5.64b) e (5.64c) das constantes de estrutura do grupo.

Sophus Lie mostrou que este procedimento pode ser invertido: se for possível empregar as propriedades (5.64b) e (5.64c) para se derivar as constantes de estrutura de um grupo de transformações, pode-se então encontrar as funções $u(x)$ e $\psi(a)$ que satisfazem (5.62) e (5.63). Com isso, pode-se encontrar finalmente as funções $x(a)$ pela integração de (5.56e), as quais irão formar o grupo de transformações.

Exemplo 5.27 (Constantes de estrutura do $SO(3)$). No exemplo 5.26 já foi mostrado que os geradores do $SO(3)$ satisfazem a relação de comutação

$$[X_i, X_j] = \epsilon_{ijk} X_k,$$

onde foi mantida a notação de Einstein de somas implícitas. Portanto, de (5.64a) segue diretamente que as constantes de estrutura do grupo são os símbolos de Levi-Civita $c_{ij}^k = \epsilon_{ijk}$.

Como os símbolos de Levi-Civita são antissimétricos, segue diretamente a propriedade (5.64b), $c_{ij}^k = -c_{ji}^k$. Já para verificar que a identidade (5.64c) é satisfeita, observa-se

$$\begin{aligned} c_{ij}^k c_{ks}^r + c_{js}^k c_{ki}^r + c_{si}^k c_{kj}^r &= \epsilon_{ijk} \epsilon_{ksr} + \epsilon_{jsk} \epsilon_{kir} + \epsilon_{sik} \epsilon_{kjr} \\ &= \delta_{is} \delta_{jr} - \delta_{ir} \delta_{js} + \delta_{ij} \delta_{rs} - \delta_{jr} \delta_{is} + \delta_{js} \delta_{ir} - \delta_{rs} \delta_{ij} \\ &= 0, \quad (\checkmark) \end{aligned}$$

onde a propriedade (6.3g) foi empregada.

Exemplo 5.28 (Geradores e constantes de estrutura do $SU(2)$). Do exemplo 5.22, observa-se que o $SU(2)$ é um grupo de $2^2 - 1 = 3$ parâmetros reais, composto por matrizes unitárias $UU^\dagger = I_2$ unimodulares ($\det U = +1$) que atuam sobre um espaço complexo de dimensão 2. Sendo $x = (x_1 \ x_2)^T$ a matriz dos componentes de um vetor, então

$$x' = f(x_1, x_2; a_1, a_2, a_3) = Ux \implies \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \eta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \alpha x_1 + \beta x_2 \\ \gamma x_1 + \eta x_2 \end{pmatrix},$$

onde $\{\alpha, \beta, \gamma, \eta\} \subset \mathbb{C}$.

Como U é unitária, então

$$UU^\dagger = \begin{pmatrix} \alpha & \beta \\ \gamma & \eta \end{pmatrix} \begin{pmatrix} \alpha^* & \gamma^* \\ \beta^* & \eta^* \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \implies \begin{cases} |\alpha|^2 + |\beta|^2 = |\gamma|^2 + |\eta|^2 = 1 \\ \alpha\gamma^* + \beta\eta^* = \alpha^*\gamma + \beta^*\eta = 0. \end{cases}$$

Além disso,

$$\det U = 1 \implies \alpha\eta - \beta\gamma = 1.$$

Observa-se que as condições acima são satisfeitas se

$$\eta = \alpha^*, \quad \gamma = -\beta^*; \quad \text{ou seja, } U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}.$$

A forma mais geral de uma matriz $U \in SU(2)$ ficou reduzida às escolhas de dois números complexos (α e β), os quais ainda estão sujeitos às condições impostas às matrizes do grupo. Por isso, dos quatro parâmetros reais ainda restantes (as partes reais e imaginárias de α e β) somente três serão independentes e os valores possíveis que estes podem assumir estão restritos a intervalos finitos no eixo real, uma vez que o grupo é compacto. Uma forma geral para os três parâmetros independentes é $\alpha = e^{i\xi} \cos \eta$ e $\beta = -e^{i\zeta} \sin \eta$, sendo $\{\xi, \zeta, \eta\} \subset \mathbb{R}$. Assim, a forma mais geral para a matriz U é

$$U = \begin{pmatrix} e^{i\xi} \cos \eta & -e^{i\zeta} \sin \eta \\ e^{-i\zeta} \sin \eta & e^{-i\xi} \cos \eta \end{pmatrix}.$$

Uma transformação infinitesimal será então obtida a partir da matriz U se

$$U = I_2 + \begin{pmatrix} \delta\alpha & \delta\beta \\ -\delta\beta^* & \delta\alpha^* \end{pmatrix} = \begin{pmatrix} 1 + \delta\alpha & \delta\beta \\ -\delta\beta^* & 1 + \delta\alpha^* \end{pmatrix}.$$

Mas, a condição $|\alpha|^2 + |\beta|^2 = 1$ implica que, em primeira ordem nos infinitésimos, $1 + \delta\alpha + \delta\alpha^* = 1 \Rightarrow \delta\alpha = -\delta\alpha^*$; ou seja, $\delta\alpha = \frac{1}{2}i\delta a_3$, ($\delta a_3 \in \mathbb{R}$). Portanto, os parâmetros infinitesimais do $SU(2)$ são $\{\delta a_1, \delta a_2, \delta a_3\} \subset \mathbb{R}$ tais que

$$\delta\beta = \frac{1}{2}\delta a_2 + \frac{i}{2}\delta a_1, \quad \delta\alpha = \frac{i}{2}\delta a_3.$$

Com as escolhas feitas, pode-se finalmente escrever, empregando a notação de Einstein de somas implícitas,

$$U(\delta a_1, \delta a_2, \delta a_3) = I_2 + \begin{pmatrix} \frac{i}{2}\delta a_3 & \frac{i}{2}\delta a_1 + \frac{1}{2}\delta a_2 \\ \frac{i}{2}\delta a_1 - \frac{1}{2}\delta a_2 & -\frac{i}{2}\delta a_3 \end{pmatrix} = I_2 + \frac{i}{2}\sigma_\alpha \delta a_\alpha,$$

sendo $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

as matrizes de Pauli.

Em termos das matrizes de Pauli, as transformações infinitesimais ficam

$$x + dx = U(\delta a_1, \delta a_2, \delta a_3)x \Rightarrow x_j + dx_j = U_{jk}x_k, \quad (j = 1, 2) \Rightarrow dx_j = \frac{i}{2}(\sigma_\alpha)_{jk} x_k \delta a_\alpha.$$

Assim, de (5.56b) resulta que

$$u_{j\alpha}(x_1, x_2) = \frac{i}{2}(\sigma_\alpha)_{jk} x_k, \quad \left(\begin{matrix} j=1,2 \\ \alpha=1,2,3 \end{matrix} \right)$$

e, portanto, de (5.57) resulta que os geradores infinitesimais do $SU(2)$ são

$$X_\alpha = u_{j\alpha}(x) \frac{\partial}{\partial x_j} = \frac{i}{2}(\sigma_\alpha)_{jk} x_k \frac{\partial}{\partial x_j}.$$

É comum afirmar-se que as matrizes de Pauli são os geradores do grupo.

Para determinar as constantes de estrutura, ao invés de partir da definição (5.61), é mais prático calcular os comutadores dos geradores. Para tanto, nota-se primeiro que as matrizes de Pauli satisfazem a relação de comutação

$$[\sigma_\alpha, \sigma_\beta] = 2i\epsilon_{\alpha\beta\gamma}\sigma_\gamma.$$

Dessa maneira,

$$\begin{aligned} [X_\alpha, X_\beta] &= X_\alpha X_\beta - X_\beta X_\alpha \\ &= -\frac{1}{4}(\sigma_\alpha)_{jk}(\sigma_\beta)_{\ell m} \left[x_k \frac{\partial}{\partial x_j} \left(x_m \frac{\partial}{\partial x_\ell} \right) - x_m \frac{\partial}{\partial x_\ell} \left(x_k \frac{\partial}{\partial x_j} \right) \right] \end{aligned}$$

$$= -\frac{1}{4} \left[(\sigma_\alpha)_{jk} (\sigma_\beta)_{\ell j} x_k \frac{\partial}{\partial x_\ell} - (\sigma_\alpha)_{j\ell} (\sigma_\beta)_{\ell k} x_k \frac{\partial}{\partial x_j} \right],$$

$$[X_\alpha, X_\beta] = \frac{1}{4} [\sigma_\alpha, \sigma_\beta]_{jk} x_k \frac{\partial}{\partial x_j}.$$

Ou seja,

$$[X_\alpha, X_\beta] = \epsilon_{\alpha\beta\gamma} X_\gamma.$$

Imediatamente, conclui-se que as constantes de estrutura do $SU(2)$ são $c_{\alpha\beta}^\gamma = \epsilon_{\alpha\beta\gamma}$. Observa-se que a relação de comutação do $SU(2)$ é idêntica à relação do $SO(3)$.

5.10.4 PARAMETRIZAÇÃO DAS TRANSFORMAÇÕES DO GRUPO

Com a definição dos geradores infinitesimais do grupo em (5.57), é possível escrever (5.56e) como

$$\frac{\partial x_i}{\partial a_\lambda} = \psi_{\kappa\lambda}(a) X_\kappa x_i, \text{ sendo } \psi_{\kappa\lambda}(0) = \delta_{\kappa\lambda}. \quad (5.65)$$

Qualquer ponto no espaço de parâmetros do grupo pode ser atingido a partir de um ponto inicial ao longo de uma reta que conecta esses pontos. Essa reta pode, por sua vez, ser parametrizada por relações do tipo

$$a_\lambda = s_\lambda \tau, \quad (\lambda = 1, \dots, r), \quad (5.66)$$

sendo $\tau \in \mathbb{R}$ o parâmetro que gera a reta e $s_\lambda \in \mathbb{R}$ as declividades das projeções da reta sobre cada coordenada no espaço de parâmetros. Com esta parametrização, o valor $\tau = 0$ gera a transformação identidade $x_i = x_i(a = 0)$. Já para $\tau \neq 0$, as componentes x_i são transformadas como

$$x_i(\tau) = S(\tau) x_i(0), \text{ sendo } S(0) = 1.$$

A quantidade $S(\tau)$ é um operador parametrizado por τ que atua sobre o valor inicial $x_i(0)$ da coordenada. Inserindo esta expressão em (5.65) obtém-se

$$\frac{dx_i}{d\tau} = \frac{\partial x_i}{\partial a_\lambda} \frac{da_\lambda}{d\tau} = s_\lambda \psi_{\kappa\lambda}(s\tau) X_\kappa x_i(\tau) \implies \frac{dS}{d\tau} x_i(0) = s_\lambda \psi_{\kappa\lambda}(s\tau) X_\kappa S(\tau) x_i(0),$$

sendo $\psi_{\kappa\lambda}(s\tau) = \psi_{\kappa\lambda}(s_1\tau, \dots, s_r\tau)$; ou seja, o operador $S(\tau)$ satisfaz

$$\frac{dS}{d\tau} = s_\lambda \psi_{\kappa\lambda}(s\tau) X_\kappa S(\tau).$$

Então, em $\tau = 0$,

$$\left. \frac{dS}{d\tau} \right|_{\tau=0} = s_\kappa X_\kappa.$$

Com os resultados recém obtidos, conclui-se que o operador $S(\tau)$ pode ser formalmente desenvolvido em uma série de potências de τ como

$$S(\tau) = S(0) + \tau \left. \frac{dS}{d\tau} \right|_{\tau=0} + \mathcal{O}(\tau^2) \implies S(\tau) = 1 + \tau s_\kappa X_\kappa + \dots \quad (5.67a)$$

Portanto, a transformação infinitesimal sobre x_i pode ser formalmente escrita como

$$x_i(\tau) = (1 + \tau s_\kappa X_\kappa + \dots) x_i(0). \quad (5.67b)$$

Ou seja, o operador $s_\kappa X_\kappa$ também gera a transformação sobre x_i , parametrizada por τ . Por outro lado, os mesmos resultados permitem concluir que

$$s_\kappa X_\kappa = \lim_{\tau \rightarrow 0} \frac{S(\tau) - 1}{\tau}. \quad (5.67c)$$

Ou seja, o operador $S(\tau)$ determina o gerador $s_\kappa X_\kappa$.

Exemplo 5.29 (Rotação em torno de um eixo arbitrário). Deseja-se realizar uma rotação dos eixos de um sistema de coordenadas Cartesianas por um ângulo ϕ em torno de um eixo que cruza a origem e está orientado na direção do vetor \hat{n} . Uma rotação qualquer em um espaço de três dimensões é realizada pelos geradores do $SO(3)$, derivados no exemplo 5.26. Contudo, como a rotação irá ocorrer em torno do eixo fixo na direção \hat{n} , o operador de rotação desejado pertence ao subgrupo $SO(2) \subset SO(3)$ correspondente às rotações que ocorrem no plano perpendicular a \hat{n} . Por esta razão, o operador de rotação é obtido de acordo com o procedimento realizado no exemplo 5.25. Empregando a forma dos geradores em termos dos operadores de momento angular, resulta que o operador de rotação desejado pode ser escrito de forma parametrizada pelo ângulo ϕ como

$$R_{\hat{n}}(\phi) = 1 - i \frac{\phi}{\hbar} (\hat{n} \cdot \mathbf{L}) + \dots = e^{-i\phi(\hat{n} \cdot \mathbf{L})/\hbar}.$$

Dada agora a transformação implementada por $S(\tau)$, a sua transformação inversa $S^{-1}(\tau)$ também pode ser desenvolvida em série como

$$S^{-1}(\tau) = 1 + \tau \left. \frac{dS^{-1}}{d\tau} \right|_{\tau=0} + \dots$$

e é tal que, na sequência de transformações $x_i \rightarrow x'_i = S(\tau) x_i \rightarrow x''_i = x_i = S^{-1}(\tau) S(\tau) x_i$,

$$x_i = \left(1 + \tau \left. \frac{dS^{-1}}{d\tau} \right|_{\tau=0} + \dots \right) (1 + \tau s_{\kappa} X_{\kappa} + \dots) x_i = \left[1 + \tau \left(s_{\kappa} X_{\kappa} + \left. \frac{dS^{-1}}{d\tau} \right|_{\tau=0} \right) + \dots \right] x_i.$$

Ou seja,

$$S^{-1}(\tau) = 1 - \tau s_{\kappa} X_{\kappa} + \dots \quad (5.67d)$$

Portanto, o operador inverso de $S(\tau)$ é obtido simplesmente pela inversão $\tau \rightarrow -\tau$.

Seja agora um outro conjunto de declividades $t = (t_1, \dots, t_r)$ de uma reta no espaço de parâmetros. Este conjunto serve para definir um outro operador de transformação parametrizado $T(\tau) = 1 + \tau t_{\lambda} X_{\lambda} + \dots$. Este operador possui as mesmas propriedades genéricas do operador $S(\tau)$. Com ambos os operadores, pode-se realizar a sequência de transformações sobre x_i dada por

$$\begin{aligned} x_i \rightarrow x'_i = T(\tau) x_i \rightarrow x''_i = S(\tau) x'_i = S(\tau) T(\tau) x_i \\ = [(1 + \tau t_{\kappa} X_{\kappa} + \dots) (1 + \tau s_{\kappa} X_{\kappa} + \dots)] x_i = [1 + \tau (s_{\kappa} + t_{\kappa}) X_{\kappa} + \dots] x_i. \end{aligned}$$

Portanto, $S(\tau) T(\tau) = 1 + \tau (s_{\kappa} + t_{\kappa}) X_{\kappa} + \dots$. A outra sequência possível de transformações é

$$x_i \rightarrow x'''_i = T(\tau) S(\tau) x_i = [1 + \tau (s_{\kappa} + t_{\kappa}) X_{\kappa} + \dots] x_i,$$

como antes. Porém, embora

$$[T(\tau) S(\tau)]^{-1} x'''_i = x_i \implies [T(\tau) S(\tau)]^{-1} = S^{-1}(\tau) T^{-1}(\tau),$$

pode-se mostrar que

$$\begin{aligned} [T(\tau) S(\tau)]^{-1} x'''_i &= [T(\tau) S(\tau)]^{-1} [S(\tau) T(\tau)] x_i \\ &= S^{-1}(\tau) T^{-1}(\tau) S(\tau) T(\tau) x_i = \{1 + \tau^2 [s_{\kappa} X_{\kappa}, t_{\lambda} X_{\lambda}] + \dots\} x_i; \end{aligned}$$

ou seja,

$$S^{-1}(\tau) T^{-1}(\tau) S(\tau) T(\tau) = 1 + \tau^2 [s_{\kappa} X_{\kappa}, t_{\lambda} X_{\lambda}] + \dots \quad (5.68)$$

Nota-se também que $[S(\tau), T(\tau)] = \tau^2 [s_{\kappa} X_{\kappa}, t_{\lambda} X_{\lambda}] + \dots$. Portanto,

$$S^{-1}(\tau) T^{-1}(\tau) S(\tau) T(\tau) = 1 + [S(\tau), T(\tau)] + \dots$$

Em geral, $S^{-1}(\tau) T^{-1}(\tau) S(\tau) T(\tau) x_i \neq x_i$. Para que a transformações operadas por $S(\tau)$ e $T(\tau)$ possam ser realizadas em qualquer ordem para se obter os mesmos valores finais das componentes, isto é, para que $x'''_i = x''_i$, é necessário que

$$[S(\tau), T(\tau)] = 0 \implies [s_{\kappa} X_{\kappa}, t_{\lambda} X_{\lambda}] = 0.$$

Diversas propriedades do grupo de Lie podem ser deduzidas a partir dos geradores infinitesimais e constantes de estrutura:

- Se o grupo \mathcal{G} é Abeliano, de tal forma que todas as transformações contidas no mesmo comutam, então

$$[X_\kappa, X_\lambda] = 0, \text{ para } \kappa, \lambda = 1, \dots, r.$$

Isto implica em que as constantes de estrutura são todas nulas,

$$c_{\kappa\lambda}^\rho = 0, \text{ para } \kappa, \lambda, \rho = 1, \dots, r. \quad (5.69a)$$

- Se \mathcal{G} for um grupo de r parâmetros e $\mathcal{H} \subset \mathcal{G}$ for um subgrupo próprio de p parâmetros ($p < r$), pode-se escolher os p primeiros operadores infinitesimais correspondentes aos elementos de \mathcal{H} , enquanto que os $r - p$ operadores restantes correspondem aos elementos de $\mathcal{G} \setminus \mathcal{H}$ (ou $\mathcal{G} - \mathcal{H}$). Como \mathcal{H} é um grupo *per se*, então os comutadores dos seus geradores devem ser expressos por (5.64a) em termos de seus próprios geradores. Portanto,

$$c_{\kappa\lambda}^\rho = 0, \text{ para } \kappa, \lambda = 1, \dots, p \text{ e } \rho = p + 1, \dots, r. \quad (5.69b)$$

- Se \mathcal{H} for um subgrupo próprio invariante de \mathcal{G} e $S \in \mathcal{H}$ é um operador, então $T^{-1}ST \in \mathcal{H}$, para qualquer operador $T \in \mathcal{G}$. Neste caso, $S^{-1}T^{-1}ST$ também pertence a \mathcal{H} e, de acordo com (5.68), os comutadores $[s_\kappa X_\kappa, t_\lambda X_\lambda]$ devem ser expressíveis como uma combinação linear dos p geradores de \mathcal{H} . Portanto, para um subgrupo invariante,

$$c_{\kappa\lambda}^\rho = 0, \text{ para } \kappa = 1, \dots, p, \lambda = 1, \dots, r \text{ e } \rho = p + 1, \dots, r. \quad (5.69c)$$

- Se o grupo \mathcal{G} é o produto direto de \mathcal{H} por $\mathcal{G} - \mathcal{H}$, conforme definido na seção 3.5, então

$$c_{\kappa\lambda}^\rho = 0, \text{ para } \begin{cases} \kappa = 1, \dots, p, & \rho = p + 1, \dots, r, \\ \kappa = p + 1, \dots, r, & \rho = 1, \dots, p. \end{cases} \quad (5.69d)$$

- O grupo \mathcal{G} será simples³⁷ se a condição (5.69c) não puder ser satisfeita para nenhum subconjunto de geradores infinitesimais do mesmo. Da mesma maneira, \mathcal{G} será semissimples se

$$c_{\kappa\lambda}^\rho = 0, \text{ para } \kappa, \lambda, \rho = 1, \dots, p \quad (5.69e)$$

e (5.69c) não puderem ser satisfeitas simultaneamente.

- O grupo de transformações \mathcal{G} é composto por operadores que agem sobre o espaço \mathcal{V}^n . Este espaço pode ser complexo e, assim, os componentes $\{x_i\}$ também podem ser complexos. Como consequência, os geradores infinitesimais podem ser quantidades complexas. Contudo, como os r parâmetros do grupo são reais, a lei de composição dos elementos de \mathcal{G} (as funções $\phi(a; b)$ definidas em 5.55b) também deve ser real. Isto implica em que as constantes de estrutura, definidas em (5.56c), (5.56d) e (5.61) também devem ser reais.

A definição a seguir, sobre o posto do grupo, também está relacionada com geradores que comutam entre si.

Definição 5.15 (Posto do grupo). O maior número de geradores de um grupo de Lie que comutam entre si é denominado o *posto* (em inglês, *rank*) do grupo.

O posto do grupo $SO(3)$ é igual a 1 porque nenhum dos geradores L_1, L_2 e L_3 comutam entre si. O mesmo ocorre com o grupo $SU(2)$.

5.10.5 ÁLGEBRAS DE LIE

A relação de independência linear entre as funções $u_{ik}(x)$ estabelecida por (5.56b) implica em que os geradores infinitesimais de um grupo de Lie de transformações de r parâmetros, definidos por (5.57), também são linearmente independentes entre si. Isto permite que o conjunto de geradores venha a servir como base de um espaço vetorial \mathcal{L}^r , de dimensão r . Assumindo que vetores desse espaço sejam formados por combinações lineares da base, *i. e.*, por $\sum_\rho a_\rho X_\rho$, onde $a_\rho \in \mathbb{R}$, o espaço vetorial é formado sobre o corpo dos reais.

³⁷Ver definição 3.19.

Sobre este espaço vetorial, a relação de comutação (5.64a) pode ser interpretada como um mapa bilinear sobre o próprio espaço. Pode-se verificar³⁸ que essa relação de comutação satisfaz as condições para a formação de uma *álgebra de Lie* sobre o corpo \mathbb{R} e serve como o “produto vetorial” em \mathcal{L}^r . A álgebra de Lie formada nestas condições é genericamente identificada pelo símbolo \mathfrak{g} .

Tomando como a base de \mathfrak{g} o conjunto de geradores $B_{\mathfrak{g}} \doteq \{X_{\kappa}\}$, uma álgebra real de Lie é formada quando qualquer vetor $A \in \mathfrak{g}$ pode ser decomposto na base $B_{\mathfrak{g}}$ como

$$A = a_{\kappa} X_{\kappa}.$$

Se a decomposição é realizada com o intuito de estudar a estrutura do grupo, os coeficientes $\{a_{\kappa}\}$ da expansão acima devem ser reais, uma vez que as constantes de estrutura são reais. Neste caso a álgebra é real. Contudo, a álgebra também pode ser complexa.

Pode-se realizar também mudanças de base em \mathfrak{g} a partir de uma matriz não singular $A = [a_{\kappa\lambda}]$. Este procedimento cria uma outra base $B'_{\mathfrak{g}} = \{X'_{\kappa}\}$ através das relações

$$X'_{\kappa} = a_{\kappa\lambda} X_{\lambda}, \quad (\kappa = 1, \dots, r). \quad (5.70a)$$

Na nova base $B'_{\mathfrak{g}}$, existe um novo conjunto de constantes de estrutura $c'_{\kappa\lambda\rho}$ determinadas por

$$[X'_{\kappa}, X'_{\lambda}] = c'_{\kappa\lambda\rho} X'_{\rho},$$

as quais se relacionam com as constantes em $B_{\mathfrak{g}}$ por

$$c'_{\kappa\lambda\rho} = a_{\kappa\mu} a_{\lambda\nu} c_{\mu\nu\rho}^{-1}, \quad (\kappa, \lambda, \rho = 1, \dots, r). \quad (5.70b)$$

As constantes $c'_{\kappa\lambda\rho}$ obedecem as mesmas propriedades das constantes originais, *i. e.*, as relações (5.64b) e (5.64c).

Discussões a respeito das propriedades das álgebras de Lie podem, em muitos casos, repetir *ipsis litteris* as propriedades de grupos de Lie apresentadas na página 253: (i) uma álgebra \mathfrak{g} é Abeliãna se (5.69a) for satisfeita; (ii) uma álgebra \mathfrak{h} é uma subálgebra de \mathfrak{g} se suas constantes de estrutura satisfizerem (5.69b); (iii) a subálgebra \mathfrak{h} é invariante se (5.69c) for satisfeita; (iv) a álgebra \mathfrak{g} é a *soma direta*³⁹ da subálgebra \mathfrak{h} com $\mathfrak{g} - \mathfrak{h}$ se (5.69d) for satisfeita; finalmente, (v) a álgebra pode ser simples ou semissimples se as mesmas condições impostas aos grupos forem satisfeitas.

Retoma-se agora a argumentação delineada na página 250: dado um grupo de Lie de transformações conhecido, pode-se obter os seus geradores infinitesimais e constantes de estrutura, o que permite construir a álgebra associada ao grupo de Lie. Sophus Lie mostrou que o procedimento pode ser invertido, uma vez que a toda álgebra de Lie há pelo menos um grupo de transformações de Lie associado por um homomorfismo. Esta é a essência do terceiro teorema de Lie, enunciado a seguir.

Teorema 5.12 (Terceiro teorema de Lie). *A toda álgebra de Lie finita \mathfrak{g} sobre o corpo \mathbb{R} existe um grupo de Lie \mathcal{G} associado por um homomorfismo.*

A inversão do problema é realizada da seguinte maneira: dada uma álgebra de Lie de r dimensões \mathfrak{g} com constantes de estrutura $\{c_{\kappa\lambda\rho}\}$ previamente determinadas e que satisfazem, juntamente com os respectivos geradores, as propriedades (5.64a-c), e um grupo de transformações de Lie de r parâmetros \mathcal{G} , cujos parâmetros $\{a_{\kappa}\}$ são parametrizados por (5.66) e cujos operadores de transformação são assim parametrizados na vizinhança da identidade por (5.67), então, devido ao terceiro teorema de Lie, o grupo \mathcal{G} é determinado localmente por uma base de \mathfrak{g} e suas constantes de estrutura. A partir disso, o grupo \mathcal{G} é construído para qualquer transformação finita por um processo de integração de seus operadores infinitesimais, determinados na vizinhança da identidade pela estrutura da álgebra \mathfrak{g} . A vantagem deste procedimento, quando factível, está no fato de que toda a informação contida em um grupo contínuo, composto por infinitos elementos, é obtida a partir do conhecimento prévio de uma álgebra de dimensão finita, bastando para isso conhecer uma base (com um número finito r de elementos) e suas constantes $\{c_{\kappa\lambda\rho}\}$.

A determinação de todas as possíveis estruturas de uma álgebra de Lie de r dimensões é um problema que se torna quadraticamente mais difícil à medida que r aumenta. Isto porque as

³⁸Ver seção 3.7.3.

³⁹Ver definição 4.4.

equações básicas que determinam as constantes $\{c_{\kappa\lambda}^{\rho}\}$ (as equações 5.64c) são quadráticas nas mesmas, formando assim um sistema algébrico não linear de equações, o qual possui mais do que uma solução possível para $r \geq 2$.

Para $r = 1$, a álgebra é unidimensional e todos os vetores são formados por múltiplos do gerador X . Por isso, todos os comutadores são nulos e, assim, devido a (5.69a), todo grupo de Lie de um parâmetro é Abelian. Esta conclusão já havia sido obtida na discussão realizada na página 242, onde mostrou-se também que todos estes grupos são isomórficos ao grupo de translações unidimensionais.

Para $r = 2$, os elementos da base são $\{X_1, X_2\}$, os quais satisfazem $[X_1, X_2] = aX_1 + bX_2$, com $\{a, b\} \in \mathbb{R}$. Neste caso há duas possibilidades:

1. $a = b = 0$. Neste caso, a álgebra é Abelian e é a soma direta das subálgebras geradas por X_1 e X_2 . Em consequência, o grupo de Lie é Abelian, formado pelo produto direto dos subgrupos invariantes.

O grupo de Lie que é gerado localmente por uma álgebra com $[X_1, X_2] = 0$ é sempre isomórfico ao grupo de translações bidimensionais, onde $x' = x + a$ e $y' = y + b$. Esta conclusão pode ser estendida a outras dimensões. Uma álgebra de r dimensões Abelian, para a qual $[X_i, X_j] = 0, \forall i, j = 1, \dots, r$, gera localmente um grupo isomórfico ao grupo de translações em r dimensões.

2. $a \neq 0$ e/ou $b \neq 0$. Se for assumido que $a \neq 0$, então, de acordo com (5.70a), pode-se fazer uma mudança de base sem alterar a álgebra. Se os elementos da nova base forem escolhidos como $X'_1 = aX_1 + bX_2$ e $X'_2 = (1/a)X_2$, resulta que $[X'_1, X'_2] = X'_1$. Isto significa, de acordo com (5.69c), que a subálgebra gerada por X'_1 é invariante e Abelian e, em consequência, uma álgebra de Lie de 2 dimensões não é semissimples.

Um grupo que é localmente gerado por esta álgebra é o grupo de alongamento/deslocamento unidimensional discutido nos exemplos 5.16 e 5.23, onde $x' = ax + b$. Neste caso, $X_1 = \partial/\partial x$ e $X_2 = x\partial/\partial x$ e as translações $x' = x + b$ formam um subgrupo invariante.

Para $r = 3$ a derivação é mais complicada e não será reproduzida aqui. Uma álgebra de Lie de três dimensões possui quatro estruturas distintas possíveis:

$$[X_1, X_2] = [X_1, X_3] = [X_2, X_3] = 0, \quad (5.71a)$$

$$[X_1, X_2] = X_1, \quad [X_1, X_3] = [X_2, X_3] = 0, \quad (5.71b)$$

$$[X_1, X_2] = X_3, \quad [X_1, X_3] = -X_2, \quad [X_2, X_3] = X_1, \quad (5.71c)$$

$$[X_1, X_2] = X_3, \quad [X_1, X_3] = X_2, \quad [X_2, X_3] = -X_1. \quad (5.71d)$$

Um grupo de Lie que possui a estrutura (5.71b) possui um subgrupo invariante gerado por X_3 (pois este comuta com os outros). Portanto, este grupo não é semissimples. Já o grupo com a estrutura (5.71c) é o grupo $SO(3)$, cujos geradores infinitesimais foram obtidos no exemplo 5.26. Por sua vez, um grupo cuja estrutura é (5.71d) é o grupo de Lorentz homogêneo em duas dimensões, o qual contém operadores que mantêm invariante a forma $x^2 + y^2 - z^2$ e cujos geradores são

$$X_1 = z \frac{\partial}{\partial y} + y \frac{\partial}{\partial z}, \quad X_2 = x \frac{\partial}{\partial z} + z \frac{\partial}{\partial x}, \quad X_3 = y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}.$$

Discussões a respeito das estruturas de álgebras de dimensões $r \geq 4$ podem ser obtidas nos textos relacionados na bibliografia.

5.10.6 REPRESENTAÇÕES DE GRUPOS DE LIE

O formalismo envolvido na obtenção de representações de grupos de Lie é fantásticamente grande e uma abordagem geral sobre o assunto excede em muito o escopo deste texto. No lugar de uma discussão ampla e genérica, serão apresentadas apenas algumas das propriedades e técnicas existentes.

Grande parte dos teoremas que abordam a derivação das representações de um grupo de Lie se baseiam no homomorfismo existente entre o grupo e sua álgebra de Lie associada. Esse homomorfismo, também percebido inicialmente por Sophus Lie, será apresentado agora com os seguintes teoremas e definições.

Definição 5.16 (Homomorfismo entre álgebras de Lie). Se \mathfrak{g} e \mathfrak{h} são álgebras de Lie, então um mapa linear $\Phi : \mathfrak{g} \rightarrow \mathfrak{h}$ é denominado um *homomorfismo entre álgebras de Lie* se, dados $\{X_\alpha, X_\beta\} \in \mathfrak{g}$ e $\{Y_\alpha, Y_\beta\} \in \mathfrak{h}$ tais que $Y_{\alpha,\beta} = \Phi(X_{\alpha,\beta})$, ocorrer

$$\Phi([X_\alpha, X_\beta]) = [Y_\alpha, Y_\beta].$$

Se, adicionalmente, o mapeamento for bijetivo, então Φ é denominado um *isomorfismo entre álgebras de Lie*. O isomorfismo $\Phi : \mathfrak{g} \rightarrow \mathfrak{g}$ é denominado *automorfismo da álgebra de Lie*.

Observa-se que o isomorfismo entre diferentes álgebras de Lie implica em uma forte relação entre suas constantes de estrutura. Se $[X_\alpha, X_\beta] = c_{\alpha\beta}^\gamma X_\gamma$ e $[Y_\alpha, Y_\beta] = d_{\alpha\beta}^\gamma Y_\gamma$, sendo $\{c_{\alpha\beta}^\gamma\}$ e $\{d_{\alpha\beta}^\gamma\}$ respectivamente as constantes de estrutura de \mathfrak{g} e \mathfrak{h} , então a condição de isomorfismo entre essas álgebras implica em,

$$\Phi([X_\alpha, X_\beta]) = \Phi(c_{\alpha\beta}^\gamma X_\gamma) = c_{\alpha\beta}^\gamma \Phi(X_\gamma) = c_{\alpha\beta}^\gamma Y_\gamma = [Y_\alpha, Y_\beta] = d_{\alpha\beta}^\gamma Y_\gamma.$$

Portanto, a condição de isomorfismo implica que $d_{\alpha\beta}^\gamma = c_{\alpha\beta}^\gamma (\forall \alpha, \beta, \gamma)$.

Os seguintes teoremas de Lie estabelecem então as relações entre homomorfismos de álgebras e homomorfismos de grupos.

Teorema 5.10 (Primeiro teorema de Lie). Se \mathcal{G} e \mathcal{H} são grupos de Lie localmente isomórficos, então as suas álgebras associadas \mathfrak{g} e \mathfrak{h} são isomórficas.

Teorema 5.11 (Segundo teorema de Lie). Se \mathfrak{g} e \mathfrak{h} são álgebras de Lie isomórficas, então \mathcal{G} e \mathcal{H} são grupos de Lie localmente isomórficos.

Deve-se notar que estes teoremas somente garantem o isomorfismo entre os grupos nas vizinhanças da identidade. Contudo, um outro teorema garante que se os grupos são localmente isomórficos então globalmente estes serão, no mínimo, homomórficos.

Teorema 5.13. Sejam \mathcal{G} e \mathcal{H} grupos de Lie, com álgebras associadas \mathfrak{g} e \mathfrak{h} , respectivamente. Se $\Phi : \mathfrak{g} \rightarrow \mathfrak{h}$ é um homomorfismo entre os grupos, então há um único mapa linear $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$.

Apresentaremos alguns exemplos a seguir ilustrando algumas representações de grupos de mais baixa dimensão.

Exemplo 5.30 (Representação do $SO(3)$ em termos dos ângulos de Euler). Nos exemplos 5.26 e 5.29 foi mostrado que o operador de rotação em torno de um eixo fixo \hat{n} pode ser escrito como em termos dos geradores do $SO(3)$ como

$$R_{\hat{n}}(\phi) = 1 + \phi(\hat{n} \cdot \mathbf{X}) + \dots = e^{\phi(\hat{n} \cdot \mathbf{X})}.$$

Os operadores $R_{\hat{n}}(\phi)$ pertencem a um subgrupo $SO(2) \subset SO(3)$, correspondente às rotações que ocorrem no plano perpendicular a \hat{n} .

Por sua vez, a representação de um operador de rotação no $SO(2)$ foi obtida no exemplo 5.25, tanto para rotações infinitesimais quanto para rotações finitas.

Se o eixo de rotação na direção e sentido de \hat{n} for sucessivamente tomado como os eixos x_1 , x_2 e x_3 de um sistema de coordenadas Cartesianas, então $\hat{n} = \hat{x}_i$ para $i = 1, 2, 3$, sucessivamente. Assim, se $\mathbf{x} = (x_1 \ x_2 \ x_3)^T$, a rotação em torno do i -ésimo eixo coordenado executará a transformação $\mathbf{x}' = f^{(i)}(x_1, x_2, x_3; \phi)\mathbf{x} = R^{(i)}(\phi)\mathbf{x}$, cuja expressão na vizinhança da identidade é

$$x_j + \delta x_j = (1 + \delta\phi X_i)x_j = x_j + \delta\phi \epsilon_{ijk} x_k \Rightarrow \delta x_j = \delta\phi \epsilon_{ijk} x_k.$$

Portanto, pode-se escrever

$$\mathbf{x}'(\delta\phi) = R^{(i)}(\delta\phi)\mathbf{x} = \left[I_3 + \delta\phi X^{(i)} \right] \mathbf{x}, \text{ sendo } X_{jk}^{(i)} = \epsilon_{ijk}.$$

Observando que

$$\left[X^{(i)} \right]_{jk}^2 = \sum_{\ell=1}^3 \epsilon_{ij\ell} \epsilon_{i\ell k} = \delta_{ij} \delta_{ik} - \delta_{jk}, \quad \left[X^{(i)} \right]_{jk}^3 = \sum_{\ell=1}^3 (\delta_{ij} \delta_{i\ell} - \delta_{j\ell}) \epsilon_{i\ell k} = -\epsilon_{ijk} = - \left[X^{(i)} \right]_{jk}^3,$$

conclui-se que

$$\begin{aligned} [X^{(1)}]^{2k} &= (-1)^k \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & [X^{(2)}]^{2k} &= (-1)^k \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & [X^{(3)}]^{2k} &= (-1)^k \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ [X^{(i)}]^{2k+1} &= (-1)^k X^{(i)}. \end{aligned}$$

Procedendo com o mesmo processo de limite empregado no exemplo 5.25, a matriz $R^{(1)}(\phi)$ resulta

$$R^{(1)}(\phi) = I_3 - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \cos \phi \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sin \phi X^{(1)}.$$

Juntando com os respectivos resultados para $R^{(2)}(\phi)$ e $R^{(3)}(\phi)$, resulta finalmente

$$R^{(1)}(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix}, \quad R^{(2)}(\phi) = \begin{pmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{pmatrix}, \quad R^{(3)}(\phi) = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Os *ângulos de Euler* consistem em uma das infinitas possíveis combinações de três rotações em torno de diferentes eixos fixos de modo a se obter uma rotação genérica no espaço de três dimensões. Estas rotações estão ilustradas na figura 5.5. A primeira rotação ocorre por um ângulo ϕ em torno de x_3 , operada por $R^{(3)}(\phi)$. A segunda rotação ocorre por um ângulo θ em torno de x'_1 , operada por $R^{(1)}(\theta)$. Finalmente, a terceira rotação ocorre por um ângulo ψ em torno de x''_3 , operada por $R^{(3)}(\psi)$.

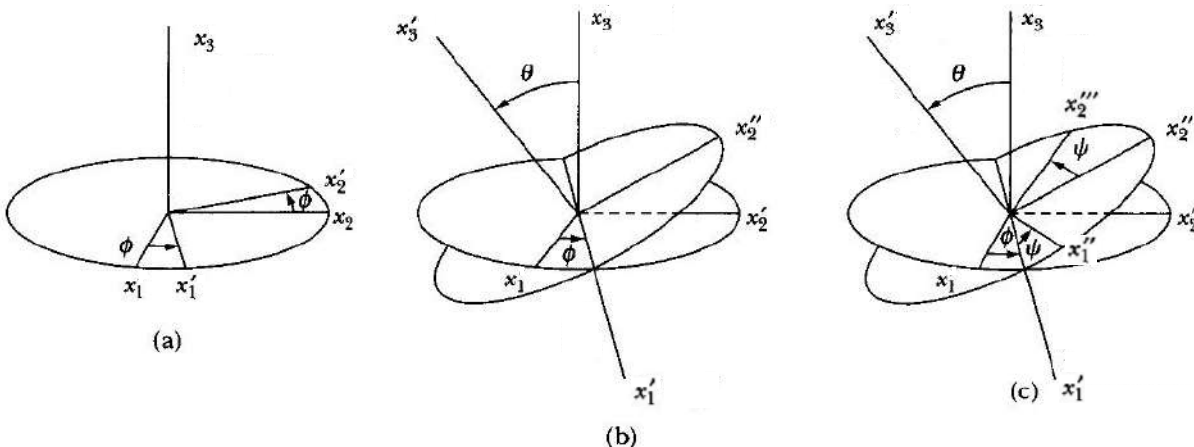


Figura 5.5: Ângulos de Euler usualmente empregados para executar uma rotação arbitrária no sistema de coordenadas. (a) Primeira rotação: sentido anti-horário em torno de x_3 , por um ângulo ϕ . (b) Segunda rotação: sentido anti-horário por um ângulo θ em torno do eixo x'_1 . (c) Terceira rotação: sentido anti-horário por um ângulo ψ em torno de x''_3 .

Portanto, a representação de uma rotação em três dimensões em termos dos ângulos de Euler é realizada pela matriz

$$\begin{aligned} R(\phi, \theta, \psi) &= R^{(3)}(\psi) R^{(1)}(\theta) R^{(3)}(\phi) \\ &= \begin{pmatrix} \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & \sin \theta \sin \psi \\ -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \sin \theta \cos \psi \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{pmatrix}. \end{aligned}$$

A matriz $R(\phi, \theta, \psi)$ recém obtida é uma das possíveis escolhas para os ângulos de Euler. Esta forma para a matriz de rotação no \mathbb{E}^3 é comumente realizada na dinâmica de corpos rígidos. Com esta escolha, o eixo Ox'_1 é denominado *linha dos nodos*.

Uma outra escolha para os ângulos de Euler, comum para representações de grupos, consiste em tomar o eixo Ox'_2 como a linha dos nodos. Com esta escolha, a matriz de rotação fica escrita

$$\bar{R}(\phi, \theta, \psi) = R^{(3)}(\psi) R^{(2)}(\theta) R^{(3)}(\phi)$$

$$= \begin{pmatrix} \cos \psi \cos \theta \cos \phi - \operatorname{sen} \phi \operatorname{sen} \psi & \cos \psi \cos \theta \operatorname{sen} \phi + \operatorname{sen} \psi \cos \phi & -\operatorname{sen} \theta \cos \psi \\ -\operatorname{sen} \psi \cos \theta \cos \phi - \operatorname{sen} \phi \cos \psi & -\operatorname{sen} \psi \cos \theta \operatorname{sen} \phi + \cos \psi \cos \phi & \operatorname{sen} \theta \operatorname{sen} \psi \\ \operatorname{sen} \theta \cos \phi & \operatorname{sen} \theta \operatorname{sen} \phi & \cos \theta \end{pmatrix}.$$

Exemplo 5.31 (Homomorfismo entre os grupos $SU(2)$ e $SO(3)$). Como as constantes de estrutura dos grupos $SO(3)$ (obtidas no exemplo 5.27) e $SU(2)$ (exemplo 5.28) coincidem, estes grupos são localmente isomórficos, segundo os teoremas de Lie. Para verificar a existência de um homomorfismo global entre os mesmos, parte-se novamente do eixo de rotação orientado na direção e sentidos do vetor \hat{n} considerado nos exemplos 5.29 e 5.30. Empregando um sistema de coordenadas polares esféricas,⁴⁰ escreve-se $\hat{n} = \operatorname{sen} \theta \cos \varphi \hat{x}_1 + \operatorname{sen} \theta \operatorname{sen} \varphi \hat{x}_2 + \cos \theta \hat{x}_3$ e, dadas as matrizes de Pauli obtidas no exemplo 5.28, escreve-se

$$\hat{n} \cdot \sigma = \begin{pmatrix} \cos \theta & \operatorname{sen} \theta e^{-i\varphi} \\ \operatorname{sen} \theta e^{i\varphi} & -\cos \theta \end{pmatrix}.$$

Esta quantidade gera uma transformação parametrizada $e^{i\hat{n} \cdot \sigma / 2}$ sobre os vetores de um espaço de Hilbert. Mudanças de base adequadas (i. e., “rotações”) envolvem então transformações de similaridade do tipo

$$\hat{n} \cdot \sigma \longrightarrow (\hat{n} \cdot \sigma)' = U^\dagger (\hat{n} \cdot \sigma) U,$$

sendo $U \in SU(2)$, dada, em princípio, pela forma generalizada obtida no exemplo 5.28.

Tomando primeiro a matriz U com $\xi = \alpha/2$ e $\eta = \zeta = 0$, temos

$$U = U_3(\alpha) = \begin{pmatrix} e^{i\alpha/2} & 0 \\ 0 & e^{-i\alpha/2} \end{pmatrix} \implies (\hat{n} \cdot \sigma)' = \begin{pmatrix} \cos \theta & \operatorname{sen} \theta e^{-i(\varphi+\alpha)} \\ \operatorname{sen} \theta e^{i(\varphi+\alpha)} & -\cos \theta \end{pmatrix}.$$

Portanto, A ação de $U_3(\alpha)$ sobre $\hat{n} \cdot \sigma$ equivale a uma rotação do vetor \hat{n} por um ângulo α em torno do eixo x_3 , pois se $\hat{n} \xrightarrow{R^{(3)}(\alpha)} \hat{n}'$, então

$$(\hat{n} \cdot \sigma)' = U_3^\dagger (\hat{n} \cdot \sigma) U_3 = \hat{n}' \cdot \sigma.$$

Por esta razão, identifica-se

$$U_3(\alpha) \mapsto R^{(3)}(\alpha),$$

sendo esta última a matriz de rotação em torno de x_3 obtida no exemplo 5.30.

Tomando agora U com $\xi = \zeta = 0$ e $\eta = \beta/2$,

$$U = U_2(\beta) = \begin{pmatrix} \cos(\beta/2) & -\operatorname{sen}(\beta/2) \\ \operatorname{sen}(\beta/2) & \cos(\beta/2) \end{pmatrix} \implies$$

$$(\hat{n} \cdot \sigma)' = \begin{pmatrix} \cos \theta \cos \beta + \operatorname{sen} \theta \cos \varphi \operatorname{sen} \beta & \operatorname{sen} \theta \cos \varphi \cos \beta - \cos \theta \operatorname{sen} \beta - i \operatorname{sen} \theta \operatorname{sen} \varphi \\ \operatorname{sen} \theta \cos \varphi \cos \beta - \cos \theta \operatorname{sen} \beta + i \operatorname{sen} \varphi \operatorname{sen} \theta & -\cos \theta \cos \beta - \operatorname{sen} \theta \cos \varphi \operatorname{sen} \beta \end{pmatrix},$$

o qual é o resultado obtido após executada uma rotação de \hat{n} em torno do eixo x_2 por um ângulo β . Por isso, identifica-se

$$(\hat{n} \cdot \sigma)' = U_2^\dagger (\hat{n} \cdot \sigma) U_2 = \hat{n}' \cdot \sigma \implies U_2(\beta) \mapsto R^{(2)}(\beta).$$

Desta maneira a matriz $U(\phi, \theta, \psi) \in SU(2)$ que corresponde à matriz de rotação em termos dos ângulos de Euler $\bar{R}(\phi, \theta, \psi) \in SO(3)$ é dada por

$$U^{(1/2)}(\phi, \theta, \psi) = U_3(\psi) U_2(\theta) U_3(\phi) = \begin{pmatrix} \cos(\theta/2) e^{i(\alpha+\psi)/2} & -\operatorname{sen}(\theta/2) e^{-i(\alpha-\psi)/2} \\ \operatorname{sen}(\theta/2) e^{i(\alpha-\psi)/2} & \cos(\theta/2) e^{-i(\alpha+\psi)/2} \end{pmatrix},$$

a qual é uma representação do $SU(2)$ em termos dos ângulos de Euler.

Além disso, é fácil mostrar que o mapeamento $SU(2) \mapsto SO(3)$ recém delineado corresponde a um homomorfismo, pois se as matrizes $A, B \in SU(2)$ são mapeadas nas matrizes de rotação $R_A, R_B \in SO(3)$ por

$$\hat{n} \xrightarrow{R_A} \hat{n}' \implies A^\dagger (\hat{n} \cdot \sigma) A = \hat{n}' \cdot \sigma, \quad \hat{n} \xrightarrow{R_B} \hat{n}'' \implies B^\dagger (\hat{n} \cdot \sigma) B = \hat{n}'' \cdot \sigma,$$

⁴⁰Seção 1.5.2.

então

$$(AB)^\dagger (\hat{n} \cdot \sigma) (AB) = B^\dagger [A^\dagger (\hat{n} \cdot \sigma) A] B = B^\dagger (\hat{n}' \cdot \sigma) B = \hat{n}''' \cdot \sigma, \text{ onde } \hat{n}''' \xleftarrow{R_B R_A} \hat{n}.$$

Observa-se também que o mapeamento não é isomórfico, pois

$$\left. \begin{aligned} U_3(\phi + 2\pi) &= -U_3(\phi) \\ U_2(\theta + 2\pi) &= -U_2(\theta) \end{aligned} \right\}, \text{ enquanto que } \begin{cases} R^{(3)}(\phi + 2\pi) = R^{(3)}(\phi) \\ R^{(2)}(\theta + 2\pi) = R^{(2)}(\theta). \end{cases}$$

Ou seja, dois elementos do $SU(2)$ são mapeados no mesmo elemento do $SO(3)$; em outras palavras, há um homomorfismo 2-para-1 do $SU(2)$ para o $SO(3)$. O $SU(2)$ é denominado o *grupo de cobertura* do $SO(3)$.

As representações obtidas nos exemplos anteriores não são as únicas possível e, em geral, são redutíveis. Para se obter representações irredutíveis do $SU(2)$, o procedimento consagrado é discutido a seguir.

5.10.7 REPRESENTAÇÕES IRREDUTÍVEIS DO $SU(2)$

Sejam u e v variáveis complexas as quais podem ser vetores em um espaço complexo de dimensão 2 ou componentes da uma função de onda de spin (o *espinor*) de uma partícula de spin $1/2$ como o elétron, por exemplo. Sob uma transformação do $U \in SU(2)$, o vetor $(u \ v)^T$ é transformado como

$$\begin{pmatrix} u \\ v \end{pmatrix} \longrightarrow \begin{pmatrix} u' \\ v' \end{pmatrix} = U \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \implies \begin{cases} u' = \alpha u + \beta v \\ v' = -\beta^* u + \alpha^* v, \end{cases} \quad (5.72)$$

onde foi empregada a forma genérica para U obtida no exemplo 5.28. Nota-se que as quantidades transformadas são combinações lineares das quantidades originais e, além disso, $|u'|^2 + |v'|^2 = |u|^2 + |v|^2$, uma vez que $|\alpha|^2 + |\beta|^2 = 1$.

Em se tratando de sistemas de partículas idênticas, termos contendo produtos e potências das quantidades u e v se tornam presentes. Neste caso, o espaço vetorial que contém o vetor de estado completo do sistema possui uma dimensão maior que o espaço de uma única partícula. O procedimento padrão adotado neste caso consiste em partir dos produtos u^{2j} , $u^{2j-1}v$, $u^{2j-2}v^2$, ..., u^2v^{2j-2} , uv^{2j-1} , v^{2j} e definir um conjunto de *produtos simétricos* $\{f_j^m\}$, sendo

$$f_j^m = \frac{u^{j+m}v^{j-m}}{\sqrt{(j+m)!(j-m)!}}, \text{ com } m = -j, -j+1, \dots, j-1, j,$$

os quais formam uma base de um espaço de $n = 2j + 1$ dimensões. Nota-se também que $j = 1/2, 1, 3/2, \dots$. Por exemplo, se $j = 1/2$, então $n = 1$, $f_{1/2}^{1/2} = u$ e $f_{1/2}^{-1/2} = v$. Já se $j = 3$, então $n = 3$, $f_3^3 = u^2/\sqrt{2}$, $f_3^0 = uv$ e $f_3^{-1} = v^2/\sqrt{2}$.

Como as quantidades u e v em f_j^m transformam-se como (5.72) frente a $U \in SU(2)$, a ação de U sobre os f_j^m irá corresponder a uma simples mudança de base no mesmo espaço, a qual irá fornecer a representação desejada. Aplicando então o operador $U(\alpha, \beta)$, correspondente à matriz U , sobre f_j^m , resulta

$$U(\alpha, \beta) f_j^m = \frac{(u')^{j+m} (v')^{j-m}}{\sqrt{(j+m)!(j-m)!}} = \frac{(\alpha u + \beta v)^{j+m} (-\beta^* u + \alpha^* v)^{j-m}}{\sqrt{(j+m)!(j-m)!}}.$$

De acordo com o teorema binomial,

$$U(\alpha, \beta) f_j^m = \sum_{k=0}^{j+m} \sum_{\ell=0}^{j-m} \frac{\sqrt{(j+m)!(j-m)!}}{k!\ell!(j+m-k)!(j-m-\ell)!} \alpha^{j+m-k} (\alpha^*)^\ell \beta^k (-\beta^*)^{j-m-\ell} u^{2j-k-\ell} v^{k+\ell}.$$

Definindo a nova variável $m' = j - k - \ell$, observa-se que $2j - k - \ell = j + m'$ e $k + \ell = j - m'$ e troca-se $\sum_{k=0}^{j+m} \sum_{\ell=0}^{j-m} \rightarrow \sum_{k=0}^{j+m} \sum_{m'=-j}$, resultando

$$U(\alpha, \beta) f_j^m = \sum_{m'=-j}^j \sum_{k=0}^{j+m} \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{k!(j-m'-k)!(j+m-k)!(k+m'-m)!} \times \alpha^{j+m-k} (\alpha^*)^{j-m'-k} \beta^k (-\beta^*)^{k+m'-m} \frac{u^{j+m'} v^{j-m'}}{\sqrt{(j+m')!(j-m')!}},$$

o qual é identificado como

$$U(\alpha, \beta) f_j^m = \sum_{m'=-j}^j \Gamma_{m,m'}^{(j)}(\alpha, \beta) f_j^{m'}, \tag{5.73a}$$

$$\Gamma_{m,m'}^{(j)}(\alpha, \beta) = \sum_{k=0}^{j+m} \frac{\sqrt{(j+m)!(j-m)!(j+m')!(j-m')!}}{k!(j-m'-k)!(j+m-k)!(k+m'-m)!} \alpha^{j+m-k} (\alpha^*)^{j-m'-k} \beta^k (-\beta^*)^{k+m'-m}. \tag{5.73b}$$

Observando que:

1. O conjunto $\{f_j^m\}$ é linearmente independente, formando assim uma base de um espaço de Hilbert de dimensão $2j + 1$.

2. Como

$$\sum_{m=-j}^j |f_j^m|^2 = \sum_{m=-j}^j \frac{|u^{j+m} v^{j-m}|^2}{(j+m)!(j-m)!} = \frac{(|u|^2 + |v|^2)^{2j}}{(2j)!},$$

a norma $\sum_{m=-j}^j |f_j^m|^2$ é invariante frente a uma transformação unitária,

as matrizes $\Gamma_{m,m'}^{(j)}(\alpha, \beta)$ formam uma representação unitária de dimensão $2j+1$ do $SU(2)$. Pode-se mostrar também que as representações assim obtidas são irredutíveis.

O número de distintas irreps do $SU(2)$ que podem ser obtidas com este método é extremamente alto, uma vez que as matrizes $\Gamma_{m,m'}^{(j)}(\alpha, \beta)$ dependem dos valores de α e β . Um caso especial é $\alpha = e^{i\phi/2}$ e $\beta = 0$, para o qual resulta

$$\Gamma_{m,m'}^{(j)}(e^{i\phi/2}, 0) = e^{im\phi} \delta_{m,m'}.$$

Representações irredutíveis do $SO(3)$ podem ser obtidas a partir de (5.73), empregando o homomorfismo entre os grupos discutido no exemplo 5.31.

5.10.8 O TEOREMA DE CASIMIR

Um outro resultado importante para as representações dos grupos de Lie é o teorema de Casimir. Dada uma álgebra de Lie de dimensão r , sendo $\{X_\kappa\}$ uma base da mesma. O elemento quadrático de Casimir ou simplesmente o elemento de Casimir é o operador

$$C = \sum_{\kappa=1}^r X_\kappa^2. \tag{5.74}$$

Será mostrado agora que o operador C comuta com todos os geradores do grupo se a álgebra for semissimples. Para qualquer gerador X_λ na álgebra, tal que a relação de comutação deste com X_κ seja $[X_\lambda, X_\kappa] = c_{\lambda\kappa}^\rho X_\rho$,

$$\begin{aligned} [X_\lambda, C] &= \sum_{\kappa=1}^r [X_\lambda, X_\kappa^2] \\ &= [X_\lambda, X_\kappa] X_\kappa + X_\kappa [X_\lambda, X_\kappa] \\ &= c_{\lambda\kappa}^\rho X_\rho X_\kappa + c_{\lambda\kappa}^\rho X_\kappa X_\rho. \end{aligned}$$

No segundo termo do lado direito do último resultado, trocam-se os índices mudos $\kappa \leftrightarrow \rho$, resultando então

$$[X_\lambda, C] = c_{\lambda\kappa}^\rho X_\rho X_\kappa + c_{\lambda\rho}^\kappa X_\rho X_\kappa.$$

Mas, pode-se mostrar que se a álgebra for semissimples, então as constantes de estrutura são totalmente antissimétricas: $c_{\lambda\kappa}^{\rho} = -c_{\lambda\rho}^{\kappa}$. Portanto,

$$[X_{\lambda}, C] = 0.$$

Para o $SO(3)$, o elemento de casimir é, simplesmente,

$$C = L_1^2 + L_2^2 + L_3^2,$$

isto é, o operador momento angular total.

Dependendo do grupo, existem outras maneiras de se construir um operador de Casimir, a partir de combinações dos geradores, que também comutam os esses geradores. Qualquer operador $C_i = C_i(\{X_{\kappa}\})$ que satisfaça

$$[C_i, X_{\kappa}] = 0, \quad (\kappa = 1, \dots, r)$$

será um múltiplo da identidade. Uma maneira de construir um operador de Casimir distinto de (5.74) é o operador cúbico

$$C^{(3)} = c_{\alpha_1\beta_1}^{\gamma_1} c_{\alpha_2\beta_2}^{\gamma_2} c_{\alpha_3\beta_3}^{\gamma_3} X_{\alpha_1} X_{\alpha_2} X_{\alpha_3}.$$

De acordo com um teorema do físico-matemático Giulio Racah (1909–1965), o número de operadores de Casimir independentes de um dado grupo de Lie é igual ao seu posto.⁴¹ Portanto, os grupos $SO(3)$ e $SU(2)$ somente possuem os elementos quadráticos (5.74) como operadores de Casimir.

REFERÊNCIAS

- BARATA, João Carlos Alves. **Notas para um Curso de Física-Matemática**. São Paulo: USP. Versão de 28/02/2018. Disponível em: <http://denebola.if.usp.br/~jbarata/Notas_de_aula/>.
- BASSALO, José Maria F.; CATTANI, Mauro Sérgio D. **Teoria de Grupos**. Segunda edição. São Paulo: Livraria da Física, 2008. 286 + ix pp. ISBN 9788578610005.
- CHEN, J. Q.; PING, J.; WANG, F. **Group Representation Theory for Physicists**. Singapore: World Scientific, 2002. ISBN 9789813106000. Disponível em: <<https://books.google.com.br/books?id=yWhIDQAAQBAJ>>.
- FAZZIO, A.; WATARI, K. **Introdução à Teoria de Grupos: com Aplicações em Moléculas e Sólidos**. Santa Maria: UFSM, 1998. 239 pp. ISBN 85-7391-010-0.
- GOLDSTEIN, H.; POOLE, C.; SAFKO, J. **Classical Mechanics**. Third Ed. New York: Addison Wesley, 2001. 638 + xi pp. ISBN 978-0201657029.
- HALL, B. **Lie Groups, Lie Algebras, and Representations: An Elementary Introduction**. Heidelberg: Springer, 2015. (Graduate Texts in Mathematics). ISBN 9783319134673. DOI: 10.1007/978-3-319-13467-3.
- HAMERMESH, Morton. **Group Theory and Its Application to Physical Problems**. New York: Dover Publications, 2012. (Dover Books on Physics). ISBN 9780486140391. Disponível em: <<http://books.google.com.br/books?id=w6PDAgAAQBAJ>>.
- JACKSON, J. D. **Classical Electrodynamics**. Third. New York: John Wiley & Sons, 1999. 808 + xxi pp. ISBN 0-471-30932-X.
- JOSHI, A. W. **Elements of Group Theory for Physicists**. Third Ed. New Delhi: John Wiley & Sons, 1982. 348 + xv pp. ISBN 0-470-27306-2.
- MESSIAH, Albert. **Quantum Mechanics: Two Volumes Bound as One**. Mineola: Dover, 1999. 1136 + xxii pp. ISBN 0-486-40924-4.
- RILEY, K. F.; HOBSON, M. P.; BENCE, S. J. **Mathematical Methods for Physics and Engineering**. Third Ed. Cambridge: Cambridge University Press, 2006. 1333 + xxvii pp. ISBN 9781139450997. Disponível em: <<http://books.google.com.br/books?id=Mq1nLEKhNcsC>>.

⁴¹Ver definição 5.15.

- SAKURAI, J. J.; NAPOLITANO, J. **Modern Quantum Mechanics**. San Francisco: Addison-Wesley, 2011. 550 + xviii pp. ISBN 9780805382914. Disponível em: <<https://books.google.com.br/books?id=N4I-AQAACAAJ>>.
- SCHIFF, Leonard I. **Quantum mechanics**. Second. New York: McGraw-Hill, 1955. (International series in pure and applied physics). Disponível em: <<https://books.google.com.br/books?id=XujUaYBU5oEC>>.
- SERRE, J. P. **Lie Algebras and Lie Groups: 1964 Lectures given at Harvard University**. Berlin: Springer, 2009. (Lecture Notes in Mathematics). ISBN 9783540706342. Disponível em: <<https://books.google.com.br/books?id=ChJqCQAAQBAJ>>.
- SZEKERES, Peter. **A Course in Modern Mathematical Physics: Groups, Hilbert Space, and Differential Geometry**. Cambridge: Cambridge University Press, 2004. ISBN 9780521536455. Disponível em: <<http://books.google.com.br/books?id=pRgjkgEACAAJ>>.
- TINKHAM, M. **Group Theory and Quantum Mechanics**. [S.l.]: Dover Publications, 2012. (Dover Books on Chemistry). 340 + xii pp. ISBN 9780486131665. Disponível em: <<https://books.google.com.br/books?id=oe7DAgAAQBAJ>>.
- YOSIDA, Kôsaku. **Functional Analysis**. Sixth. Berlin: Springer-Verlag, 1980. (Grundlehren der mathematischen Wissenschaften). ISBN 9783662117910. Disponível em: <<https://books.google.com.br/books?id=xWrvCAAQBAJ>>.

ÁLGEBRA E ANÁLISE TENSORIAIS

A LO LONGO DO PROCESSO HISTÓRICO de desenvolvimento das ciências da natureza, observou-se com frequência a necessidade da definição e do uso de estruturas matemáticas com graus crescentes de generalidade e abstração. Um exemplo disso é a evolução da mecânica newtoniana.

Conforme é a prática usual em cursos contemporâneos de física básica, a mecânica é inicialmente apresentada e aplicada a sistemas físicos compostos por um número pequeno de partículas que interagem entre si através de forças conservativas, tais como as interações gravitacional e eletrostática. Para tais sistemas simples, o arcabouço matemático usualmente necessário limita-se ao cálculo infinitesimal e aos conceitos básicos de espaços vetoriais, nos quais as definições abstratas de vetores, produto interno e operadores lineares são apresentadas. Na mecânica newtoniana, estas entidades matemáticas abstratas (os vetores) são então identificadas com grandezas físicas mensuráveis tais como posição, velocidade, aceleração e força.

Contudo, quando o formalismo da mecânica foi aplicado ao estudo de sistemas com um grau de complexidade física maior, tais como sistemas com vínculos, meios contínuos e transformações de referenciais, e também com a descoberta e pesquisa de outros tipos de interações e objetos que levaram à criação de novas teorias físicas, tais como o eletromagnetismo, a mecânica quântica e a teoria da relatividade, constatou-se a necessidade do uso de estruturas matemáticas mais abrangentes do que aquelas oferecidas pelas definições de um espaço vetorial, conforme usualmente apresentadas em disciplinas básicas de álgebra linear. Uma classe destas novas estruturas (ou objetos) matemáticos são os *tensores*.

De uma forma simplista, pode-se caracterizar um tensor como um conjunto de objetos matemáticos (em geral) ou físicos (em particular) que estão relacionados entre si e que determinam algum tipo de relação entre duas outras entidades matemáticas (ou físicas). Estes diferentes objetos podem ser identificados de forma prática com o uso de um número finito de índices e possuem leis de transformação bem definidas quando sua representação é alterada de um dado sistema de coordenadas ou referencial para outros.

Neste capítulo, será realizada uma introdução à definição, à álgebra e à análise de tensores.

6.1 INTRODUÇÃO E DEFINIÇÕES

Partindo das definições de espaços vetoriais e operadores lineares,¹ neste capítulo será realizada uma investigação mais aprofundada das propriedades de tensores, operadores e outros objetos relacionados, os quais existem em espaços vetoriais específicos.

O termo *tensor* foi utilizado pela primeira vez em 1846 por William Rowan Hamilton (1805 – 1865), matemático bastante conhecido também por suas contribuições para a mecânica newtoniana. Porém, este termo foi utilizado em um contexto distinto daquele empregado atualmente. Historicamente, as primeiras concepções do que viria a posteriori ser denominado de análise tensorial foram introduzidas por Carl Friedrich Gauss (1777 – 1855), em seu pioneiro trabalho a respeito de geometria diferencial. Estes conceitos foram então desenvolvidos ao longo do século XIX, sendo que o termo *tensor*, neste contexto, foi introduzido em 1898 por Woldemar Voigt (1850 – 1919).

O cálculo tensorial foi desenvolvido em definitivo a partir de 1890 com os trabalhos de Woldemar Voigt, Elwin Bruno Christoffel (1829 – 1900), Gregorio Ricci-Curbastro (1853 – 1925) e

¹Discutidos na seção 4.4.

Tullio Levi-Civita (1873 – 1941), com a designação inicial *cálculo diferencial absoluto*. Já no século XX, esta área da álgebra e análise matemáticas passou a ser denominada em definitivo como *análise tensorial*, tendo a sua popularização muito a dever com a proposta, em 1915, da teoria da relatividade geral formulada por Albert Einstein (1879 – 1955). Durante a formulação da teoria da relatividade geral, inteiramente baseada no conceito de tensores, Einstein contou com a colaboração intensa tanto de Levi-Civita quanto do matemático Marcel Grossmann (1878 – 1936).

O desenvolvimento da análise tensorial também está ligada ao estudo da mecânica de meios contínuos, sendo o *tensor de stress*, o qual determina as tensões internas que surgem no meio quando este é submetido a esforços aplicados em diferentes direções do espaço, uma das quantidades físicas que surgiram desde os primeiros estudos nesta área da física e engenharia.

De uma forma genérica, tensores são objetos geométricos que descrevem relações lineares entre escalares, vetores e outros tensores. Na formulação da álgebra e análise tensoriais, as quantidades físicas identificadas usualmente como escalares e vetores são elas próprias casos particulares de tensores. Porém, esta formulação realiza a extensão lógica destes conceitos, permitindo o tratamento de estruturas matemáticas mais abstratas e complexas. Esta extensão possibilitou o desenvolvimento posterior das áreas da física mencionadas acima (entre outras).

Como já mencionado, tensores são importantes em muitas áreas da física, as quais estudam diferentes tipos de sistemas, tais como sólidos, gases ionizados, teoria eletromagnética, relatividade restrita e geral e mecânica quântica. Um meio contínuo inhomogêneo e/ou anisotrópico é um exemplo típico de sistema físico onde o conceito e a necessidade do uso de tensores ocorrem de forma natural.

Um exemplo inicial pode ser mencionado a partir do problema do fluxo de corrente elétrica em um meio anisotrópico. A Lei de Ohm é apresentada em textos de física básica como

$$\mathbf{J} = \sigma \mathbf{E}, \quad (6.1)$$

sendo as quantidades vetoriais \mathbf{J} e \mathbf{E} respectivamente a densidade de corrente elétrica e o campo elétrico. Em um meio homogêneo e isotrópico, a relação entre estes campos é determinada pela quantidade escalar σ , denominada a *condutividade* do meio.

Se o meio for anisotrópico, por outro lado, a relação empírica (6.1) não é válida em geral, pois o agente que gera a anisotropia do meio determina a existência de pelo menos uma direção preferencial no espaço. Então, verifica-se empiricamente que, como consequência, a densidade de corrente resulta distinta caso o campo \mathbf{E} esteja orientado na direção da anisotropia ou contido sobre o plano perpendicular a esta direção. Neste caso, a expressão correta para o i -ésimo componente do vetor \mathbf{J} é

$$J_i = \sum_{j=1}^3 \sigma_{ij} E_j, \quad (i = 1, 2, 3), \quad (6.2a)$$

sendo que o conjunto de 9 valores $\{\sigma_{11}, \sigma_{12}, \dots, \sigma_{32}, \sigma_{33}\}$ pode ser expresso na forma matricial. Assim, a forma generalizada (6.2a) para a Lei de Ohm pode ser expressa também como uma multiplicação matricial,

$$\begin{pmatrix} J_1 \\ J_2 \\ J_3 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix}. \quad (6.2b)$$

Os 9 elementos da matriz de condutividade $\{\sigma_{ij}\}$ ($i, j = 1, \dots, 3$) estão relacionados entre si tanto do ponto de vista do seu significado físico, sendo os valores da condutividade do meio em função da orientação relativa entre os campos \mathbf{E} e \mathbf{J} , quanto do ponto de vista matemático, uma vez que a matriz de condutividade obedece as regras da álgebra de matrizes.

Porém, verifica-se, também de forma empírica, que os elementos da matriz $\{\sigma_{ij}\}$ obedecem a um conjunto de regras mais amplo que a simples álgebra matricial. Estas regras determinam, por exemplo, como as quantidades físicas em (6.2) devem se alterar quando é realizada uma mudança no sistema de referências do laboratório, ou quando se muda o sistema de coordenadas em um dado referencial. Observa-se então que os elementos do conjunto $\{\sigma_{ij}\}$ devem se transformar de forma a satisfazer a definição de um *tensor*, sendo que estas regras de transformação também são verificadas empiricamente. Portanto, a denominação mais correta para o conjunto $\{\sigma_{ij}\}$ é *tensor de condutividade* do meio anisotrópico, o qual pode ser expresso na forma matricial. Usualmente, empregando as propriedades microscópicas do meio ou fazendo uso de propriedades de simetria, é possível se mostrar que diferentes componentes do tensor de condutividade estão relacionadas entre si.

Generalizando e expandindo o argumento acima, espera-se que a descrição quantitativa dos processos físicos, ou seja, as leis físicas que descrevem a evolução espaço-temporal das quantidades físicas mensuráveis, não dependam do sistema de coordenadas empregado em um dado referencial, nem da transformação de um referencial (inercial) a outro. A aplicação deste princípio às leis físicas, dentre as quais a Lei de Ohm (6.1) ou (6.2) é apenas um exemplo, irá determinar a natureza e a classificação das quantidades matemáticas envolvidas na descrição dessas leis.

Antes, porém, de se entrar em maiores detalhes neste ponto, serão introduzidos tanto a notação básica a ser empregada ao longo deste capítulo, quanto dois símbolos tensoriais empregados amiúde em todos os textos de física-matemática: os símbolos da delta de Kronecker e de Levi-Civita.

6.1.1 CONVENÇÃO DE SOMA DE ÍNDICES E SÍMBOLOS AUXILIARES

Para evitar o acúmulo de símbolos de soma nas expressões apresentadas, será utilizada a *convenção de somas implícitas*, usualmente atribuída a Einstein.

Qualquer índice *minúsculo* que apareça *exatamente duas vezes* em quaisquer termos de uma expressão algébrica é assumido implicitamente como sendo somado sobre todos os valores possíveis que aquele índice possa assumir. Por outro lado, o caso particular em que esta convenção não deve ser empregada deve ser considerado como uma exceção, sendo feita então uma observação explícita deste caso. Usualmente, inclui-se o termo “NS” (não somado) nas proximidades da expressão.

Alguns exemplos desta convenção, válidos para expressões no espaço vetorial \mathcal{R}^3 definido no exemplo 4.3, e empregando um sistema de coordenadas Cartesiano, são os seguintes:

- Produto escalar dos vetores a e b :

$$a_i b_i \equiv \sum_{i=1}^3 a_i b_i = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

- Produto matricial entre as matrizes A e B :

$$(AB)_{ik} = A_{ij} B_{jk} \equiv \sum_{j=1}^3 A_{ij} B_{jk} = A_{i1} B_{1k} + A_{i2} B_{2k} + A_{i3} B_{3k}, \quad (i, k = 1, 2, 3).$$

- Produto misto entre matrizes e vetores:

$$a_{ij} b_{jk} c_k \equiv \sum_{j=1}^3 \sum_{k=1}^3 a_{ij} b_{jk} c_k = a_{i1} b_{11} c_1 + a_{i1} b_{12} c_2 + a_{i1} b_{13} c_3 + a_{i2} b_{21} c_1 + \cdots + a_{i3} b_{32} c_2 + a_{i3} b_{33} c_3.$$

- Divergente do campo vetorial $v(\mathbf{r})$:

$$\frac{\partial v_i}{\partial x_i} \equiv \sum_{i=1}^3 \frac{\partial v_i}{\partial x_i} = \nabla \cdot \mathbf{v}.$$

- Laplaciano do campo escalar $\phi(\mathbf{r})$:

$$\frac{\partial^2 \phi}{\partial x_i \partial x_i} \equiv \sum_{i=1}^3 \frac{\partial^2 \phi}{\partial x_i^2} = \nabla^2 \phi.$$

Algumas observações adicionais a respeito desta convenção devem ser realizadas.

- Índices repetidos são denominados *índices mudos*, enquanto que os demais são chamados *índices livres*. Uma característica de índices (ou de variáveis) mudos está no fato de que o resultado final não mais depende dos mesmos; portanto, o resultado independe do caractere alfabético realmente empregado para representar o índice mudo. Ou seja,

$$a_i b_i = a_j b_j \text{ ou } a_{ij} b_{jk} c_k = a_{ik} b_{kj} c_j,$$

onde no último exemplo acima realizou-se a troca $j \leftrightarrow k$, o que não afeta o resultado final da operação algébrica.

- Quando há índices mudos e livres em uma expressão, não é permitida a permuta de caracteres entre os mesmos. Ou seja, em geral

$$a_{ij}b_{jk}c_k \neq a_{ji}b_{ik}c_k.$$

- Ao se introduzir novos índices mudos em uma expressão, deve-se tomar cuidado para não repetir índices já presentes, quer sejam estes mudos ou livres. Ou seja, a troca

$$a_{ij}b_{jk}c_k \longrightarrow a_{ij}b_{jj}c_j$$

não é permitida, pois gera uma ambiguidade na maneira como as somas implícitas devem ser realizadas.

6.1.2 SÍMBOLOS AUXILIARES: KRONECKER E LEVI-CIVITA

Dois símbolos tensoriais muito úteis para as expressões empregadas neste capítulo serão introduzidos agora.

Delta de Kronecker. Trata-se da quantidade δ_{ij} , definida por

$$\delta_{ij} = \begin{cases} 1, & \text{se } i = j \\ 0, & \text{se } i \neq j. \end{cases}$$

Símbolo (ou tensor) de Levi-Civita. Trata-se da quantidade ϵ_{ijk} , com três índices, a qual estabelece uma relação totalmente antissimétrica entre os índices. A sua definição é

$$\epsilon_{ijk} = \begin{cases} +1, & \text{se } \{i, j, k\} \text{ é uma permutação par de } \{1, 2, 3\} \\ -1, & \text{se } \{i, j, k\} \text{ é uma permutação ímpar de } \{1, 2, 3\} \\ 0, & \text{se dois ou mais índices são repetidos} \end{cases}$$

$$= \frac{1}{2} [(j-i)(k-i)(k-j)].$$

Alguns exemplos: $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1$, $\epsilon_{213} = \epsilon_{132} = \epsilon_{321} = -1$ e $\epsilon_{113} = \epsilon_{122} = \epsilon_{111} = 0$.

O símbolo ϵ_{ijk} possui ao todo $3^3 = 27$ valores que podem ser organizados em uma matriz $3 \times 3 \times 3$. Se o símbolo for escrito $\epsilon_{i_1 i_2 i_3}$ ($1 \leq i_j \leq 3$, $1 \leq j \leq 3$), este pode também ser calculado como

$$\epsilon_{i_1 i_2 i_3} = \prod_{1 \leq j < k \leq 3} \text{sgn}(i_k - i_j) = \text{sgn}(i_2 - i_1) \text{sgn}(i_3 - i_1) \text{sgn}(i_3 - i_2),$$

sendo $\text{sgn}(x)$ a função sinal de x .

Algumas propriedades matemáticas importantes que fazem uso dos símbolos acima são as seguintes:

$$a_i = \delta_{ij} a_j \quad (6.3a)$$

$$a_{ij} \delta_{jk} = a_{ij} \delta_{kj} = a_{ik} \quad (6.3b)$$

$$a_{ij} b_{ji} = a_{ij} b_{jk} \delta_{kl} \quad (6.3c)$$

$$\delta_{ij} \delta_{ik} = \delta_{jk} \quad (6.3d)$$

$$\delta_{ii} = 3 \quad (6.3e)$$

$$\epsilon_{ijk} \epsilon_{lmn} = \det \begin{vmatrix} \delta_{il} & \delta_{im} & \delta_{in} \\ \delta_{jl} & \delta_{jm} & \delta_{jn} \\ \delta_{kl} & \delta_{km} & \delta_{kn} \end{vmatrix} \quad (6.3f)$$

$$\epsilon_{ijk} \epsilon_{lmk} = \epsilon_{kij} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \quad (6.3g)$$

$$\epsilon_{ijk} \epsilon_{ljk} = 2\delta_{il} \quad (6.3h)$$

$$\epsilon_{ijk} \epsilon_{ijk} = 6 \quad (6.3i)$$

$$\epsilon_{ijk} \delta_{ij} = 0. \quad (6.3j)$$

Generalizações do símbolo de Levi-Civita, como sendo o objeto matricial totalmente antissimétrico em um espaço de n dimensões, também existem. Por exemplo, se $n = 4$,

$$\epsilon_{i_1 i_2 i_3 i_4} = \begin{cases} +1, & \text{se } \{i_1, i_2, i_3, i_4\} \text{ é uma permutação par de } \{1, 2, 3, 4\} \\ -1, & \text{se } \{i_1, i_2, i_3, i_4\} \text{ é uma permutação ímpar de } \{1, 2, 3, 4\} \\ 0, & \text{se dois ou mais índices são repetidos} \end{cases} = \prod_{1 \leq j < k \leq 4} \text{sgn}(i_k - i_j),$$

existindo ao todo $4^4 = 256$ valores que podem ser organizados em uma matriz $4 \times 4 \times 4 \times 4$.

Em geral, em um espaço de n dimensões,

$$\epsilon_{i_1 \dots i_n} = \begin{cases} +1, & \text{se } \{i_1, \dots, i_n\} \text{ é uma permutação par de } \{1, \dots, n\} \\ -1, & \text{se } \{i_1, \dots, i_n\} \text{ é uma permutação ímpar de } \{1, \dots, n\} \\ 0, & \text{se dois ou mais índices são repetidos} \end{cases} = \prod_{1 \leq j < k \leq n} \text{sgn}(i_k - i_j),$$

sendo que os n^n valores podem ser organizados em uma matriz $n \times \dots \times n$.

6.2 PROPRIEDADES DE TRANSFORMAÇÃO DE ESCALARES, VETORES E TENSORES

Retorna-se agora à discussão envolvendo a invariância das leis físicas frente a diferentes transformações realizadas no sistema de referências e a sua relação com as leis de transformação dos tensores empregados na descrição dessas leis físicas. Relações semelhantes às formas da Lei de Ohm (6.1) ou (6.2) podem ser generalizadas considerando-se a relação entre dois campos vetoriais quaisquer $A = A(\mathbf{r})$ e $B = B(\mathbf{r})$, dada por

$$A_i = \alpha_{ij} B_j,$$

sendo $\alpha_{ij} = \alpha_{ij}(\mathbf{r})$ a matriz que descreve a **relação constitutiva** entre os campos.

Nesta seção será iniciada a discussão de *tensores Cartesianos* e suas transformações. Serão considerados objetos matemáticos denominados *escalares*, *vetores* e *tensores* em geral, aos quais serão atribuídas as noções intuitivas de *campos escalares*, *vetoriais* ou *tensoriais*, respectivamente. Uma definição mais rigorosa destes campos será apresentada posteriormente. Para o presente momento, será assumido que todos os campos envolvidos (escalares, vetoriais e tensoriais) existem em um espaço vetorial particular, o *espaço vetorial Euclidiano de dimensão três*, denotado por \mathbb{E}^3 . O espaço \mathbb{E}^3 é, na verdade, um espaço afim que é também um espaço vetorial métrico no qual a norma induz a métrica. A definição do \mathbb{E}^3 foi realizada no exemplo 4.5. Um espaço Euclidiano é aquele no qual as noções geométricas intuitivas de espaço, dimensão e deslocamento em um *sistema de coordenadas Cartesiano* ou *retangular*² são respeitadas, em conjunto com as noções algébricas de vetores posição que ligam a origem do sistema com um dado ponto do espaço e de vetores deslocamento que são setas orientadas que ligam dois pontos quaisquer no referencial. Por ser um espaço afim, o \mathbb{E}^3 identifica tanto “pontos” do espaço, os quais podem ser ocupados por partículas, por exemplo, como “linhas orientadas”, que são os vetores posição e deslocamento. Por ser também um espaço métrico, o comprimento de qualquer segmento de reta pode ser obtido pela fórmula de Pitágoras, concordando assim com medidas experimentais das posições e deslocamentos das partículas. As definições realizadas no exemplo 4.5 para um espaço de dimensão 3 pode ser automaticamente generalizada para o \mathbb{E}^n , o qual é o espaço Euclidiano de dimensão (finita) n .

Sobre o substrato algébrico/geométrico fornecido pelo espaço \mathbb{E}^3 , atribui-se agora a cada ponto do mesmo um *campo escalar* $\phi = \phi(\mathbf{r}) \equiv \phi(x_1, x_2, x_3)$ ou um *campo vetorial* $A = A(\mathbf{r}) \equiv A(x_1, x_2, x_3)$. Portanto, os campos escalares e os componentes dos campos vetoriais que pertencem a este espaço são funções do vetor posição

$$\mathbf{r} = x_i \hat{e}_i \equiv x_1 \hat{e}_1 + x_2 \hat{e}_2 + x_3 \hat{e}_3, \quad (6.4)$$

ou seja,

$$A = A(\mathbf{r}) \equiv A(x_1, x_2, x_3) = A_i(\mathbf{r}) \hat{e}_i.$$

²Ver figuras 4.1.

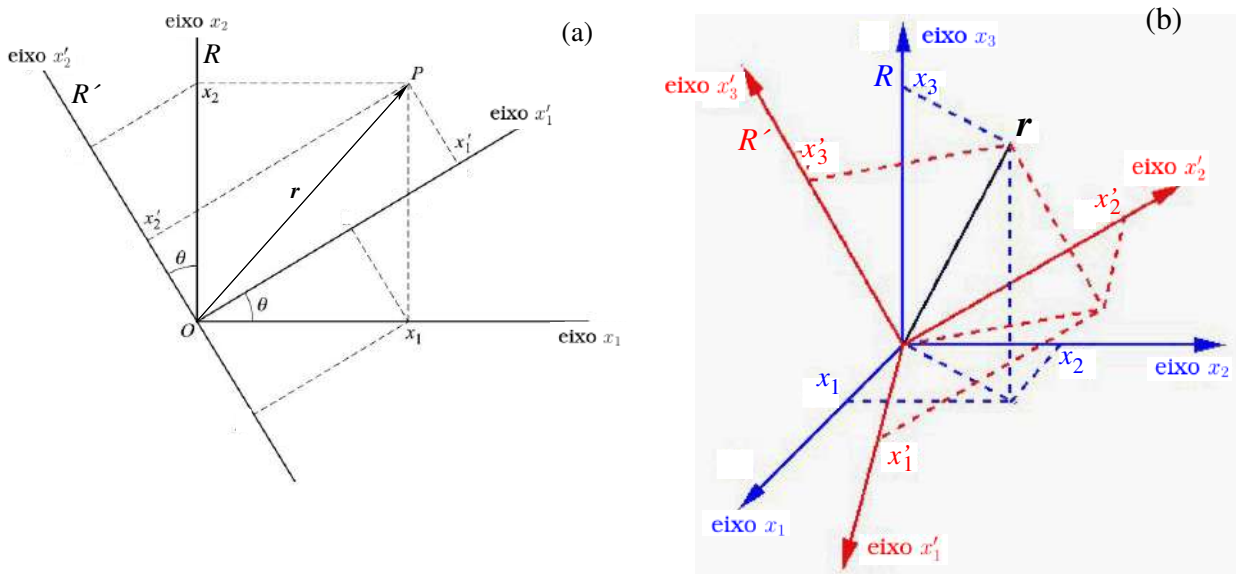


Figura 6.1: (a) Rotação sobre o plano (x_1, x_2) (em torno do eixo x_3) por um ângulo θ . (b) Rotação arbitrária de eixos em torno da origem do sistema de coordenadas. Pode-se observar que em ambos os casos as rotações mantêm a norma do vetor r invariante.

Em (6.4), o conjunto de vetores *ortonormais* $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ forma uma base do espaço \mathbb{E}^3 . Neste caso, em se tratando de um sistema Cartesiano, emprega-se a base canônica $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\} = \{\hat{i}, \hat{j}, \hat{k}\}$, onde $\hat{i} = (1, 0, 0)$, $\hat{j} = (0, 1, 0)$ e $\hat{k} = (0, 0, 1)$. Como é usual, assume-se que o sistema de coordenadas Cartesiano é *dextrógiro*, isto é, a base canônica obedece a relação³

$$\hat{e}_i \times \hat{e}_j = \epsilon_{ijk} \hat{e}_k. \quad (6.5)$$

Algumas propriedades básicas de escalares e vetores que pertencem ao espaço Euclidiano frente a transformações no sistema de coordenadas serão brevemente discutidas agora. Em seguida, esta discussão será generalizada para tensores Cartesianos em geral.

A hipótese de que as quantidades físicas apresentem propriedades matemáticas bem definidas frente a certos tipos de transformação de coordenadas impõe limitações e exigências importantes às leis físicas. É necessário, portanto, discutir em algum detalhe as propriedades de transformação de algumas quantidades físicas escalares ou vetoriais.

Alguns dos tipos mais importantes de transformações de coordenadas para a física são *rotações*, *reflexão espacial* ou *transformação de paridade* e *reversão temporal*. Já para a relatividade restrita, também são fundamentais as mudanças entre diferentes *referenciais inerciais*, dadas por uma *translação* entre os dois referenciais. Embora a rotação seja o tipo de transformação para o qual será dada a maior atenção neste capítulo, é importante realizar também uma breve discussão a respeito das outras transformações.

6.2.1 ROTAÇÕES

Uma rotação no \mathbb{E}^3 é uma transformação linear das coordenadas, realizada em torno de um ponto fixo e de tal forma que a *norma* do espaço permaneça invariante. Este ponto fixo usualmente é a origem do sistema de coordenadas e pode ocorrer sobre um determinado plano (subespaço) de \mathbb{E}^3 ou sobre todo o espaço. A figura 6.1 ilustra rotações de eixos sobre o plano ou no espaço.

O caso particular ilustrado na figura 6.1a, onde é realizada uma rotação do sistema de coordenadas sobre o plano em torno do eixo x_3 por um ângulo θ , será considerado como referência na discussão a seguir. As conclusões obtidas serão então generalizadas para uma rotação arbitrária dos eixos no \mathbb{E}^3 , representada na figura 6.1b. A discussão feita neste ponto pode ser então facilmente generalizada para rotações em um espaço vetorial de n dimensões.

O vetor posição r é representado no sistema de coordenadas Cartesiano original por $r = x_i \hat{e}_i$, onde $\{x_1, x_2, x_3\}$ são as **componentes** (ou **projeções**) de r que constituem a sua **representação**

³Esta relação segue a definição de produto vetorial entre vetores do \mathcal{R}^3 dada em (4.7).

no sistema de coordenadas R . Realizando então uma rotação de eixos arbitrária, ilustrada na figura 6.1b, o vetor posição passa a ser, em princípio, representado por

$$\mathbf{r}' = x'_i \hat{e}'_i,$$

onde $\{x'_i\}$ são suas componentes no sistema de coordenadas R' e $\{\hat{e}'_i\}$ são os vetores de base no sistema rotado.

Atribuindo agora uma realidade física ao vetor posição r , assume-se que o espaço é isotrópico; isto é, não há direção preferencial ou, em outras palavras, todas as direções são equivalentes. Então, o sistema físico em estudo, para o qual uma lei física está sendo aplicada fazendo-se uso de quantidades vetoriais, não pode depender da orientação do sistema de coordenadas. Como consequência desta exigência, necessariamente

$$\mathbf{r}' = \mathbf{r} \implies x_i \hat{e}_i = x'_i \hat{e}'_i. \quad (6.6)$$

Observando a figura 6.1a, percebe-se a seguinte relação imediata entre $\{\hat{e}_1, \hat{e}_2\}$ e $\{\hat{e}'_1, \hat{e}'_2\}$:

$$\begin{aligned} \hat{e}'_1 &= \cos \theta \hat{e}_1 + \sin \theta \hat{e}_2 \\ \hat{e}'_2 &= -\sin \theta \hat{e}_1 + \cos \theta \hat{e}_2. \end{aligned}$$

Generalizando para o caso de uma rotação arbitrária dos eixos no \mathbb{E}^3 , como ilustrado na figura 6.1b, pode-se escrever, em geral,

$$\hat{e}'_i = S_{ji} \hat{e}_j, \quad (6.7a)$$

onde $\{S_{ij}\}$ são os elementos da *matriz de rotação* S . No caso particular da rotação em torno de x_3 , mostrado na figura 6.1a, esta matriz tem ordem 2 e seus elementos são

$$S = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix},$$

ou seja, estes elementos dependem somente de um parâmetro (ou ângulo) fixo.

Retornando ao caso geral, a relação (6.7a) pode ser escrita na forma matricial ao se definir as seguintes matrizes coluna e linha

$$\hat{\mathbf{e}} = \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{pmatrix}, \quad \tilde{\hat{\mathbf{e}}} \equiv \hat{\mathbf{e}}^T = (\hat{e}_1 \hat{e}_2 \hat{e}_3)$$

para os vetores unitários, sendo que os símbolos “ \sim ” ou “ T ” indicam a transposição da matriz. Neste caso, a transformação (6.7a) pode ser representada pela multiplicação matricial

$$\hat{\mathbf{e}}'^T = \hat{\mathbf{e}}^T S, \text{ ou } \hat{\mathbf{e}}' = \tilde{S} \hat{\mathbf{e}}.$$

Para uma rotação em torno de um eixo, somente um ângulo se faz necessário. Já no caso geral, é necessário pelo menos um ângulo adicional, mas esta rotação arbitrária sempre pode ser escrita em termos da matriz quadrada S de ordem 3, cujos elementos são funções de um ou mais parâmetros (ângulos) fixos. Para se construir essa matriz de rotação S para o caso geral, retorna-se inicialmente à figura 6.1a e define-se o **cosseno diretor**

$$S_{ji} \doteq \hat{e}'_i \cdot \hat{e}_j \doteq \cos \theta_{ji},$$

o qual é a projeção do vetor \hat{e}'_i sobre o vetor \hat{e}_j . Este cosseno diretor é simplesmente o cosseno do ângulo θ_{ij} entre os vetores de base. Ou seja, pode-se escrever

$$\begin{aligned} \hat{e}'_1 &= \cos \theta_{11} \hat{e}_1 + \cos \theta_{21} \hat{e}_2 = \cos \theta \hat{e}_1 + \sin \theta \hat{e}_2 \\ \hat{e}'_2 &= \cos \theta_{12} \hat{e}_1 + \cos \theta_{22} \hat{e}_2 = -\sin \theta \hat{e}_1 + \cos \theta \hat{e}_2, \end{aligned}$$

onde foi chamado $\theta_{11} = \theta$ e, por consequência, $\theta_{21} = \frac{\pi}{2} - \theta$, $\theta_{22} = \theta$ e $\theta_{12} = \frac{\pi}{2} + \theta$, de onde resultam as últimas expressões. Fazendo referência agora à figura 6.1b, observa-se que uma rotação geral do referencial R' em relação a R pode sempre ser expressa em termos dos 9 cossenos diretores $\{\theta_{ij}\}$. Contudo, será mostrado em seguir que somente 3 desses ângulos são independentes entre si.

A expressão (6.7a) descreve uma mudança de bases na transformação $\{\hat{e}_i\} \rightarrow \{\hat{e}'_i\}$ de um sistema de coordenadas Cartesiano para outro, rotado em relação ao original. A transformação inversa $\{\hat{e}'_i\} \rightarrow \{\hat{e}_i\}$ consiste na rotação inversa $\theta \rightarrow -\theta$ aplicada sobre a base $\{\hat{e}'_i\}$; esta transformação é escrita como

$$\hat{e}_i = (S^{-1})_{ji} \hat{e}'_j, \quad (6.7b)$$

onde S^{-1} é a matriz de rotação inversa.

O objetivo agora consiste em tomar a expressão para a mudança de bases $\{\hat{e}_i\} \leftrightarrow \{\hat{e}'_i\}$ para se obter uma relação entre as coordenadas dos sistemas, ou seja, para escrever $x'_i = x'_i(\{x_j\})$ ou $x_i = x_i(\{x'_j\})$. Antes, porém, com o intuito de simplificar a notação, introduz-se a *matriz de transformação* L tal que $L = S^{-1}$. Introduzindo então as relações (6.7a,b) em (6.6), resulta

$$\begin{cases} \hat{e}'_i = (L^{-1})_{ji} \hat{e}_j \\ \hat{e}_i = L_{ji} \hat{e}'_j \end{cases} \implies \begin{cases} x'_i = (S^{-1})_{ij} x_j = L_{ij} x_j \\ x_i = S_{ij} x'_j = (L^{-1})_{ij} x'_j. \end{cases} \quad (6.7c)$$

Cabe ressaltar agora que, definindo a matriz coluna $r = (x_1 \ x_2 \ x_3)^T$, as transformações de coordenadas acima podem ser representadas pelas multiplicações matriciais

$$r' = Lr, \quad r = L^{-1}r'.$$

Levando em conta agora a exigência adicional de invariância da *norma* do espaço Euclidiano \mathbb{E}^3 , isto é, $\|r\| = \|r'\|$, resulta

$$x_i x_i = x'_i x'_i \implies x_i x_i = L_{ij} L_{ik} x_j x_k.$$

A identidade somente pode ser obedecida em geral se a seguinte **condição de ortogonalidade** é satisfeita:

$$L_{ki} L_{kj} = \delta_{ij} \text{ ou } L_{ik} L_{jk} = \delta_{ij}. \quad (6.8)$$

Comparando esta condição com a definição da matriz inversa de L , $L_{ij} (L^{-1})_{ki} = \delta_{jk}$, resulta que $L^{-1} = \tilde{L}$, sendo esta última a *transposta* da matriz L . A segunda expressão em (6.8) é consequência disto. Portanto, as leis desejadas para a transformação dos sistemas de coordenadas podem ser escritas como

$$\begin{cases} \hat{e}'_i = L_{ij} \hat{e}_j \\ \hat{e}_i = L_{ji} \hat{e}'_j \end{cases} \iff \begin{cases} x'_i = L_{ij} x_j \\ x_i = L_{ji} x'_j. \end{cases} \quad (6.9a)$$

A condição de ortogonalidade (6.8) também leva à seguinte classificação para a classe de rotações executadas no sistema de coordenadas. Escrevendo (6.8) na forma matricial, sendo I_3 a matriz identidade, e calculando o determinante, resulta

$$L\tilde{L} = I_3 \implies \det(L\tilde{L}) = \det(L) \det(\tilde{L}) = [\det(L)]^2 = 1.$$

Classifica-se, então,

$$\begin{aligned} \det(L) = +1 &\rightsquigarrow \text{Rotações próprias} \\ \det(L) = -1 &\rightsquigarrow \text{Rotações impróprias.} \end{aligned} \quad (6.9b)$$

Em (6.9b), uma rotação própria é aquela em que a transformação $R \rightarrow R'$ pode ser obtida por uma sequência de rotações infinitesimais. Já uma rotação imprópria corresponde a uma reflexão dos eixos (transformação de paridade), seguida por uma rotação própria.

A partir das relações de mudança de bases (6.7a,b) e lembrando que a matriz L depende somente de parâmetros fixos, as seguintes relações também podem ser deduzidas,

$$\begin{aligned} \hat{e}'_i = L_{ij} \hat{e}_j &\implies L_{ij} = \hat{e}'_i \cdot \hat{e}_j \\ \hat{e}_i = L_{ji} \hat{e}'_j &\implies L_{ji} = \hat{e}_i \cdot \hat{e}'_j \end{aligned} \quad (6.9c)$$

$$\begin{aligned} x'_i = L_{ij} x_j &\implies L_{ij} = \frac{\partial x'_i}{\partial x_j} \\ x_i = L_{ji} x'_j &\implies L_{ji} = \frac{\partial x_i}{\partial x'_j}. \end{aligned} \quad (6.9d)$$

Lembrando finalmente dos cossenos diretores, escreve-se $L_{ij} = \hat{e}'_i \cdot \hat{e}_j = \cos \theta_{ij}$ e então, a partir da condição de ortogonalidade (6.8), resultam as equações

$$\begin{aligned} L_{1i}^2 + L_{2i}^2 + L_{3i}^2 &= 1 \quad (i = 1, 2, 3) \\ L_{1i}L_{1j} + L_{2i}L_{2j} + L_{3i}L_{3j} &= 0 \quad (i \neq j, i, j = 1, 2, 3). \end{aligned}$$

Ou seja, a condição de ortogonalidade implica em 6 equações distintas que relacionam os diferentes valores de ângulo.

Por conseguinte, dos 9 cossenos diretores, somente 3 ângulos são realmente independentes. Uma das definições com frequência empregada para esses ângulos será discutida na seção 6.5.2.

A discussão recém realizada, sob o ponto de vista geométrico, acerca das propriedades das matrizes de rotação L e S ($= \tilde{L}$) já foi abordada nos exemplos 5.19, 5.20, 5.25 e 5.29 no contexto de grupos de Lie. Ou seja, a matriz de rotação L (própria ou imprópria) é uma representação do grupo $O(3)$.

Exercício 6.1. Encontre a matriz L (3×3) que realiza a rotação (própria) de eixos representada na figura 6.1a.

Solução. Usando os elementos da matriz S apresentados acima e sabendo que $S = L^{-1} = \tilde{L}$, resulta

$$S = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \tilde{L} \implies L = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Observa-se que a matriz L obtida corresponde à representação $R^{(3)}(\theta) \in SO(3)$ derivada no exemplo 5.30.

Com base na lei de transformação (6.9a-e), é possível agora realizar uma classificação das quantidades físicas de acordo com o seu comportamento frente a uma rotação arbitrária no sistema de coordenadas adotado para representá-las.

Escalares (tensores de posto zero). Seja $\phi(\mathbf{r})$ uma quantidade física que, em geral, pode depender da posição de observação da mesma em relação à origem do sistema de coordenadas, mas que não é um vetor, ou seja, esta pode ser caracterizada por um único número. Se $\phi(\mathbf{r})$ for invariante frente a uma rotação arbitrária do sistema de coordenadas, descrita pela matriz L , então esta quantidade é denominada um *escalar* ou *tensor de posto* (ou *ordem*) *zero*.

Exemplos de quantidades físicas escalares são: massa, carga elétrica, potencial elétrico e energia. Certos produtos escalares de vetores e divergentes de campos vetoriais também são quantidades escalares.

Vetores (tensores de posto um). Seja $A \equiv A(\mathbf{r}) = A_i(\mathbf{r}) \hat{e}_i$ uma quantidade física representada por um conjunto de três quantidades escalares $\{A_i(\mathbf{r})\}$ no \mathbb{E}^3 , quando medidas em relação ao sistema de coordenadas Cartesiano R . Ao se aplicar uma rotação arbitrária $R \rightarrow R'$ ao sistema de coordenadas, descrita pela matriz L , as componentes desta quantidade física passam a ser representadas por $\{A'_i(\mathbf{r})\}$. A quantidade $A(\mathbf{r})$ é, então, um *vetor* ou um *tensor de posto um* se e somente se a relação equivalente a (6.9a), ou seja,

$$A'_i = L_{ij}A_j, \quad (6.10)$$

também se aplica entre suas componentes nos respectivos sistemas de coordenadas.

Posteriormente serão apresentadas algumas quantidades físicas vetoriais.

Tensores de posto (ou ordem) dois. Antes de se introduzir uma definição formal de tensores, será feita uma breve menção sobre tensores de posto dois, uma vez que estes aparecem com frequência em problemas físicos. A relação constitutiva (Lei de Ohm) (6.2) mostra que em um meio anisotrópico a relação entre os vetores \mathbf{J} e \mathbf{E} é determinada pela matriz de condutividade $\{\alpha_{ij}\}$, a qual possui 9 elementos no \mathbb{E}^3 . Esta matriz será a representação de um *tensor de posto dois* (tensor de condutividade elétrica) se e somente se, após realizada a transformação $R \rightarrow R'$, quando seus elementos serão então transformados $\{\alpha_{ij}\} \rightarrow \{\alpha'_{ij}\}$, estes elementos se relacionarem por

$$\alpha'_{ij} = L_{ik}L_{j\ell}\alpha_{k\ell}.$$

Além do tensor de condutividade elétrica, outros exemplos de tensores de posto dois são o tensor de campo eletromagnético e o tensor energia-momento, também denominado tensor de stress de Maxwell.

Finalmente, pode-se também classificar rotações e outras transformações como *passivas* ou *ativas*.

Transformação passiva. Em uma transformação passiva o sistema físico é mantido inalterado e somente o sistema de coordenadas é alterado. As rotações ilustradas na figura 6.1 são exemplos de transformações passivas.

Transformação ativa. Já em uma transformação ativa, o referencial é mantido fixo e a transformação é aplicada ao sistema físico em estudo. A figura 6.2 ilustra uma rotação ativa realizada sobre um sistema composto por duas cargas elétricas interagentes.

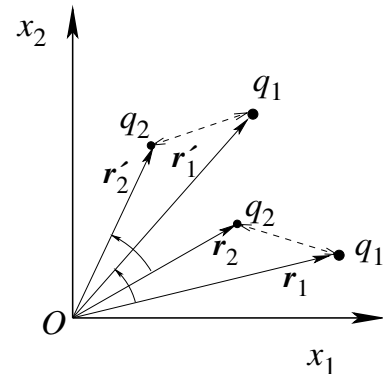


Figura 6.2: Exemplo de uma transformação ativa.

Na seção seguinte será realizada uma definição formal de tensores, partindo das leis de transformação (6.9a-e) aqui deduzidas. Antes, porém, outros tipos de transformações serão brevemente discutidos.

6.2.2 TRANSFORMAÇÕES DE PARIDADE OU REFLEXÕES

Uma *transformação de paridade*, também denominada *reflexão espacial* ou *inversão de paridade*, é uma transformação em que uma ou mais coordenadas do sistema de referência são invertidas, ou seja, $x_i \rightarrow -x_i$. Uma reflexão espacial em um plano consiste em inverter o sinal da coordenada normal ao plano, mantendo as coordenadas sobre o plano inalteradas. Assim, uma reflexão no plano $x_1 - x_2$ consiste em realizar a transformação

$$(x_1, x_2, x_3) \rightarrow (x_1, x_2, -x_3).$$

Já uma *inversão espacial* ou transformação de paridade é realizada invertendo-se os sinais de todas as coordenadas do referencial, ou seja,

$$\mathbf{r} \rightarrow \mathbf{r}' = -\mathbf{r}. \quad (6.11)$$

A matriz de transformação para (6.11) pode ser escrita simplesmente como $L_{ij} = -\delta_{ij}$. Ou seja, $\det(L) = -1$, implicando que uma inversão espacial é sempre imprópria.

Esta transformação permite a definição de uma nova classificação das quantidades físicas em função de seu comportamento frente a inversões espaciais.

Escalares ou pseudoescalares. Dado o campo $\phi(\mathbf{r})$, esta quantidade é um *escalar* se permanecer inalterado frente a uma operação de inversão espacial. Caso contrário, esta quantidade é um *pseudoescalar*.

Um exemplo de um pseudoescalar é obtido a partir do produto misto $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, desde que \mathbf{a} , \mathbf{b} e \mathbf{c} sejam todos *vetores polares*, os quais serão definidos a seguir.

Vetores polares ou vetores axiais (pseudovetores). Um *vetor polar* ou simplesmente vetor é aquela quantidade $\mathbf{A} = \mathbf{A}(\mathbf{r})$ que continua se transformando de acordo com a relação (6.10), mesmo quando realizada uma reflexão espacial, ou seja, se

$$\mathbf{A} \rightarrow \mathbf{A}' = -\mathbf{A} \text{ quando } \mathbf{r} \rightarrow -\mathbf{r}.$$

Um *vetor axial* ou *pseudovetor* é aquela quantidade $\mathbf{B} = \mathbf{B}(\mathbf{r})$ que frente a uma transformação de paridade comporta-se como

$$\mathbf{B} \rightarrow \mathbf{B}' = \mathbf{B}.$$

Um exemplo de um vetor axial é obtido partindo-se de dois vetores polares \mathbf{a} e \mathbf{b} e obtendo-se o vetor \mathbf{c} a partir de $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. Neste caso, o vetor \mathbf{c} é um pseudovetor.

Tensores ou pseudotensores. As propriedades de transformação de tensores de posto N arbitrário podem ser deduzidas facilmente se estes são puderem ser construídos a partir de produtos de vetores polares e/ou axiais. Se for realizada uma transformação de paridade sobre um tensor de posto N e este se transformar com o fator $(-1)^N$, então este é denominado um *tensor verdadeiro* ou simplesmente *tensor*. Contudo, se a inversão espacial levar ao fator $(-1)^{N+1}$, então este é denominado de *pseudotensor* de posto N .

6.2.3 REVERSÃO TEMPORAL

Um outro tipo de transformação relevante aos sistemas físicos é a *transformação de reversão temporal* $t \rightarrow t' = -t$. Embora esta transformação não se aplica a sistemas descritos pela mecânica newtoniana, na qual a dependência temporal no comportamento do sistema físico é considerada de forma distinta da sua dependência espacial, mesmo assim é importante que esta seja discutida.

As leis básicas da física comportam-se de maneira bem determinada frente a inversão no sentido de evolução do tempo, e esse comportamento permite classificar as quantidades físicas como *pares* ou *ímpares* frente a uma reversão temporal.

Transformação par. Uma determinada quantidade física é *par* frente a uma reversão temporal se a lei física que a determina não muda de sinal frente à transformação $t \rightarrow t' = -t$.

Um exemplo simples de uma quantidade par é o vetor posição, $\mathbf{r} \xrightarrow{t \rightarrow -t} \mathbf{r}$. Um outro vetor par é a aceleração de uma partícula, pois

$$\mathbf{a} = \frac{d^2 \mathbf{r}}{dt^2} \xrightarrow{t \rightarrow -t} \frac{d^2 \mathbf{r}}{dt^2} = \mathbf{a}.$$

Transformação ímpar. Uma quantidade física é *ímpar* frente a uma reversão temporal se a lei física que a determina muda de sinal frente a esta transformação.

Um exemplo de quantidade ímpar é o momentum de uma partícula, pois (para massa constante)

$$\mathbf{p} = m \frac{d\mathbf{r}}{dt} \xrightarrow{t \rightarrow -t} -m \frac{d\mathbf{r}}{dt} = -\mathbf{p}.$$

As propriedades de algumas quantidades fundamentais na mecânica clássica e no eletromagnetismo frente às transformações discutidas nas seções 6.2.1 – 6.2.3 são apresentadas na tabela 6.1.

As definições e o comportamento das quantidades físicas frente a transformações passivas do sistema de coordenadas, discutidos nesta seção, serão desenvolvidos em maiores detalhes nas seções posteriores. Para esta discussão aprofundada acerca dos tensores Cartesianos, será considerada de forma preponderante a transformação de rotação (própria) do sistema de referência.

6.3 TENSORES CARTESIANOS

Nesta seção serão realizadas definições um pouco mais rigorosas dos campos escalares, vetoriais e tensoriais e suas propriedades sob transformações em geral. A discussão ainda estará restrita aos chamados *tensores Cartesianos*. Esta restrição será posteriormente eliminada na seção 6.7.

Uma vez que os campos de interesse na física dependem de forma contínua nas coordenadas do vetor posição, será realizada inicialmente uma breve discussão a respeito de espaços funcionais e suas classes. Essa discussão servirá como uma continuação à definição 3.22 de uma função.

6.3.1 ESPAÇOS FUNCIONAIS

Um *espaço funcional* é formado por um conjunto de funções $f : X \mapsto Y$ (definição 3.22) de um determinado tipo ou *classe*, que estabelecem um mapeamento do conjunto X ao conjunto Y .

Tabela 6.1: Propriedades de transformação de algumas quantidades físicas na mecânica clássica e no eletromagnetismo.

Quantidade Física		Rotação (posto do tensor)	Inversão Espacial	Reversão Temporal
Mecânica Clássica				
Posição	r	1	Polar	Par
Momentum linear	p	1	Polar	Ímpar
Momentum angular	$L = r \times p$	1	Axial	Ímpar
Força	F	1	Polar	Par
Torque	$\tau = r \times F$	1	Axial	Par
Energia Cinética	$p^2/2m$	0	Escalar	Par
Energia potencial	$U(r)$	0	Escalar	Par
Eletromagnetismo				
Densidade de carga	$\rho(r)$	0	Escalar	Par
Densidade de corrente	$J(r)$	1	Polar	Ímpar
Campo elétrico	$E(r)$	1	Polar	Par
Deslocamento elétrico	$D(r)$	1	Polar	Par
Polarização	$P(r)$	1	Polar	Par
Indução magnética	$B(r)$	1	Axial	Ímpar
Campo magnético	$H(r)$	1	Axial	Ímpar
Vetor de Poynting	$S = E \times B$	1	Polar	Ímpar
Tensor de stress	T_{ij}	2	Tensor	Par

Este conjunto de funções é denominado um *espaço* porque em muitas aplicações de interesse para a física esse conjunto forma um *espaço topológico*⁴ (incluindo espaços métricos⁵), um espaço vetorial,⁶ ou ambos. Por exemplo, o conjunto de todas as transformações lineares⁷ (ou funções) do espaço vetorial \mathcal{V} ao espaço vetorial \mathcal{W} sobre o mesmo corpo \mathbb{K} é, em si mesmo, um espaço vetorial sobre o corpo \mathbb{K} .

Alguns exemplos relevantes destes espaços funcionais são:

- $\mathcal{C}[a, b]$, o conjunto de todas as funções reais $f : \mathbb{R} \mapsto \mathbb{R}$ contínuas no intervalo (fechado) $[a, b] \subset \mathbb{R}$.
- $\mathcal{C}^r[a, b]$, o conjunto de todas as funções reais que são contínuas até a derivada de ordem r no intervalo $[a, b] \subset \mathbb{R}$.
- $\mathcal{C}_0(\mathbb{R})$, o conjunto de todas as funções reais contínuas que são nulas no infinito.
- $\mathcal{C}^r(\mathbb{R})$, o conjunto de todas as funções reais que são contínuas até a derivada de ordem r .
- $\mathcal{C}^\infty(\mathbb{R})$, o conjunto de todas as funções reais que possuem derivadas em todas as ordens. Estas funções também são denominadas de *funções suaves*.
- $\mathcal{L}_1[a, b]$, o conjunto de todas as funções reais cujo valor absoluto é integrável no intervalo $[a, b] \subset \mathbb{R}$.
- $\mathcal{L}_2[a, b]$, o conjunto de todas as funções reais quadraticamente integráveis no intervalo $[a, b] \subset \mathbb{R}$.

A partir desta definição de espaços funcionais é possível prosseguir com a definição de campos tensoriais em geral. Inicialmente serão tratados os campos escalares e vetoriais, os quais serão em seguida generalizados.

6.3.2 TENSORES CARTESIANOS DE POSTOS ZERO E UM

Nesta seção serão apresentadas as definições dos campos escalares e vetoriais, bem como suas propriedades de transformação.

⁴Definição 4.39.

⁵Definição 4.35.

⁶Capítulo 4.

⁷Seção 4.4.

Sejam os conjuntos \mathbb{R}^n e \mathbb{C}^n , respectivamente formados pelos produtos Cartesianos⁸ de \mathbb{R} ou \mathbb{C} consigo mesmos $n - 1$ vezes. Uma **região** Ω de \mathbb{R}^n ou de \mathbb{C}^n consiste em um subconjunto dos mesmos que é não vazio, aberto e conectado.⁹ O subconjunto Ω é uma **região fechada** se este contiver o *fecho* do mesmo, isto é, seus pontos-limite. Esta definição de região de um conjunto numérico pode ser imediatamente estendida para referir a subconjuntos de qualquer estrutura mais complexa, como corpos, espaços vetoriais ou espaços afim.

Nas definições a seguir, os campos escalares e vetoriais são definidos sobre espaços vetoriais Euclidianos.¹⁰ A generalização para espaços não Euclidianos será apresentada na seção 6.7.

Definição 6.1 (Campo escalar). Seja $U \subset \mathbb{E}^n$ uma região do espaço Euclidiano \mathbb{E}^n de dimensão n . Um *campo escalar* ϕ da classe C^r sobre U é o mapeamento

$$\phi : U \mapsto \mathbb{K}$$

sobre o corpo \mathbb{K} , o qual atribui a cada ponto $r \equiv (x_1, x_2, \dots, x_n) \in U$ a função $\phi = \phi(r) \equiv \phi(x_1, x_2, \dots, x_n) \in C^r(U)$.

Nas aplicações para a física, usualmente $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . Alguns dos exemplos mais conhecidos de campos escalares:

- *Campos potenciais*, tais como o potencial escalar elétrico ou o potencial escalar gravitacional.
- Em hidrodinâmica, *temperatura*, *humidade* e *pressão* também são descritos como campos escalares.
- Na teoria quântica de campos, os campos escalares são atribuídos a partículas de spin 0.

Já um campo vetorial é apresentado desde as disciplinas básicas de física e matemática como um conjunto de setas orientadas, tangenciais às *linhas de força* oriundas do(s) agente(s) que gera(m) o campo. Um exemplo típico de visualização de um campo vetorial é apresentado na figura 6.3. Duas cargas elétricas pontuais geram o campo elétrico representado em alguns pontos da figura como setas orientadas que são tangenciais às linhas de força. Uma terceira carga elétrica posicionada em qualquer ponto do espaço irá sentir uma força elétrica cuja direção será paralela à reta tangente à linha de força nessa posição.

A aparente simplicidade desta concepção visual do que é um campo vetorial esconde uma complexidade intrínseca devido ao conjunto de objetos matemáticos distintos envolvidos e nas suas inter-relações.

Um campo vetorial é composto por *vetores*, que são os componentes de um espaço vetorial. Observando com atenção a definição do que é um espaço vetorial, realizada na seção 4.1, nota-se que os vetores são, estritamente, objetos algébricos, *i. e.*, não lhes é atribuída *a priori* nenhuma estrutura geométrica ou analítica.

Porém, um campo vetorial também possui uma concepção geométrica inerente, uma vez que o campo se distribui no espaço a partir do(s) agente(s) gerador(es). Além disso, para o seu cálculo, o campo vetorial também necessita de um sistema de referências, com uma escala de medidas bem definida. Por isso, um campo vetorial também deve ser um tipo de espaço métrico (definição 4.35).

A estrutura matemática que estabelece a interrelação entre um espaço vetorial e um espaço métrico é o espaço afim (definição 4.45), o qual estabelece de forma rigorosa os conceitos de *vetores posição* e *deslocamento*, os quais localizam os pontos do espaço métrico através de setas

⁸Definição 3.27.

⁹Um conjunto ou *espaço conectado* é um espaço topológico que não pode ser formado pela união de dois ou mais subconjuntos não vazios disjuntos.

¹⁰Ver exemplo 4.5.

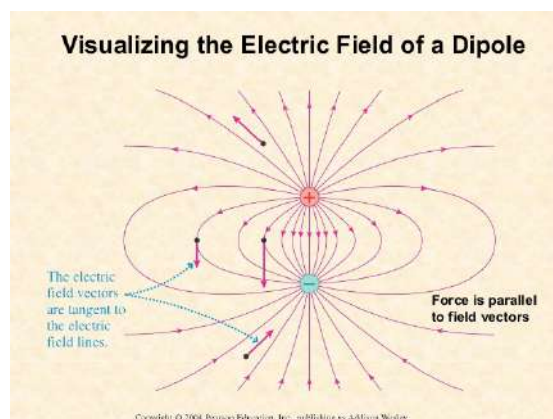


Figura 6.3: Visualização das linhas de força de um dipolo elétrico e do campo elétrico resultante em alguns pontos particulares.

orientadas que partem da origem do sistema de coordenadas. Qualquer transformação de um ponto do espaço a outro é então realizada pelo vetor deslocamento, o qual é uma seta orientada que parte do primeiro a termina no segundo (nesta seção serão considerados campos vetoriais sobre o espaço vetorial Euclidiano \mathbb{E}^n , exemplo 4.5).

Finalmente, sobre cada ponto do espaço \mathbb{E}^n localizado pelo vetor \mathbf{r} será atribuído um vetor do campo vetorial, vinculado a esse ponto. Esse vetor terá a característica adicional de possuir uma topologia, *i. e.*, os conceitos de limite e continuidade são supostos válidos em relações aos vetores do campo em pontos da vizinhança imediata de \mathbf{r} , de tal forma que seja possível realizar operações do cálculo infinitesimal (derivações ou integrações), pelo menos na vizinhança do ponto.

Definição 6.2 (Campo vetorial). Seja $\mathbb{E}^n = (\mathbb{E}^n, \mathcal{R}^n)$ o espaço vetorial Euclidiano e seja $U \subset \mathbb{E}^n$ uma região deste espaço. Seja \mathcal{K}^n um espaço vetorial sobre o corpo \mathbb{K} de dimensão n , o qual pode ser o espaço real \mathcal{R}^n ou o espaço complexo \mathcal{C}^n (definições 4.3 e 4.4). Um *campo vetorial A da classe \mathcal{C}^r sobre U* é o mapeamento

$$\mathbf{A} : U \mapsto \mathcal{K}^n,$$

o qual atribui a cada ponto $\mathbf{r} \equiv (x_1, \dots, x_n) \in \mathbb{E}^n$ a *função vetorial*

$$\mathbf{A}(\mathbf{r}) \equiv \mathbf{A}(x_1, \dots, x_n) \doteq (A_1(\mathbf{r}), A_2(\mathbf{r}), \dots, A_n(\mathbf{r})) \in \mathcal{K}^n,$$

para a qual cada componente é o campo escalar

$$A_k(\mathbf{r}) \in \mathcal{C}^r(U), \quad \forall k = 1, \dots, n.$$

O conjunto de todos os campos vetoriais da classe $\mathcal{C}^r(U)$ sobre U é denotado por $\mathfrak{A}^r(U)$. O conjunto de todos os espaços vetoriais da classe $\mathcal{C}^\infty(U)$ sobre U é denotado por $\mathfrak{A}^\infty(U)$ ou simplesmente por $\mathfrak{A}(U)$.

Retomando então as definições e considerações realizadas na seção 6.2.1, serão impostas agora condições adicionais às componentes de um campo vetorial. A quantidade $\mathbf{A}(\mathbf{r}) = A_i(\mathbf{r}) \hat{e}_i$ é um **vetor** ou um **tensor Cartesiano de posto (ou ordem) um** de um campo vetorial se e somente se suas componentes comportam-se frente a uma rotação do sistema de coordenadas de acordo com as relações (6.9a-d) e (6.10), repetidas aqui juntamente com a respectiva transformação inversa,

$$A'_i = L_{ij} A_j = \frac{\partial x'_i}{\partial x_j} A_j \quad (6.12a)$$

$$A_i = L_{ji} A'_j = \frac{\partial x_i}{\partial x'_j} A'_j, \quad (6.12b)$$

uma vez que a matriz de rotação é ortogonal, de acordo com (6.8).

Exercício 6.2. Considere a matriz de rotação L obtida no exercício 6.1, a qual executa uma rotação passiva de R por um ângulo θ em torno de x_3 . Verifique se as triplas ordenadas $\mathbf{v}(\mathbf{r}) = (v_1(\mathbf{r}), v_2(\mathbf{r}), v_3(\mathbf{r}))$, dadas por

$$(i) \mathbf{v}(\mathbf{r}) = (x_2, -x_1, x_3) \quad (ii) \mathbf{v}(\mathbf{r}) = (x_2, x_1, x_3) \quad (iii) \mathbf{v}(\mathbf{r}) = (x_1^2, x_2^2, x_3)$$

correspondem a campos vetoriais no \mathbb{E}^3 .

Solução. Dada a matriz

$$L = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

como o eixo x_3 não é transformado pela rotação, este não precisa ser considerado. Por outro lado, de acordo com a relação (6.9a), $x'_i = L_{ij} x_j$, ou seja,

$$\begin{aligned} x'_1 &= L_{1j} x_j = x_1 \cos \theta + x_2 \sin \theta \\ x'_2 &= L_{2j} x_j = -x_1 \sin \theta + x_2 \cos \theta. \end{aligned}$$

Tripla (i). As componentes de $\mathbf{v}(\mathbf{r})$ são $v_1(\mathbf{r}) = x_2$, $v_2(\mathbf{r}) = -x_1$ e $v_3(\mathbf{r}) = x_3$.

Se esta tripla realmente é um vetor, então espera-se que suas componentes se transformem como

$$\begin{aligned}v_1 &\longrightarrow v'_1 = x'_2 = -x_1 \operatorname{sen} \theta + x_2 \operatorname{cos} \theta \\v_2 &\longrightarrow v'_2 = -x'_1 = -x_1 \operatorname{cos} \theta - x_2 \operatorname{sen} \theta.\end{aligned}$$

Para verificar se isto ocorre, aplica-se a relação (6.12a), ou seja, $v'_i = L_{ij}v_j$, de onde se obtém

$$\begin{aligned}v'_1 &= v_1 \operatorname{cos} \theta + v_2 \operatorname{sen} \theta = x_2 \operatorname{cos} \theta - x_1 \operatorname{sen} \theta \\v'_2 &= -v_1 \operatorname{sen} \theta + v_2 \operatorname{cos} \theta = -x_2 \operatorname{sen} \theta - x_1 \operatorname{cos} \theta.\end{aligned}$$

Como ambas as expressões são as mesmas, a quantidade $v = (x_2, -x_1, x_3)$ de fato é um vetor.

Tripla (ii). As componentes desta tripla devem se transformar como

$$\begin{aligned}v'_1 &= x'_2 = -x_1 \operatorname{sen} \theta + x_2 \operatorname{cos} \theta \\v'_2 &= x'_1 = x_1 \operatorname{cos} \theta + x_2 \operatorname{sen} \theta.\end{aligned}$$

Verificando, de $v'_i = L_{ij}v_j$ obtém-se

$$\begin{aligned}v'_1 &= v_1 \operatorname{cos} \theta + v_2 \operatorname{sen} \theta = x_2 \operatorname{cos} \theta + x_1 \operatorname{sen} \theta \\v'_2 &= -v_1 \operatorname{sen} \theta + v_2 \operatorname{cos} \theta = -x_2 \operatorname{sen} \theta + x_1 \operatorname{cos} \theta,\end{aligned}$$

as quais são distintas das expressões acima. Portanto este objeto não é um vetor.

Tripla (iii). Para este objeto, basta verificar que a primeira componente,

$$v'_1 = (x'_1)^2 = x_1^2 \operatorname{cos}^2 \theta + x_2^2 \operatorname{sen}^2 \theta + 2x_1x_2 \operatorname{sen} \theta \operatorname{cos} \theta$$

não satisfaz a condição, uma vez que deveria resultar

$$v'_1 = v_1 \operatorname{cos} \theta + v_2 \operatorname{sen} \theta = x_1^2 \operatorname{cos} \theta + x_2^2 \operatorname{sen} \theta.$$

Portanto, este objeto também não é um vetor.

Um vetor, portanto, é um tensor cujos componentes são identificados com um único índice, por enquanto posicionado somente na posição inferior (subscrito). Os componentes deste tensor devem se alterar, frente a uma rotação do sistema de coordenadas, de acordo com as relações (6.12).

Já um *campo escalar* ou, simplesmente, um *escalar*, por sua vez, é descrito por uma única função $\phi(\mathbf{r})$ e, por esta razão, é também denominado um *tensor de posto zero*, pois não necessita de índices para identificação de componentes. Por conseguinte, um objeto matemático somente será classificado como um escalar se este for invariante frente à rotação do referencial.

Objetos escalares também podem ser contruídos a partir do produto escalar de dois vetores. Sendo $\mathbf{A}(\mathbf{r})$ e $\mathbf{B}(\mathbf{r})$ dois vetores quaisquer, então, usando (6.12a) e (6.8),

$$\phi(\mathbf{r}) = \mathbf{A} \cdot \mathbf{B} \xrightarrow{R \rightarrow R'} \phi'(\mathbf{r}) = \mathbf{A}' \cdot \mathbf{B}' = L_{ij}L_{ik}A_jB_k = A_jB_j = \phi(\mathbf{r}).$$

Esta propriedade possui aplicações físicas importantes, pois diversas quantidades escalares como trabalho, energia potencial e densidade de energia nos campos eletromagnéticos são obtidas a partir de produtos escalares de vetores. Estas propriedades são, respectivamente, proporcionais a $\mathbf{F} \cdot d\mathbf{r}$, $q\mathbf{E} \cdot d\mathbf{r}$, $\mathbf{D} \cdot \mathbf{E}$ e $\mathbf{B} \cdot \mathbf{H}$.

Outra maneira de se obter um escalar a partir de um vetor é através do operador divergente. Se $\mathbf{A}(\mathbf{r})$ é um vetor, então

$$\phi(\mathbf{r}) = \nabla \cdot \mathbf{A} = \frac{\partial A_i}{\partial x_i} \xrightarrow{R \rightarrow R'} \phi'(\mathbf{r}) = \left(\frac{\partial A_i}{\partial x_i} \right)' = \frac{\partial A'_i}{\partial x'_i}.$$

Levando em conta agora que deve existir uma lei de transformação $x'_i = x'_i(\{x_j\})$ bem definida, pode-se usar a regra da cadeia e escrever, usando (6.9d), (6.12a) e (6.8),

$$\phi'(\mathbf{r}) = \frac{\partial x_j}{\partial x'_i} \frac{\partial A'_i}{\partial x_j} = L_{ij}L_{ik} \frac{\partial A_k}{\partial x_j} = \frac{\partial A_j}{\partial x_j} = \phi(\mathbf{r}).$$

De forma recíproca, é possível obter-se um vetor a partir de um escalar. Uma maneira de realizar este feito consiste na aplicação do operador gradiente sobre um campo escalar $\phi(\mathbf{r})$, resultando assim

$$\mathbf{A}(\mathbf{r}) = \nabla\phi = \frac{\partial\phi}{\partial x_i} \hat{e}_i.$$

Dada a lei de transformação $x_i = x_i(\{x'_j\})$, emprega-se a derivação em cadeia, realiza-se a mudança de base $\{\hat{e}_i\} \rightarrow \{\hat{e}'_i\}$, e, empregando (6.9d), resulta

$$\mathbf{A}(\mathbf{r}) \rightarrow \frac{\partial x'_j}{\partial x_i} \frac{\partial\phi}{\partial x'_j} L_{ki} \hat{e}'_k = L_{ji} L_{ki} \frac{\partial\phi}{\partial x'_j} \hat{e}'_k = \frac{\partial\phi}{\partial x'_j} \hat{e}'_j,$$

a qual é justamente a expressão para o vetor $\mathbf{A}' = \nabla'\phi$ no referencial R' .

6.3.3 TENSORES CARTESIANOS DE POSTO DOIS OU SUPERIOR

Seguindo na discussão iniciada na seção 6.2.1 acerca de tensores de posto dois, a definição prévia então realizada será agora generalizada.

Tensores são objetos geométricos que estabelecem relações lineares entre vetores, escalares e outros tensores. Um exemplo que já foi empregado é o tensor de condutividade (6.2) que relaciona o campo elétrico com o vetor densidade de corrente elétrica.

Um tensor pode ser representado por uma matriz multidimensional de valores numéricos. O **posto** ou **ordem** de um tensor é a dimensão da matriz necessária para a sua representação. Desta maneira, um *escalar* é um tensor de posto zero e um *vetor* é um vetor de posto um.

Por expressar uma relação entre vetores, a representação empregada para um tensor irá depender da base do sistema de coordenadas e do referencial empregado. Grande parte dos tensores empregados na física estabelecem relações entre objetos definidos no espaço vetorial Euclidiano \mathbb{E}^n usando uma base ortonormal. Estes são usualmente denominados de tensores Cartesianos.

Definição 6.3 (Tensor Cartesiano). Um *tensor Cartesiano* é aquele tensor cuja representação é obtida a partir de uma base ortonormal do espaço vetorial Euclidiano \mathbb{E}^n , no qual é empregado o sistema Cartesiano de coordenadas. O número de índices necessário para identificar todos os seus componentes determina o *posto* ou a *ordem* do tensor.

Da mesma maneira como foi argumentado para vetores na seção 6.2.1, o *tensor* como uma entidade fisico-matemática deve permanecer invariante frente a uma transformação no sistema de coordenadas. Esta exigência fundamental estabelece simultaneamente a lei de transformação de seus elementos ou componentes, bem como a própria definição de um tensor de um determinado posto. Será apresentada primeiro a definição de um tensor do posto dois.

Definição 6.4 (Tensor Cartesiano de posto 2). Seja \mathbb{E}^3 o espaço vetorial Euclidiano de dimensão 3 e $\hat{e} \doteq \{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ uma base ortonormal do mesmo. Seja $T[\hat{e}]$ a matriz quadrada 3×3 cujos elementos $T_{ij}[\hat{e}]$ estão representados na base \hat{e} . Seja também a transformação de coordenadas $\hat{e} \rightarrow \hat{e}'$, em conjunto com a transformação inversa $\hat{e}' \rightarrow \hat{e}$, tais que

$$\hat{e}'_i = L_{ij} \hat{e}_j = \frac{\partial x'_i}{\partial x_j} \hat{e}_j, \quad \hat{e}_i = L_{ji} \hat{e}'_j = \frac{\partial x_i}{\partial x'_j} \hat{e}'_j.$$

Se os elementos da matriz $T[\hat{e}]$ se transformarem frente às transformações $\hat{e} \leftrightarrow \hat{e}'$ de acordo com a lei de transformação

$$T'_{ij} = L_{ik} L_{jl} T_{kl} = \frac{\partial x'_i}{\partial x_k} \frac{\partial x'_j}{\partial x_l} T_{kl} \quad (6.13a)$$

$$T_{ij} = L_{ki} L_{lj} T'_{kl} = \frac{\partial x_i}{\partial x'_k} \frac{\partial x_j}{\partial x'_l} T'_{kl}, \quad (6.13b)$$

então a matriz $T[\hat{e}]$ representa um *tensor Cartesiano de posto dois*.

Da mesma maneira como foi realizado com vetores, os índices que identificam os distintos componentes do tensor estão na posição inferior. Posteriormente, no contexto de tensores generalizados, será permitido o posicionamento de índices tanto na posição inferior (subíndices) quanto na superior (superíndices). Isto será discutido na seção 6.7.

Como a extensão lógica da definição 6.13, considera-se agora uma matriz $T[\hat{e}]$ de n dimensões, isto é, a coleção de 3^n quantidades identificadas pelo símbolo $T_{i_1 \dots i_n}[\hat{e}]$, o qual contém n índices. Então, a seguinte definição é realizada.

Definição 6.5 (Tensor Cartesiano de posto n). Seja \mathbb{E}^3 o espaço vetorial Euclidiano de dimensão 3 e $\hat{e} \doteq \{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ uma base ortonormal do mesmo. Seja $T[\hat{e}]$ a matriz cujos elementos

$$T_{i_1 i_2 \dots i_n}[\hat{e}], \quad (i_1, i_2, \dots, i_n = 1, 2, 3)$$

estão representados na base \hat{e} . A matriz $T[\hat{e}]$ representa um *tensor Cartesiano de posto n* se e somente se suas coordenadas se transformarem de acordo com a lei de transformação

$$T'_{i_1 i_2 \dots i_n} = L_{i_1 j_1} L_{i_2 j_2} \dots L_{i_n j_n} T_{j_1 j_2 \dots j_n} = \frac{\partial x'_{i_1}}{\partial x_{j_1}} \frac{\partial x'_{i_2}}{\partial x_{j_2}} \dots \frac{\partial x'_{i_n}}{\partial x_{j_n}} T_{j_1 j_2 \dots j_n} \quad (6.14a)$$

$$T_{i_1 \dots i_n} = L_{j_1 i_1} L_{j_2 i_2} \dots L_{j_n i_n} T'_{j_1 j_2 \dots j_n} = \frac{\partial x_{i_1}}{\partial x'_{j_1}} \frac{\partial x_{i_2}}{\partial x'_{j_2}} \dots \frac{\partial x_{i_n}}{\partial x'_{j_n}} T'_{j_1 j_2 \dots j_n}. \quad (6.14b)$$

Obviamente, a definição (6.14) contém os tensores de ordem zero, um e dois como casos particulares.

Uma das desvantagens das definições 6.4 e 6.5 está no fato de não ser evidente a invariância do tensor frente a transformação de coordenadas. Embora essa invariância possa ser demonstrada, é interessante apresentar também uma definição mais moderna, a qual independe do sistema de coordenadas e da base adotados para representar o tensor.

A definição a seguir interpreta um tensor de posto n como uma *forma multilinear*¹¹ que realiza a projeção do produto Cartesiano \mathcal{V}^n do espaço vetorial \mathcal{V} sobre o corpo subjacente \mathbb{K} .

Definição 6.6 (Espaço tensorial (Cartesiano)). Seja \mathcal{V} um espaço vetorial sobre o corpo \mathbb{K} . Seja

$$\mathcal{V}^n \doteq \underbrace{\mathcal{V} \times \dots \times \mathcal{V}}_{n \text{ vezes}}$$

o produto Cartesiano de \mathcal{V} . A forma multilinear

$$T : \mathcal{V}^n \longrightarrow \mathbb{K}$$

é denominada um *tensor Cartesiano sobre \mathcal{V}* . A potência n do produto \mathcal{V}^n determina o *posto* do tensor. Conjunto de tensores sobre \mathcal{V} podem formar espaços vetoriais denominados *espaços tensoriais Cartesianos*.

Retornando agora à discussão feita na página 277 sobre a concepção intuitiva de um campo vetorial, dado o espaço Euclidiano, sobre cada ponto do mesmo atribui-se um tensor Cartesiano, cujos elementos são contínuos e diferenciáveis, de tal forma que seja possível a aplicação de operadores íntegro-diferenciais sobre os mesmos.

Antes de se realizar a definição de um campo tensorial, é necessário definir um campo vetorial generalizado.

Definição 6.7 (Campo vetorial generalizado). Seja $\mathbb{E}^m = (\mathbb{E}^m, \mathcal{R}^m)$ o espaço vetorial Euclidiano de dimensão m e seja $U \subset \mathbb{E}^m$ uma região deste espaço. Sejam $A_1(\mathbf{r}), A_2(\mathbf{r}), \dots, A_n(\mathbf{r})$ um conjunto de n campos vetoriais de dimensão m e da classe $\mathcal{C}^r(U)$. Um *campo vetorial generalizado de ordem n e da classe \mathcal{C}^r sobre U* é a m^n -upla ordenada

$$\mathcal{A}(\mathbf{r}) \doteq (A_{11}(\mathbf{r}), \dots, A_{1m}(\mathbf{r}), A_{21}(\mathbf{r}), \dots, A_{ij}(\mathbf{r}), \dots, A_{nm}(\mathbf{r})),$$

formada pelos campos escalares $\{A_{ij}(\mathbf{r}) \in \mathbb{K}, \forall i = 1, \dots, n \text{ e } j = 1, \dots, m\}$.

Definição 6.8 (Campo tensorial Cartesiano). Seja $\mathbb{E}^m = (\mathbb{E}^m, \mathcal{R}^m)$ o espaço vetorial Euclidiano de dimensão m e seja $U \subset \mathbb{E}^m$ uma região deste espaço. Seja $\mathcal{T}(\mathbf{r})$ um campo vetorial generalizado de ordem n e da classe $\mathcal{C}^r(U)$. Sendo

$$T_{i_1, \dots, i_n}(\mathbf{r}), \quad (i_j = 1, \dots, m; j = 1, \dots, n),$$

uma componente deste espaço vetorial, se para todos $\{i_1, \dots, i_n = 1, \dots, m\}$ esta componente se transformar como um tensor Cartesiano de posto n (equações 6.14) frente a uma transformação de coordenadas em \mathbb{E}^m , então $\mathcal{T}(\mathbf{r})$ forma um *campo tensorial Cartesiano de posto n e da classe \mathcal{C}^r sobre U* .

¹¹Ver definição 4.12.

Serão apresentados agora alguns exercícios referentes a campos tensoriais.

Exercício 6.3. Mostre que os componentes da matriz

$$T = \begin{pmatrix} x_2^2 & -x_1x_2 \\ -x_1x_2 & x_1^2 \end{pmatrix}$$

formam um tensor de posto 2 frente a uma rotação dos eixos coordenados.

Solução. Dadas as relações entre as coordenadas $x'_1 = x_1 \cos \theta + x_2 \sin \theta$ e $x'_2 = -x_1 \sin \theta + x_2 \cos \theta$, espera-se que

$$\begin{aligned} T' &= \begin{pmatrix} (x'_2)^2 & -x'_1x'_2 \\ -x'_1x'_2 & (x'_1)^2 \end{pmatrix} \\ &= \begin{pmatrix} (-x_1 \sin \theta + x_2 \cos \theta)^2 & (x_1 \cos \theta + x_2 \sin \theta)(x_1 \sin \theta - x_2 \cos \theta) \\ (x_1 \cos \theta + x_2 \sin \theta)(x_1 \sin \theta - x_2 \cos \theta) & (x_1 \cos \theta + x_2 \sin \theta)^2 \end{pmatrix}. \end{aligned}$$

Para verificar, emprega-se a lei de transformação (6.13a), resultando

$$\begin{aligned} T'_{11} &= x_2^2 \cos^2 \theta - 2x_1x_2 \sin \theta \cos \theta + x_1^2 \sin^2 \theta \\ T'_{12} &= -x_2^2 \sin \theta \cos \theta - x_1x_2 \cos^2 \theta + x_1x_2 \sin^2 \theta + x_1^2 \sin \theta \cos \theta \\ T'_{21} &= T'_{12} \\ T'_{22} &= x_2^2 \sin^2 \theta + 2x_1x_2 \sin \theta \cos \theta + x_1^2 \cos^2 \theta. \end{aligned}$$

Ou seja, as expressões para os componentes são idênticas, o que confirma que T é de fato um tensor de posto 2.

Exercício 6.4 (Símbolos de Kronecker e Levi-Civita). Mostre que os símbolos de Kronecker (δ_{ij}) e de Levi-Civita (ϵ_{ijk}) apresentados na seção 6.1.2 são, na verdade, tensores de postos dois e três, respectivamente, sob o ponto de vista de rotações próprias no espaço.

Solução. Dados os símbolos δ_{ij} e ϵ_{ijk} , definidos conforme é descrito na seção 6.1.2 em um determinado sistema de coordenadas. Se estes objetos forem realmente tensores, respectivamente de segunda e terceira ordens, então a aplicação da lei de transformação geral (6.14) a ambos irá gerar outros tensores que possuem as mesmas propriedades no outro sistema de coordenadas.

Verificando esta propriedade para a delta de Kronecker, de (6.14) e (6.8) resulta

$$\delta'_{ij} = L_{ik}L_{jm}\delta_{km} = L_{ik}L_{jk} = \delta_{ij}.$$

Ou seja, a lei de transformação $R \rightarrow R'$ gera um símbolo de Kronecker δ'_{ij} que possui os mesmos componentes do sistema original. Portanto, δ_{ij} transforma-se da maneira esperada para um tensor de posto dois.

Verificando agora para o símbolo de Levi-Civita, novamente de (6.14) resulta

$$\epsilon'_{ijk} = L_{im}L_{jn}L_{kl}\epsilon_{mnl}.$$

Emprega-se agora a seguinte fórmula para o cálculo do determinante de uma matriz quadrada de ordem 3: sendo A uma matriz qualquer, então

$$\det(A) \epsilon_{ijk} = A_{im}A_{jn}A_{kl}\epsilon_{mnl}. \quad (6.15)$$

Substituindo na expressão acima os elementos de A pelos respectivos elementos da matriz de transformação L , resulta então que

$$\epsilon'_{ijk} = \det(L) \epsilon_{ijk}.$$

Portanto, de acordo com (6.9c), se a transformação realizada em R for própria, como é o caso de uma rotação dos eixos, então, $\epsilon'_{ijk} = \epsilon_{ijk}$ e o símbolo de Levi-Civita satisfaz a lei de transformação.

Exercício 6.5. Mostre que o tensor de stress de Maxwell, definido por

$$T_{ij} = \epsilon_0 \left[E_i E_j + c^2 B_i B_j - \frac{1}{2} (\mathbf{E} \cdot \mathbf{E} + c^2 \mathbf{B} \cdot \mathbf{B}) \delta_{ij} \right],$$

sendo \mathbf{E} e \mathbf{B} respectivamente os campos elétrico e de indução magnética, ϵ_0 a constante de permissividade elétrica do vácuo e c a velocidade da luz no vácuo, é realmente um tensor de posto dois.

Solução. Como ϵ_0 e c são escalares e \mathbf{E} e \mathbf{B} vetores por hipótese, então na transformação $R \rightarrow R'$, $E'_i = L_{ij}E_j$, $B'_i = L_{ij}B_j$ e $\mathbf{E} \cdot \mathbf{E}$ e $\mathbf{B} \cdot \mathbf{B}$ também são escalares. Portanto, se T_{ij} é um tensor de posto 2, então este deve satisfazer (6.13),

$$\begin{aligned} T'_{ij} &= L_{ik}L_{jm}T_{km} \\ &= \epsilon_0 \left[L_{ik}L_{jm}E_kE_m + c^2L_{ik}L_{jm}B_kB_m - \frac{1}{2}(\mathbf{E} \cdot \mathbf{E} + c^2\mathbf{B} \cdot \mathbf{B})L_{ik}L_{jm}\delta_{km} \right] \\ &= \epsilon_0 \left[E'_iE'_j + c^2B'_iB'_j - \frac{1}{2}(\mathbf{E}' \cdot \mathbf{E}' + c^2\mathbf{B}' \cdot \mathbf{B}')\delta_{ij} \right]. \end{aligned}$$

Ou seja, T_{ij} possui a mesma estrutura em qualquer sistema de coordenadas.

Um comentário final importante se refere à simbologia utilizada para representar um tensor de posto N . Um tensor de posto 2 pode ser representado naturalmente por uma matriz; assim, se o conjunto $\{T_{ij}\}$ compreende os 9 componentes de um tensor deste posto, este pode ser representado coletivamente por \mathbf{T} a qual é a notação empregada neste texto para uma matriz genérica.

Contudo, tensores de postos mais altos não podem ser representados por matrizes. Assim, se $\{T_{ijk}\}$ contém os 27 componentes de um tensor de posto 3, este pode ser representado de forma genérica como \mathcal{T} , sem mostrar explicitamente o seu posto.

A notação introduzida acima será empregada para representar um tensor genérico, quando o seu posto não for importante. Esta notação permite também tratar o tensor como um objeto geométrico genérico, onde não é feita menção do sistema de coordenadas em particular onde seus componentes são calculados.

Assim, o tensor \mathcal{T} tem os seus componentes dados por $\{T_{ijk}\}$ no sistema de coordenadas R , ao passo que os componentes do mesmo tensor no referencial rotado R' são dados por $\{T'_{ijk}\}$. Desta forma, o tensor pode ser representado de uma maneira semelhante aos operadores diferenciais gradiente (∇), divergente ($\nabla \cdot$) e rotacional ($\nabla \times$), cujos símbolos são os mesmos, independente do sistema adotado. Esta notação será aplicada na próxima seção.

6.4 ÁLGEBRA TENSORIAL

Como os tensores consistem em extensões lógicas dos conceitos de vetores e matrizes, é natural que as regras algébricas impostas a estes objetos sejam equivalentes àquelas aplicadas às estruturas mais simples citadas, e se reduzam a estas regras como casos particulares. Estas regras de álgebra serão discutidas nesta seção.

6.4.1 ADIÇÃO DE TENSORES

A adição (e também a subtração) de tensores é definida naturalmente como a extensão lógica das adições de vetores e matrizes. Para que esta generalização tenha sentido, contudo, deve-se impor como condição necessária que todos os tensores sendo adicionados tenham o mesmo posto.

Se $\{T_{ij\dots r}\}$ e $\{U_{ij\dots r}\}$ são os componentes dos tensores \mathcal{T} e \mathcal{U} , ambos de posto N , então a soma e a diferença destes geram, respectivamente, os tensores \mathcal{S} e \mathcal{D} de mesmo posto, cujos componentes são dados por

$$\begin{aligned} S_{ij\dots r} &= T_{ij\dots r} + U_{ij\dots r} \\ D_{ij\dots r} &= T_{ij\dots r} - U_{ij\dots r}. \end{aligned}$$

Estas operações podem ser representadas de forma genérica como $\mathcal{S} = \mathcal{T} + \mathcal{U}$ e $\mathcal{D} = \mathcal{T} - \mathcal{U}$, respectivamente.

6.4.2 SIMETRIA E ANTISSIMETRIA

É simples demonstrar que, se $\{T_{ij\dots k}\}$ são os componentes de um tensor, então o conjunto de funções obtido pela permutação de quaisquer dois índices do conjunto original, ou seja, $\{T_{ji\dots k}\}$, também será um tensor. Contudo, este tensor não irá apresentar, em geral, uma simetria bem

definida frente a permuta quaisquer dois de seus índices. Somente uma classe particular de tensores apresenta tais propriedades.

Restringindo inicialmente a discussão para tensores de posto 2, dado o \mathcal{T} , cujos componentes são $\{T_{ij}\}$. Se suas componentes apresentarem uma simetria bem definida frente a troca $i \leftrightarrow j$, esta pode se manifestar de duas maneiras:

$$\begin{aligned} \text{se } T_{ij} &= T_{ji} \rightsquigarrow \text{tensor } \mathbf{simétrico}; \\ \text{se } T_{ij} &= -T_{ji} \rightsquigarrow \text{tensor } \mathbf{antissimétrico}. \end{aligned}$$

Embora um outro tensor \mathcal{U} não necessariamente apresente propriedades de simetria bem definidas, este sempre pode ser escrito como uma combinação de suas partes simétrica e antissimétrica através da identidade

$$U_{ij} = \frac{1}{2}(U_{ij} + U_{ji}) + \frac{1}{2}(U_{ij} - U_{ji}),$$

sendo que o primeiro termo a *parte simétrica* de U_{ij} , enquanto que o segundo termo é a sua *parte antissimétrica*.

Considerando agora um tensor de posto N qualquer, seja $\{T_{ijk\dots r}\}$ o conjunto de componentes de \mathcal{T} . Se este tensor apresentar propriedade de simetria, então há duas possibilidades:

$$T_{ijk\dots r} = T_{jik\dots r}, \quad T_{ijk\dots r} = T_{kji\dots r}, \text{ etc}$$

é **simétrico** com respeito aos índices i e j ou i e k , *etc*; ou,

$$T_{ijk\dots r} = -T_{jik\dots r}, \quad T_{ijk\dots r} = -T_{kji\dots r}, \text{ etc}$$

é **antissimétrico** com respeito aos mesmos índices.

Da mesma maneira, um tensor genérico \mathcal{U} pode sempre ser escrito como uma combinação de suas partes simétrica e antissimétrica, frente uma permutação de dois índices quaisquer, como

$$U_{ijk\dots r} = \frac{1}{2}(U_{ijk\dots r} + U_{jik\dots r}) + \frac{1}{2}(U_{ijk\dots r} - U_{jik\dots r}),$$

ou

$$U_{ijk\dots r} = \frac{1}{2}(U_{ijk\dots r} + U_{kji\dots r}) + \frac{1}{2}(U_{ijk\dots r} - U_{kji\dots r}),$$

etc.

A separação de tensores ou operadores em suas partes simétrica e antissimétrica é de extrema importância em diversas disciplinas, tais como mecânica e eletromagnetismo de meios contínuos e mecânica quântica.

6.4.3 TENSORES HERMITIANOS OU ANTI-HERMITIANOS

Uma propriedade adicional relacionada a tensores de posto dois e cujas componentes são funções analíticas complexas é a hermiticidade dos mesmos.

Usualmente representados por matrizes, tensores de posto dois também são *Hermitianos* ou *anti-Hermitianos* se apresentarem simetria frente a operação de **conjugação Hermitiana**, definida da seguinte maneira. Seja A uma matriz quadrada cujos elementos $\{A_{ij}\}$ são complexos. O seu *Hermitiano conjugado* ou *adjunto*, denotado por A^\dagger , é obtido a partir da conjugação complexa dos elementos da matriz transposta, isto é,

$$A^\dagger = (A^*)^T = (A^T)^*,$$

sendo que os elementos de A^\dagger são dados por

$$(A^\dagger)_{ij} = A_{ji}^*.$$

Como é sempre possível usar uma representação matricial para um tensor de posto dois, em muitas aplicações físicas é aplicada a operação de conjugação Hermitiana ao mesmo.

Da mesma forma como é definido para matrizes, portanto, um tensor de posto dois \mathcal{A} é *Hermitiano* ou *autoadjunto* se satisfaz a propriedade

$$\mathcal{A} = \mathcal{A}^\dagger.$$

Por outro lado, este tensor é *anti-Hermitiano* se

$$\mathcal{A} = -\mathcal{A}^\dagger.$$

É possível então definir as partes Hermitiana e anti-Hermitiana de um tensor \mathcal{A} respectivamente por

$$\begin{aligned}\mathcal{A}^H &= \frac{1}{2} (\mathcal{A} + \mathcal{A}^\dagger) \\ \mathcal{A}^{aH} &= \frac{1}{2i} (\mathcal{A} - \mathcal{A}^\dagger).\end{aligned}$$

Desta forma, sempre se pode escrever \mathcal{A} como

$$\mathcal{A} = \mathcal{A}^H + i\mathcal{A}^{aH}.$$

Em muitos meios descritos por tensores complexos, as respectivas partes Hermitiana e anti-Hermitiana estão relacionadas com processos físicos distintos que ocorrem neste meio.

Exemplo 6.1. Em um meio contínuo, linear e homogêneo, mas anisotrópico e dissipativo, a relação constitutiva entre o campo elétrico \mathbf{E} e o vetor deslocamento elétrico \mathbf{D} pode ser escrita, no espaço de Fourier, como

$$D_i(\mathbf{k}, \omega) = \varepsilon_{ij}(\mathbf{k}, \omega) E_j(\mathbf{k}, \omega),$$

onde \mathbf{k} e ω são, respectivamente, o vetor de onda e a frequência angular das ondas que se propagam neste meio e $\{\varepsilon_{ij}\}$ são as componentes do tensor dielétrico do meio. Para este meio, o teorema de Poynting, que descreve a conservação de energia entre campos e partículas, pode ser escrito como

$$\frac{\partial U}{\partial t} + \nabla \cdot (\mathbf{v}_g U) = -\frac{\omega}{8\pi} \mathbf{E} \cdot \varepsilon^{aH} \cdot \mathbf{E},$$

sendo $\mathbf{v}_g = \partial\omega/\partial\mathbf{k}$ a velocidade de grupo das ondas e¹²

$$U = \frac{1}{8\pi} \left[\frac{\partial (\omega \varepsilon^H)}{\partial \omega} \right] : \mathbf{E} \mathbf{E}$$

a densidade de energia contida nos campos.

Na equação acima, o termo $\mathbf{v}_g U = \mathbf{S}$ corresponde ao vetor de Poynting, ou seja, ao fluxo local de energia. Conclui-se, portanto, que a parte Hermitiana de ε_{ij} está relacionada com a refração ou ao fluxo da energia transportada pelas ondas, ao passo que a parte anti-Hermitiana está relacionada com a dissipação irreversível de energia devido a interação dos campos com o meio.

6.4.4 PRODUTO EXTERNO DE TENSORES

O *produto externo* de dois ou mais tensores, também denominado *produto direto* ou *produto tensorial*, consiste em uma operação que permite a construção de um tensor com posto mais alto que os tensores-pais. De fato, o posto do tensor resultante é exatamente igual à soma dos postos de seus progenitores.

O produto externo é realizado de diferentes maneiras. A maneira usual consiste na multiplicação direta de componentes dos tensores envolvidos. Uma outra maneira de se construir um tensor de posto mais alto consiste em se aplicar os componentes de um operador sobre os componentes de outro tensor. Ambas as maneiras serão abordadas nesta seção.

6.4.4.1 PRODUTO EXTERNO DE DOIS TENSORES

Considera-se inicialmente um caso mais simples. Dados os vetores $\mathbf{a} = a_i \hat{e}_i$ e $\mathbf{b} = b_i \hat{e}_i$, deseja-se contruir a partir destes um tensor de posto dois \mathcal{T} , cujos componentes são dados pelo produto externo dos vetores acima, ou seja,

$$T_{ij} = a_i b_j.$$

¹²Ver notação de diádicas e seus produtos nas seções 6.4.4.2 e 6.4.5.3.

Os componentes T_{ij} resultantes desta construção de fato pertencem ao tensor \mathcal{T} , o que pode ser facilmente demonstrado. Se isto for verdade, então na transformação $R \rightarrow R'$ devem resultar os componentes

$$T'_{ij} = a'_i b'_j.$$

Usando a lei de transformação para os vetores, resulta

$$T'_{ij} = L_{ik} L_{j\ell} a_k b_\ell = L_{ik} L_{j\ell} T_{k\ell},$$

a qual é justamente a lei de transformação de um tensor do posto 2.

Na sua representação matricial, o tensor \mathcal{T} pode ser escrito como a matriz 3×3

$$\mathbb{T} = \begin{pmatrix} a_1 b_1 & a_1 b_2 & a_1 b_3 \\ a_2 b_1 & a_2 b_2 & a_2 b_3 \\ a_3 b_1 & a_3 b_2 & a_3 b_3 \end{pmatrix}.$$

Este tipo de produto externo é representado de diferentes maneiras, todas independentes do sistema de coordenadas adotado. Uma maneira usual consiste em empregar o símbolo \otimes . Nesta notação, dados os vetores \mathbf{a} e \mathbf{b} acima, o tensor \mathcal{T} é representado como

$$\mathcal{T} = \mathbf{a} \otimes \mathbf{b} = a_i b_j \hat{e}_i \otimes \hat{e}_j, \quad (6.16a)$$

e as componentes resultantes deste tensor também podem ser representadas por

$$T_{ij} = (\mathbf{a} \otimes \mathbf{b})_{ij}. \quad (6.16b)$$

6.4.4.2 DIÁDICAS

No caso particular de tensores de posto dois, uma outra representação, um tanto ultrapassada mas ainda empregada em alguns textos, é a de uma diádica. Esta é uma outra maneira de se representar o produto externo de dois vetores, estendendo a álgebra vetorial usual e resultando em um tensor. Dados os vetores \mathbf{a} e \mathbf{b} , a diádica \mathcal{T} é obtida pela simples justaposição destes vetores, ou seja,

$$\mathcal{T} = \mathbf{a}\mathbf{b} = a_i b_j \hat{e}_i \hat{e}_j.$$

Uma vantagem que esta notação possui está na praticidade de realização de produtos de diádicas entre si ou com vetores, como será mostrado a seguir. Contudo, em se tratando de uma prática em desuso, será dada preferência para a notação de produto externo introduzida em (6.16).

6.4.4.3 GRADIENTE DE UM VETOR

Uma outra maneira de se construir um tensor de posto 2 consiste em aplicar o operador gradiente sobre um vetor, operação esta que não está definida na álgebra vetorial ordinária.

Dado o operador $\nabla = \hat{e}_i \partial / \partial x_i$ e o vetor $\mathbf{v} = v_i \hat{e}_i$, o componente T_{ij} do tensor \mathcal{T} é definido como

$$T_{ij} = \frac{\partial v_j}{\partial x_i}.$$

Esta expressão pode ser escrita de uma forma genérica, independente do sistema de coordenadas adotado, como

$$\mathcal{T} = \nabla \mathbf{v} = \frac{\partial v_j}{\partial x_i} \hat{e}_i \hat{e}_j.$$

Neste caso em particular, a notação empregada é em muito semelhante ao de uma diádica, exceto que é feito uso de um operador diferencial no lugar de um vetor ordinário.

6.4.4.4 PRODUTO EXTERNO EM GERAL

Considera-se agora o caso mais geral de produto externo entre dois tensores arbitrários. Dado o tensor \mathcal{T} de posto M e o tensor \mathcal{U} de posto N , o produto externo destes gera o tensor \mathcal{V} com posto $M + N$, cujas componentes são

$$\underbrace{V_{i \dots pq \dots}}_{M+N \text{ índices}} = \underbrace{T_{i \dots p}}_{M \text{ índices}} \underbrace{U_{q \dots}}_{N \text{ índices}}.$$

A notação para este produto externo é, então

$$\mathcal{V} = \mathcal{T} \otimes \mathcal{U}.$$

6.4.5 CONTRAÇÃO E PRODUTO INTERNO

A definição do produto externo, na seção 6.4.4, possibilitou a construção de um tensor de um dado posto a partir de outros tensores de ordem mais baixa. Processos inversos também são definidos na álgebra tensorial, ou seja, operações aplicadas sobre um determinado tensor de posto N que reduzem o seu posto, resultando em outro tensor de posto $N - 2$. Uma destas operações é denominada de *contração de índices* ou, simplesmente, *contração*. Uma outra operação, denominada *produto interno*, está relacionada com a contração dos índices e será também discutida. Obviamente, para que esta operação tenha sentido, é necessário que $N \geq 2$.

6.4.5.1 CONTRAÇÃO

Dado um tensor \mathcal{T} de posto $N \geq 2$, uma contração deste tensor consiste em tomar qualquer par de seus índices, representá-los com o mesmo símbolo e executar então a sua soma implícita. A resultante desta operação será um novo tensor \mathcal{U} de posto $N - 2$. Dado então o conjunto $\{T_{ijkl\dots r}\}$, as expressões a seguir apresentam algumas de suas possíveis contrações:

$$U_{kl\dots r} = T_{iikl\dots r}, \quad V_{il\dots r} = T_{ijjl\dots r}, \quad W_{jl\dots r} = T_{ijil\dots r}, \quad X_{ikl\dots} = T_{ijkl\dots j}, \quad \text{etc.}$$

Claramente, o número de contrações possíveis depende do posto do tensor. Se $N = 2$, somente uma contração é possível. Neste caso, a operação é equivalente ao cálculo do traço da matriz que representa o tensor e o resultado da operação é uma quantidade escalar. Se o posto for $N = 3$, há 3 contrações possíveis, gerando o mesmo número de vetores distintos. Em geral, um tensor de posto $N \geq 2$ pode gerar por contrações um total de $N!/2 (N - 2)!$ tensores de posto $N - 2$.

Para mostrar que o objeto resultante da contração é realmente um tensor de posto $N - 2$, parte-se da lei de transformação (6.14a) aplicada a \mathcal{T} ,

$$T'_{ij\dots\ell\dots m\dots k} = L_{ip}L_{jq}\dots L_{lr}\dots L_{ms}\dots L_{kn}T_{pq\dots r\dots s\dots n}.$$

Contraindo os índices ℓ e m no componente transformado à esquerda e utilizando a condição de ortogonalidade (6.8), resulta

$$\begin{aligned} T'_{ij\dots\ell\dots\ell\dots k} &= L_{ip}L_{jq}\dots L_{lr}\dots L_{ls}\dots L_{kn}T_{pq\dots r\dots s\dots n} \\ &= L_{ip}L_{jq}\dots \delta_{rs}\dots L_{kn}T_{pq\dots r\dots s\dots n} \\ &= \underbrace{L_{ip}L_{jq}\dots L_{kn}}_{N-2 \text{ fatores}} T_{pq\dots r\dots r\dots n}, \end{aligned}$$

ou seja,

$$U'_{ij\dots k} = L_{ip}L_{jq}\dots L_{kn}U_{pq\dots n},$$

ou que mostra que \mathcal{U} é realmente um tensor de posto $N - 2$.

6.4.5.2 PRODUTO INTERNO

O produto interno, conforme definido em espaços vetoriais, é bem conhecido. Se a e b são membros do espaço vetorial \mathcal{V} , então o produto interno entre ambos, usualmente representado por $\langle a, b \rangle$, terá como resultado uma quantidade escalar.

A forma mais empregada do produto interno para aplicações físicas é o *produto escalar* entre dois vetores a e b do \mathbb{E}^3 ,¹³ resultando no escalar c ,

$$c = a \cdot b. \tag{6.17}$$

De acordo com a terminologia aqui empregada, trata-se de um produto entre dois tensores de posto um, resultando em um tensor de posto zero.

Estendendo o conceito de produto interno a tensores de posto mais alto e relacionando esta operação com a definição de contração de índices anteriormente apresentada, pode-se dizer que a operação realizada em (6.17) consiste, inicialmente, no produto externo dos vetores a e b , resultando em um tensor de posto 2, seguida posteriormente pela contração dos índices deste tensor, o que resulta finalmente em um escalar:

$$C = a \otimes b \rightsquigarrow c = C_{ii} = a_i b_i.$$

¹³Exemplo 4.3.

A única característica própria nesta operação em relação ao processo geral de contração de índices está no fato de que no produto interno a contração é realizada sobre os índices de tensores-pais (de postos M e N) distintos, os quais foram inicialmente justapostos por uma operação de produto externo, resultando em um tensor-filho de ordem $M + N$. Como resultado do produto interno, igualmente resulta um tensor de posto $M + N - 2$.

Assim, se \mathcal{A} , \mathcal{B} , \mathcal{C} e \mathcal{D} são tensores de postos M , N , $M + N$ e $M + N - 2$, respectivamente a seguinte cadeia de operações ilustra a o produto interno entre \mathcal{A} e \mathcal{B} , resultando no tensor \mathcal{D} ,

$$\begin{aligned} \mathcal{C} = \mathcal{A} \otimes \mathcal{B} &\rightsquigarrow \underbrace{C_{ijk\dots r}}_M \underbrace{\ell mn\dots s}_N = A_{ijk\dots r} B_{\ell mn\dots s} \\ &\implies \underbrace{D_{ik\dots r mn\dots s}}_{M+N-2} \equiv C_{ijk\dots r} B_{mn\dots s} = A_{ijk\dots r} B_{jmn\dots s}. \end{aligned}$$

No caso acima, foi realizada a contração no índices j de cada tensor. Obviamente, a contração em outros índices irá gerar tensores de posto $M + N - 2$ em geral distintos de \mathcal{D} . Por conseguinte, há diversos produtos internos possíveis entre tensores de ordem maior que dois.

6.4.5.3 PRODUTOS COM DIÁDICAS

A definição de produto interno acima realizada é particularmente útil na notação de diádicas (seção 6.4.4.2). Sendo a e b vetores e \mathcal{C} uma diádica, formada pela justaposição dos vetores,

$$\mathcal{C} = \mathbf{ab} = a_i b_j \hat{e}_i \hat{e}_j,$$

Esta poderá sofrer produto interno com um outro vetor c , tanto à esquerda quanto à direita, resultando nos vetores d e e , em geral distintos entre si. Além disso, o produto interno envolvendo diádicas utiliza a mesma notação do produto escalar entre vetores. Assim,

$$\begin{aligned} d &= c \cdot \mathcal{C} = c_i a_i b_j \hat{e}_j \\ e &= \mathcal{C} \cdot c = a_i b_j c_j \hat{e}_i. \end{aligned}$$

Na verdade, no caso específico de produtos envolvendo diádicas é possível também definir o *produto vetorial* de uma diádica por um vetor, com a mesma definição empregada na álgebra vetorial.¹⁴ Agora, o resultando desta operação é uma outra diádica e também são possíveis produtos à esquerda ou à direita. Assim, sendo \mathcal{F} e \mathcal{G} diádicas, o produto vetorial de c por \mathcal{C} pela esquerda ou pela direita resultam em

$$\begin{aligned} \mathcal{F} = c \times \mathcal{C} &= (c \times a) b = c_i a_j b_k (\hat{e}_i \times \hat{e}_j) \hat{e}_k = \epsilon_{ilk} a_k b_j c_\ell \hat{e}_i \hat{e}_j \\ \mathcal{G} = \mathcal{C} \times c &= a (b \times c) = a_i b_j c_k \hat{e}_i (\hat{e}_j \times \hat{e}_k) = \epsilon_{jkl} a_i b_k c_\ell \hat{e}_i \hat{e}_j, \end{aligned}$$

onde se fez uso de (6.5). Portanto, há quatro operações possíveis envolvendo uma diádica e um vetor.

Existem também 7 possíveis produtos, tanto escalares quanto vetoriais, envolvendo duas diádicas. Considerando agora os vetores a , b , c e d e as diádicas $\mathcal{A} = ab$ e $\mathcal{B} = cd$, estas são apresentadas na tabela 6.2.

6.4.6 REGRA DO QUOCIENTE

Como consequência da definição de produto externo na seção 6.4.4, foi demonstrado que se \mathcal{A} e \mathcal{B} são tensores de postos M e N , respectivamente, então o produto externo destes gera o tensor $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$, de posto $M + N$.

A regra do quociente consiste na propriedade recíproca do produto externo. Assumindo agora que \mathcal{B} e \mathcal{C} são tensores de postos N e $M + N$, respectivamente, caso exista um conjunto de objetos $\{A_{ij\dots k\dots m}\}$, identificados por M índices, tal que

$$\underbrace{A_{ij\dots k\dots m}}_{M \text{ índices}} \underbrace{B_{np\dots q\dots r}}_{N \text{ índices}} = \underbrace{C_{ij\dots k\dots mn p\dots q\dots r}}_{M+N \text{ índices}}$$

então $\{A_{ij\dots k\dots m}\}$ necessariamente compõe os componentes do tensor \mathcal{A} de posto M .

¹⁴Exemplo 4.6.

Tabela 6.2: Possíveis operações de produto entre duas diádicas.

	Produtos escalares	Produtos vetoriais
Operação única	Produto escalar $\mathcal{A} \cdot \mathcal{B} = (ab) \cdot (cd) = (b \cdot c) ad$	N/A
Contração de índices	O fator de expansão consiste na contração de índices usual: $\text{Tr}(\mathcal{A}) = a \cdot b = a_i b_i$	O fator de rotação origina-se do produto vetorial entre os vetores da diádica: $A \equiv \langle \mathcal{A} \rangle = a \times b$ $= a_i b_j \hat{e}_i \times \hat{e}_j$ $= \epsilon_{ikj} a_k b_j \hat{e}_i.$
Operação dupla incluindo produto escalar	Produto escalar duplo $\mathcal{A} : \mathcal{B} = (ab) : (cd)$ $= a_i b_j c_k d_\ell (\hat{e}_i \hat{e}_j) : (\hat{e}_k \hat{e}_\ell)$ $= a_i b_j c_k d_\ell (\hat{e}_j \cdot \hat{e}_k) (\hat{e}_i \cdot \hat{e}_\ell)$ $= (a \cdot d) (b \cdot c)$	Produto escalar-vetorial $\mathcal{A} \dot{\times} \mathcal{B} = (ab) \dot{\times} (cd) = (a \cdot c) (b \times d)$
Operação dupla incluindo produto vetorial	Produto vetorial-escalar $\mathcal{A} \times \mathcal{B} = (ab) \times (cd) = (a \times c) (b \cdot d)$	Duplo produto vetorial $\mathcal{A} \times \times \mathcal{B} = (ab) \times \times (cd) = (a \times c) (b \times d)$

Demonstração. A demonstração deste teorema segue da lei de transformação de tensores. Como \mathcal{B} (posto N) e \mathcal{C} (posto $M + N$) são, por hipótese, tensores, então a transformação $R \rightarrow R'$ leva os componentes do conjunto $\{A_{\dots}\}$ a um outro conjunto de valores $\{A'_{\dots}\}$, ainda não determinados. Como consequência, existe a seguinte relação no sistema R' ,

$$A'_{ij\dots k\dots m} B'_{np\dots q\dots r} = C'_{ij\dots k\dots mnp\dots q\dots r}.$$

Dada agora a lei de transformação (6.14a),

$$A'_{ij\dots k\dots m} L_{na} L_{pb} \dots L_{qc} \dots L_{rd} B_{ab\dots c\dots d} = L_{ie} L_{jh} \dots L_{ks} \dots L_{mv} L_{na} L_{pb} \dots L_{qc} \dots L_{rd} C_{eh\dots s\dots vab\dots c\dots d},$$

$$L_{na} L_{pb} \dots L_{qc} \dots L_{rd} (A'_{ij\dots k\dots m} B_{ab\dots c\dots d} - L_{ie} L_{jh} \dots L_{ks} \dots L_{mv} C_{eh\dots s\dots vab\dots c\dots d}) = 0.$$

Como a transformação é arbitrária, a equação acima somente pode ser satisfeita para todos os componentes se

$$A'_{ij\dots k\dots m} B_{ab\dots c\dots d} - L_{ie} L_{jh} \dots L_{ks} \dots L_{mv} C_{eh\dots s\dots vab\dots c\dots d} = 0.$$

Introduzindo agora a relação original entre os conjuntos $\{A_{\dots}\}$, $\{B_{\dots}\}$ e $\{C_{\dots}\}$,

$$A'_{ij\dots k\dots m} B_{ab\dots c\dots d} - L_{ie} L_{jh} \dots L_{ks} \dots L_{mv} A_{eh\dots s\dots v} B_{ab\dots c\dots d} = 0,$$

$$(A'_{ij\dots k\dots m} - L_{ie} L_{jh} \dots L_{ks} \dots L_{mv} A_{eh\dots s\dots v}) B_{ab\dots c\dots d} = 0.$$

Como esta identidade deve ser satisfeita para qualquer tensor \mathcal{B} , então, necessariamente,

$$A'_{ij\dots k\dots m} = L_{ie} L_{jh} \dots L_{ks} \dots L_{mv} A_{eh\dots s\dots v}.$$

Ou seja, a relação entre os conjuntos $\{A_{\dots}\}$ e $\{A'_{\dots}\}$ é idêntica à lei de transformação (6.14a). Portanto, \mathcal{A} é realmente um tensor de posto M . \square

O uso da regra do quociente para determinar se um determinado conjunto de objetos é de fato um tensor consiste, muitas vezes, em uma maneira mais conveniente de se cumprir este objetivo do que a aplicação direta das leis de transformação. Uma maneira particular de se realizar esta tarefa consiste em efetuar contração de índices de forma a obter-se um escalar ou um tensor de posto mais baixo.

A mesma conclusão obtida acima com relação ao produto externo de dois tensores também é válida para o seu produto interno. Sendo novamente $\{A_{ij\dots k\dots m}\}$ um conjunto de M objetos e \mathcal{B} e \mathcal{D} tensores de postos N e $M + N - 2$, respectivamente, então a relação

$$A_{ij\dots k\dots m} B_{np\dots k\dots r} = D_{ij\dots mnp\dots r}$$

implica em que $\{A_{\dots}\}$ forma o conjunto de componentes de um tensor de posto M .

Demonstração. A demonstração segue novamente da relação equivalente no sistema transformado,

$$\begin{aligned} A'_{ij\dots\ell\dots m} B'_{np\dots\ell\dots r} &= D'_{ij\dots mnp\dots r}, \\ L_{na} L_{pb} \dots L_{rd} (L_{\ell k} A'_{ij\dots\ell\dots m} B_{ab\dots k\dots d} - L_{ie} L_{jh} \dots L_{mv} D_{eh\dots vab\dots d}) &= 0, \\ L_{\ell k} A'_{ij\dots\ell\dots m} B_{ab\dots k\dots d} - L_{ie} L_{jh} \dots L_{mv} D_{eh\dots vab\dots d} &= 0, \end{aligned}$$

$$\begin{aligned} L_{\ell k} A'_{ij\dots\ell\dots m} B_{ab\dots k\dots d} - L_{ie} L_{jh} \dots L_{mv} A_{eh\dots k\dots v} B_{ab\dots k\dots d} &= 0, \\ (L_{\ell k} A'_{ij\dots\ell\dots m} - L_{ie} L_{jh} \dots L_{mv} A_{eh\dots k\dots v}) B_{ab\dots k\dots d} &= 0, \\ L_{\ell k} A'_{ij\dots\ell\dots m} - L_{ie} L_{jh} \dots L_{mv} A_{eh\dots k\dots v} &= 0. \end{aligned}$$

Usando a condição de ortogonalidade (6.8) no resultado acima, resulta

$$\begin{aligned} L_{nk} L_{\ell k} A'_{ij\dots\ell\dots m} - L_{ie} L_{jh} \dots L_{nk} \dots L_{mv} A_{eh\dots k\dots v} &= 0, \\ A'_{ij\dots n\dots m} &= L_{ie} L_{jh} \dots L_{nk} \dots L_{mv} A_{eh\dots k\dots v}, \end{aligned}$$

a qual é a lei de transformação do tensor \mathcal{A} de posto M . □

Obviamente, como há mais de uma contração de índices possível para tensores de posto maior que um, o produto interno ilustrado acima poderia ser realizado em qualquer par de índices dos tensores \mathcal{A} e \mathcal{B} .

Exercício 6.6. Use a lei do quociente para mostrar que a matriz \mathbb{T} do exercício 6.3 é realmente a representação de um tensor de posto dois.

Solução. Dado que a diádica rr é um tensor de posto dois, realiza-se a contração dupla de índices das componentes desta diádica com a matriz \mathbb{T} , o que pode ser representado pela seguinte multiplicação matricial,

$$T_{ij} x_i x_j = (x_1 \ x_2) \begin{pmatrix} x_1^2 & -x_1 x_2 \\ -x_1 x_2 & x_2^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \ x_2) \begin{pmatrix} x_1 x_2^2 - x_1 x_2^2 \\ -x_1^2 x_2 + x_1^2 x_2 \end{pmatrix} = 0.$$

Como o resultado desta contração é um escalar e rr é um tensor, então, pela lei do quociente, \mathcal{T} também deve ser um tensor.

Exemplo 6.2. Considera-se uma partícula de massa m rigidamente conectada à origem do referencial. Se r e p forem, respectivamente, a posição e momentum linear instantâneos da partícula, então a i -ésima componente de seu momentum angular é dada por

$$L_i = (\mathbf{r} \times \mathbf{p})_i = \epsilon_{ijk} x_j p_k.$$

Como a partícula está presa a uma distância fixa da origem O , então

$$\mathbf{p} = m\mathbf{v} = m(\boldsymbol{\omega} \times \mathbf{r}),$$

sendo $\boldsymbol{\omega}$ a sua velocidade angular instantânea. Portanto, L_i pode ser escrito, fazendo uso de (6.3), como

$$L_i = \epsilon_{ijk} x_j p_k = m \epsilon_{ijk} \epsilon_{klm} x_j x_m \omega_l = m (r^2 \omega_i - x_i x_j \omega_j) = m (r^2 \delta_{ij} - x_i x_j) \omega_j \equiv I_{ij} \omega_j.$$

A quantidade

$$I_{ij} = m (r^2 \delta_{ij} - x_i x_j)$$

é denominada *tensor de inércia* da partícula em relação a O .

O conjunto $\{I_{ij}\}$ ($i, j = 1, 2, 3$) realmente compõe um tensor de posto dois, uma vez que L e $\boldsymbol{\omega}$ são tensores de posto um e, de acordo com a lei do quociente, segue então que \mathcal{I} é necessariamente um tensor de posto dois.

6.5 COMPOSIÇÃO DE TRANSFORMAÇÕES, ROTAÇÕES INFINITESIMAIS E TENSORES ISOTRÓPICOS

Nesta seção serão discutidos alguns tópicos de interesse geral.

6.5.1 COMPOSIÇÃO DE TRANSFORMAÇÕES

Quando mais de uma transformação é realizada sobre o sistema de coordenadas de forma concomitante, a transformação resultante sempre pode ser escrita em termos de uma única matriz de transformação.

Restringindo a tensores de posto um, realizam-se duas transformações concomitantes ao referencial,

$$R \xrightarrow{L^{(1)}} R' \xrightarrow{L^{(2)}} R''.$$

Dado então um vetor a , este será transformado como

$$a'_i = L^{(1)}_{ij} a_j \rightsquigarrow a''_i = L^{(2)}_{ij} a'_j = L^{(2)}_{ij} L^{(1)}_{jk} a_k.$$

Ou seja, as transformações concomitantes irão corresponder a uma única transformação

$$a''_i = L_{ij} a_j,$$

cujas matriz de transformação é

$$L_{ij} = L^{(2)}_{ik} L^{(1)}_{kj} \longrightarrow L = L^{(2)} L^{(1)}.$$

Em particular, duas rotações seguidas em torno do eixo x_3 , de ângulos θ_1 e θ_2 , consecutivamente, correspondem a uma única rotação por um ângulo $\theta_1 + \theta_2$. Pode-se verificar facilmente que, neste caso, estas rotações podem ser realizadas em qualquer ordem.

O exemplo acima, contudo, é um caso particular onde as transformações são permutáveis. Em geral, a multiplicação matricial não comuta, isto é, $L^{(2)} L^{(1)} \neq L^{(1)} L^{(2)}$.

6.5.2 ROTAÇÕES INFINITESIMAIS

Rotações arbitrárias são um exemplo de transformações não comutáveis. Dadas duas matrizes de rotações gerais $L^{(1)}$ e $L^{(2)}$, em geral a ordem de realização das mesmas é importante.

Uma forma comum de se realizar uma rotação arbitrária de eixos consiste no emprego dos *ângulos de Euler*, os quais são ilustrados na figura 6.4. Com o uso destes ângulos, pode-se construir uma matriz genérica de rotação, a qual é uma função somente dos parâmetros ϕ , θ e ψ . Fazendo isso, a expressão obtida para a matriz de rotação E é

$$\begin{aligned} E_{11} &= \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & E_{12} &= \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & E_{13} &= \sin \psi \sin \theta \\ E_{21} &= -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & E_{22} &= -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & E_{23} &= \cos \psi \sin \theta \\ E_{31} &= \sin \theta \sin \phi & E_{32} &= -\sin \theta \cos \phi & E_{33} &= \cos \theta. \end{aligned}$$

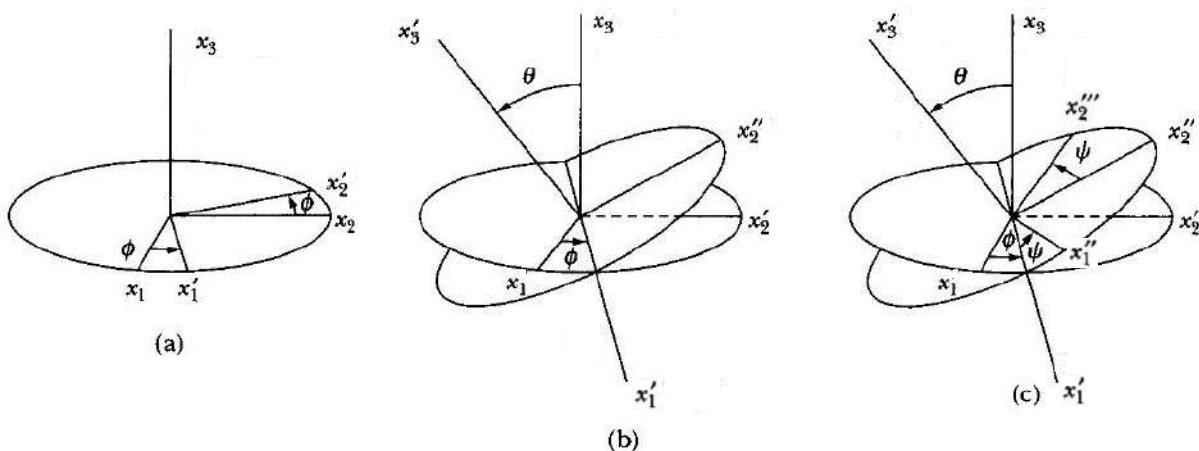


Figura 6.4: Ângulos de Euler usualmente empregados para executar uma rotação arbitrária no sistema de coordenadas. (a) Primeira rotação: sentido anti-horário em torno de x_3 , por um ângulo ϕ . (b) Segunda rotação: sentido anti-horário por um ângulo θ em torno do eixo x'_1 . (c) Terceira rotação: sentido anti-horário por um ângulo ψ em torno de x''_3 .

Na técnica dos ângulos de Euler, a matriz de rotação geral é construída a partir de três rotações consecutivas, realizadas sempre no sentido anti-horário e em torno de um determinado eixo de rotação.

Os ângulos de Euler são apenas uma das formas de se implementar uma rotação arbitrária em três dimensões. De uma forma genérica, estas rotações parciais são realizadas com as seguintes matrizes

$$L^{(1)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_1 & \sin \theta_1 \\ 0 & -\sin \theta_1 & \cos \theta_1 \end{pmatrix} \quad L^{(2)} = \begin{pmatrix} \cos \theta_2 & 0 & -\sin \theta_2 \\ 0 & 1 & 0 \\ \sin \theta_2 & 0 & \cos \theta_2 \end{pmatrix} \quad L^{(3)} = \begin{pmatrix} \cos \theta_3 & \sin \theta_3 & 0 \\ -\sin \theta_3 & \cos \theta_3 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

as quais correspondem a rotações em torno dos eixos x_1 , x_2 e x_3 , respectivamente. Pode-se verificar facilmente que a composição de rotações arbitrárias não é comutável. Por exemplo, $L^{(1)}L^{(2)} \neq L^{(2)}L^{(1)}$.

As formas de todas estas matrizes já foram derivadas no exemplo 5.30 como representações do grupo $SO(3)$. As derivações se basearam nas formas das matrizes frente a rotações infinitesimais, conforme discutido a seguir.

Quando as rotações são *infinitesimais*, a comutatividade nas rotações é obtida. Considera-se como exemplo uma rotação infinitesimal $\delta\theta_3$ em torno de x_3 . Neste caso, pode-se substituir em $L^{(3)}$, $\cos \theta_3 \simeq 1$ e $\sin \theta_3 \simeq \delta\theta_3$ e esta matriz pode ser decomposta como

$$L^{(3)} \simeq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & \delta\theta_3 & 0 \\ -\delta\theta_3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ou seja, as componentes de um vetor a transformam-se de acordo com

$$a'_i = L_{ij}^{(3)} a_j \simeq (\delta_{ij} + \epsilon_{ij3} \delta\theta_3) a_j.$$

Realizando a mesma consideração para as outras matrizes de rotação, observa-se que a forma infinitesimal das mesmas é

$$L_{ij}^{(1)} = \delta_{ij} + \epsilon_{ij1} \delta\theta_1, \quad L_{ij}^{(2)} = \delta_{ij} + \epsilon_{ij2} \delta\theta_2, \quad L_{ij}^{(3)} = \delta_{ij} + \epsilon_{ij3} \delta\theta_3.$$

Assim, é possível compor uma rotação infinitesimal arbitrária, dada por $L = L^{(3)}L^{(2)}L^{(1)}$ a qual, em ordem mais baixa nas rotações, é dada por

$$L_{ij} = L_{ik}^{(3)} L_{kl}^{(2)} L_{lj}^{(1)} \simeq \delta_{ij} + \epsilon_{ijk} \delta\theta_k, \quad (6.18)$$

a qual é visivelmente comutável nesta ordem. Pode-se facilmente inverter a relação acima com o auxílio de (6.3h,j) para obter

$$\delta\theta_i = \frac{1}{2} \epsilon_{ijk} L_{jk}.$$

Qualquer rotação arbitrária pode então ser a princípio realizada com a composição de um número grande de rotações arbitrárias executadas pela matriz (6.18).

6.5.3 TENSORES ISOTRÓPICOS

Um *tensor isotrópico* (ou *invariante*) é aquele que mantém a forma matemática de seus componentes em qualquer sistema de coordenadas. Determinar quais são os possíveis tensores isotrópicos com um determinado posto é importante, não somente do ponto de vista matemático, mas também físico, uma vez que este conhecimento terá consequências importantes para diversas propriedades físicas de meios contínuos, propriedades tais como suas condutividades elétrica e térmica, susceptibilidade magnética e resistência a tensões, para citar somente alguns exemplos, quando esses meios são uniformes.

No exemplo 6.4 foi mostrado que os tensores de Kronecker e Levi-Civita são tensores isotrópicos de postos dois e três, sob o ponto de vista de rotações próprias arbitrárias, uma vez que suas componentes possuem os mesmos valores para qualquer orientação do sistema de coordenadas. A questão que surge neste ponto é se existem outros tensores isotrópicos de mesmos postos e de postos superiores.

Iniciando por tensores de posto dois, se T_{ij} é isotrópico frente a rotações próprias arbitrárias, então

$$T'_{ij} = L_{ik}L_{j\ell}T_{k\ell} = T_{ij}.$$

Como foi demonstrado na seção 6.5.2, uma rotação arbitrária pode sempre ser composta por um número grande de rotações infinitesimais. Para a dedução da forma geral de um tensor isotrópico, é conveniente então utilizar a matriz de rotação infinitesimal (6.18). Desta forma, em ordem mais baixa nas rotações,

$$\begin{aligned} (\delta_{ik} + \epsilon_{ikm}\delta\theta_m)(\delta_{j\ell} + \epsilon_{j\ell n}\delta\theta_n)T_{k\ell} &= T_{ij}, \\ (\epsilon_{j\ell m}T_{i\ell} + \epsilon_{ikm}T_{kj})\delta\theta_m &= 0. \end{aligned}$$

Como as rotações $\{\delta\theta_m\}$ são arbitrárias, então

$$\epsilon_{j\ell m}T_{i\ell} + \epsilon_{ikm}T_{kj} = 0.$$

Multiplicando ambos os lados por ϵ_{jmr} e somando sobre j e m , obtém-se

$$2T_{ij} + T_{ji} = T_{kk}\delta_{ij}.$$

Permutando os índices livres $i \leftrightarrow j$ e subtraindo ambas as expressões, conclui-se que

$$T_{ij} = T_{ji}.$$

Ou seja, um tensor isotrópico é também simétrico. Retornando então à expressão anterior, obtém-se que

$$T_{ij} = \frac{1}{3}T_{kk}\delta_{ij}.$$

Ou seja, a forma mais geral para um tensor isotrópico de posto dois é

$$T_{ij} = \lambda\delta_{ij},$$

sendo $\lambda = \text{Tr}(T)/3$ uma constante.

Considerando agora tensores de posto três, parte-se novamente de

$$T'_{ijk} = L_{i\ell}L_{jm}L_{kn}T_{\ell mn} = T_{ijk}.$$

Inserindo novamente (6.18), obtém-se em ordem mais baixa nas rotações,

$$\epsilon_{mis}T_{sjk} + \epsilon_{mjs}T_{isk} + \epsilon_{mks}T_{ijs} = 0. \quad (6.19)$$

Multiplicando a equação acima sucessivamente por ϵ_{mir} , ϵ_{mjr} e ϵ_{mkr} e somando,

$$\begin{aligned} 2T_{ijk} + T_{jik} + T_{kji} &= \delta_{ij}T_{ssk} + \delta_{ik}T_{sjs} \\ 2T_{ijk} + T_{jik} + T_{ikj} &= \delta_{ji}T_{ssk} + \delta_{jk}T_{iss} \\ 2T_{ijk} + T_{kji} + T_{ikj} &= \delta_{ki}T_{sjs} + \delta_{kj}T_{iss}. \end{aligned}$$

Multiplicando agora as equações acima respectivamente por δ_{jk} , δ_{ik} e δ_{ij} e somando, obtém-se

$$\begin{aligned} 2T_{iss} + T_{sis} + T_{ssi} &= T_{ssi} + T_{sis} & T_{iss} &= 0 \\ 2T_{sis} + T_{iss} + T_{ssi} &= T_{ssi} + T_{iss} & \implies T_{sis} &= 0 \\ 2T_{ssi} + T_{iss} + T_{sis} &= T_{sis} + T_{iss} & T_{ssi} &= 0. \end{aligned}$$

Portanto, retornando ao sistema acima, conclui-se que

$$\begin{aligned} 2T_{ijk} + T_{jik} + T_{kji} &= 0 \\ 2T_{ijk} + T_{jik} + T_{ikj} &= 0 \\ 2T_{ijk} + T_{kji} + T_{ikj} &= 0. \end{aligned}$$

O que implica em que as soluções deste sistema devem satisfazer as relações

$$T_{jik} + T_{kji} = T_{jik} + T_{ikj} \implies T_{kji} = T_{ikj}$$

$$\begin{aligned} T_{jik} + T_{kji} &= T_{kji} + T_{ikj} \implies T_{jik} = T_{ikj} \\ T_{jik} + T_{ikj} &= T_{kji} + T_{ikj} \implies T_{jik} = T_{kji}, \end{aligned}$$

ou seja,

$$T_{ikj} = T_{jik} = T_{kji} = -T_{ijk}.$$

Neste ponto, conclui-se que T_{ijk} é antissimétrico ante uma permutação ímpar dos índices. Em consequência, será simétrico frente a uma permutação par de índices, pois

$$T_{ijk} = -T_{jik} = T_{jki} \text{ ou } T_{ijk} = -T_{kji} = T_{kij} \text{ ou } \dots$$

Retornando a (6.19), multiplica-se agora por δ_{mi} e, somando, obtém-se

$$\epsilon_{ijs}T_{isk} + \epsilon_{iks}T_{ijs} = 0.$$

Como j e k são parâmetros livres, no primeiro termo somente são não nulos os termos onde $i \neq s$, $i \neq j$ e $s \neq j$, restando somente 2 termos. No segundo, somente são não nulos aqueles onde $i \neq s$, $i \neq k$ e $s \neq k$, restando também 2 termos. Pode-se escrever então, sem somas implícitas,

$$\epsilon_{ijs}T_{iks} - \epsilon_{iks}T_{ijs} = 0.$$

Se $k = j$, a igualdade é trivialmente satisfeita. Por outro lado, se $k \neq j$, considerando os símbolos de Levi-Civita, ou $i \neq j \neq s$, em cuja situação k deve ser necessariamente igual a i ou s e o segundo termo é nulo, ou $i \neq k \neq s$, e então j deve ser igual a i ou s , quando então o primeiro termo é nulo. Considerando todas as possibilidades, conclui-se então que

$$T_{ijj} = T_{iji} = T_{jii} = 0,$$

ou seja, quando 2 índices são iguais, o componente do tensor é nulo.

Portanto, o tensor isotrópico \mathcal{T} possui as mesmas propriedades do símbolo de Levi-Civita. Conclui-se então que a forma geral de um tensor de posto 3 isotrópico deve ser dada por

$$T_{ijk} = \mu\epsilon_{ijk},$$

onde μ é uma constante.

Finalmente, a forma geral de um tensor isotrópico de posto quatro é,

$$T_{ijkl} = \alpha\delta_{ij}\delta_{kl} + \beta\delta_{ik}\delta_{jl} + \gamma\delta_{il}\delta_{jk},$$

sendo α , β e γ constantes, forma esta apresentada sem demonstração.

6.6 ROTAÇÕES IMPRÓPRIAS, PSEUDOTENSORES E TENSORES DUAIS

Até este ponto, considerou-se sempre *transformações próprias*, no sentido atribuído pela expressão (6.9c), isto é, as transformações são tais que $\det(L) = +1$. Na maior parte dos casos considerados, assumiu-se explicitamente que a transformação em questão era uma *rotação passiva própria*.

Contudo, outras transformações de interesse para a física existem. Dentre essas, uma de particular importância são as *rotações impróprias*, para as quais $\det(L) = -1$. Esta e outros tipos de transformações impróprias são importantes para determinar a *quiralidade* de um determinado referencial, ou seja, se este é invariante ou não frente a uma reflexão planar (ou se este é idêntico ou não à sua imagem especular, ver seção 6.2.2).

Esta seção irá discutir aspectos relacionados com a transformação de tensores quando a lei de transformação abrange tanto transformações próprias quanto impróprias.

6.6.1 ROTAÇÕES IMPRÓPRIAS E PSEUDOTENSORES

Uma rotação imprópria de um sistema de coordenadas pode ser executada através da composição de duas transformações: uma transformação de paridade ou inversão espacial, para a qual $\{x_i\} \rightarrow \{x'_i\} : x'_i = -x_i (i = 1, 2, 3)$, cuja matriz de transformação P é dada simplesmente por

$$P_{ij} = -\delta_{ij},$$

seguida por uma rotação própria, executada pela matriz de rotação $L^{(p)}$. Pode-se então construir a *matriz de rotação imprópria* $L^{(i)}$ pela composição

$$L^{(i)}_{ij} = L^{(p)}_{ik} P_{kj},$$

onde se verifica facilmente que $\det(L^{(i)}) = \det(P) = -1$.

Conforme já foi introduzido na seção 6.2.2, um objeto de ordem zero é classificado como um *escalar* ou *pseudoescalar* se este é, respectivamente, invariante ou muda de sinal frente a uma transformação imprópria. Objetos físicos classificados como escalares consistem usualmente nas propriedades intrínsecas da matéria tais como massa de repouso e carga elétrica, as quais são assumidas invariantes frente a transformações arbitrárias por princípios que devem ser continuamente corroborados experimentalmente, sob as mais diversas condições. Adicionalmente, outras quantidades físicas escalares, tais como trabalho mecânico e potencial elétrico, são obtidas via produtos internos entre vetores. Por sua vez, objetos classificados como pseudoescalares são usualmente construídos através de produtos internos entre vetores e pseudovetores; estes objetos serão mencionados novamente a seguir.

Prosseguindo, um objeto de primeira ordem será denominado *vetor polar* ou, simplesmente, *vetor* se satisfizer as leis de transformação (6.12), quer ela seja própria ou imprópria. Já um *vetor axial*, que é na maior parte dos textos denominado também como *pseudovetor*, pode ser definido então como uma tripla ordenada $\mathbf{B} = (B_1, B_2, B_3)$ que obedece as leis de transformação

$$B'_i = \det(L) L_{ij} B_j, \quad B_i = \det(L) L_{ji} B'_j, \tag{6.20}$$

onde $\det(L) \neq +1$. Para transformações ortogonais, $\det(L) = -1$.

A figura 6.5 ilustra uma transformação de paridade ($P : P_{ij} = -\delta_{ij}$) ou inversão espacial (passiva) aplicada a um referencial dextrógiro R . O primeiro aspecto digno de nota é que a transformação P altera a quiralidade do referencial empregado, isto é, o referencial R dextrógiro transformou-se no referencial R' levógiro. Esta alteração de quiralidade não pode ser realizada por nenhuma composição de rotações próprias.

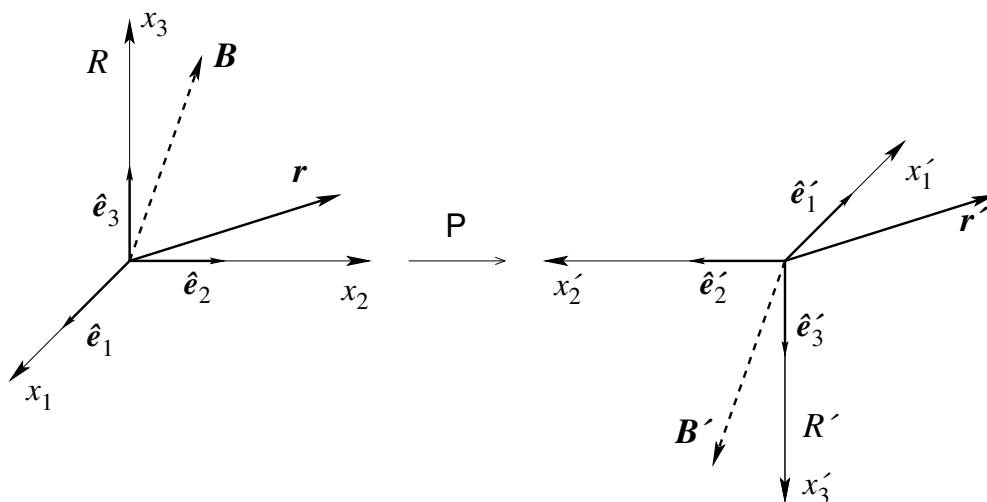


Figura 6.5: Um sistema dextrógiro de coordenadas (R) é transformado em um sistema levógiro (R') através de uma transformação de paridade P . Observa-se também os comportamentos de um vetor (\mathbf{r}) e de um pseudovetor (\mathbf{B}) frente a esta transformação.

Um vetor polar, tal como o vetor \mathbf{r} representado na figura 6.5 tem seus componentes transformados de acordo com (6.12); simultaneamente, os vetores de base $\{\hat{e}_i\}$ também são transfor-

mados, de acordo com (6.7c), como $\hat{e}'_i = P_{ij} \hat{e}_j$. Portanto,

$$\mathbf{r} \xrightarrow{P} \mathbf{r}' : \begin{cases} x_i \rightarrow x'_i = P_{ij} x_j = -x_i \\ \hat{e}_i \rightarrow \hat{e}'_i = P_{ij} \hat{e}_j = -\hat{e}_i \end{cases} \text{ mas } \mathbf{r}' = x'_i \hat{e}'_i = P_{ij} P_{ik} x_j \hat{e}_k = x_j \hat{e}_j = \mathbf{r},$$

uma vez que a transformação P é ortogonal. Ou seja, um vetor polar é um objeto geométrico cujas características (módulo, direção e sentido) não são alteradas por uma transformação (própria ou imprópria).

Por outro lado, um vetor axial transforma-se de acordo com a lei (6.20), sendo o vetor axial \mathbf{B} representado na figura 6.5 um exemplo. Frente a uma inversão espacial, este se transforma como

$$\mathbf{B} \xrightarrow{P} \mathbf{B}' : B_i \rightarrow B'_i = -P_{ij} B_j = B_i, \text{ mas } \mathbf{B}' = B'_i \hat{e}'_i = -P_{ij} P_{ik} B_j \hat{e}_k = -\mathbf{B}.$$

Ou seja, uma transformação de paridade inverte o sentido de um vetor axial, embora ele se comporte como um vetor frente a transformações próprias. Diz-se então que os componentes desta terna ordenada compõem um *pseudovetor* ou um *pseudotensor* Cartesiano de posto um.

Estendendo então o conceito de um pseudotensor a objetos de ordens mais altas, a definição de um *tensor* Cartesiano de ordem N é mantida como aquele objeto cujos componentes transformam-se de acordo com a lei (6.14) frente a qualquer transformação (própria ou imprópria). Por sua vez, um **pseudotensor de posto N** é agora definido como aquele objeto cujos componentes transformam-se de acordo com a lei de transformação

$$T'_{ij\dots n} = \det(\mathbf{L}) L_{ip} L_{jq} \dots L_{nr} T_{pq\dots r} \quad (6.21a)$$

$$T_{ij\dots n} = \det(\mathbf{L}) L_{pi} L_{qj} \dots L_{rn} T'_{pq\dots r}. \quad (6.21b)$$

Ou seja, comporta-se como um tensor somente frente a transformações próprias.

Exemplo 6.3 (Símbolo de Levi-Civita). No exercício 6.4, mostrou-se que frente a uma transformação L qualquer, o símbolo de Levi-Civita transforma-se como $\epsilon'_{ijk} = \det(\mathbf{L}) \epsilon_{ijk}$. Para transformações próprias, como foi assumido no exemplo, este se comporta como um tensor. Entretanto, para uma transformação imprópria $\mathbf{L}^{(i)}$,

$$\{\epsilon_{ijk}\} \xrightarrow{\mathbf{L}^{(i)}} \{\epsilon'_{ijk}\} / \epsilon'_{ijk} = \det(\mathbf{L}^{(i)}) \epsilon_{ijk} = -\epsilon_{ijk}.$$

Portanto, a lei de transformação apropriada para o símbolo de Levi-Civita é a relação (6.21),

$$\epsilon'_{ijk} = \det(\mathbf{L}) L_{il} L_{jm} L_{kn} \epsilon_{lmn},$$

sendo assim identificado como um pseudotensor de posto três.

Existem vários exemplos de vetores axiais na física. Em particular, vetores obtidos a partir de produtos vetoriais ou rotacionais de vetores polares são os mais comuns. Assim, se \mathbf{a} e \mathbf{b} são ambos polares, então o vetor $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ é axial, uma vez que este se comporta, frente a $\mathbf{L}^{(i)}$ como

$$\begin{aligned} \mathbf{c} \xrightarrow{\mathbf{L}^{(i)}} \mathbf{c}' = \mathbf{a}' \times \mathbf{b}' &\implies c'_i = \epsilon'_{ijk} a'_j b'_k = \det(\mathbf{L}^{(i)}) L_{il}^{(i)} L_{jm}^{(i)} L_{jq}^{(i)} L_{kn}^{(i)} L_{kr}^{(i)} \epsilon_{lmn} a_q b_r \\ &= \det(\mathbf{L}^{(i)}) L_{il}^{(i)} \epsilon_{lmn} a_m b_n, \\ c'_i &= \det(\mathbf{L}^{(i)}) L_{il}^{(i)} c_l, \end{aligned}$$

obedecendo a (6.20). A tabela 6.1 mostra exemplos de vetores axiais na mecânica clássica e no eletromagnetismo.

Em termos das operações algébricas definidas na seção 6.4, a possibilidade de ocorrência de operações mistas envolvendo tensores e pseudotensores merecem alguma consideração. Em primeiro lugar, a operação de adição não faz sentido neste caso. Assumindo que \mathbf{a} é um vetor polar e \mathbf{c} um vetor axial, então um possível objeto \mathbf{b} decorrente da operação $\mathbf{b} = \mathbf{a} + \mathbf{c}$ seria transformado, de acordo com (6.12a) e (6.20) como

$$\mathbf{b} \xrightarrow{\mathbf{L}} \mathbf{b}' = \mathbf{a}' + \mathbf{c}' \implies b'_i = a'_i + c'_i = L_{ij} a_j + \det(\mathbf{L}) L_{ij} c_j = L_{ij} [a_j + \det(\mathbf{L}) c_j],$$

o que não satisfaz nenhuma das leis de transformação. Em particular, para transformações impróprias, $|\mathbf{b}'| = |\mathbf{a} - \mathbf{c}|$

Já o produto externo envolvendo tensores e pseudotensores pode resultar em objetos de ordens mais altas de diferentes naturezas. Na construção do objeto \mathcal{Z} realizada por

$$\mathcal{Z} = \mathcal{A} \otimes \mathcal{B} \otimes \dots \otimes \mathcal{Y},$$

se houver N pseudotensores-pais, a lei de transformação de \mathcal{Z} será

$$Z'_{i\dots k} = [\det(L)]^N L_{iq} \dots L_{kr} Z_{q\dots r},$$

ou seja, \mathcal{Z} será um tensor se N for par ou um pseudotensor se N for ímpar.

Observa-se que a mesma propriedade é válida para o caso do produto interno entre tensores e/ou pseudotensores. De particular importância para a teoria da interação eletrofraca, a *helicidade* de uma partícula é determinada pelo produto interno $h = \mathbf{p} \cdot \mathbf{S}$, sendo \mathbf{p} o momentum linear transportado pela mesma e \mathbf{S} o seu spin. Como o momentum linear é um vetor polar e o spin é um vetor axial, a quantidade h é um pseudotensor de posto zero ou um pseudoescalar. Se a helicidade da partícula é positiva, então seu spin tem orientação dextrógiara. Se $h < 0$, a partícula tem spin com orientação levógiara. Uma transformação de paridade inverte os sinais de h .

Inversões espaciais, como exemplo de transformações impróprias, não são a única possibilidade coletivamente denominada de transformações de paridade. Comportamentos semelhantes são obtidos por operações de reflexão no plano, representadas na figura 6.6. Uma reflexão sobre o plano $x_1 - x_3$ é realizada por intermédio da matriz de transformação

$$P^{(r)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Neste caso, um vetor polar como a polarização elétrica \mathbf{P} transforma-se, de acordo com (6.12a), como

$$\mathbf{P} \xrightarrow{P^{(r)}} \mathbf{P}' = (P_1, -P_2, P_3),$$

ou seja, somente a componente P_2 muda de sinal, o que corresponde à imagem especular de \mathbf{P} sobre um espelho colocado no plano $x_1 - x_3$. Esta operação está representada na figura 6.6a.

Já um vetor axial, como a indução magnética \mathbf{B} , por exemplo, transforma-se, de acordo com (6.20), como

$$\mathbf{B} \xrightarrow{P^{(r)}} \mathbf{B}' = (-B_1, B_2, -B_3),$$

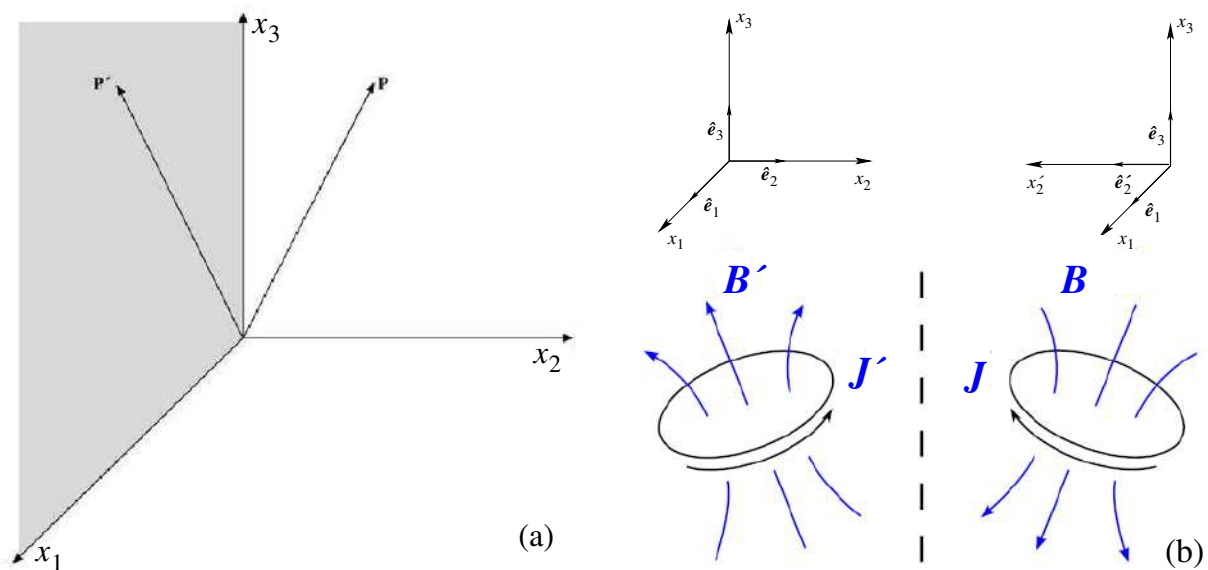


Figura 6.6: Reflexão sobre o plano $x_1 - x_3$. (a) O vetor polarização elétrica \mathbf{P} (polar) transforma-se em \mathbf{P}' , o qual é sua imagem especular. (b) O vetor indução magnética \mathbf{B} (axial), frente a uma reflexão espacial, resulta com a sua imagem especular invertida.

o que corresponde a uma reflexão especular seguida por uma inversão de sentido. Esta operação está ilustrada na figura 6.6b. Nesta, uma espira circular que conduz uma corrente elétrica com densidade J , gerando assim o campo B de acordo com a lei de Biot-Savart, também é refletida sobre o plano $x_1 - x_3$. Como posição e J são polares, os vetores transformados são as imagens especulares destes. Porém, o campo B sofre adicionalmente uma inversão, resultando em uma componente B_2 inalterada.

A análise do comportamento das leis físicas frente a transformações de paridade é um tópico importante na atualidade. Dentre as quatro interações fundamentais, as equações da gravitação, do eletromagnetismo e das interações fortes são invariantes frente a transformações de paridade.

Por exemplo, a imagem especular do campo B gerado pela corrente J na espira, ilustrada na figura 6.6b, está orientada de tal forma que as leis de Biot-Savart e Faraday permanecem válidas em um mundo virtual que seria a reflexão especular do mundo real. A operação de reflexão, ditada por $P^{(r)}$, transforma-se em uma inversão espacial completa, ditada por P , se a imagem especular for girada ativamente em 180° em torno do eixo x_2 (figura 6.5). Neste caso, a imagem refletida seria indistinguível do sistema físico real. Portanto, para uma interação invariante com a transformação de paridade, o mundo virtual é indistinguível do mundo real, o que está em acordo com a suposição de que leis físicas devem ser invariantes frente a transformações realizadas sobre o referencial utilizado.

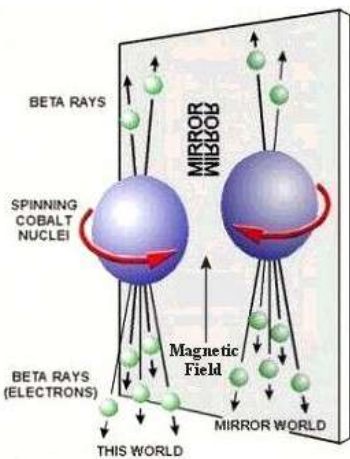
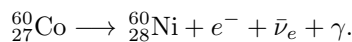


Figura 6.7: Violação de paridade no decaimento β de núcleos de cobalto-60.

Já a interação fraca apresenta uma quebra de simetria no decaimento β . Dentre os léptons sujeitos à interação fraca, somente partículas cujos spins estão no sentido levógiro e anti-partículas com spins no sentido dextrógiro interagem através da força fraca. A combinação inversa é proibida pelo modelo padrão, violando a paridade da interação.

Esta quebra de paridade, denominada *violação CP*, foi sugerida por T. D. Lee e C. N. Yang em 1956 e verificada no ano seguinte por dois experimentos independentes. O primeiro experimento consistiu na observação de elétrons emitidos pelo decaimento β de núcleos de cobalto, conforme está representado na figura 6.7. A baixas temperaturas, a maior parte dos núcleos podem ser mantidos com seus spins alinhados com um campo magnético externo. O elipsóide da esquerda na figura ilustra um grande número de núcleos de ${}^{60}_{27}\text{Co}$, os quais decaem de acordo com a fórmula



A seta vermelha indica o sentido de giro dos núcleos, o que significa que estes têm seu spin no sentido dextrógiro. Já a imagem da direita representa a imagem especular e, portanto, o sentido do spin nuclear S_N é invertido por ser um vetor axial. No experimento, os elétrons emitidos por decaimento β com momentum linear p_e mostraram uma correlação notável com o pseudoescalar $h_{Ne} = S_N \cdot p_e$, sendo emitidos preferencialmente na direção e sentido determinados por $h_{Ne} < 0$. De acordo com a teoria eletrofraca, os elétrons emitidos são levógiros ($h_e < 0$) e os anti-neutrinos do elétron ($\bar{\nu}_e$) são dextrógiros ($h_{\bar{\nu}} > 0$).

Na imagem especular à direita na figura 6.7, o spin nuclear S_N resulta invertido, mas p_e não o é, resultando em uma inversão de sinal dos pseudoescalares h_{Ne} , h_e e $h_{\bar{\nu}}$. A inversão espacial torna-se completa se o espelho for girado em 180° , quando então o spin nuclear volta a sua orientação original, mas os elétrons seriam observados sendo emitidos preferencialmente para cima.

Em outras palavras, a imagem no mundo virtual é distinguível do sistema físico real, ao contrário do que ocorre com as outras interações fundamentais. A dependência da força fraca com o pseudoescalar h_{Ne} cria uma violação do princípio de invariância na transformação de paridade que não é observada nas outras interações fundamentais. Adicionalmente, a helicidade dos léptons envolvidos no decaimento β também seria invertida em comparação com o mundo real, o que ressalta ainda mais a violação de paridade nesta interação fundamental.

O outro experimento foi realizado em um anel de cíclotron e investigou a helicidade dos léptons emitidos nos decaimentos

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu$$

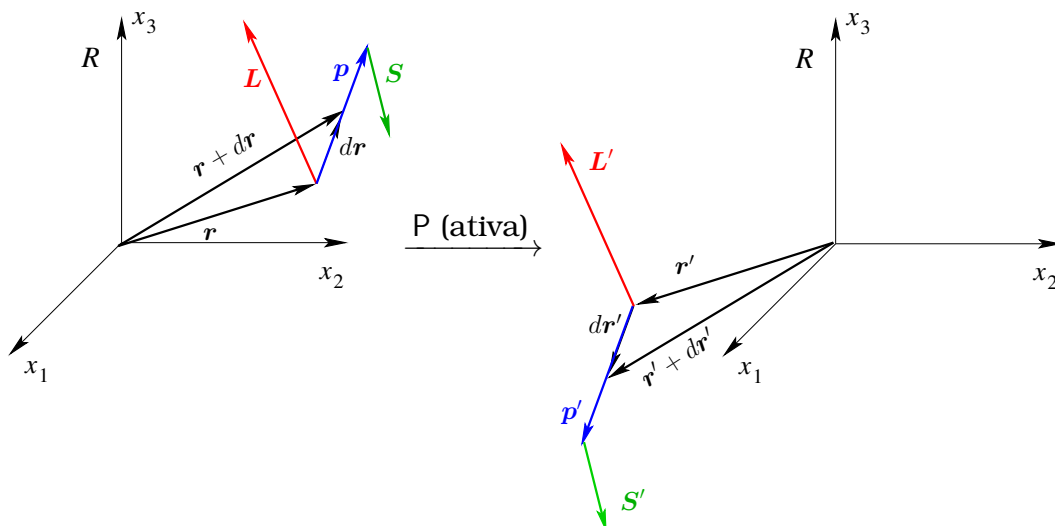


Figura 6.8: Inversão espacial ativa realizada sobre um sistema físico. Vetores polares são invertidos, mas vetores axiais permanecem invariantes.

$$\mu^+ \longrightarrow e^+ + \nu_e + \bar{\nu}_\mu.$$

Este experimento verificou a polarização sugerida por Lee & Yang, segundo a qual as partículas são emitidas no sentido levógiro enquanto que as anti-partículas são dextróginas. Esta preferência na helicidade das partículas e anti-partículas que interagem via força fraca evidenciam a quebra de paridade na interação.¹⁵

A descrição aqui apresentada para a violação CP é apenas qualitativa. Uma descrição mais detalhada, envolvendo o formalismo da teoria quântica de campos, é fornecida, por exemplo, no capítulo 3 de (SILVEIRA, 2019).

Uma observação importante deve ser feita agora relativa à realidade física dos pseudotensores. Alguns textos argumentam que objetos geométricos tais como vetores e tensores somente podem ser atribuídos a quantidades que possuem realidade física se não forem alterados por transformações arbitrárias no sistema de coordenadas. Neste caso, pseudovetores e pseudotensores deveriam ser considerados meramente objetos matemáticos abstratos que não poderiam descrever quantidades tais como momentum angular e indução magnética.

A solução para este aparente dilema é encontrada se as transformações de paridade não forem realizadas de forma passiva, como foi assumido até este momento, mas sim de forma ativa, isto é, o sistema físico é fisicamente rotado em torno da origem do referencial por 180°, de tal forma que

$$r \xrightarrow{P \text{ (ativa)}} r' = -r.$$

A figura 6.8 ilustra uma inversão espacial ativa em torno da origem. Neste caso, o sistema de coordenadas permanece sempre dextrógiro, mas as quantidades vetoriais são transformadas de acordo com sua paridade frente a esta inversão ativa. Qualquer rotação passiva posteriormente aplicada após a inversão será sempre própria, e assim a quiralidade do referencial não é alterada e, portanto, todas os objetos geométricos são tensores que satisfazem as leis de transformação (6.14).

Realizando esta inversão ativa, os vetores podem então ser classificados como *vetores polares*, os quais são ímpares frente à inversão, tais como os vetores r e p na figura 6.8 e *vetores axiais*, os quais são pares frente à inversão, tais como os vetores $L = r \times p$ e S (spin). Agora, porém, ambos os tipos de vetores podem ser classificados como *tensores de posto um*, uma vez que todas as transformações doravante aplicadas ao sistema de coordenadas serão próprias.

Contudo, é importante mencionar também que pseudoescalares ainda são necessários para teorias físicas que prevêem quebra de paridade, tal como a teoria eletrofraca. Na interpretação de que transformações de paridade devem ser ativas, estes pseudoescalares continuam existindo. Um exemplo é fornecido pelos vetores p e S na figura 6.8, onde se verifica que a helicidade $h = p \cdot S$ continua mudando de sinal, mesmo frente a uma transformação de paridade ativa.

¹⁵A violação da simetria quiral possui uma importância vital na bioquímica e na genética. Uma discussão atual a respeito deste assunto pode ser encontrada no doi: [10.1103/RevModPhys.85.603](https://doi.org/10.1103/RevModPhys.85.603).

Pode-se verificar também que o produto externo misto entre vetores polares e axiais continua gerando tensores de postos mais altos, mas o produto interno entre dois tensores criados desta maneira pode resultar em um pseudotensor.

6.6.2 TENSORES DUAIS

Independente da polêmica com relação a pseudotensores, estes têm sido empregados em diversas teorias físicas, mesmo em teorias invariantes frente a transformações de paridade.

Um tipo de pseudotensor com frequência empregado é o *tensor dual*. Na sua definição, usualmente faz-se uso do símbolo de Levi-Civita. Realizando produtos internos deste com um tensor totalmente antissimétrico conhecido, pode-se criar um tensor dual de posto igual ou inferior, dependendo da necessidade.

Um exemplo pode ser dado tomando o tensor \mathcal{A} , totalmente antissimétrico, isto é, $A_{ij} = -A_{ji}$, o que implica que $\text{Tr}(\mathcal{A}) = 0$. Com ϵ_{ijk} , pode-se definir então o pseudovetor p , cujos componentes são

$$p_i = \frac{1}{2} \epsilon_{ijk} A_{jk}. \quad (6.22a)$$

Ou, reciprocamente, multiplicando a expressão acima por ϵ_{ijk} e contraíndo os índices, obtém-se

$$\epsilon_{ijk} p_i = \frac{1}{2} \epsilon_{ijk} \epsilon_{ilm} A_{lm} = \frac{1}{2} (A_{jk} - A_{kj}) \implies A_{ij} = \epsilon_{ijk} p_k. \quad (6.22b)$$

Pode-se criar também um pseudoescalar s a partir agora de um tensor \mathcal{C} de posto 3 totalmente antissimétrico,

$$s = \frac{1}{3!} \epsilon_{ijk} C_{ijk}.$$

Então, multiplicando ambos os lados por ϵ_{ijk} e usando (6.3), resulta

$$\epsilon_{ijk} s = \frac{1}{3!} \epsilon_{ijk} \epsilon_{lmn} C_{lmn} \implies C_{ijk} = s \epsilon_{ijk}.$$

Tensores duais são empregados, por exemplo, na formulação covariante do eletromagnetismo, onde o dual do tensor de campo é empregado para expressar a forma covariante das equações homogêneas de Maxwell. Para tanto, faz-se uso de um pseudotensor de posto quatro, o qual é a extensão do símbolo de Levi-Civita para um espaço vetorial de dimensão quatro.

6.6.3 TENSORES IRREDUTÍVEIS

Um uso para tensores duais está na representação de um tensor qualquer em formas irredutíveis. Para exemplificar, emprega-se um tensor de posto dois qualquer \mathcal{T} . Da seção 6.4.2, sabe-se que é possível representar todos os seus componentes em partes simétrica (T_{ij}^S) e antissimétrica (T_{ij}^A). Para a parte antissimétrica, cria-se o pseudovetor u de acordo com (6.22),

$$u_i = \frac{1}{2} \epsilon_{ijk} T_{ij}^A = \frac{1}{4} \epsilon_{ijk} (T_{ij} - T_{ji}) \iff T_{ij}^A = \epsilon_{ijk} u_k.$$

Já da parte simétrica, subtrai-se o traço da matriz T , definindo então o tensor

$$S_{ij} = T_{ij}^S - \frac{1}{3} \text{Tr}(T) \delta_{ij} = \frac{1}{2} (T_{ij} + T_{ji}) - \frac{1}{3} \text{Tr}(T) \delta_{ij}.$$

Portanto, o tensor \mathcal{T} pode agora ser representado pelas suas partes irredutíveis,

$$T_{ij} = \frac{1}{3} \text{Tr}(T) \delta_{ij} + \epsilon_{ijk} u_k + S_{ij},$$

as quais são compostas por tensores de posto zero, um e dois.

6.7 TENSORES GENERALIZADOS

As seções anteriores restringiram-se ao estudo de tensores descritos em termos de coordenadas Cartesianas e os tipos de transformação passiva aplicados a este sistema aos quais foi dedicada a maior atenção foram as rotações rígidas dos eixos coordenados e as transformações de paridade.

A partir desta seção, os conceitos e propriedades considerados anteriormente serão generalizados para transformações arbitrárias aplicadas sobre um sistema de coordenadas geral e que levam a outro sistema. Antes de se abordar tensores propriamente ditos, será realizada uma revisão e adaptação das propriedades matemáticas de sistemas de coordenadas curvilíneas, discutidas no capítulo 1. Posteriormente, essas propriedades serão generalizadas e aplicadas a tensores de postos arbitrários.

6.7.1 COORDENADAS CURVILÍNEAS GENERALIZADAS

Fazendo referência novamente à figura 1.2, a posição de um ponto P arbitrário no espaço \mathbb{E}^3 pode ser expressa em termos de uma terna ordenada de números (q_1, q_2, q_3) , as quais são as coordenadas deste ponto dentro do sistema de coordenadas curvilíneas adotado.

Qualquer vetor neste espaço, sendo o vetor posição \mathbf{r} um caso particular, pode ser expresso em termos de uma combinação linear de duas bases apresentadas na seção 1.1, as bases $\{e_i\}$ e $\{\epsilon_i\}$, dadas por

$$e_i = \frac{\partial \mathbf{r}}{\partial q_i} \text{ e } \epsilon_i = \nabla q_i,$$

sendo $i = 1, 2, 3$. Em geral estes vetores não formam bases ortonormais. Por outro lado, ambos os conjuntos sempre formam duas bases recíprocas no \mathbb{E}^3 , isto é,

$$e_i \cdot \epsilon_j = \delta_{ij}.$$

Dentro do contexto de tensores generalizados, uma das modificações necessárias está na posição dos índices, tanto livres quanto mudos, que identificam os componentes dos mesmos. Devido a esta modificação, a base $\{\epsilon_i\}$ será doravante identificada por $\{e^i\}$, com a equivalência $\{\epsilon_1, \epsilon_2, \epsilon_3\} \leftrightarrow \{e^1, e^2, e^3\}$.

Adicionalmente, a convenção de somas implícitas introduzida na seção 6.1.1 será modificada. Doravante, qualquer índice minúsculo que apareça exatamente duas vezes em qualquer componente ou expressão, *uma vez como subscrito e uma vez como superescrito*, será somado sobre todos os valores que um índice naquela posição pode assumir, exceto quando for realizada uma observação explícita ao contrário.

Com esta nova convenção de posicionamento dos índices, a relação de reciprocidade entre as bases possíveis para um dado sistema de coordenadas curvilíneas passa a ser escrita como

$$e_i \cdot e^j = \delta_i^j, \tag{6.23}$$

sendo δ_i^j o mesmo tensor delta de Kronecker discutido em detalhes na seção 6.1.2.

Pela mesma razão, as coordenadas curvilíneas do ponto P serão também doravante identificadas por índices superescritos $P = (q^1, q^2, q^3)$, de tal forma que as bases passam a ser obtidas pelas expressões

$$e_i = \frac{\partial \mathbf{r}}{\partial q^i} \text{ e } e^i = \nabla q^i. \tag{6.24}$$

Nestas expressões nota-se que a derivação em relação a uma coordenada cujo índice está na posição *superior* resulta em um objeto com o mesmo índice na posição *inferior*.

Posteriormente, será demonstrado que o operador derivação em relação a coordenada *contravariante* q^i ($\partial/\partial q^i$) comporta-se, frente a uma transformação qualquer, como o i -ésimo componente de um vetor *covariante*. Em muitos textos, costuma-se escrever, de forma resumida,

$$\partial_i \equiv \frac{\partial}{\partial q^i}.$$

De forma equivalente,

$$\partial^i \equiv \frac{\partial}{\partial q_i},$$

o qual se comporta como a i -ésima coordenada de um vetor *contravariante*.

Dadas bases (6.24), um vetor $\mathbf{a} \in \mathbb{E}^3$ qualquer será escrito em termos de combinações lineares das mesmas de duas formas equivalentes,

$$\begin{aligned} \mathbf{a} &= a^1 \mathbf{e}_1 + a^2 \mathbf{e}_2 + a^3 \mathbf{e}_3 = a^i \mathbf{e}_i \\ \mathbf{a} &= a_1 \mathbf{e}^1 + a_2 \mathbf{e}^2 + a_3 \mathbf{e}^3 = a_i \mathbf{e}^i. \end{aligned}$$

Da mesma forma como foi mencionado para a expressão (1.9), os números no conjunto $\{a^i\}$ são denominados as componentes **contravariantes** do vetor \mathbf{a} , ao passo que $\{a_i\}$ contém as componentes **covariantes** do mesmo. A posição dos índices distingue os dois tipos de componentes entre si. Por conseguinte, os vetores $\{\mathbf{e}_i\}$ formam uma base covariante no \mathbb{E}^3 , ao passo que $\{\mathbf{e}^i\}$ formam uma base contravariante no mesmo espaço vetorial. Obviamente, os diferentes componentes do vetor \mathbf{a} são dados por

$$a^i = \mathbf{a} \cdot \mathbf{e}^i \text{ e } a_i = \mathbf{a} \cdot \mathbf{e}_i.$$

A distinção entre coordenadas covariantes e contravariantes não foi realizada para tensores Cartesianos porque neste sistema ambas as bases são idênticas. Contudo, em sistemas generalizados isto não é necessariamente verdade.

A partir deste momento, a designação de tensores também deverá ser alterada, em função das distintas posições que seus índices podem adotar. Uma definição formal de um tensor de posto arbitrário será realizada posteriormente. Neste momento, será empregado somente um tensor de posto dois, para exemplificar a identificação. Um tensor \mathcal{T} de posto dois pode ser escrito e/ou construído de diferentes maneiras. Se for empregado o produto externo de dois vetores \mathbf{a} e \mathbf{b} , escritos em termos de uma ou ambas as bases em (6.24), as possibilidades são:

1. Sendo $\mathbf{a} = a^i \mathbf{e}_i$ e $\mathbf{b} = b^j \mathbf{e}_j$, então $\mathcal{T} = \mathbf{a} \otimes \mathbf{b} = a^i b^j \mathbf{e}_i \otimes \mathbf{e}_j$, de onde se identifica $T^{ij} = a^i b^j$, os quais são os *componentes contravariantes* do tensor \mathcal{T} .
2. Sendo $\mathbf{a} = a^i \mathbf{e}_i$ e $\mathbf{b} = b_i \mathbf{e}^i$, então $\mathcal{T} = \mathbf{a} \otimes \mathbf{b} = a^i b_j \mathbf{e}_i \otimes \mathbf{e}^j$, identificando $T^i_j = a^i b_j$ como os *componentes mistos* do tensor \mathcal{T} . Outra possibilidade é $\mathcal{T} = T_i^j \mathbf{e}^i \otimes \mathbf{e}_j$.
3. Sendo $\mathbf{a} = a_i \mathbf{e}^i$ e $\mathbf{b} = b_j \mathbf{e}^j$, então $\mathcal{T} = a_i b_j \mathbf{e}^i \otimes \mathbf{e}^j$, identificando $T_{ij} = a_i b_j$ como os *componentes covariantes* do tensor \mathcal{T} .

É importante ressaltar que $\{T^{ij}\}$, $\{T^i_j\}$ ou $\{T_{ij}\}$ formam três conjuntos distintos de componentes do mesmo tensor \mathcal{T} , porém correspondentes a diferentes construções dos vetores de base do sistema de coordenadas.

As expressões acima envolvendo o produto escalar entre dois vetores merece uma maior atenção neste momento. Na notação tensorial covariante empregada nesta seção, uma contração de índices (ou seja, a soma implícita dos mesmos) somente será realizada se um estiver na posição contravariante e o outro na posição covariante. De acordo com esta convenção, o produto escalar entre os vetores \mathbf{a} e \mathbf{b} somente pode ser escrito na forma usual, ou seja, em termos da soma dos produtos das respectivas coordenadas dos mesmos, da seguinte maneira,

$$\mathbf{a} \cdot \mathbf{b} = a_i b^i \text{ ou } \mathbf{a} \cdot \mathbf{b} = a^i b_i. \quad (6.25a)$$

Este resultado está garantido graças à condição de reciprocidade (6.23), pois com esta pode-se escrever

$$\mathbf{a} \cdot \mathbf{b} = (a_i \mathbf{e}^i) \cdot (b^j \mathbf{e}_j) = a_i b^j \mathbf{e}^i \cdot \mathbf{e}_j = a_i b^j \delta_j^i = a_i b^i, \quad (6.25b)$$

ou vice-versa. Esta discussão terá prosseguimento na próxima seção, quando for discutida a operação de elevação ou rebaixamento de índices.

6.7.2 O ESPAÇO DE RIEMANN E O TENSOR DE MÉTRICA

Conforme já foi discutido na seção 1.1, a métrica de qualquer espaço Riemanniano é completamente determinada pelo tensor de métrica g , o qual é um tensor de posto dois. Antes de se prosseguir com a apresentação das propriedades do tensor de métrica, é conveniente definir de forma mais rigorosa o *espaço de Riemann* ou o *espaço Riemanniano*.

O espaço (métrico) de Riemann é a generalização do espaço métrico Euclidiano E^n definido no exemplo 4.2. Esta generalização é realizada tanto na dimensão arbitrária do espaço quanto na definição de sua métrica, a qual passa a descrever tanto espaços planos, como o Euclidiano, como espaços curvos, importantes para as teorias modernas em física e matemática.

Definição 6.9 (Espaço métrico de Riemann). Seja

$$\mathbb{R}^n \equiv \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ vezes}} = \{(q^1, \dots, q^n) \mid q^j \in \mathbb{R}, j = 1, \dots, n\}$$

o conjunto de todas as n -uplas ordenadas obtidas a partir do produto Cartesiano do corpo dos números reais e com valores determinados por um sistema de coordenadas X . Seja $d\ell$ um elemento de arco medido em X e cujo valor é determinado pela *métrica Riemanniana*

$$d\ell^2 = g_{ij} dq^i dq^j, \tag{6.26}$$

onde $g = [g_{ij}]$ é o *tensor de métrica* que satisfaz as condições:

(ER₁) As componentes $g_{ij}(q^1, \dots, q^n)$ do tensor de métrica pertencem à classe $C^2(\mathbb{R}^n)$.

(ER₂) O tensor $g = [g_{ij}]$ é simétrico, *i. e.*, $g_{ij} = g_{ji}$.

(ER₃) O tensor $g = [g_{ij}]$ é não singular, *i. e.*, $\det [g_{ij}] \neq 0$.

(ER₄) A forma diferencial (6.26) é invariante frente a uma troca arbitrária de coordenadas.

Então, estrutura $R^n \doteq \langle \mathbb{R}^n, g \rangle$ forma o *Espaço Riemanniano de dimensão n* .

Rigorosamente, as condições (ER₁)–(ER₄) impostas ao tensor de métrica definem um espaço *pseudo-Riemanniano*, pois o elemento $d\ell^2$ obtido por (6.26) não é positivo-definido. Para se definir um *espaço Riemanniano*, a condição (ER₃) deve ser substituída pela condição mais restritiva

(ER'₃) O tensor $g = [g_{ij}]$ é *positivo-definido*, isto é, $g_{ij}v^i v^j > 0$ para todo vetor $v = (v^1, \dots, v^n) \in \mathbb{R}^n$ não nulo.

A distinção entre um espaço Riemanniano e um espaço pseudo-Riemanniano não afeta a discussão realizada nesta e nas próximas seções. Um exemplo importante de um espaço pseudo-Riemanniano é o espaço-tempo de Mikowski, o qual será discutido nas seções 6.15.1 e 6.15.2.

As componentes covariantes do tensor de métrica são obtidos a partir da base covariante $\{e_i\}$ e são dados por (1.8),

$$g_{ij} = e_i \cdot e_j, \tag{6.27a}$$

o que mostra claramente que o tensor de métrica é simétrico.

Se o sistema de coordenadas é ortogonal, $e_i \cdot e_j = 0$ para $j \neq i$. Neste caso, é conveniente empregar mais uma vez os fatores de escala h_i , definidos em (1.11), com os quais se pode escrever a base ortonormal $\{\hat{e}_i\}$ como $\hat{e}_i = e_i/h_i$, ressaltando que neste caso i é um índice livre. Assim, o tensor de métrica fica escrito simplesmente como

$$g_{ij} = h_i^2 \delta_{ij} \quad (i : \text{índice livre}). \tag{6.27b}$$

6.7.2.1 OPERAÇÃO DE ELEVAÇÃO OU REBAIXAMENTO DE ÍNDICE

O tensor de métrica possui uma outra função na álgebra tensorial covariante que é de extrema importância. Ele possibilita alterar a posição (contra- ou covariante) de um determinado índice (livre ou mudo).

Para se mostrar como esta operação ocorre, retoma-se a discussão do produto escalar entre os vetores a e b realizada no final da seção anterior. Ressaltando que, como estes são tensores de posto um, o objeto resultante de seu produto interno deve ser um escalar, que possui o mesmo valor seja qual for a representação ou sistema de coordenadas adotado para os vetores. Por isto, uma expressão alternativa às obtidas acima para o produto escalar é

$$a \cdot b = (a^i e_i) \cdot (b^j e_j) = a^i b^j e_i \cdot e_j = g_{ij} a^i b^j,$$

onde se nota o surgimento do tensor de métrica, devido a sua definição. Comparando com uma das expressões (6.25), percebe-se que, necessariamente,

$$a_i = g_{ij} a^j.$$

Ou seja, uma componente *covariante* de a é obtida a partir de suas componentes *contravariantes* com o uso do tensor de métrica. Diz-se então que o índice foi *rebaixado* para a posição covariante.

Por outro lado, da mesma forma como a expressão (6.27) define o tensor de métrica covariante, pode-se também definir o mesmo tensor com seus componentes escritos na forma contravariante através de

$$g^{ij} \doteq e^i \cdot e^j. \quad (6.28)$$

Ressalta-se que os componentes de g^{ij} são, em geral, distintos dos respectivos componentes de g_{ij} .

Desta maneira, o produto escalar entre a e b também pode ser escrito como

$$a \cdot b = (a_i e^i) \cdot (b_j e^j) = g^{ij} a_i b_j.$$

Nota-se agora que $a^i = g^{ij} a_j$, ou seja, foi realizada a *elevação* o índice para a posição contravariante por intermédio também do tensor de métrica.

A capacidade do tensor de métrica de realizar a mudança na posição de um determinado índice não se restringe aos componentes de tensores de posto um. Este pode ser empregado também para mover o índice de um vetor de base. Ou seja, se $e_k \cdot e^j = \delta_k^j$, multiplicando ambos os lados por g_{ji} e realizando a soma implícita,

$$e_k \cdot e^j g_{ji} = \delta_k^j g_{ji} = g_{ki} = e_k \cdot e_i.$$

Portanto, observa-se que $e_i = g_{ij} e^j$, uma vez que o tensor de métrica é simétrico. A partir do produto externo, verifica-se facilmente que o tensor de métrica pode realizar a mesma operação em qualquer índice de um tensor de posto dois ou superior, ou seja,

$$T_j^i = T^{ik} g_{kj} \text{ ou } T^{ij} = T_k^i g^{kj}, \text{ etc.}$$

Finalmente, a relação entre g_{ij} e g^{ij} também é facilmente derivada. Dado o i -ésimo componente do vetor a , pode-se escrever

$$a^i = g^{ij} a_j, \text{ mas, } a_j = g_{jk} a^k. \text{ Portanto, } a^i = g^{ij} g_{jk} a^k.$$

Este resultado implica em que

$$g^{ik} g_{kj} = \delta_j^i, \quad (6.29)$$

ou seja, usando uma representação matricial para o tensor de métrica e denotando $g = [g_{ij}]$, a sua inversa será $g^{-1} = [g^{ij}]$.

Finalmente, a forma mista (contra- e covariante) do tensor de métrica pode ser obtida diretamente do resultado anterior, empregando o mesmo para alterar a posição de um de seus índices,

$$g_j^i = g^{ik} g_{kj} = \delta_j^i. \quad (6.30)$$

6.7.2.2 ELEMENTOS INFINITESIMAIS DE ARCO E VOLUME

Continuando com a discussão realizada na seção 1.1, se $dr = dq^i e_i$ é o vetor deslocamento infinitesimal em um espaço Riemanniano escrito em termos da base $\{e_i\}$, a norma ou o elemento infinitesimal de arco deste espaço é dado pela métrica (6.26), isto é,

$$d\ell^2 = dr \cdot dr = g_{ij} dq^i dq^j.$$

Considera-se agora a base canônica. De acordo com o que foi demonstrado no exercício 1.2, para sistemas ortogonais resulta $\{\hat{x}_i\} = \{\hat{x}^i\}$. Além disso, como os fatores de escala são todos unitários, resulta também que $x_i = x^i$. Assim, os vetores de base $\{e_i\}$ podem ser novamente expressos, conforme as definições (6.24) e (1.7a), como

$$e_i = \frac{\partial \mathbf{r}}{\partial q^i} = \frac{\partial x^j}{\partial q^i} \hat{x}_j = \frac{\partial x_j}{\partial q^i} \hat{x}^j \equiv H^j_i \hat{x}_j = H_{ji} \hat{x}^j, \quad (6.31)$$

respectivamente, sendo agora

$$H^i_j = \frac{\partial x^i}{\partial q^j} \text{ e } H_{ij} = \frac{\partial x_i}{\partial q^j},$$

onde H é a matriz de transformação do sistema Cartesiano ao curvilíneo, cujo determinante corresponde ao Jacobiano da transformação. Portanto, o tensor de métrica pode ser escrito também como

$$g_{ij} = e_i \cdot e_j = H^k_i H_{\ell j} \hat{x}_k \cdot \hat{x}^\ell = H^k_i H_{\ell j} \delta_k^\ell = H^k_i H_{kj}, \quad (6.32)$$

a qual é simplesmente a expressão (1.8) já obtida.

Por sua vez, o elemento de arco ao longo da i -ésima coordenada curvilínea é escrito como $dl_i = dq^i e_i$, sendo que neste caso não se realiza a soma implícita. Portanto, o elemento de arco na direção de q^i pode ser escrito como $dl_i = dq^i H^j_i \hat{x}_j$.

De acordo com a discussão realizada na página 9, o elemento de volume delimitado pelos arcos $\{dl_1, dl_2, dl_3\}$, independente destes serem ortogonais ou não, é sempre dado por

$$d^3r = |dl_1 \cdot (dl_2 \times dl_3)| = |e_1 \cdot (e_2 \times e_3)| dq^1 dq^2 dq^3.$$

Portanto,

$$d^3r = \left| H^i_1 H^j_2 H^k_3 \hat{x}_i \cdot (\hat{x}_j \times \hat{x}_k) \right| dq^1 dq^2 dq^3 = \left| \epsilon_{ijk} H^i_1 H^j_2 H^k_3 \right| dq^1 dq^2 dq^3.$$

É conveniente neste ponto introduzir alguns resultados oriundos da álgebra de matrizes. Sendo A e B duas matrizes 3×3 não singulares, e sendo $B = A^{-1}$, recorda-se a expressão (6.15), a qual fornece o determinante da matriz A. Esta expressão pode ser adaptada para a notação covariante empregada nesta seção de diversas maneiras, como apresentado a seguir, onde também é apresentada uma expressão para os elementos da matriz B,

$$\det(A) \epsilon_{ijk} = A_{im} A_{jn} A_{kl} \epsilon^{mnl} \quad (6.33a)$$

$$\det(A) \epsilon_{ijk} = A^m_i A^n_j A^\ell_k \epsilon_{mnl} \quad (6.33b)$$

$$2 \det(A) B^{ij} = \epsilon^{ikl} \epsilon^{jmn} A_{mk} A_{nl}. \quad (6.33c)$$

As operações de elevação ou rebaixamento de índices envolvidas nas expressões acima são realizadas com o emprego do tensor de métrica e serão discutidas em breve. Desta forma, resulta

$$e_1 \cdot (e_2 \times e_3) = \det(H)$$

e o elemento de volume fica escrito

$$d^3r = |\det(H)| dq^1 dq^2 dq^3 = |J| dq^1 dq^2 dq^3,$$

lembrando que

$$J \equiv \det(H) = \frac{\partial(x^1, x^2, x^3)}{\partial(q^1, q^2, q^3)}$$

é o Jacobiano da transformação do sistema Cartesiano ao curvilíneo. Esta é novamente a expressão para o elemento de volume já obtida em (1.21b).

Agora, o determinante do tensor de métrica também pode ser calculado com as expressões acima. Usando então a expressão (6.32) para o mesmo,

$$\det(g) = \epsilon^{ijk} g_{1i} g_{2j} g_{3k} = \epsilon^{ijk} H_{li} H_{mj} H_{nk} H^\ell_1 H^m_2 H^n_3 = \det(H) \epsilon_{lmn} H^\ell_1 H^m_2 H^n_3 = [\det(H)]^2.$$

Ou seja,

$$e_1 \cdot (e_2 \times e_3) = \sqrt{\det g}$$

e o elemento de volume pode ser escrito na forma generalizada como

$$d^3r = \sqrt{\det g} dq^1 dq^2 dq^3,$$

o qual é um resultado importante, válido também para sistemas de coordenadas não ortogonais.

Algumas relações adicionais envolvendo as bases e o tensor de métrica são úteis para aplicações posteriores. Em primeiro lugar, o produto triplo entre os vetores da base covariante pode ser escrito de uma forma geral como

$$e_i \cdot (e_j \times e_k) = \sqrt{g} \epsilon_{ijk}, \quad (6.34)$$

o que pode ser facilmente verificado. No resultado acima, definiu-se $g \equiv \det(g)$ para simplificar a notação.

Deseja-se agora escrever os vetores de uma base como combinações lineares de vetores da base recíproca. Isto é realizado através da operação de elevação ou rebaixamento de índices executada pelo tensor de métrica, isto é,

$$e^i = g^{ij} e_j \quad e_i = g_{ij} e^j.$$

Porém, empregando (6.33c) e (6.29), pode-se escrever $2gg^{ij} = \epsilon^{ikl}\epsilon^{jmn}g_{mk}g_{nl}$. Portanto, para e^i ,

$$e^i = \frac{1}{2g}\epsilon^{ikl}\epsilon^{jmn}g_{km}g_{ln}e_j.$$

Empregando novamente as relações (6.31) e (6.32), resulta então

$$e^i = \frac{1}{2g}\epsilon^{ikl}\epsilon^{jmn}H^q_k H^r_\ell H_{qm} H_{rn} H_{pj} \hat{x}^p = \frac{1}{2\sqrt{g}}\epsilon^{ikl}\epsilon_{qrp} H^q_k H^r_\ell \hat{x}^p,$$

onde foi usado também (6.33a). Comparando este resultado com o produto

$$e_i \times e_j = H^k_i H^\ell_j \hat{x}_k \times \hat{x}_\ell = \epsilon_{klr} H^k_i H^\ell_j \hat{x}^r,$$

conclui-se que

$$e^i = \frac{1}{2\sqrt{g}}\epsilon^{ikl}(e_k \times e_\ell). \quad (6.35a)$$

Uma outra expressão útil pode ser agora obtida a partir deste resultado,

$$e^i \times e^j = \frac{1}{4g}\epsilon^{ikl}\epsilon^{jmn}(e_k \times e_\ell) \times (e_m \times e_n).$$

Como $(a \times b) \times (c \times d) = d \cdot (a \times b)c - c \cdot (a \times b)d$, o produto triplo (6.34) mostra que

$$e^i \times e^j = \frac{1}{2g}\epsilon^{ikl}\epsilon^{jmn}[e_n \cdot (e_k \times e_\ell)]e_m = \frac{1}{\sqrt{g}}\epsilon^{ijk}e_k, \quad (6.35b)$$

$$e_i = \frac{\sqrt{g}}{2}\epsilon_{ijk}e^j \times e^k. \quad (6.35c)$$

Finalmente, se o sistema de coordenadas for ortogonal, então de (6.27b) e de (6.29) pode-se deduzir a forma para os componentes contravariantes do tensor de métrica. Se $j \neq i$, resulta

$$g_{ik}g^{kj} = 0 \implies h_i^2 g^{ij} = 0 \implies g^{ij} = 0, \quad (j \neq i).$$

Por outro lado, se $j = i$, então

$$g_{ik}g^{ki} = 1 \implies g^{ii} = 1/h_i^2.$$

Portanto,

$$g^{ij} = \frac{\delta^{ij}}{h_i^2}$$

para um sistema ortogonal.

Da mesma forma,

$$g = h_1^2 h_2^2 h_3^2 \implies d^3 r = h_1 h_2 h_3 dq^1 dq^2 dq^3.$$

6.8 TRANSFORMAÇÕES GENERALIZADAS DE COORDENADAS E TENSORES GENERALIZADOS

A discussão realizada no capítulo 1 e na seção 6.7 restringiu-se a transformações realizadas do sistema Cartesiano para um sistema curvilíneo. Nesta seção esta discussão será generalizada para incluir uma transformação realizada entre dois sistemas de coordenadas quaisquer no R^n .

Parte-se então de um sistema de coordenadas no R^n previamente estabelecido, de tal forma que um ponto P neste espaço é localizado pela n -upla $P = (q^1, \dots, q^n)$. Generalizando as definições realizadas no capítulo 1 e na seção anterior, o ponto P consiste na intersecção de n hipersuperfícies coordenadas no R^n , sendo cada uma destas um subespaço de $n - 1$ dimensões no R^n determinado pela equação $q^i = c^i$ ($i = 1, \dots, n$), sendo $c^i = \text{cte}$. A curva coordenada de q^i consiste naquela intersecção das demais $n - 1$ hipersuperfícies; isto é, as coordenadas $\{q^j\}$ ($j \neq i$) permanecem constantes ao longo desta curva. Desta maneira, pode-se definir um conjunto de n vetores tangentes $e_i = \partial r / \partial q^i$, os quais formam a base covariante do sistema de coordenadas pois estes são linearmente independentes entre si.

Dada a base covariante, pode-se definir a *base contravariante* $\{e^i\}$, exigindo que esta base seja recíproca à base covariante; isto é, estes satisfazem a generalização da condição (6.23),

$$e_i \cdot e^j = \delta_i^j.$$

Desta maneira, um vetor a qualquer no R^n pode ser decomposto por meio das duas bases como $a = a^i e_i = a_i e^i$, sendo que agora $i = 1, \dots, n$ e o conjunto $\{a^i\}$ consiste nas *componentes contravariantes* de a , enquanto que $\{a_i\}$ contém as *componentes covariantes* do vetor.

Identificando por $Q = \{q^1, \dots, q^n\}$ o conjunto das coordenadas originais do ponto $P \in R^n$, considera-se agora o novo sistema de coordenadas curvilíneas $Q' = \{q'^1, \dots, q'^n\}$ para P , obtido a partir da lei de transformação

$$q'^i = q'^i(q^1, \dots, q^n), \quad (6.36a)$$

sendo que a transformação inversa

$$q^i = q^i(q'^1, \dots, q'^n) \quad (6.36b)$$

também deve estar definida.

Duas observações importantes a serem realizadas neste ponto são que as leis de transformação (6.36a,b) não são necessariamente lineares; ou seja, os elementos das matrizes de transformação podem ser funções das coordenadas, ao contrário do que foi considerado na seção 6.3, onde as matrizes de rotação, por exemplo, dependem somente de um ângulo fixo de rotação. Também devido a isso, as matrizes de transformação não são necessariamente ortogonais.

Como Q' contém por hipótese um conjunto de coordenadas válidas para um ponto $P \in R^n$ qualquer, pode-se definir então a base covariante neste novo sistema por

$$e'_i \doteq \frac{\partial \mathbf{r}}{\partial q'^i}, \quad (i = 1, \dots, n).$$

Contudo, dada a lei de transformação (6.36b), pode-se escrever

$$e'_i = \frac{\partial q^j}{\partial q'^i} \frac{\partial \mathbf{r}}{\partial q^j} = \frac{\partial q^j}{\partial q'^i} e_j, \quad (6.37a)$$

o que estabelece a mudança de base covariante decorrente da transformação $Q \rightarrow Q'$. Por sua vez, dada a transformação (6.36a), obtém-se mudança de base inversa associada

$$e_i = \frac{\partial q'^j}{\partial q^i} e'_j. \quad (6.37b)$$

De (6.37b) obtém-se imediatamente a lei de transformação das componentes contravariantes do vetor a , exigindo que o mesmo permaneça invariante na transformação $Q \rightarrow Q'$, i. e.,

$$a = a^i e_i = a'^i e'_i = \frac{\partial q'^i}{\partial q^j} a^j e_i \implies a'^i = \frac{\partial q'^i}{\partial q^j} a^j, \quad (i = 1, \dots, n). \quad (6.38a)$$

Por sua vez, (6.37a) fornece a lei de transformação inversa $Q' \rightarrow Q$,

$$a^i e_i = a'^i e'_i = \frac{\partial q^i}{\partial q'^j} a'^j e_i \implies a^i = \frac{\partial q^i}{\partial q'^j} a'^j, \quad (i = 1, \dots, n). \quad (6.38b)$$

As leis (6.38a,b) consistem nas condições necessárias e suficientes para que um conjunto de objetos $\{a^i\}$ contenha as componentes contravariantes de um tensor de posto um.

Seja agora o conjunto $\{e'^i\}$, o qual é recíproco à base covariante $\{e'_i\}$, i. e., $e'_i \cdot e'^j = \delta_i^j$. Observa-se que a condição de reciprocidade entre as bases em ambos os sistemas, juntamente com (6.37a), implica que a mudança de base contravariante frente a transformação $Q \rightarrow Q'$ deve ser dada por

$$e'^i = \frac{\partial q'^i}{\partial q^j} e^j, \quad (6.39a)$$

pois

$$\delta_i^j = e'_i \cdot e'^j = \frac{\partial q^k}{\partial q'^i} \frac{\partial q'^j}{\partial q^l} e_k \cdot e^l = \frac{\partial q^k}{\partial q'^i} \frac{\partial q'^j}{\partial q^l} \delta_k^l = \frac{\partial q^k}{\partial q'^i} \frac{\partial q'^j}{\partial q^k} = \frac{\partial q'^j}{\partial q'^i}.$$

Já frente a transformação $Q' \rightarrow Q$, a mudança de base é

$$e^i = \frac{\partial q^i}{\partial q'^j} e'^j, \tag{6.39b}$$

pois de (6.37b),

$$e_i \cdot e^j = \frac{\partial q'^k}{\partial q^i} \frac{\partial q^j}{\partial q'^\ell} e'_k \cdot e'^\ell = \frac{\partial q'^k}{\partial q^i} \frac{\partial q^j}{\partial q'^k} = \delta_i^j.$$

Assim, dada a decomposição do vetor a pelas bases contravariantes, $a = a_i e^i = a'_i e'^i$, juntamente com (6.39a,b), obtém-se as leis de transformação das componentes covariantes:

$$a'_i = \frac{\partial q^j}{\partial q'^i} a_j \tag{6.40a}$$

$$a_i = \frac{\partial q'^j}{\partial q^i} a'_j. \tag{6.40b}$$

Partindo das leis de transformação (6.38) e (6.40), definem-se agora os componentes contravariante, mistos e covariantes de um tensor generalizado de posto dois \mathcal{T} respectivamente como

$$T'^{ij} = \frac{\partial q'^i}{\partial q^k} \frac{\partial q^j}{\partial q'^\ell} T^{k\ell}, \quad T'^i_j = \frac{\partial q'^i}{\partial q^k} \frac{\partial q^\ell}{\partial q'^j} T^k_\ell, \quad T'_{ij} = \frac{\partial q^k}{\partial q'^i} \frac{\partial q^\ell}{\partial q'^j} T_{k\ell}.$$

Pode-se generalizar ainda mais a definição de um tensor com um número arbitrário de índices contravariantes e/ou covariantes. Esta forma generalizada possui várias designações na literatura. Assumindo que o tensor \mathcal{T} possua r índices na posição contravariante e s índices na posição covariante, o seu posto é $N = r + s$ e este é denominado um **tensor do tipo** (r, s) ou **tensor de valência** (r, s) ou ainda **tensor r -vezes contravariante e s -vezes covariante**. A notação empregada também varia na literatura. De acordo com a notação empregada neste texto até o presente momento, os índices mistos de um tensor devem apresentar-se espaçados. Contudo, esta regra pode ser flexibilizada para facilitar a notação e visualização das expressões.

Portanto, um tensor \mathcal{T} é do tipo (r, s) se os seus componentes transformarem-se de acordo com

$$T'^{i_1 \dots i_r}_{j_1 \dots j_s} = \frac{\partial q'^{i_1}}{\partial q^{k_1}} \dots \frac{\partial q'^{i_r}}{\partial q^{k_r}} \frac{\partial q^{\ell_1}}{\partial q'^{j_1}} \dots \frac{\partial q^{\ell_s}}{\partial q'^{j_s}} T^{k_1 \dots k_r}_{\ell_1 \dots \ell_s}. \tag{6.41}$$

Por conveniência de notação, estes componentes também podem ser representados simplesmente por $T'^{i_1 \dots i_r}_{j_1 \dots j_s}$.

Os componentes do tensor generalizado definido por (6.41) seguem as mesmas regras algébricas dos tensores Cartesianos discutidas na seção 6.4.

Exercício 6.7. Verifique se a delta de Kronecker satisfaz a lei de transformação (6.41).

Solução. Dado δ^i_j , se este símbolo é de fato um tensor misto de posto dois, então frente a uma transformação de coordenadas generalizada,

$$\delta^{i_i}_{j_j} = \frac{\partial q'^i}{\partial q^k} \frac{\partial q^\ell}{\partial q'^j} \delta^k_\ell = \frac{\partial q'^i}{\partial q^k} \frac{\partial q^k}{\partial q'^j} = \frac{\partial q'^i}{\partial q'^j} = \delta^i_j,$$

onde foi empregada derivação em cadeia. Ou seja, a delta de Kronecker possui os mesmos componentes em qualquer sistema de coordenadas, demonstrando ser um tensor misto de posto dois. A expressão (6.30) já havia mostrado que o tensor de métrica na forma mista é justamente dado por esta delta.

Exercício 6.8. Mostre que $g_{ij} = e_i \cdot e_j$ são os componentes covariantes de um tensor de posto dois.

Demonstração. Em uma transformação generalizada, a nova base $\{e'_i\}$ deve levar aos componentes

$$g'_{ij} = e'_i \cdot e'_j.$$

Usando a mudança de base em (6.37a), resulta

$$g'_{ij} = \frac{\partial q^k}{\partial q'^i} \frac{\partial q^\ell}{\partial q'^j} e_k \cdot e_\ell = \frac{\partial q^k}{\partial q'^i} \frac{\partial q^\ell}{\partial q'^j} g_{k\ell}.$$

Ou seja, os componentes de $\{g_{ij}\}$ transformam-se como um tensor covariante de posto dois. \square

6.9 TENSORES RELATIVOS

Na seção 6.6.1 discutiu-se a diferença entre um tensor Cartesiano e um pseudotensor. Esta distinção será novamente discutida em um contexto mais amplo, generalizando esta distinção para qualquer transformação entre sistemas de coordenadas.

Um **tensor relativo** ou **densidade de tensor** é aquele objeto que se transforma como um tensor entre um sistema de coordenadas a outro, exceto que sua lei de transformação difere de (6.41) pela multiplicação de um *peso*, correspondente a uma determinada potência do Jacobiano da transformação.

Observando a lei de transformação dos componentes contravariantes de um vetor (6.38a), percebe-se que as quantidades $\{\partial q'^i/\partial q^j\}$ podem ser escritas na forma de uma matriz J , cujos elementos são definidos por

$$J^i_j = \frac{\partial q'^i}{\partial q^j}, \quad (6.42a)$$

e cujo determinante $J \equiv \det(J)$ é justamente o Jacobiano da transformação $\{q^i\} \rightarrow \{q'^i\}$. Assim, pode-se escrever de forma compacta $a'^i = J^i_j a^j$.

Por sua vez, a transformação inversa (6.38b) também pode ser escrita em termos de uma matriz K , definida por

$$K^i_j = \frac{\partial q^i}{\partial q'^j}, \quad (6.42b)$$

ou seja, $a^i = K^i_j a'^j$. É fácil mostrar que $K = J^{-1}$ e, portanto, $K \equiv \det(K) = J^{-1}$.

Com a mesma notação, as leis de transformação dos componentes covariantes de um vetor, dadas por (6.40a,b), podem ser escritas de forma compacta como

$$a'_i = K^j_i a_j \quad \text{e} \quad a_i = J^j_i a'_j.$$

Usando então o determinante K , define-se o **tensor relativo de peso** w como o conjunto $\{T^{i_1 \dots i_r}_{j_1 \dots j_s}\}$, composto por n^{r+s} objetos matemáticos identificados por r índices contravariantes e s índices covariantes que se transformam de acordo com a lei

$$T'^{i_1 \dots i_r}_{j_1 \dots j_s} = K^w \frac{\partial q'^{i_1}}{\partial q^{k_1}} \dots \frac{\partial q'^{i_r}}{\partial q^{k_r}} \frac{\partial q^{\ell_1}}{\partial q'^{j_1}} \dots \frac{\partial q^{\ell_s}}{\partial q'^{j_s}} T^{k_1 \dots k_r}_{\ell_1 \dots \ell_s}. \quad (6.43)$$

A partir da definição (6.43), os diferentes valores que o peso w pode assumir estabelecem classificações dos tensores relativos que muitas vezes depende do texto consultado na literatura. Uma classificação distingue entre **densidades de tensores (autênticos)** ou **densidades tensoriais** e **densidades de pseudotensores**, quando o peso w é uma quantidade inteira. De acordo com esta classificação, chamam-se:

Tensores verdadeiros ou absolutos: quando o peso é $w = 0$, ou seja, o tensor segue a lei (6.41).

Densidade de tensor (autêntico) de peso w (inteiro): quando o peso é $w \geq 1$. Em particular, quando $w = +1$, este objeto é chamado simplesmente de *densidade de tensor* ou *densidade tensorial*.

Densidade de pseudotensor de peso w (inteiro): quando o peso é $w \leq -1$. Em particular, quando $w = -1$, este objeto é chamado simplesmente de *pseudotensor*. Cabe mencionar aqui que tanto objetos com peso $w = -1$ quanto com peso $w = +1$ são classificados como pseudotensores escalares, de acordo com a lei de transformação (6.21).

Esta classificação exige que o peso w seja inteiro, uma vez que no caso contrário a transformação (6.43) não é necessariamente unívoca.

Uma outra classificação que pode ser encontrada, para a qual o peso w não necessita ser necessariamente inteiro. Segundo esta classificação, as densidades de tensores podem ser:

Densidade par de tensor de peso w : quando o objeto se transforma de acordo com

$$T'^{i_1 \dots i_r}_{j_1 \dots j_s} = |K|^w \frac{\partial q'^{i_1}}{\partial q^{k_1}} \dots \frac{\partial q'^{i_r}}{\partial q^{k_r}} \frac{\partial q^{\ell_1}}{\partial q'^{j_1}} \dots \frac{\partial q^{\ell_s}}{\partial q'^{j_s}} T^{k_1 \dots k_r}_{\ell_1 \dots \ell_s}.$$

Densidade ímpar de tensor de peso w : quando o objeto se transforma de acordo com

$$T'^{i_1 \dots i_r}_{j_1 \dots j_s} = \text{sgn}(K) |K|^w \frac{\partial q'^{i_1}}{\partial q^{k_1}} \dots \frac{\partial q'^{i_r}}{\partial q^{k_r}} \frac{\partial q^{\ell_1}}{\partial q'^{j_1}} \dots \frac{\partial q^{\ell_s}}{\partial q'^{j_s}} T^{k_1 \dots k_r}_{\ell_1 \dots \ell_s},$$

onde $\text{sgn}(K)$ é o sinal de K .

Neste texto, a lei de transformação geral a ser adotada para tensores relativos continuará a ser (6.43).

Exercício 6.9. Sendo $\{g_{ij}\}$ as componentes covariantes do tensor de métrica, mostre que o determinante $g = \det(g)$ é uma densidade escalar de peso $w = +2$.

Solução. Os componentes do tensor de métrica transformam-se de acordo com (6.41),

$$g'_{ij} = \frac{\partial q^k}{\partial q'^i} \frac{\partial q^\ell}{\partial q'^j} g_{k\ell}.$$

A representação matricial desta lei de transformação é

$$g' = K^T g K.$$

Portanto,

$$g' = \det(g') = \det(K^T g K) = K^2 g,$$

o que mostra que g é uma densidade escalar (ou escalar relativo) de peso $w = +2$.

Algumas propriedades de tensores relativos podem ser agora apresentadas:

1. A transformação tensorial é **transitiva**, ou seja, ao se realizarem duas transformações seguidas, representadas esquematicamente por $\{q_i\} \rightarrow \{q'_i\} \rightarrow \{q''_i\}$, o tensor resultante pode ser descrito por uma única transformação $\{q_i\} \rightarrow \{q''_i\}$, a qual é dada pela composição das transformações anteriores. Esta propriedade já havia sido discutida no contexto de tensores Cartesianos na seção 6.5.1.
2. Tensores relativos de mesmo posto, tipo e peso podem ser adicionados e a soma resulta em um outro tensor relativo com as mesmas características.
3. Se dois tensores relativos possuem pesos w_1 e w_2 , respectivamente, então os produtos externo ou interno dos mesmos resulta em um novo tensor relativo com peso $w_1 + w_2$.
4. A operação de contração aplicada a um tensor relativo resulta em um outro tensor relativo com o mesmo peso que o tensor original.
5. O resultado obtido no exercício 6.9 mostra que qualquer tensor relativo \mathcal{T} de peso w pode gerar um tensor absoluto de mesmo posto através do produto externo

$$U^{i_1 \dots i_r}_{j_1 \dots j_s} = g^{-w/2} T^{i_1 \dots i_r}_{j_1 \dots j_s},$$

pois ao se realizar a transformação, resulta

$$U'^{i_1 \dots i_r}_{j_1 \dots j_s} = (g')^{-w/2} T'^{i_1 \dots i_r}_{j_1 \dots j_s} = (K^2 g)^{-w/2} K^w \frac{\partial q'^{i_1}}{\partial q^{k_1}} \dots \frac{\partial q'^{i_r}}{\partial q^{k_r}} \frac{\partial q^{\ell_1}}{\partial q'^{j_1}} \dots \frac{\partial q^{\ell_s}}{\partial q'^{j_s}} T^{k_1 \dots k_r}_{\ell_1 \dots \ell_s},$$

o qual se comporta como um tensor verdadeiro.

Exercício 6.10. Calcule os pesos dos símbolos de Levi-Civita ϵ_{ijk} e ϵ^{ijk} .

Solução. Partindo de ϵ_{ijk} , válido em um determinado sistema de coordenadas, deseja-se aplicar sobre o mesmo a transformação $\{q_i\} \rightarrow \{q'_i\}$. Fazendo isso, resulta o objeto

$$\frac{\partial q^\ell}{\partial q'^i} \frac{\partial q^m}{\partial q'^j} \frac{\partial q^n}{\partial q'^k} \epsilon_{\ell mn} = K^\ell_i K^m_j K^n_k \epsilon_{\ell mn}.$$

Contudo, de acordo com a identidade matricial (6.33b), $K^\ell_i K^m_j K^n_k \epsilon_{\ell mn} = K \epsilon_{ijk}$. Portanto, a lei de transformação correta para ϵ_{ijk} , a qual irá resultar em um tensor ϵ'_{ijk} cujos componentes têm os valores esperados para o símbolo de Levi-Civita em qualquer sistema de coordenadas, é

$$\epsilon'_{ijk} = K^{-1} \frac{\partial q^\ell}{\partial q'^i} \frac{\partial q^m}{\partial q'^j} \frac{\partial q^n}{\partial q'^k} \epsilon_{\ell mn}.$$

Portanto, $\{\epsilon_{ijk}\}$ é uma densidade tensorial de peso $w = -1$.

Já para o símbolo ϵ^{ijk} , a mesma transformação de coordenadas resulta no objeto

$$\frac{\partial q^i}{\partial q^\ell} \frac{\partial q^j}{\partial q^m} \frac{\partial q^k}{\partial q^n} \epsilon^{\ell mn} = J_\ell^i J_m^j J_n^k \epsilon^{\ell mn} = J \epsilon^{ijk} = K^{-1} \epsilon^{ijk}.$$

Portanto, o símbolo ϵ^{ijk} deve se transformar como

$$\epsilon^{ijk} = K \frac{\partial q^i}{\partial q^\ell} \frac{\partial q^j}{\partial q^m} \frac{\partial q^k}{\partial q^n} \epsilon^{\ell mn},$$

correspondendo a uma densidade tensorial de peso $w = +1$.

6.10 DERIVADAS DOS VETORES DE BASE E OS SÍMBOLOS DE CHRISTOFFEL

No sistema Cartesiano, os vetores da base canônica $\{\hat{x}_i\}$ são constantes e por isso a aplicação de qualquer operador diferencial sobre um campo vetorial irá atuar somente sobre os componentes do campo. Por outro lado, em um sistema de coordenadas generalizado os vetores de base $\{e_i\}$ e $\{e^i\}$ são funções das próprias coordenadas curvilíneas. Portanto, qualquer operador diferencial irá atuar também sobre estes vetores.

Considera-se então uma derivada qualquer atuando sobre um determinado vetor de base; por exemplo, a operação $\partial e_i / \partial q^j$. O resultado desta derivação é também um vetor, o qual pode, portanto, ser expresso em termos da própria base $\{e_i\}$. Dado então o vetor $e_i = e_i(q^1, q^2, q^3)$, escreve-se

$$\frac{\partial e_i}{\partial q^j} = \Gamma^k_{ij} e_k, \tag{6.44a}$$

onde Γ^k_{ij} é denominado **símbolo de Christoffel (de três índices) do segundo tipo**. Os símbolos de Christoffel podem ser visualizados como matrizes $n \times n \times n$, possuindo portanto n^3 elementos ao todo no R^n . Na literatura, uma outra notação empregada com frequência para os símbolos de Christoffel do segundo tipo é:

$$\Gamma^i_{jk} \equiv \left\{ \begin{matrix} i \\ j \quad k \end{matrix} \right\}.$$

Usando a relação de reciprocidade (6.23), os valores de Γ^k_{ij} podem ser obtidos por

$$\Gamma^k_{ij} = e^k \cdot \frac{\partial e_i}{\partial q^j}. \tag{6.44b}$$

Com a mesma relação pode-se obter também a expressão para as derivadas de $\{e^i\}$. Escrevendo

$$\frac{\partial e^i}{\partial q^j} = \alpha^i_{jk} e^k,$$

observa-se que

$$e_k \cdot \frac{\partial e^i}{\partial q^j} = \alpha^i_{jk}.$$

Porém, derivando a relação de reciprocidade, obtém-se que

$$e_k \cdot \frac{\partial e^i}{\partial q^j} = -\Gamma^i_{kj}.$$

Ou seja, $\alpha^i_{jk} = -\Gamma^i_{kj}$ e

$$\frac{\partial e^i}{\partial q^j} = -\Gamma^i_{kj} e^k. \tag{6.44c}$$

Exercício 6.11. Mostre que o conjunto dos símbolos de Christoffel $\{\Gamma^i_{kj}\}$ não são componentes de um tensor misto de terceira ordem.

Solução. Aplicando uma transformação arbitrária sobre (6.44b), no novo sistema de coordenadas resulta

$$\Gamma'^k_{ij} = e'^k \cdot \frac{\partial e'_i}{\partial q'^j}.$$

Mas, conforme as leis de transformação (6.37a) e (6.39a), pode-se escrever

$$\begin{aligned} \Gamma'^k_{ij} &= \frac{\partial q'^k}{\partial q^\ell} e^\ell \cdot \frac{\partial}{\partial q'^j} \left(\frac{\partial q^m}{\partial q'^i} e_m \right) \\ &= \frac{\partial q'^k}{\partial q^\ell} \left(\frac{\partial^2 q^m}{\partial q'^j \partial q'^i} e^\ell \cdot e_m + \frac{\partial q^m}{\partial q'^i} e^\ell \cdot \frac{\partial e_m}{\partial q'^j} \right) \\ &= \frac{\partial q'^k}{\partial q^\ell} \left(\frac{\partial^2 q^\ell}{\partial q'^j \partial q'^i} + \frac{\partial q^m}{\partial q'^i} \frac{\partial q^n}{\partial q'^j} e^\ell \cdot \frac{\partial e_m}{\partial q^n} \right), \\ \Gamma'^k_{ij} &= \frac{\partial q'^k}{\partial q^\ell} \frac{\partial^2 q^\ell}{\partial q'^j \partial q'^i} + \frac{\partial q'^k}{\partial q^\ell} \frac{\partial q^m}{\partial q'^i} \frac{\partial q^n}{\partial q'^j} \Gamma^\ell_{mn}. \end{aligned} \quad (6.45)$$

Observa-se que somente o segundo termo do lado direito está de acordo com a lei de transformação de um tensor de posto 3. A presença do termo adicional mostra que o símbolo de Christoffel não é um tensor.

Uma expressão alternativa para o cálculo de Γ^k_{ij} , em termos do tensor de métrica, pode ser obtida. Inicialmente, observa-se que o símbolo de Christoffel é simétrico frente a permutação $i \leftrightarrow j$. Isto é facilmente demonstrado a partir da definição (6.24). Uma vez que

$$\frac{\partial e_i}{\partial q^j} = \frac{\partial^2 \mathbf{r}}{\partial q^j \partial q^i} = \frac{\partial^2 \mathbf{r}}{\partial q^i \partial q^j} = \frac{\partial e_j}{\partial q^i},$$

resulta de (6.44a,b) que

$$\frac{\partial e_i}{\partial q^j} = \Gamma^k_{ij} e_k = \frac{\partial e_j}{\partial q^i} \implies \Gamma^k_{ij} = \frac{\partial e_j}{\partial q^i} \cdot e^k = \Gamma^k_{ji}.$$

Considera-se então a definição (6.27a). Derivando a mesma, obtém-se

$$\begin{aligned} \frac{\partial g_{ij}}{\partial q^k} &= \frac{\partial e_i}{\partial q^k} \cdot e_j + e_i \cdot \frac{\partial e_j}{\partial q^k} \\ &= \Gamma^\ell_{ik} e_\ell \cdot e_j + \Gamma^\ell_{jk} e_i \cdot e_\ell \\ &= \Gamma^\ell_{ik} g_{\ell j} + \Gamma^\ell_{jk} g_{i \ell}. \end{aligned}$$

Realizando duas permutações cíclicas dos índices livres i, j e k neste resultado, resultam as expressões

$$\begin{aligned} \frac{\partial g_{jk}}{\partial q^i} &= \Gamma^\ell_{ji} g_{\ell k} + \Gamma^\ell_{ki} g_{j \ell} \\ \frac{\partial g_{ki}}{\partial q^j} &= \Gamma^\ell_{kj} g_{\ell i} + \Gamma^\ell_{ij} g_{k \ell}. \end{aligned}$$

Somando as duas últimas expressões e subtraindo a primeira e fazendo uso da simetria de g_{ij} e Γ^k_{ij} , resulta

$$\frac{\partial g_{jk}}{\partial q^i} + \frac{\partial g_{ki}}{\partial q^j} - \frac{\partial g_{ij}}{\partial q^k} = 2\Gamma^\ell_{ij} g_{k \ell}.$$

Realizando o produto interno com g^{mk} e usando (6.29), resulta finalmente

$$\Gamma^k_{ij} = \frac{1}{2} g^{k \ell} \left(\frac{\partial g_{j \ell}}{\partial q^i} + \frac{\partial g_{\ell i}}{\partial q^j} - \frac{\partial g_{ij}}{\partial q^\ell} \right). \quad (6.46a)$$

Uma maneira prática de se visualizar os símbolos de Christoffel é oferecida definindo três matrizes (no R^3) 3×3 simétricas, cada uma correspondendo a um valor do índice k em (6.46a), ou seja,

$$\Gamma^k \equiv [\Gamma^k_{ij}] = \begin{pmatrix} \Gamma^k_{11} & \Gamma^k_{12} & \Gamma^k_{13} \\ \Gamma^k_{21} & \Gamma^k_{22} & \Gamma^k_{23} \\ \Gamma^k_{31} & \Gamma^k_{32} & \Gamma^k_{33} \end{pmatrix}, \quad (k = 1, 2, 3). \quad (6.46b)$$

Exercício 6.12. Calcule os símbolos de Christoffel para o sistema de coordenadas cilíndricas.

Solução. Usando os resultados da seção 1.5.1, o tensor de métrica e sua inversa são dados por

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \implies g^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Portanto, a primeira conclusão é que $g_{ij} = h_i^2 \delta_{ij}$, resultando para (6.46),

$$\Gamma^k_{ij} = \frac{1}{2} \left(h_j^{-2} \frac{\partial h_j^2}{\partial q^i} \delta^{kj} + h_i^{-2} \frac{\partial h_i^2}{\partial q^j} \delta^{ki} - h_k^{-2} \frac{\partial h_i^2}{\partial q^k} \delta_{ij} \right),$$

sendo que não existem somas implícitas nos índices duplos. Observa-se então que somente são não nulos os símbolos onde $i = 2$ e/ou $j = 2$ e/ou $k = 2$. Ou seja,

$$\Gamma^{\phi}_{\rho\phi} = \Gamma^{\phi}_{\phi\rho} = \frac{1}{\rho}, \quad \Gamma^{\rho}_{\phi\phi} = -\rho.$$

Todas as outras possibilidades resultam nulas.

Em termos das representações matriciais,

$$\Gamma^{\rho} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\rho & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Gamma^{\phi} = \begin{pmatrix} 0 & 1/\rho & 0 \\ 1/\rho & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \Gamma^z = 0.$$

Os **símbolos de Christoffel (de três índices) do primeiro tipo** correspondem à expansão equivalente a (6.44a), porém em termos dos vetores da base contravariante $\{e^i\}$, ou seja,

$$\frac{\partial e_i}{\partial q^j} = [ij, k] e^k, \quad (6.47a)$$

onde $[ij, k]$ é a notação usualmente empregada na literatura para os símbolos de Christoffel do primeiro tipo.

Da mesma forma como se procedeu com Γ^i_{jk} , as expressões para estes símbolos são obtidas a partir da relação de reciprocidade, resultando em

$$[ij, k] = \frac{\partial e_i}{\partial q^j} \cdot e_k. \quad (6.47b)$$

Como $e_i = g_{i\ell} e^{\ell}$, resulta de (6.44b) a seguinte relação entre ambos os tipos de símbolos,

$$[ij, k] = g_{k\ell} \Gamma^{\ell}_{ij} = g_{k\ell} \left\{ \begin{matrix} \ell \\ i & j \end{matrix} \right\} \iff \Gamma^k_{ij} = \left\{ \begin{matrix} k \\ i & j \end{matrix} \right\} = g^{k\ell} [ij, \ell]. \quad (6.47c)$$

Este último resultado mostra também que $[ij, k] = [ji, k]$.

Finalmente, usando (6.46) obtém-se a seguinte expressão para $[ij, k]$ em termos do tensor de métrica,

$$[ij, k] = \frac{1}{2} \left(\frac{\partial g_{ik}}{\partial q^j} + \frac{\partial g_{jk}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^k} \right). \quad (6.47d)$$

6.11 DIFERENCIAÇÃO COVARIANTE

Nesta seção introduz-se o conceito de *derivada covariante*, a qual está envolvida no cálculo de operadores diferenciais aplicados a campos tensoriais em sistemas de coordenadas generalizadas. Para iniciar a discussão, são considerados os objetos matemáticos resultantes das variações arbitrárias de campos tensoriais de mais baixa ordem, ou seja, campos escalares e vetoriais.

Para o caso de um campo escalar $\psi = \psi(\mathbf{r})$, sua variação infinitesimal, quando calculada no sistema Cartesiano, pode sempre ser escrita como $d\psi = \nabla\psi \cdot d\mathbf{r}$, sendo o gradiente escrito em coordenadas Cartesianas. No sistema curvilíneo, a mesma variação resulta em

$$d\psi = \frac{\partial\psi}{\partial q^i} dq^i.$$

Como $\{dq^i\}$ são os componentes de um vetor contravariante e $d\psi$ é um escalar, resulta que, de acordo com a lei do quociente, $\{\partial\psi/\partial q^i\}$ formam os componentes de um vetor covariante.

Por outro lado, para o campo vetorial $v = v^i e_i$, embora em coordenadas Cartesianas $\{\partial v^i/\partial x^j\}$ contém os componentes de um tensor de posto dois, esta regra não é geral para qualquer sistema de coordenadas, pois $\partial v^i/\partial q^j$ se transforma de acordo com

$$\begin{aligned} \left(\frac{\partial v^i}{\partial q^j}\right)' &= \frac{\partial v'^i}{\partial q'^j} = \frac{\partial q^k}{\partial q'^j} \frac{\partial}{\partial q^k} \left(\frac{\partial q^i}{\partial q^\ell} v^\ell\right) \\ &= \frac{\partial q^k}{\partial q'^j} \frac{\partial^2 q^i}{\partial q^k \partial q^\ell} v^\ell + \frac{\partial q^i}{\partial q^\ell} \frac{\partial q^k}{\partial q'^j} \frac{\partial v^\ell}{\partial q^k}, \end{aligned}$$

o qual não segue a lei de transformação de um tensor generalizado de segunda ordem. Este objeto, a propósito, transforma-se de maneira semelhante aos símbolos de Christoffel. Este resultado não segue a lei de transformação de tensores porque a matriz de transformação J é, em geral, uma função das coordenadas, ou seja, $J = J(q^1, \dots, q^n)$.

Ao se calcular a derivada de um campo vetorial v em relação a uma coordenada q^j qualquer, é necessário levar em conta também a geometria do sistema adotado, a qual é descrita pelos símbolos de Christoffel. Neste caso, pode-se escrever

$$\frac{\partial v}{\partial q^j} = \frac{\partial}{\partial q^j} (v^i e_i) = \frac{\partial v^i}{\partial q^j} e_i + v^i \frac{\partial e_i}{\partial q^j} = \frac{\partial v^i}{\partial q^j} e_i + v^i \Gamma^k_{ij} e_k = \left(\frac{\partial v^i}{\partial q^j} + v^k \Gamma^i_{kj}\right) e_i.$$

Agora o objeto obtido na expressão acima comporta-se como um tensor de posto dois. Este objeto é denominado a *derivada covariante* do vetor v e é representado por

$$v^i_{;j} \equiv \frac{\partial v^i}{\partial q^j} + \Gamma^i_{kj} v^k \equiv v^i_{,j} + \Gamma^i_{kj} v^k. \quad (6.48a)$$

Em (6.48a), a notação “;j” denota a diferenciação covariante em relação a coordenada q^j , enquanto que “,j” indica a diferenciação usual na mesma coordenada, i. e., $v^i_{,j} = \partial v^i/\partial q^j$. Usando o símbolo da diferenciação covariante, pode-se escrever então

$$\frac{\partial v}{\partial q^j} = v^i_{;j} e_i. \quad (6.48b)$$

Exercício 6.13. Mostre que $\{v^i_{;j}\}$ forma os componentes de um tensor misto de posto dois.

Solução. Realizando a transformação $\{q^i\} \rightarrow \{q'^i\}$,

$$(v^i_{;j})' = v'^i_{;j} = \frac{\partial v'^i}{\partial q'^j} + v'^k \Gamma'^i_{kj}.$$

Usando a lei de transformação de Γ^k_{ij} dada por (6.45), pode-se escrever

$$\begin{aligned} v'^i_{;j} &= \frac{\partial q^\ell}{\partial q'^j} \frac{\partial}{\partial q^\ell} \left(\frac{\partial q'^i}{\partial q^m} v^m\right) + \frac{\partial q'^k}{\partial q^r} v^r \left(\frac{\partial q'^i}{\partial q^\ell} \frac{\partial^2 q^\ell}{\partial q'^j \partial q'^k} + \frac{\partial q'^i}{\partial q^\ell} \frac{\partial q^m}{\partial q'^k} \frac{\partial q^n}{\partial q'^j} \Gamma^{\ell}_{mn}\right) \\ &= \frac{\partial q^\ell}{\partial q'^j} \frac{\partial^2 q'^i}{\partial q^\ell \partial q^m} v^m + \frac{\partial q'^k}{\partial q^r} \frac{\partial q'^i}{\partial q^\ell} \frac{\partial^2 q^\ell}{\partial q'^j \partial q'^k} v^r + \frac{\partial q^\ell}{\partial q'^j} \frac{\partial q'^i}{\partial q^m} \frac{\partial v^m}{\partial q^\ell} + \frac{\partial q'^i}{\partial q^\ell} \frac{\partial q^n}{\partial q'^j} v^m \Gamma^{\ell}_{mn}. \end{aligned}$$

Mas os dois primeiros termos resultam

$$\begin{aligned} \frac{\partial q^\ell}{\partial q'^j} \frac{\partial^2 q'^i}{\partial q^\ell \partial q^m} v^m + \frac{\partial q'^k}{\partial q^r} \frac{\partial q'^i}{\partial q^\ell} \frac{\partial^2 q^\ell}{\partial q'^j \partial q'^k} v^r &= \left[\frac{\partial q^\ell}{\partial q'^j} \frac{\partial^2 q'^i}{\partial q^\ell \partial q^m} + \frac{\partial q'^i}{\partial q^\ell} \left(\frac{\partial q'^k}{\partial q^m} \frac{\partial}{\partial q'^k}\right) \frac{\partial q^\ell}{\partial q'^j} \right] v^m \\ &= \left(\frac{\partial^2 q'^i}{\partial q^m \partial q^\ell} \frac{\partial q^\ell}{\partial q'^j} + \frac{\partial q'^i}{\partial q^\ell} \frac{\partial^2 q^\ell}{\partial q^m \partial q'^j} \right) v^m = \frac{\partial}{\partial q^m} \left(\frac{\partial q'^i}{\partial q^\ell} \frac{\partial q^\ell}{\partial q'^j} \right) v^m = \frac{\partial \delta^i_j}{\partial q^m} v^m = 0. \end{aligned}$$

Portanto,

$$v'^i_{;j} = \frac{\partial q'^i}{\partial q^m} \frac{\partial q^\ell}{\partial q'^j} \left(\frac{\partial v^m}{\partial q^\ell} + v^n \Gamma^m_{n\ell} \right) = \frac{\partial q'^i}{\partial q^k} \frac{\partial q^\ell}{\partial q'^j} v^k_{;\ell},$$

o qual segue a lei de transformação de um tensor misto de posto dois.

Exercício 6.14. Calcule a derivada covariante $v^i_{;j}$ em coordenadas cilíndricas. Em seguida, calcule a contração $v^i_{;i}$.

Solução. Da definição (6.48) e dos valores de Γ^k_{ij} para as coordenadas cilíndricas obtidos no exercício 6.12, resulta, para $\{i, j\} = \{1, 2, 3\} \equiv \{\rho, \phi, x^3\}$,

$$v^i_{;j} = \frac{\partial v^i}{\partial q^j} + \frac{\delta^{i\phi}}{\rho} (v^\rho \delta_{j\phi} + v^\phi \delta_{j\rho}) - \rho v^\phi \delta^{i\rho} \delta_{j\phi}.$$

Realizando agora contração,

$$v^i_{;i} = \frac{\partial v^i}{\partial q^i} + \frac{v^\rho}{\rho} = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho v^\rho) + \frac{\partial v^\phi}{\partial \phi} + \frac{\partial v^3}{\partial x^3},$$

a qual é exatamente a expressão para o divergente do campo v em coordenadas cilíndricas, obtido na seção 1.5.1.

A expressão equivalente a (6.48), porém para os componentes covariantes do vetor v , pode ser obtida derivando agora $v = v_i e^i$, com o emprego de (6.44c), ou seja,

$$\frac{\partial}{\partial q^j} (v_i e^i) = \left(\frac{\partial v_i}{\partial q^j} - v_k \Gamma^k_{ij} \right) e^i.$$

Chamando $v_{i;j}$ a derivada covariante da i -ésima componente covariante do vetor v , obtém-se

$$v_{i;j} = \frac{\partial v_i}{\partial q^j} - \Gamma^k_{ij} v_k = v_{i,j} - \Gamma^k_{ij} v_k. \quad (6.49a)$$

Portanto, pode-se escrever

$$\frac{\partial v}{\partial q^j} = v_{i;j} e^i. \quad (6.49b)$$

Derivadas covariantes de tensores de ordens mais altas podem ser obtidas seguindo o mesmo procedimento adotado acima. Por exemplo, um tensor de posto dois possui ao todo 4 representações: 1 contravariante, 1 covariante e 2 mistas. Para realizar em detalhes a derivação covariante para os componentes contravariantes do tensor \mathcal{T} , escreve-se o mesmo como $\mathcal{T} = T^{ij} e_i \otimes e_j$. Derivando então em relação a q^k ,

$$\begin{aligned} \frac{\partial \mathcal{T}}{\partial q^k} &= \frac{\partial}{\partial q^k} (T^{ij} e_i \otimes e_j) \\ &= \frac{\partial T^{ij}}{\partial q^k} e_i \otimes e_j + T^{ij} \frac{\partial e_i}{\partial q^k} \otimes e_j + T^{ij} e_i \otimes \frac{\partial e_j}{\partial q^k} \\ &= \left(\frac{\partial T^{ij}}{\partial q^k} + \Gamma^i_{\ell k} T^{\ell j} + \Gamma^j_{\ell k} T^{i \ell} \right) e_i \otimes e_j. \end{aligned}$$

Chamando então $T^{ij}_{;k}$ como a derivada covariante dos componentes contravariantes do tensor \mathcal{T} , identifica-se

$$T^{ij}_{;k} = \frac{\partial T^{ij}}{\partial q^k} + \Gamma^i_{\ell k} T^{\ell j} + \Gamma^j_{\ell k} T^{i \ell}, \quad (6.50a)$$

podendo escrever

$$\frac{\partial \mathcal{T}}{\partial q^k} = T^{ij}_{;k} e_i \otimes e_j.$$

Pode-se mostrar facilmente que o conjunto $\{T^{ij}_{;k}\}$ forma os componentes de um tensor de posto três.

Expressões equivalentes a (6.50a) para as outras representações de \mathcal{T} podem ser derivadas da mesma maneira. Estas são:

$$T^i_{j;k} = T^i_{j,k} + \Gamma^i_{\ell k} T^\ell_j - \Gamma^\ell_{jk} T^i_\ell \quad (6.50b)$$

$$T_{ij;k} = T_{ij,k} - \Gamma^{\ell}_{ik} T_{\ell j} - \Gamma^{\ell}_{jk} T_{i \ell}, \quad (6.50c)$$

onde o índice “ k ” indica agora a derivação usual nesta coordenada, i. e., $T_{ij,k} = \partial T_{ij} / \partial q^k$.

Expressões para as derivadas covariantes de tensores de postos maiores que dois podem ser derivadas seguindo os modelos fornecidos pelas expressões (6.48), (6.49) e (6.50a-c): a derivada covariante de cada componente de um tensor de posto N é composta por $N+1$ termos. O primeiro termo é sempre a derivada direta do componente em relação à coordenada; os demais N termos são combinações de produtos dos símbolos de Christoffel com os componentes do tensor. Para cada índice na posição contravariante, o termo correspondente na derivada covariante é somado, enquanto que para cada índice covariante, o termo correspondente é subtraído. Em cada um destes termos, o último índice do símbolo de Christoffel é sempre a coordenada que está sendo derivada. Por sua vez, os demais índices são sempre tais que a convenção de somas implícitas é respeitada para a particular representação do tensor que está sendo derivada.

A derivação covariante também serve para atribuir significado à aplicação do operador nabla (∇) sobre um tensor de posto N , de tal forma que o objeto matemático resultante seja um tensor de posto $N + 1$ que satisfaz a lei de transformação. Por exemplo, para o tensor \mathcal{T} em (6.50a-c), a operação $\mathcal{U} = \nabla\mathcal{T}$ resulta em um tensor de posto três, sendo que algumas de suas representações possíveis são

$$\mathcal{U} = T^{ij}_{;k} e_i \otimes e_j \otimes e^k = T^i_{j;k} e_i \otimes e^j \otimes e^k = T_{ij;k} e^i \otimes e^j \otimes e^k.$$

Retornando rapidamente aos tensores de posto zero, ou seja, a escalares, como estes não fazem uso de vetores de base para a sua representação, resulta que a derivada covariante de $\psi(q^1, \dots, q^n)$ é idêntica à derivação direta, ou seja,

$$\nabla\psi = \psi_{;j} e^j = \frac{\partial\psi}{\partial q^j} e^j = \psi_{,j} e^j.$$

Para a física em espaços curvos, em particular para a relatividade geral e gravitação, a derivada covariante possui uma importância ímpar, pois

...(a) substituição consistente das derivadas parciais regulares por derivadas covariantes carrega as leis da física (na forma de componentes) a partir do espaço-tempo plano para o espaço-tempo (Riemanniano) curvo da relatividade geral. De fato, esta substituição pode ser tomada como uma expressão matemática do Princípio da Equivalência de Einstein.

Finalmente, um teorema importante para a derivada covariante do tensor de métrica é apresentado a seguir.

Teorema 6.1 (Teorema de Ricci). *A derivada covariante do tensor de métrica é identicamente nula.*

Demonstração. Este teorema será inicialmente demonstrado para a forma covariante do tensor de métrica. De acordo com (6.50c),

$$g_{ij;k} = g_{ij,k} - \Gamma^{\ell}_{ik} g_{\ell j} - \Gamma^{\ell}_{jk} g_{i\ell}.$$

Então, introduzindo (6.46) na expressão acima e usando também (6.47c,d), resulta

$$\begin{aligned} g_{ij;k} &= g_{ij,k} - [ik, j] - [jk, i] \\ &= g_{ij,k} - \frac{1}{2} g_{ji,k} - \frac{1}{2} g_{ij,k} - \frac{1}{2} g_{kj,i} + \frac{1}{2} g_{jk,i} + \frac{1}{2} g_{ik,j} - \frac{1}{2} g_{ki,j} = 0. \end{aligned}$$

Empregando as relações entre as formas contravariante e mista do tensor de métrica, é fácil demonstrar que este teorema é válido também para estas formas. \square

Em particular, o teorema de Ricci implica na seguinte expressão para a derivada usual do tensor de métrica,

$$\frac{\partial g_{ij}}{\partial q^k} \equiv g_{ij,k} = [ik, j] + [jk, i], \quad (6.51a)$$

a qual, com o uso de $g^{ik} g_{kj} = \delta^i_j$, resulta em

$$\frac{\partial g^{ij}}{\partial q^k} \equiv g^{ij}_{,k} = -g^{i\ell} \Gamma^j_{\ell k} - g^{j\ell} \Gamma^i_{\ell k}. \quad (6.51b)$$

6.12 OPERADORES VETORIAIS NA FORMA TENSORIAL

Na seção 1.4 foram derivadas as expressões para os operadores diferenciais gradiente, rotacional, laplaciano e rotacional, válidas para sistemas de coordenadas ortogonais no espaço Euclidiano \mathbb{E}^3 . Em seguida, na seção 1.5 foram derivadas expressões explícitas destes operadores para alguns exemplos de sistemas de coordenadas ortogonais. Nesta seção, os mesmos operadores diferenciais serão novamente obtidos, porém agora para *qualquer* sistema de coordenadas, inclusive para aqueles que não são ortogonais.

Para fins de posterior comparação entre as expressões que serão derivadas nesta seção com as apresentadas na seção 1.5, deve-se lembrar que as bases aqui adotadas, $\{e_i\}$ e $\{e^i\}$ não são necessariamente normalizadas. Contudo, é sempre possível normalizar a primeira base, lembrando da definição (1.11) para os fatores de escala h_i . Desta forma, pode-se adotar a base $\{\hat{e}_i\} = \{e_i/h_i\}$ para fins de comparação. Nesta base, o vetor v , por exemplo, pode ser escrito $v = v^i e_i = \hat{v}^i \hat{e}_i$, sendo $\hat{v}^i \stackrel{\text{NS}}{=} h_i v^i$, o que possibilita uma comparação direta entre os resultados. Adicionalmente, deve-se lembrar também que para um sistema ortogonal as representações do tensor de métrica são $g_{ij} = h_i^2 \delta_{ij}$ e $g^{ij} = h_i^{-2} \delta^{ij}$, enquanto que o seu determinante é $g = \det([g_{ij}]) = h_1^2 h_2^2 h_3^2$.

6.12.1 GRADIENTE DE CAMPO ESCALAR

Como o campo escalar $\psi(q^1, \dots, q^n)$ não depende dos vetores de base, o seu gradiente é obtido simplesmente pela aplicação do operador vetorial ∇ covariante,

$$\nabla \psi = \frac{\partial \psi}{\partial q^i} e^i = \sum_{i=1}^3 \frac{1}{h_i} \frac{\partial \psi}{\partial q^i} \hat{e}_i. \quad (6.52)$$

Pode-se ver que neste caso simples a expressão resultante coincide com (1.23).

6.12.2 DIVERGENTE DE CAMPO VETORIAL

Dado o campo vetorial $v = v(q^1, \dots, q^n)$, o seu divergente consiste na aplicação sobre o mesmo de um operador diferencial de tal forma que resulte um campo escalar. O divergente de um campo vetorial generalizado é representado pela operação $\text{div} v$.

Para resultar um tensor de posto zero, esta operação pode ser concebida como composta por duas etapas: inicialmente, o operador ∇ é multiplicado diretamente ao vetor v , resultando em um tensor de posto dois. Em seguida, é realizada uma contração entre os índices do tensor, tendo como resultado um escalar. A primeira etapa destas operações pode ser representada por

$$v = v^i e_i \implies \nabla v = v^i_{;j} e_i \otimes e^j,$$

onde foi introduzida a derivada covariante (6.48). Portanto, o divergente de v é dado pela contração

$$\text{div} v \doteq v^i_{;i} = \frac{\partial v^i}{\partial q^i} + \Gamma^i_{ki} v^k.$$

Usando agora a expressão (6.46) para os símbolos de Christoffel, observa-se que

$$\Gamma^i_{ki} = \frac{1}{2} g^{i\ell} \left(\frac{\partial g_{i\ell}}{\partial q^k} + \frac{\partial g_{k\ell}}{\partial q^i} - \frac{\partial g_{ki}}{\partial q^\ell} \right) = \frac{1}{2} g^{i\ell} \frac{\partial g_{i\ell}}{\partial q^k}, \quad (6.53)$$

uma vez que os últimos dois termos se cancelam.

Por outro lado, o termo restante em (6.53) pode ser simplificado fazendo uso do seguinte lema.

Lema. Sendo as matrizes $a(q^1, \dots, q^n) = [a_{ij}]$ e $b(q^1, \dots, q^n) = [b^{ij}]$ tais que $b = a^{-1}$, e sendo $a = \det(a)$, então

$$\frac{\partial a}{\partial q^k} = a b^{ji} \frac{\partial a_{ij}}{\partial q^k}. \quad (6.54)$$

Demonstração. Para se demonstrar este lema no R^3 , emprega-se novamente a identidade matricial (6.33a), a qual pode ser escrita como

$$a = \frac{1}{6} \epsilon^{ijk} \epsilon^{abc} a_{ai} a_{bj} a_{ck}.$$

Por outro lado, os elementos da matriz b podem ser expressos em termos dos elementos de a como

$$b^{ij} = \frac{1}{2a} \epsilon^{ikl} \epsilon^{jmn} a_{mk} a_{nl}.$$

Então, derivando a em relação a q^ℓ ,

$$\begin{aligned} \frac{\partial a}{\partial q^\ell} &= \frac{1}{6} \epsilon^{ijk} \epsilon^{abc} \left(a_{bj} a_{ck} \frac{\partial a_{ai}}{\partial q^\ell} + a_{ai} a_{ck} \frac{\partial a_{bj}}{\partial q^\ell} + a_{ai} a_{bj} \frac{\partial a_{ck}}{\partial q^\ell} \right) \\ &= \frac{1}{3} a \left(b^{ia} \frac{\partial a_{ai}}{\partial q^\ell} + b^{jb} \frac{\partial a_{bj}}{\partial q^\ell} + b^{kc} \frac{\partial a_{ck}}{\partial q^\ell} \right) = ab^{ji} \frac{\partial a_{ij}}{\partial q^\ell}, \end{aligned}$$

resultando assim (6.54).

Embora o lema tenha sido demonstrado para uma matriz 3×3 , pode-se mostrar que este resultado independe da ordem da matriz. \square

Aplicando então a identidade (6.54) para o tensor de métrica e lembrando que o mesmo é simétrico, obtém-se

$$\frac{\partial g}{\partial q^k} = g g^{ij} \frac{\partial g_{ij}}{\partial q^k}. \quad (6.55)$$

Usando então este resultado em (6.53), conclui-se que

$$\Gamma^i_{ki} = \frac{1}{2g} \frac{\partial g}{\partial q^k} = \frac{1}{\sqrt{|g|}} \frac{\partial \sqrt{|g|}}{\partial q^k} = \frac{\partial}{\partial q^k} \ln(\sqrt{|g|}).$$

Portanto, o divergente de um campo vetorial pode ser sempre expresso em termos do determinante do tensor de métrica como

$$\operatorname{div} v = v^i_{;i} = \frac{\partial v^i}{\partial q^i} + \frac{1}{\sqrt{|g|}} \frac{\partial \sqrt{|g|}}{\partial q^i} v^i = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial q^i} (\sqrt{|g|} v^i) \quad (6.56a)$$

$$= \frac{1}{\sqrt{|g|}} \sum_{i=1}^3 \frac{\partial}{\partial q^i} \left(\sqrt{|g|} \hat{h}_i^i \right). \quad (6.56b)$$

6.12.3 LAPLACIANO DE UM CAMPO ESCALAR

Dado o campo escalar $\psi = \psi(q^1, \dots, q^n)$, o seu laplaciano pode ser obtido a partir da operação usual, $\nabla^2 \psi = \operatorname{div}(\nabla \psi)$. Inicialmente, percebe-se em (6.52) que $(\nabla \psi)_i = \partial \psi / \partial q^i$. Então, com o emprego do tensor de métrica, pode-se escrever

$$v^i = (\nabla \psi)^i = g^{ij} \frac{\partial \psi}{\partial q^j}.$$

Introduzindo este resultado em (6.56), obtém-se

$$\nabla^2 \psi = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial q^i} \left(\sqrt{|g|} g^{ij} \frac{\partial \psi}{\partial q^j} \right). \quad (6.57)$$

Exercício 6.15. Mostre que (6.57) se reduz a (1.25) no caso particular de um sistema de coordenadas ortogonal.

Demonstração. Para um sistema ortogonal, $g^{ij} = h_i^{-2} \delta^{ij}$ e $\sqrt{|g|} = h_1 h_2 h_3$. Então, em (6.57),

$$\nabla^2 \psi = \frac{1}{h_1 h_2 h_3} \sum_{i=1}^3 \frac{\partial}{\partial q^i} \left(\frac{h_1 h_2 h_3}{h_i^2} \frac{\partial \psi}{\partial q^i} \right),$$

o qual é exatamente a expressão (1.25). \square

6.12.4 ROTACIONAL DE UM CAMPO VETORIAL

A definição usual do rotacional de um campo vetorial está associada ao conceito do produto vetorial entre dois vetores. Contudo, é interessante obter-se neste momento uma expressão mais geral, a qual pode ser facilmente estendida a espaços vetoriais de dimensões mais altas, uma vez que o produto vetorial somente existe no \mathbb{E}^3 .

Considerando inicialmente o produto vetorial entre os vetores $\mathbf{a}, \mathbf{b} \in \mathbb{E}^3$, pode-se escrever

$$\mathbf{a} \times \mathbf{b} = a_i b_j \mathbf{e}^i \times \mathbf{e}^j = \frac{1}{\sqrt{|g|}} \epsilon^{ijk} a_i b_j \mathbf{e}_k,$$

onde foi empregado (6.35b). Definindo então o tensor antissimétrico

$$M_{ij} = \epsilon_{ijk} \epsilon^{klm} a_\ell b_m = a_i b_j - a_j b_i \iff \epsilon^{ijk} a_i b_j = \frac{1}{2} \epsilon^{ijk} M_{ij},$$

observa-se que é possível escrever

$$\mathbf{a} \times \mathbf{b} = \frac{1}{2\sqrt{|g|}} \epsilon^{ijk} M_{ij} \mathbf{e}_k.$$

Ou seja, é sempre possível escrever os componentes do produto vetorial em termos do tensor antissimétrico M_{ij} . É vantajoso definir este tensor, uma vez que se pode então considerar a sua extensão para espaços vetoriais de dimensão mais alta.

Focando agora no rotacional de um campo $\mathbf{v}(q^1, \dots, q^n)$, identificado por $\text{rot} \mathbf{v}$, este será definido pela generalização do produto vetorial do operador ∇ com \mathbf{v} . Escreve-se então

$$\text{rot} \mathbf{v} = -\frac{1}{\sqrt{|g|}} \epsilon^{ijk} v_{ij} \mathbf{e}_k,$$

onde foram empregados (6.49b) e (6.35b). Define-se então o tensor antissimétrico

$$(\text{rot} \mathbf{v})_{ij} = \epsilon_{ijk} \epsilon^{klm} v_{\ell,m} = v_{i,j} - v_{j,i} \iff \epsilon^{ijk} v_{i,j} = \frac{1}{2} \epsilon^{ijk} (\text{rot} \mathbf{v})_{ij},$$

o qual está relacionado ao i -ésimo componente do rotacional de \mathbf{v} . Contudo, de (6.49), observa-se que $(\text{rot} \mathbf{v})_{ij} = \epsilon_{ijk} \epsilon^{klm} (v_{\ell,m} - \Gamma_{\ell m}^n v_n) = \epsilon_{ijk} \epsilon^{klm} v_{\ell,m} = v_{i,j} - v_{j,i}$, devido à simetria de Γ_{ij}^k . Portanto,

$$(\text{rot} \mathbf{v})_{ij} = \epsilon_{ijm} \epsilon^{klm} v_{k,\ell} = v_{i,j} - v_{j,i}.$$

Então, para reproduzir o resultado (1.26) conhecido mas, ao mesmo tempo, generalizando para sistemas não ortogonais e, eventualmente, para espaços de dimensões superiores, o campo vetorial $\mathbf{w} = \text{rot} \mathbf{v}$, desenvolvido como $\mathbf{w} = w^i \mathbf{e}_i$, tem seus componentes dados por

$$w^i = \frac{1}{\sqrt{|g|}} \epsilon^{ijk} \frac{\partial v_k}{\partial q^j}. \quad (6.58)$$

6.13 DIFERENCIAÇÃO ABSOLUTA E CURVAS GEODÉSICAS

Nesta seção serão introduzidos alguns conceitos mais avançados da análise tensorial, com frequência empregados em trabalhos envolvendo geometria diferencial e relatividade geral.

6.13.1 DIFERENCIAÇÃO ABSOLUTA OU INTRÍNSECA

Uma aplicação importante das derivadas covariantes consiste na derivação de um campo tensorial ao longo de uma curva no R^n , a qual é parametrizada por $\mathbf{r} = \mathbf{r}(t)$, ou seja, em termos de um parâmetro t .

Considera-se inicialmente o campo vetorial $\mathbf{v} = \mathbf{v}(q^1, \dots, q^n)$. Escrevendo $\mathbf{v} = v^i \mathbf{e}_i$, a derivada absoluta deste campo em relação a t fica

$$\frac{d\mathbf{v}}{dt} = \frac{dv^i}{dt} \mathbf{e}_i + v^i \frac{d\mathbf{e}_i}{dt}$$

$$= \frac{dv^i}{dt} e_i + v^i \frac{dq^j}{dt} \frac{\partial e_i}{\partial q^j}.$$

Introduzindo o símbolo de Christoffel (6.44), resulta

$$\frac{dv}{dt} = \left(\frac{dv^i}{dt} + \Gamma^i_{jk} v^j \frac{dq^k}{dt} \right) e_i \doteq \frac{\delta v^i}{\delta t} e_i.$$

O termo entre parênteses é denominada a **derivada absoluta** (ou **intrínseca**) $\delta v^i / \delta t$ do componente v^i ao longo da curva $r(t)$. Esta derivada é usualmente representada, de acordo com (6.48), como

$$\frac{\delta v^i}{\delta t} \equiv \frac{dv^i}{dt} + \Gamma^i_{jk} v^j \frac{dq^k}{dt} = v^i_{;j} \frac{dq^j}{dt}. \quad (6.59)$$

Ou seja, pode-se escrever

$$\frac{dv}{dt} = v^i_{;j} \frac{dq^j}{dt} e_i.$$

De forma semelhante, as seguintes expressões para as derivadas absolutas das diferentes formas de um tensor de posto dois podem ser deduzidas,

$$\begin{aligned} \frac{\delta T^{ij}}{\delta t} &= T^{ij}_{;k} \frac{dq^k}{dt} \\ \frac{\delta T_{ij}}{\delta t} &= T_{ij;k} \frac{dq^k}{dt} \\ \frac{\delta T^i_j}{\delta t} &= T^i_{j;k} \frac{dq^k}{dt}. \end{aligned}$$

As expressões acima atribuem significado à derivação absoluta deste tensor como um todo,

$$\frac{dT}{dt} = \frac{\delta T^{ij}}{\delta t} e_i \otimes e_j = \frac{\delta T_{ij}}{\delta t} e^i \otimes e^j = \dots$$

Finalmente, se $\psi(q^1, \dots, q^n)$ é um campo escalar, então sua derivada intrínseca é simplesmente a derivada ordinária,

$$\frac{\delta \psi}{\delta t} = \frac{d\psi}{dt}. \quad (6.60)$$

As derivadas absoluta e covariante obedecem as seguintes propriedades da diferenciação:

1. A derivada de uma soma de tensores é a soma das derivadas.
2. A derivada do produto (externo ou interno) dos tensores \mathcal{T} e \mathcal{U} é igual a $\mathcal{U}\delta\mathcal{T} + \mathcal{T}\delta\mathcal{U}$, onde o símbolo “ δ ” representa qualquer tipo de diferenciação.
3. O fato da derivada covariante do tensor de métrica ser nula implica em que as operações de elevação ou rebaixamento de índices e de diferenciação podem ser permutadas.

6.13.2 CURVAS GEODÉSICAS

Um exemplo importante de aplicação da derivada absoluta de um campo em uma determinada geometria do espaço são as *curvas geodésicas*. De maneira simples, uma geodésica é aquela curva no espaço cujos vetores tangentes permanecem paralelos ou invariantes ao longo da mesma. Em um espaço Riemanniano, a geodésica também é a curva de menor distância entre dois pontos no espaço.

Em um espaço Euclideano, a geodésica entre dois pontos é, simplesmente, a linha reta entre os mesmos. Porém em um espaço curvo, a geodésica é realmente uma curva. Uma maneira alternativa de se definir uma geodésica está relacionada com topologia e geometria diferencial. Nesta abordagem, a geodésica é a curva de menor distância entre dois pontos *sobre uma superfície curva* no R^n . Esta interpretação é particularmente importante para a teoria da relatividade geral, uma vez que a geometria do espaço-tempo, ou seja, a sua métrica, é determinada pela presença de objetos maciços. Quando uma partícula-teste está nas vizinhanças destes objetos, a sua trajetória entre dois pontos quaisquer neste espaço-tempo segue ao longo de uma geodésica.

É conveniente neste ponto introduzir a definição formal de uma curva no \mathbb{E}^3 . A definição abaixo é ilustrada pela figura 6.9.

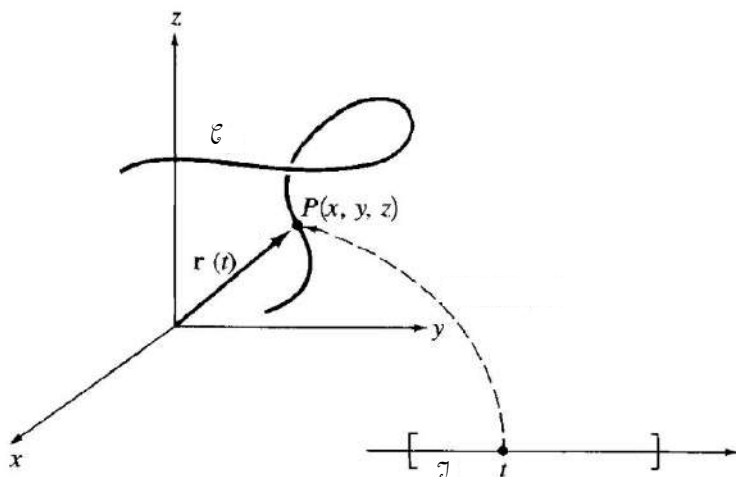


Figura 6.9: Uma curva \mathcal{C} no espaço \mathbb{E}^3 , definida como o mapeamento de um intervalo de variação do parâmetro $t \in \mathcal{J}$ ($\mathcal{J} \subset \mathbb{R}$) sobre uma coleção de pontos $P \subset \mathbb{E}^3$, determinados pela variação do vetor posição $\mathbf{r}(t)$.

Definição 6.10 (Curva no espaço \mathbb{E}^3). Seja $\mathcal{J} \subset \mathbb{R}$ um conjunto de valores do parâmetro $t \in \mathcal{J}$ e o espaço Euclidiano \mathbb{E}^3 . Uma curva \mathcal{C} no espaço \mathbb{E}^3 é o mapeamento contínuo

$$\mathcal{J} \mapsto \mathbb{E}^3$$

das funções $x(t), y(t), z(t) \in \mathcal{C}^2(\mathcal{J})$ sobre \mathbb{E}^3 tal que para cada $t \in \mathcal{J}$ existe um e somente um vetor posição $\mathbf{r} = \mathbf{r}(t) \in \mathbb{E}^3$, determinado por

$$\mathbf{r}(t) = (x(t), y(t), z(t)).$$

Para seguir a discussão, já é possível e conveniente abandonar a restrição do \mathbb{E}^3 e considerar um espaço Riemanniano qualquer de dimensão n , denotado por R^n . Uma curva no R^n é definida, de forma análoga à definição 6.10, como o conjunto de pontos $\mathcal{C} \subset R^n$ determinados pelo mapeamento da varredura do parâmetro $t \in \mathbb{R}$ em imagens das funções $q^1(t), q^2(t), \dots, q^n(t) \in \mathcal{C}^2(\mathbb{R})$, as quais irão formar as n -uplas ordenadas $(q^1(t), q^2(t), \dots, q^n(t)) \in \mathcal{C}$.

Para se determinar então a equação de uma geodésica no R^n , emprega-se com frequência o cálculo variacional. Neste caso, a geodésica é definida como a curva cujo comprimento possui um valor estacionário com respeito a variações arbitrariamente pequenas, mas com pontos extremos mantidos fixos.

Seja então $d\ell$ o comprimento elementar de arco correspondente ao deslocamento infinitesimal $d\ell$ e t_i e t_f os valores extremos do parâmetro t que determina a curva \mathcal{C} entre os pontos A e B , contidos no R^n . Então, as coordenadas dos pontos ao longo de \mathcal{C} serão dadas pelas fórmulas

$$q^i = q^i(t), \quad t_i \leq t \leq t_f, \quad q^i(t) \in \mathcal{C}^3(\mathbb{R}), \quad i = 1, 2, \dots, n,$$

sendo que $\{q^i(t_i)\} \mapsto A$ e $\{q^i(t_f)\} \mapsto B$. Portanto, a extensão da curva \mathcal{C} , denotada por ℓ , será dada por

$$\ell = \int_{A, \mathcal{C}}^B d\ell.$$

De acordo com o princípio variacional acima, a curva geodésica entre A e B será determinada então por

$$\delta \ell = \delta \int_{A, \mathcal{C}}^B d\ell = 0.$$

Mas como o deslocamento $d\ell$ ocorre ao longo da curva \mathcal{C} , o princípio variacional acima pode ser escrito também em termos de uma integração no parâmetro t . Em um espaço Riemanniano, o elemento de arco $d\ell$ é dado por $d\ell^2 = g_{ij} dq^i dq^j$. Então, ao longo de uma variação infinitesimal ao longo de \mathcal{C} , entre os valores t e $t + dt$ do parâmetro livre, a coordenada $q^i(t)$ varia por

$$dq^i = \frac{dq^i}{dt} dt,$$

ou seja, $d\ell$ varia por

$$d\ell^2 = g_{ij} \frac{dq^i}{dt} \frac{dq^j}{dt} dt^2 \implies d\ell = \sqrt{g_{ij} \frac{dq^i}{dt} \frac{dq^j}{dt}} dt.$$

Portanto, o princípio variacional pode ser expresso como

$$\delta \ell = \delta \int_{t_i}^{t_f} \sqrt{g_{ij} \frac{dq^i}{dt} \frac{dq^j}{dt}} dt = 0.$$

O integrando acima é um funcional das quantidades $\{q^i(t)\}$ (através do tensor de métrica) e $\{dq^i/dt\}$, as quais serão determinadas a partir do princípio variacional. Denota-se este funcional por

$$\mathcal{L} \left[\{q^i\}, \left\{ \frac{dq^i}{dt} \right\} \right] = \sqrt{g_{ij} \frac{dq^i}{dt} \frac{dq^j}{dt}}.$$

Portanto, aplicando a variação $\delta \ell$, mantendo os extremos fixos, resulta

$$\delta \ell = \delta \int_{t_i}^{t_f} \mathcal{L} dt = \int_{t_i}^{t_f} \delta \mathcal{L} dt = \int_{t_i}^{t_f} \left[\frac{\partial \mathcal{L}}{\partial q^i} \delta q^i + \frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \delta \left(\frac{dq^i}{dt} \right) \right] dt.$$

Como é usual, a variação δq^i corresponde à diferença entre duas formas funcionais para $q^i(t)$ que são arbitrariamente próximas entre si. Denotando estas formas funcionais como $q_{(1)}^i(t)$ e $q_{(2)}^i(t)$, escreve-se $\delta q^i = q_{(2)}^i - q_{(1)}^i$ e, portanto,

$$\frac{d}{dt} (\delta q^i) = \frac{d}{dt} [q_{(2)}^i - q_{(1)}^i] = \frac{d}{dt} q_{(2)}^i - \frac{d}{dt} q_{(1)}^i = \delta \left(\frac{dq^i}{dt} \right).$$

Então, integrando por partes o segundo termo, resulta

$$\begin{aligned} \int_{t_i}^{t_f} \frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \delta \left(\frac{dq^i}{dt} \right) dt &= \int_{t_i}^{t_f} \frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \frac{d}{dt} (\delta q^i) dt \\ &= \frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \delta q^i \Big|_{t_i}^{t_f} - \int_{t_i}^{t_f} \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \right] \delta q^i dt = - \int_{t_i}^{t_f} \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \right] \delta q^i dt, \end{aligned}$$

uma vez que os extremos são fixos. Portanto, o princípio variacional sustenta que

$$\delta \ell = \int_{t_i}^{t_f} \left\{ \frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \right] \right\} \delta q^i dt = 0.$$

Para uma variação arbitrária das coordenadas q^i , a identidade acima somente pode ser satisfeita se forem obedecidas as *equações de Euler-Lagrange*

$$\frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dt} \right)} \right] = 0, \quad (i = 1, 2, \dots, n).$$

Até este ponto, o parâmetro t foi considerado arbitrário. Porém, como se deseja relacionar a geodésica como a curva de menor extensão entre A e B , o parâmetro t será escolhido como o comprimento de arco ℓ ao longo da mesma, ou seja,

$$t = \ell, \quad \frac{dq^i}{dt} = \frac{dq^i}{d\ell} \implies \mathcal{L} = \frac{d\ell}{d\ell} = 1.$$

Calculando então as derivadas acima,

$$\frac{\partial \mathcal{L}}{\partial q^i} = \frac{\partial}{\partial q^i} \sqrt{g_{jk} \frac{dq^j}{d\ell} \frac{dq^k}{d\ell}} = \frac{1}{2} \mathcal{L}^{-1} \frac{\partial g_{jk}}{\partial q^i} \frac{dq^j}{d\ell} \frac{dq^k}{d\ell} = \frac{1}{2} \frac{\partial g_{jk}}{\partial q^i} \frac{dq^j}{d\ell} \frac{dq^k}{d\ell}.$$

Já a outra derivada fica

$$\frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{d\ell} \right)} = \frac{\partial}{\partial \left(\frac{dq^i}{d\ell} \right)} \sqrt{g_{jk} \frac{dq^j}{d\ell} \frac{dq^k}{d\ell}} = \frac{1}{2} g_{jk} \left(\delta_i^j \frac{dq^k}{d\ell} + \frac{dq^j}{d\ell} \delta_i^k \right) = g_{ij} \frac{dq^j}{d\ell}.$$

Para este termo é ainda necessário calcular a derivada absoluta

$$\frac{d}{dl} \left[\frac{\partial \mathcal{L}}{\partial \left(\frac{dq^i}{dl} \right)} \right] = \frac{d}{dl} \left(g_{ij} \frac{dq^j}{dl} \right) = \frac{dg_{ij}}{dl} \frac{dq^j}{dl} + g_{ij} \frac{d^2 q^j}{dl^2} = \frac{\partial g_{ij}}{\partial q^k} \frac{dq^k}{dl} \frac{dq^j}{dl} + g_{ij} \frac{d^2 q^j}{dl^2}.$$

Portanto, as equações de Euler-Lagrange tornam-se

$$g_{ij} \frac{d^2 q^j}{dl^2} + \left(\frac{\partial g_{ij}}{\partial q^k} - \frac{1}{2} \frac{\partial g_{jk}}{\partial q^i} \right) \frac{dq^j}{dl} \frac{dq^k}{dl} = 0.$$

Nota-se que esta equação pode ainda ser escrita como

$$g_{ij} \frac{d^2 q^j}{dl^2} + \frac{1}{2} \left(\frac{\partial g_{ij}}{\partial q^k} + \frac{\partial g_{ik}}{\partial q^j} - \frac{\partial g_{jk}}{\partial q^i} \right) \frac{dq^j}{dl} \frac{dq^k}{dl} = 0,$$

o que, de acordo com (6.47d), resulta em

$$g_{ij} \frac{d^2 q^j}{dl^2} + [jk, i] \frac{dq^j}{dl} \frac{dq^k}{dl} = 0, \quad (i = 1, 2, \dots, n), \quad (6.61a)$$

as quais são as equações da curva geodésica. Finalmente, usando (6.47c), pode-se escrever estas equações também como

$$\frac{d^2 q^i}{dl^2} + \Gamma^i_{jk} \frac{dq^j}{dl} \frac{dq^k}{dl} = 0 \quad (i = 1, 2, \dots, n). \quad (6.61b)$$

Exercício 6.16. Encontre as equações da curva geodésica em coordenadas cilíndricas.

Solução. Usando os resultados para os símbolos de Christoffel obtidos no exercício 6.12, as equações (6.61b) são escritas para $n = 3$ e $\{q^i\} \mapsto \{\rho, \phi, z\}$. Os pontos extremos entre $\ell_i = 0$ e ℓ_f são dados respectivamente pelas coordenadas $\{\rho, \phi, z\}_i$ e $\{\rho, \phi, z\}_f$. Então,

$$\begin{aligned} \frac{d^2 \rho}{dl^2} + \Gamma^1_{22} \left(\frac{d\phi}{dl} \right)^2 = 0 & \implies \frac{d^2 \rho}{dl^2} - \rho \left(\frac{d\phi}{dl} \right)^2 = 0 \\ \frac{d^2 \phi}{dl^2} + 2\Gamma^2_{12} \frac{d\rho}{dl} \frac{d\phi}{dl} = 0 & \implies \frac{d^2 \phi}{dl^2} + \frac{2}{\rho} \frac{d\rho}{dl} \frac{d\phi}{dl} = 0 \\ \frac{d^2 z}{dl^2} = 0 & \implies z(\ell) = z_i + (z_f - z_i) \frac{\ell}{\ell_f}. \end{aligned}$$

A equação para $z(\ell)$ é trivial, porém, as equações diferenciais para $\rho(t)$ e $\phi(t)$ são acopladas e não lineares, portanto muito difíceis de serem resolvidas analiticamente para dois pontos arbitrários no \mathbb{E}^3 .

Uma solução simples do sistema acima ocorre quando $\phi = \text{cte.}$, ou seja, entre dois pontos no mesmo azimute. Neste caso, o sistema se reduz a

$$\frac{d^2 \rho}{dl^2} = 0 \implies \rho(\ell) = \rho_i + (\rho_f - \rho_i) \frac{\ell}{\ell_f},$$

ou seja,

$$\frac{z - z_i}{z_f - z_i} = \frac{\rho - \rho_i}{\rho_f - \rho_i} \implies z = z_i + \frac{z_f - z_i}{\rho_f - \rho_i} (\rho - \rho_i),$$

a qual é uma linha reta no plano $\phi = \text{cte.}$ conectando os pontos extremos.

As soluções das equações (6.61) sempre fornecem a curva de menor extensão entre dois pontos para qualquer sistema de coordenadas no R^n . Contudo, em certas situações, é interessante encontrar a geodésica quando a curva \mathcal{C} possui vínculos; por exemplo, a curva de menor extensão entre os pontos A e B que se encontram sobre a superfície $F(\{q^i\}) = 0$, tal que todos os pontos ao longo de \mathcal{C} também estão sobre essa superfície. Em um espaço R^n , a equação de superfície define uma *variedade*¹⁶ em R^{n-1} . Por exemplo, a geodésica sobre uma esfera de raio a .

Neste caso, a norma da variedade R^{n-1} irá determinar um tensor de métrica representado por uma matriz $(n-1) \times (n-1)$ e as equações (6.61) ainda poderão ser aplicadas, com uma definição conveniente das coordenadas q^i sobre essa variedade.

¹⁶**Variedade** (*manifold*) é uma generalização do conceito de superfície em geometria diferencial e topologia. Uma variedade de dimensão n é um espaço topológico que nas vizinhanças de cada ponto assemelha-se a um espaço \mathbb{E}^n . Curvas e circunferências são variedades no R^1 ; esferas e cilindros são variedades no R^2 .

Exercício 6.17. Encontre a curva geodésica conectando dois pontos sobre a superfície de um cilindro $\rho = \text{cte.}$ e que está vinculada a esta superfície.

Solução. Neste caso, as equações obtidas no exercício 6.16 não podem ser aplicadas, pois mesmo que os pontos extremos $\{\rho, \phi, z\}_i$ e $\{\rho, \phi, z\}_f$ encontrem-se sobre a superfície, a curva de menor extensão não está necessariamente restrita à mesma. Isto pode ser verificado colocando $d\rho/d\ell = 0$ nas equações. Observa-se que a única solução compatível neste caso é $\rho = \text{cte.}$ e $\phi = \text{cte.}$, o que corresponde a uma reta vertical sobre a superfície do cilindro. Esta é a única solução possível para as equações obtidas.

Sobre a superfície $\rho = \text{cte.}$ pode ser realizada a parametrização $\phi = \phi(\ell)$, sendo que então o elemento de arco é dado por $d\ell^2 = ds^2 + dz^2$, sendo $ds = \rho d\phi$. Neste caso, $n = 2$, o tensor de métrica é simplesmente $g_{ij} = \delta_{ij}$ e os símbolos de Christoffel são nulos. Assim, as equações geodésicas ficam dadas simplesmente por

$$\begin{aligned} \frac{d^2 s}{d\ell^2} = 0 & \implies s(\ell) = \rho\phi(\ell) \implies \phi(\ell) = \phi_i + (\phi_f - \phi_i) \frac{\ell}{\ell_f} \\ \frac{d^2 z}{d\ell^2} = 0 & \implies z(\ell) = z_i + (z_f - z_i) \frac{\ell}{\ell_f}. \end{aligned}$$

Ou seja,

$$\frac{z - z_i}{z_f - z_i} = \frac{\phi - \phi_i}{\phi_f - \phi_i} \implies z = z_i + (z_f - z_i) \frac{\phi - \phi_i}{\phi_f - \phi_i},$$

a qual descreve uma hélice sobre a superfície cilíndrica.

Exercício 6.18. Usando o resultado do exercício 6.17, calcule a extensão da geodésica que une os pontos Cartesianos $P = (x, y, z) = (1, 0, 0)$ e $Q = (0, 1, 1)$ sobre a superfície cilíndrica de raio $\rho = 1$.

Solução. De acordo com o exemplo, a curva geodésica que une os pontos P e Q é

$$z = \frac{2}{\pi}\phi \implies dz = \frac{2}{\pi}d\phi.$$

Então, o elemento de arco é

$$d\ell^2 = d\phi^2 + \frac{4}{\pi^2}d\phi^2$$

e a extensão da geodésica é dada por

$$\ell_{P \rightarrow Q} = \int_P^Q d\ell = \sqrt{1 + \frac{4}{\pi^2}} \int_0^{\pi/2} d\phi = \sqrt{1 + \frac{\pi^2}{4}} = 1,86209588912 \dots$$

Para comparar, a linha reta unindo P a Q possui uma extensão igual a $\sqrt{3} = 1,73205080757 \dots < \ell_{P \rightarrow Q}$.

Exercício 6.19. Mostre que sobre a superfície de uma esfera todas as curvas meridianas são geodésicas. Mostre também que nenhuma outra circunferência é uma geodésica.

Solução. Dada uma esfera de raio a , esta se trata de uma variedade no R^2 com $\{q^1, q^2\} \mapsto \{\theta, \phi\}$. O elemento de arco sobre esta superfície é $d\ell^2 = a^2 (d\theta^2 + \sin^2 \theta d\phi^2)$. Então, o tensor de métrica e os símbolos de Christoffel (6.46b) são dados por

$$g = \begin{pmatrix} a^2 & 0 \\ 0 & a^2 \sin^2 \theta \end{pmatrix} \implies \Gamma^\theta = \begin{pmatrix} 0 & 0 \\ 0 & -\sin \theta \cos \theta \end{pmatrix}, \Gamma^\phi = \begin{pmatrix} 0 & \cotan \theta \\ \cotan \theta & 0 \end{pmatrix},$$

e as equações da geodésica (6.61b) se tornam

$$\begin{aligned} \frac{d^2 \theta}{d\ell^2} - \sin \theta \cos \theta \left(\frac{d\phi}{d\ell} \right)^2 &= 0 \\ \frac{d^2 \phi}{d\ell^2} + 2 \cotan \theta \frac{d\theta}{d\ell} \frac{d\phi}{d\ell} &= 0. \end{aligned}$$

Embora estas equações sejam não lineares e acopladas, pode-se usar a simetria de uma esfera para se realizar a demonstração. Assumindo que $\phi = \text{cte.}$, as equações acima reduzem-se a

$$\frac{d^2\theta}{d\ell^2} = 0 \implies \theta(\ell) = \theta_i + (\theta_f - \theta_i) \frac{\ell}{\ell_f},$$

ou seja, qualquer meridiano da esfera é uma curva geodésica.

Tomando agora $\theta_i \neq \pi/2$ e assumindo circunferências $\theta = \text{cte.}$, as equações reduzem-se a

$$\frac{d\phi}{d\ell} = 0 \implies \phi = \text{cte.},$$

o que contradiz a suposição inicial. Por outro lado, se $\theta = \pi/2 = \text{cte.}$, as equações são consistentes com $d\phi/d\ell = \text{cte.} \neq 0$. Portanto, exceto as meridianas e a linha do equador, nenhuma outra circunferência sobre a esfera é uma geodésica.

6.13.3 TRANSPORTE PARALELO DE CAMPOS VETORIAIS

Em um espaço Euclideano, um campo vetorial $v(r)$ é denominado um **campo de vetores paralelos** se os seus componentes $\{v^i\}$ são constantes, ou seja, se $\partial v^i/\partial q^j = 0$. De forma equivalente, supondo que o vetor v possui os seus componentes $\{v^i\}$ definidos em um ponto P do \mathbb{E}^n , localizado pela n -upla (q^1, q^2, \dots, q^n) , se o ponto P é deslocado por dr , localizando então no ponto $r + dr$, o campo vetorial nesta nova posição é dado por $v + dv$. Diz-se então que o campo vetorial v sofre um **transporte paralelo** se $\partial v^i/\partial q^j = 0$.

Um exemplo deste tipo de campo vetorial é fornecido pelo campo elétrico no interior de um capacitor de placas paralelas. Nesta região o campo E é uniforme, isto é, em todos os pontos este possui os mesmos módulo, direção e sentido. Outro exemplo seria o campo de aceleração gravitacional nas proximidades da superfície terrestre.

Deseja-se agora generalizar este conceito para qualquer campo vetorial em um espaço Riemanniano. Será verificado também que a definição realizada abaixo para um campo de vetores paralelos em um espaço curvo qualquer também pode ser empregada para definir as curvas geodésicas neste espaço.

Nas seções 6.11 e 6.13.1 mostrou-se que, dados os componentes $\{v^i(q^1, \dots, q^n)\}$, sua derivada coordenada usual $\partial v^i/\partial q^j$ não é um tensor, mas sim a sua derivada covariante ou a sua derivada absoluta ao longo de uma curva $q^i = q^i(t)$. Portanto, para manter a característica tensorial, o transporte paralelo de um campo vetorial em um espaço Riemanniano deve ser definido de forma distinta do que é feito no espaço Euclideano.

Considera-se então um espaço Riemanniano R^n , no qual uma curva \mathcal{C} é parametrizada por

$$\mathcal{C} \mapsto q^i = q^i(t), \quad (t_i \leq t \leq t_f, i = 1, 2, \dots, n),$$

e um campo vetorial $v = v(q^1, \dots, q^n)$ em R^n . Diz-se que v é um **campo de vetores paralelos a \mathcal{C}** , ou, de forma equivalente, que v é **transportado paralelamente ao longo de \mathcal{C}** se a derivada intrínseca de v ao longo desta curva é nula, isto é, se

$$\frac{dv}{dt} = 0. \tag{6.62}$$

O comportamento de um campo vetorial transportado paralelo a uma curva \mathcal{C} no \mathbb{E}^3 está representado na figura 6.10. No ponto P , pode-se observar também uma representação para as curvas coordenadas $\{q^1, q^2, q^3\}$ neste ponto.

De forma equivalente, pode-se definir através de (6.62) um campo vetorial transportado paralelamente ao longo de uma curva \mathcal{C} que se estende sobre uma variedade no R^{n-1} . Para exemplificar, com esta última definição é possível discutir campos de vetores paralelos sobre uma superfície curva no R^3 .

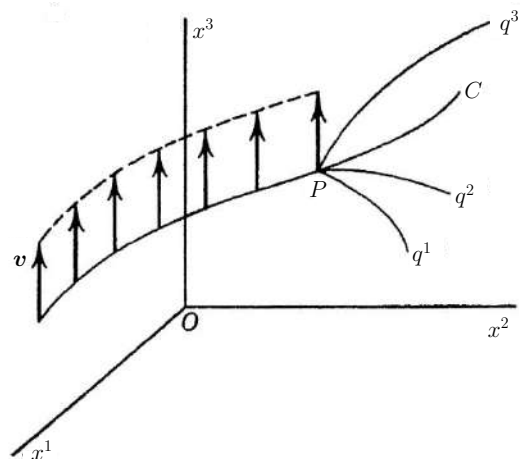


Figura 6.10: Campo vetorial v transportado paralelamente à curva \mathcal{C} em \mathbb{E}^3 . No ponto P , estão também representadas as curvas coordenadas $\{q^1, q^2, q^3\}$.

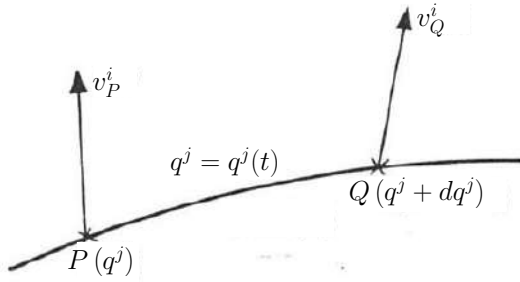


Figura 6.11: Componente v^i transportada paralelamente ao longo da curva \mathcal{C} . Quando a coordenada $q^j(t)$ varia entre os pontos P e Q , a componente varia de v_P^i a v_Q^i .

Segue então de (6.59) que a componente contravariante v^i de v satisfaz a equação diferencial

$$\frac{\delta v^i}{\delta t} = \frac{dv^i}{dt} + \Gamma^i_{jk} v^j \frac{dq^k}{dt} = 0. \quad (6.63a)$$

Ou seja, ao longo de \mathcal{C} , a componente v^i varia por

$$\frac{dv^i}{dt} = -\Gamma^i_{jk} v^j \frac{dq^k}{dt} \text{ ou } dv^i = -\Gamma^i_{jk} v^j dq^k, \quad (6.63b)$$

quando o parâmetro varia de t a $t + dt$, sendo que v^j e Γ^i_{jk} são calculados nas coordenadas correspondentes a t . A quantidade dv^i em (6.63b) mostra que à medida que o campo v é transportado paralelamente ao longo da curva \mathcal{C} , sua i -ésima componente varia por dv^i quando as coordenadas variam de q^j para $q^j + dq^j$. Este comportamento está il-

ustrado na figura 6.11 e mostra como o transporte paralelo de um campo vetorial em um espaço curvo difere de um espaço plano. Neste último caso, os símbolos de Christoffel são todos nulos e, portanto, $dv^i = 0$, ou seja, as componentes de v não se alteram ao longo da curva \mathcal{C} .

Da mesma forma, a componente covariante v_i de v satisfaz a equação diferencial

$$\frac{\delta v_i}{\delta t} = \frac{dv_i}{dt} - \Gamma^j_{ik} v_j \frac{dq^k}{dt} = 0, \quad (6.64a)$$

ou seja, a i -ésima componente covariante de v varia por

$$dv_i = \Gamma^j_{ik} v_j dq^k \quad (6.64b)$$

ao longo de \mathcal{C} quando a j -ésima coordenada varia de q^j a $q^j + dq^j$.

Dados agora dois campos vetoriais v e w que são transportados paralelamente ao longo da mesma curva \mathcal{C} , pode-se mostrar facilmente que tanto os módulos dos mesmos quanto o ângulo θ entre v e w permanecem constantes durante os seus transportes. Isto ocorre porque, em primeiro lugar, de acordo com (6.60) a derivada intrínseca de um escalar é igual a sua derivada ordinária. Em segundo lugar, de acordo com o teorema de Ricci (teorema 6.1), a derivada covariante do tensor de métrica é nula; conseqüentemente, sua derivada absoluta também o será. Assim, dado o invariante

$$\mathbf{v} \cdot \mathbf{w} = vw \cos \theta = v^i w_i = g_{ij} v^i w^j,$$

resulta

$$\frac{d}{dt} (\mathbf{v} \cdot \mathbf{w}) = \frac{\delta}{\delta t} (g_{ij} v^i w^j) = g_{ij} \frac{\delta v^i}{\delta t} w^j + g_{ij} v^i \frac{\delta w^j}{\delta t} = 0, \quad (6.65a)$$

de acordo com (6.63a). Portanto, se $w = v$, então

$$\frac{d}{dt} (\mathbf{v} \cdot \mathbf{v}) = \frac{dv^2}{dt} = 2v \frac{dv}{dt} = 0, \quad (6.65b)$$

ou seja, o módulo de v permanece constante ao longo do transporte paralelo do vetor. O mesmo acontecendo com o módulo de w . Por conseqüência, o ângulo θ entre os mesmos,

$$\frac{d}{dt} (\mathbf{v} \cdot \mathbf{w}) = \frac{d}{dt} (vw \cos \theta) = vw \frac{d}{dt} \cos \theta = -vw \sin \theta \frac{d\theta}{dt} = 0, \quad (6.65c)$$

também permanece constante ao longo de \mathcal{C} , mesmo para $\theta \neq 0$ ou $\theta \neq \pi$.

Considera-se agora o parâmetro t como sendo novamente o comprimento de arco ℓ ao longo da curva \mathcal{C} . O elemento de arco é dado sempre por $d\ell = \sqrt{g_{ij} dq^i dq^j}$, de tal forma que, de acordo com a discussão realizada na seção 6.13.2,

$$d\ell = \sqrt{g_{ij} \frac{dq^i}{d\ell} \frac{dq^j}{d\ell}} d\ell = \mathcal{L} d\ell \implies \mathcal{L} = \sqrt{g_{ij} \frac{dq^i}{d\ell} \frac{dq^j}{d\ell}} = 1 \implies g_{ij} \frac{dq^i}{d\ell} \frac{dq^j}{d\ell} = 1,$$

ao longo de \mathcal{C} . Como $\lambda \doteq dr/d\ell$ é um vetor, este resultado mostra que o mesmo é unitário. Porém, o interessante aqui é que, em primeiro lugar, λ é o vetor tangente a \mathcal{C} em todos os pontos. Em

segundo lugar, e o mais importante, de acordo com (6.63), a derivada intrínseca da i -ésima componente de λ quando este é transportado paralelamente ao longo de \mathcal{C} , resulta ser

$$\frac{\delta \lambda^i}{\delta \ell} = \frac{d\lambda^i}{d\ell} + \Gamma^i_{jk} \lambda^j \frac{dq^k}{d\ell} = \frac{d^2 q^i}{d\ell^2} + \Gamma^i_{jk} \frac{dq^j}{d\ell} \frac{dq^k}{d\ell} = 0,$$

a qual é justamente a equação diferencial satisfeita pela curva geodésica, (equação 6.61b).

Portanto, um vetor que é tangente a uma curva geodésica em um determinado ponto e que é transportado paralelamente ao longo desta curva permanecerá sempre tangente à geodésica. Segue disto que, dado um campo vetorial que é sempre tangente a alguma variedade no R^n , as curvas geodésicas sobre esta variedade podem sempre ser estabelecidas de forma inequívoca a partir deste campo vetorial.

Um exemplo de um campo vetorial deste tipo vem de uma superfície esférica no \mathbb{E}^3 . Sobre esta superfície, os vetores unitários \hat{e}_θ são sempre tangentes à mesma. Estes vetores unitários são sempre transportados paralelamente ao longo de uma curva meridiana, a qual é, de acordo com o exemplo 6.19, uma curva geodésica sobre a superfície esférica.

Dados agora os vetores λ e w que são transportados paralelamente ao longo de uma curva geodésica, sendo que λ é o campo vetorial tangente à mesma, os resultados (6.65b e c) mostram que não somente o módulo de w permanece constante ao longo da geodésica, mas que também o seu ângulo com a curva não varia durante o seu transporte.

A figura 6.12 ilustra o transporte paralelo de um vetor ao longo de geodésicas de um espaço R^2 esférico. Partindo do ponto A e seguindo até N ao longo de um meridiano (curva geodésica), o vetor permanece tangencial à curva ao longo do trajeto. Em seguida, o mesmo vetor é transportado paralelamente ao longo de outra geodésica conectando os pontos N e B . Finalmente, o vetor é paralelamente transportado de B a A ao longo da terceira geodésica (linha do equador). Ao retornar ao ponto de partida, o vetor, embora tenha sido paralelamente transportado ao longo de um caminho fechado composto somente por curvas geodésicas, resulta com uma orientação espacial que difere de sua orientação original pelo ângulo α ilustrado na figura.

Este tipo de resultado é típico de espaços denominados *curvos* e não ocorre em espaços planos, como o \mathbb{E}^n . Qualquer vetor transportado paralelamente em um espaço Euclidiano ao longo de suas geodésicas (linhas retas), sempre irá retornar ao ponto de partida com a mesma orientação inicial. Uma medida da curvatura de um espaço é fornecida pelo tensor de Riemann, discutido na próxima seção.

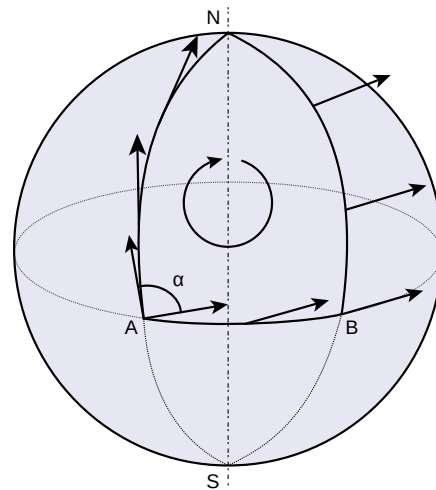


Figura 6.12: Transporte paralelo de um vetor ao longo de um caminho fechado $\mathcal{C} : (A \rightarrow N \rightarrow B \rightarrow A)$ sobre uma esfera. Ao retornar ao ponto A , o vetor teve sua orientação original alterada pelo ângulo α .

6.14 OS TENSORES DE RIEMANN, RICCI E EINSTEIN

O cálculo do *tensor de curvatura de Riemann-Christoffel* para uma determinada métrica no R^n consiste no método mais prático para determinar se o espaço descrito pela métrica é plano ou curvo e, neste último caso, para então determinar a sua curvatura. Em outras palavras, o tensor de Riemann mede o quanto o tensor de métrica é ou não localmente isométrico à métrica Euclidiana.

Nesta seção será realizada uma dedução do tensor de Riemann e discutidas algumas de suas propriedades. Em seguida, serão obtidos dois outros tensores relacionados, os *tensores de Ricci* e de *Einstein*, este último sendo fundamental na teoria da relatividade generalizada.

6.14.1 O TENSOR DE CURVATURA DE RIEMANN-CHRISTOFFEL

Uma quantidade fundamental para a geometria de espaços curvos e para a dinâmica de corpos materiais em tais espaços, é o *tensor de curvatura de Riemann-Christoffel*. Este tensor está relacionado com as diferenças obtidas nos componentes de tensores quando o deslocamento paralelo dos mesmos entre dois pontos do espaço R^n ocorre por dois caminhos distintos.

Matematicamente, esta diferença está relacionada com a ausência de simetria na ordem das derivadas mistas em espaços curvos. Em análise de funções matemáticas, se $u = u(x, y)$ é uma função das variáveis x e y e pertence à classe $C^2(\mathbb{R})$, i. e., é duas vezes diferenciável em qualquer uma das variáveis, então $\partial^2 u / \partial x \partial y = \partial^2 u / \partial y \partial x$. Contudo, em espaços curvos, a diferenciação covariante dos componentes de um tensor não apresenta, em geral, a mesma propriedade.

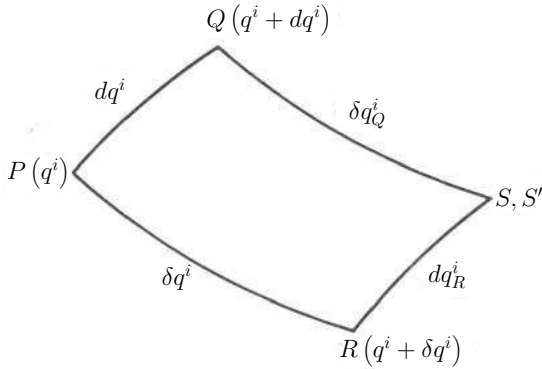


Figura 6.13: Os deslocamentos paralelos do segmento δq^i ao longo de \overline{PQ} e de dq^i ao longo de \overline{PR} resultam no mesmo ponto final $S' = S$.

Para tanto, considera-se um ponto inicial P contido no R^n , ilustrado na figura 6.13. A partir deste ponto, consideram-se dois deslocamentos distintos, ambos partindo de P ($\{q^i\}$) e chegando, respectivamente, a Q ($\{q^i + dq^i\}$) e R ($\{q^i + \delta q^i\}$), sendo que estes não são necessariamente infinitesimais. O símbolo “ d ” indica aqui um determinado incremento na coordenada, ao passo que “ δ ” representa um outro incremento.

Imagina-se agora que o segmento \overline{PR} é deslocado de tal forma que quando o ponto P coincidir com Q , o ponto R coincidirá com o ponto S na figura 6.13. Isto equivale a definir o segmento \overline{QS} , composto pelo mapeamento unívoco de cada ponto ao longo de \overline{PR} sendo transladado pela quantidade “ d ”, resultando nos mapeamentos $P \rightarrow Q : q^i \rightarrow q^i + dq^i$, $P \rightarrow R : q^i \rightarrow q^i + \delta q^i$, e então

$$R \rightarrow S : q^i + \delta q^i \rightarrow q^i + \delta q^i + dq_R^i,$$

sendo

$$S : q^i + \delta q^i + dq_R^i = q^i + \delta q^i + d(q^i + \delta q^i) = q^i + \delta q^i + dq^i + d(\delta q^i)$$

as coordenadas do ponto S . Como $\delta r = \delta q^i e_i$ é um vetor no R^n e como os deslocamentos ocorrem ao longo de geodésicas, então, se os deslocamentos “ d ” e “ δ ” tornarem-se elementares, o transporte paralelo de δq^i resulta em

$$d(\delta q^i) = -\Gamma^i_{jk(P)} \delta q^j dq^k,$$

de acordo com (6.63b), onde $\Gamma^i_{jk(P)}$ indica que os símbolos de Christoffel devem ser calculados no ponto P . De forma similar, se o segmento \overline{PQ} é deslocado de tal forma que quando P coincide com R , o ponto Q coincide com o ponto S' , resultando assim no segmento $\overline{RS'}$. Desta maneira, cada ponto ao longo de \overline{PQ} é transladado pela quantidade “ δ ”, resultando no mapeamento

$$Q \rightarrow S' : q^i + dq^i \rightarrow q^i + dq^i + \delta q_Q^i.$$

Assim, as coordenadas do ponto S' são

$$S' : q^i + dq^i + \delta q_Q^i = q^i + dq^i + \delta(q^i + dq^i) = q^i + dq^i + \delta q^i + \delta(dq^i).$$

Portanto, se o deslocamento “ δ ” for elementar, a variação na coordenada dq^i também é dada por (6.63b), resultando em

$$\delta(dq^i) = -\Gamma^i_{jk(P)} dq^j \delta q^k.$$

Comparando as expressões para $d(\delta q^i)$ e $\delta(dq^i)$, observa-se que uma simples troca de índices mudos, aliada à simetria de Γ^i_{jk} frente à permutação dos índices j e k , mostra que ambas as expressões são idênticas, resultando, portanto, que $S' = S$. Ou seja, as curvas elementares $\mathcal{C}_1 \equiv \mathcal{C}_{P \rightarrow R \rightarrow S}$ e $\mathcal{C}_2 \equiv \mathcal{C}_{P \rightarrow Q \rightarrow S'}$, partindo do mesmo ponto inicial P , conduzem ao mesmo ponto final, desde que estas sejam realizadas ao longo de curvas geodésicas no R^n .

Considera-se agora o transporte paralelo de um campo vetorial $v = v^i e_i$ ao longo de ambos os caminhos \mathcal{C}_1 e \mathcal{C}_2 da figura 6.13. Este transporte está ilustrado agora na figura 6.14. O campo vetorial é suposto estar definido no ponto P , onde suas componentes possuem valores iguais a $\{v_P^i\}$, e estes componentes são transportados até o ponto S ao longo de dois caminhos distintos:

\mathcal{C}_1 : Realiza-se o transporte paralelo de v de P até Q com um deslocamento $\{dq^i\}$, resultando então nas componentes $\{v_Q^i\}$. Em seguida, realiza-se o transporte paralelo do mesmo vetor de Q até S com o deslocamento $\{\delta q^i\}$, resultando nos valores finais $\{v_{SQ}^i\}$ para as componentes de v .

\mathcal{C}_2 : Desloca-se v paralelamente de P a R por $\{dq^i\}$. Em seguida, desloca-se o mesmo de R a S por $\{\delta q^i\}$, resultando as componentes $\{v_{SR}^i\}$.

As componentes $\{v_{SQ}^i\}$ e $\{v_{SR}^i\}$ são idênticas ou distintas entre si? Em geral, para um espaço curvo, $v_{SQ}^i \neq v_{SR}^i$ e a razão para tanto pode ser entendida calculando as variações da i -ésima componente de v ao longo dos dois caminhos distintos.

Ao longo do transporte paralelo $P \rightarrow Q$, a referida componente resulta com o valor

$$v_Q^i = v_P^i + dv_P^i,$$

onde dv_P^i é a variação de v^i no referido deslocamento. De acordo com (6.63b), para um deslocamento suficientemente pequeno,

$$dv_P^i = -\Gamma^i_{jk(P)} v_P^j dq^k,$$

onde $\{\Gamma^i_{jk(P)}\}$ são os valores assumidos pelos símbolos de Christoffel no ponto P . Ou seja,

$$v_Q^i = v_P^i - \Gamma^i_{jk(P)} v_P^j dq^k.$$

A componente v_Q^i é agora transportada ao ponto S , assumindo o valor v_{SQ}^i no mesmo, a qual é dada por

$$v_{SQ}^i = v_Q^i + \delta v_Q^i = v_Q^i - \Gamma^i_{jk(Q)} v_Q^j \delta q^k,$$

onde agora $\{\Gamma^i_{jk(Q)}\}$ são os símbolos de Christoffel em Q . Como estes são funções das coordenadas $\{q^i\}$, pode-se escrever, para dq^i pequeno o suficiente,

$$\Gamma^i_{jk} = \Gamma^i_{jk}(\{q^i\}) \implies \Gamma^i_{jk(Q)} = \Gamma^i_{jk(P)} + \left. \frac{\partial \Gamma^i_{jk}}{\partial q^\ell} \right|_P dq^\ell \equiv \Gamma^i_{jk(P)} + \Gamma^i_{jk,\ell} dq^\ell,$$

sendo $\Gamma^i_{jk,\ell} \equiv \partial \Gamma^i_{jk} / \partial q^\ell$ em P . Então, a componente v_{SQ}^i acima pode ser escrita como

$$\begin{aligned} v_{SQ}^i &= v_P^i - \Gamma^i_{jk(P)} v_P^j dq^k - \left(\Gamma^i_{jk(P)} + \Gamma^i_{jk,\ell} dq^\ell \right) \left(v_P^j - \Gamma^j_{mn(P)} v_P^m dq^n \right) \delta q^k \\ &= v_P^i - \Gamma^i_{jk(P)} v_P^j dq^k - \Gamma^i_{jk(P)} v_P^j \delta q^k - \Gamma^i_{jk,\ell} v_P^j \delta q^k dq^\ell + \Gamma^i_{jk(P)} \Gamma^j_{mn(P)} v_P^m \delta q^k dq^n + \mathcal{O} \left[(dq^i)^3 \right] \\ &= v^i - \Gamma^i_{jk} v^j dq^k - \Gamma^i_{jk} v^j \delta q^k - \Gamma^i_{jk,\ell} v^j \delta q^k dq^\ell + \Gamma^i_{jk} \Gamma^j_{mn} v^m \delta q^k dq^n + \mathcal{O} \left[(dq^i)^3 \right]. \end{aligned}$$

Na última expressão, o índice “ P ” foi removido porque todas as quantidades no lado direito são calculadas neste ponto.

Realizando agora o mesmo procedimento para o transporte $P \rightarrow R \rightarrow S$, resulta

$$v_{SR}^i = v^i - \Gamma^i_{jk} v^j \delta q^k - \Gamma^i_{jk} v^j dq^k - \Gamma^i_{jk,\ell} v^j \delta q^k \delta q^\ell + \Gamma^i_{jk} \Gamma^j_{mn} v^m \delta q^k \delta q^n + \mathcal{O} \left[(dq^i)^3 \right].$$

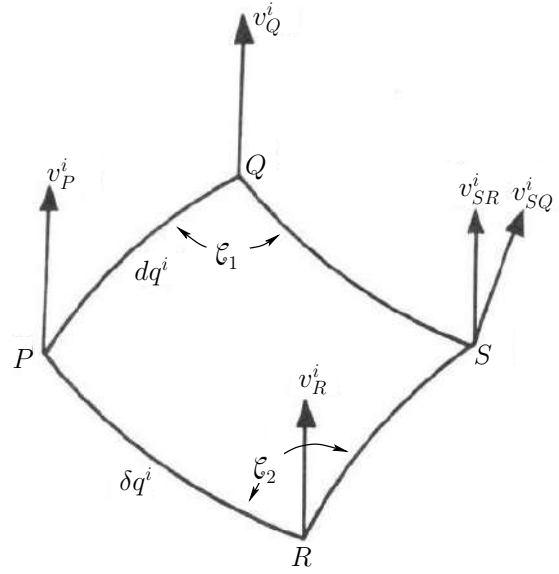


Figura 6.14: O deslocamento paralelo do vetor v do ponto P ao ponto S ao longo de dois caminhos distintos resulta em diferentes valores para a componente v^i .

Portanto, a diferença na i -ésima componente de v no ponto S , para dois caminhos arbitrários e distintos partindo de P , é igual a

$$v_{SR}^i - v_{SQ}^i = \Gamma_{jk,\ell}^i v^j \delta q^k dq^\ell - \Gamma_{jk,\ell}^i v^j dq^k \delta q^\ell + \Gamma_{jk}^i \Gamma_{mn}^j v^m dq^k \delta q^n - \Gamma_{jk}^i \Gamma_{mn}^j v^m \delta q^k dq^n,$$

onde foram mantidos somente os termos até segunda ordem em dq^i .

A diferença acima pode ser escrita como

$$v_{SR}^i - v_{SQ}^i = R^i_{jkl} v^j dq^k \delta q^\ell, \quad (6.66a)$$

onde

$$R^i_{jkl} \equiv \Gamma_{jk,\ell}^i - \Gamma_{j\ell,k}^i + \Gamma_{m\ell}^i \Gamma_{jk}^m - \Gamma_{mk}^i \Gamma_{j\ell}^m \quad (6.66b)$$

é o **tensor de curvatura de Riemann-Christoffel misto** ou **do segundo tipo**. Como v^i , dq^k e δq^ℓ em (6.66a) são todos componentes de vetores, pela regra do quociente as quantidades $\{R^i_{jkl}\}$ realmente compõe um tensor de posto quatro. Uma outra maneira de se escrever (6.66b) é na forma de determinantes,

$$R^i_{jkl} = \begin{vmatrix} \frac{\partial}{\partial q^\ell} & \frac{\partial}{\partial q^k} \\ \Gamma_{j\ell}^i & \Gamma_{jk}^i \end{vmatrix} + \begin{vmatrix} \Gamma_{m\ell}^i & \Gamma_{mk}^i \\ \Gamma_{j\ell}^m & \Gamma_{jk}^m \end{vmatrix}. \quad (6.66c)$$

Observa-se que o tensor de curvatura independe do campo v ; este depende somente do tensor de métrica e de suas derivadas, ou seja, é uma função somente da geometria do espaço curvo. Para que o valor de v_S^i independa do caminho adotado a partir de P , é necessário que $R^i_{jkl} = 0$, uma vez que o campo v é arbitrário. Em um espaço Euclideo, sempre é possível encontrar um sistema de coordenadas (Cartesiano, por exemplo), onde $\Gamma_{jk}^i = 0$. Neste sistema de coordenadas, o tensor de curvatura é identicamente nulo. Pode-se mostrar que o mesmo ocorre para qualquer outro sistema de coordenadas neste espaço.

Através do tensor de curvatura, é possível atribuir significado ao termo espaço *plano* ou *Euclideo*, como sendo aquele onde o tensor de Riemann-Christoffel é identicamente nulo em todos os pontos deste espaço. Se esta condição não for satisfeita, o espaço é *curvo* ou *não Euclideo*. Este resultado é de fundamental importância para a dinâmica de sistemas físicos em espaços curvos, descritos por teorias tais como a Relatividade Geral.

Se ao invés das componentes contravariantes de v fossem realizados os transportes paralelos das componentes covariantes $\{v_i\}$ entre os pontos P e S da figura 6.14, pode-se mostrar, com o emprego de (6.64b), que a diferença entre os valores de v_{iS} obtidos nos dois caminhos distintos seria igual a

$$v_{iSR} - v_{iSQ} = -R^j_{ilk} v_j dq^k \delta q^\ell.$$

Estes resultados mostram que o deslocamento paralelo de um vetor e, em geral, de um tensor, entre dois pontos de um espaço Riemanniano depende do caminho escolhido. Segue disto que se um tensor é deslocado paralelamente ao longo de uma curva fechada formada por geodésicas, ao retornar ao ponto de partida os seus componentes não irão possuir em geral os mesmos valores que possuíam originalmente. Este fato, característico de espaços curvos, tem consequências importantes para a física acerca do conceito de campos conservativos em espaços Riemannianos curvos.

Um tensor associado a R^i_{jkl} é

$$R_{ijkl} = g_{im} R^m_{jkl}, \quad (6.67a)$$

denominado o **tensor de curvatura de Riemann-Christoffel covariante** ou **do primeiro tipo**. Não é difícil verificar que este tensor pode ser escrito como

$$R_{ijkl} = [jk, i]_{,\ell} - [j\ell, i]_{,k} + [ik, m] \Gamma_{j\ell}^m - [i\ell, m] \Gamma_{jk}^m \quad (6.67b)$$

$$= [jk, i]_{,\ell} - [j\ell, i]_{,k} + g^{mn} ([ik, m] [j\ell, n] - [i\ell, m] [jk, n]) \quad (6.67c)$$

$$= \begin{vmatrix} \frac{\partial}{\partial q^\ell} & \frac{\partial}{\partial q^k} \\ [j\ell, i] & [jk, i] \end{vmatrix} + \begin{vmatrix} [ik, m] & [i\ell, m] \\ \Gamma_{jk}^m & \Gamma_{j\ell}^m \end{vmatrix} \quad (6.67d)$$

$$= \begin{vmatrix} \frac{\partial}{\partial q^\ell} & \frac{\partial}{\partial q^k} \\ [j\ell, i] & [jk, i] \end{vmatrix} + g^{mn} \begin{vmatrix} [ik, m] & [i\ell, m] \\ [jk, n] & [j\ell, n] \end{vmatrix}. \quad (6.67e)$$

PROPRIEDADES DO TENSOR DE CURVATURA

Pode-se ver, tanto a partir da definição de R^i_{jkl} em (6.66) quanto a partir da definição de R_{ijkl} em (6.67) que as seguintes propriedades de simetria são satisfeitas:

$$R^i_{jkl} = -R^i_{jlk} \qquad R_{ijkl} = -R_{ijlk} \qquad (6.68a)$$

$$R^i_{jkl} = -R^j_{ikl} \qquad R_{ijkl} = -R_{jikl} \qquad (6.68b)$$

$$R^i_{jkl} = R^k_{lij} \qquad R_{ijkl} = R_{klij}. \qquad (6.68c)$$

Uma consequência digna de nota da antissimetria do tensor de curvatura frente aos pares iniciais e finais de índices, propriedades (6.68a,b), é:

$$(NS) : R^i_{ikl} = R^i_{jkk} = R_{ikk} = R_{ijkk} = 0. \qquad (6.68d)$$

A seguinte propriedade cíclica também é válida: *se qualquer índice do tensor é mantido fixo enquanto os três índices restantes são permutados de forma cíclica e os componentes resultantes são adicionados, o resultado é nulo.* Por exemplo,

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0. \qquad (6.68e)$$

As propriedades de simetria (6.68a-e) reduzem substancialmente o número total de componentes independentes do tensor de curvatura. Em um espaço R^n , este tensor possui, ao todo, n^4 componentes. Porém, com as simetrias conhecidas, o número de componentes independentes é de “somente” $n^2(n^2 - 1)/12$, ao todo. A tabela a seguir ilustra estes números para algumas dimensionalidades:

Dimensionalidade do espaço	1	2	3	4	5
Número total de componentes	1	16	81	256	625
Número de componentes independentes	0	1	6	20	50

O fato de que o número de componentes independentes em um espaço unidimensional é nulo implica em que estes espaços são necessariamente Euclidianos. Isto pode ser entendido pelo fato de que o elemento de arco em um espaço 1D geral deve necessariamente ser do tipo

$$d\ell^2 = f(q) dq^2,$$

o qual pode ser transformado a um espaço Cartesiano do tipo $d\ell^2 = dq'^2$ através da definição $dq' = \sqrt{f(q)}dq$.

Exercício 6.20. Encontre os componentes do tensor de curvatura no espaço de dimensão 2 sobre a superfície de uma esfera de raio a .

Solução. Dados os símbolos de Christoffel para esta variedade, obtidos no exemplo 6.19, pode-se escrever os símbolos de Christoffel do primeiro tipo e o tensor de métrica na forma matricial como

$$[i] = \begin{pmatrix} [11, i] & [12, i] \\ [21, i] & [22, i] \end{pmatrix} \implies [\theta] = \begin{pmatrix} 0 & 0 \\ 0 & -a^2 \sin \theta \cos \theta \end{pmatrix} \text{ e } [\phi] = \begin{pmatrix} 0 & a^2 \sin \theta \cos \theta \\ a^2 \sin \theta \cos \theta & 0 \end{pmatrix}.$$

Então, como há somente 1 componente independente no tensor de curvatura, e como

$$g^{-1} = \begin{pmatrix} a^{-2} & 0 \\ 0 & a^{-2} \sin^{-2} \theta \end{pmatrix},$$

pode-se tomar o componente $R_{1221} = R_{\theta\phi\phi\theta}$ como este componente e calcular a partir de (6.67c),

$$\begin{aligned} R_{1221} &= [\phi\phi, \theta]_{,\theta} + g^{\phi\phi} [\theta\phi, \phi] [\phi\theta, \phi] \\ &= a^2 \sin^2 \theta. \end{aligned}$$

Os outros componentes de R_{ijkl} são dados por (6.68a-e). Ou seja, este de fato é um espaço curvo.

6.14.2 O TENSOR DE RICCI E O ESCALAR DE CURVATURA

Dois outros tensores importantes para a dinâmica de partículas e campos em espaços curvos são os tensores de Ricci e de Einstein. Ambos são obtidos a partir do tensor de curvatura de Riemann-Christoffel.

O **tensor de Ricci** é obtido a partir da contração de R^i_{jkl} dado por (6.66b,c) como

$$R_{ij} = R^k_{ijk} = g^{pk} R_{pijk} \quad (6.69a)$$

$$= \Gamma^k_{ij,k} - \Gamma^k_{ik,j} + \Gamma^\ell_{ij} \Gamma^k_{k\ell} - \Gamma^\ell_{ik} \Gamma^k_{j\ell} \quad (6.69b)$$

$$= \left[\begin{array}{c} \frac{\partial}{\partial q^k} \\ \frac{\partial}{\partial q^j} \end{array} \right] + \left[\begin{array}{c} \Gamma^k_{k\ell} \\ \Gamma^\ell_{ik} \end{array} \right] \left[\begin{array}{c} \Gamma^k_{j\ell} \\ \Gamma^\ell_{ij} \end{array} \right]. \quad (6.69c)$$

Uma outra expressão importante para o tensor de Ricci pode ser obtida a partir da identidade (6.55). Empregando (6.51a) e (6.47c), esta pode ser escrita como

$$\frac{\partial g}{\partial q^i} = gg^{mn} ([mi, n] + [ni, m]) = 2g\Gamma^m_{im} \implies \Gamma^m_{im} = \frac{\partial}{\partial q^i} \ln \sqrt{|g|}.$$

Portanto, (6.69b) pode ser expresso em termos de g como

$$R_{ij} = \Gamma^k_{ij,k} - \Gamma^\ell_{ik} \Gamma^k_{j\ell} - \frac{\partial^2 \ln \sqrt{|g|}}{\partial q^i \partial q^j} + \Gamma^\ell_{ij} \frac{\partial}{\partial q^\ell} \ln \sqrt{|g|}. \quad (6.69d)$$

A partir da expressão (6.69d), pode-se ver facilmente que o tensor de Ricci é simétrico, $R_{ij} = R_{ji}$. Devido a isso, o número de componentes independentes de R_{ij} é igual a $n(n+1)/2$. Em uma variedade no R^4 , se a métrica for determinada pelas equações diferenciais parciais $R_{ij} = 0$, isto resultará em um sistema de 10 equações, as quais foram adotadas por Einstein como as equações do campo gravitacional no espaço livre, na teoria da relatividade geral.

Uma contração adicional de índices sobre o tensor de Ricci (6.69) determina o **escalar de curvatura ou de Ricci**,

$$R = R^i_i = g^{ij} R_{ij} = g^{ij} g^{pk} R_{pijk}. \quad (6.70)$$

O escalar de Ricci será sempre não nulo em um espaço curvo. Esta quantidade irá aparecer logo abaixo, nas equações de Einstein para o campo gravitacional.

6.14.3 O TENSOR DE EINSTEIN E AS EQUAÇÕES DO CAMPO GRAVITACIONAL

As equações de campo de Einstein, ou simplesmente as equações de Einstein consistem em um conjunto de equações diferenciais que descrevem a interação gravitacional entre corpos físicos como o resultado da curvatura do espaço-tempo causada pela presença de matéria e energia. As equações de Einstein formam a parte quantitativa da *teoria da relatividade generalizada* e foram publicados pela primeira vez em 1915,¹⁷ sendo divulgadas novamente em 1916, quando então a teoria completa foi publicada.¹⁸

Para se obter o tensor derivado por Einstein na relatividade geral, pode-se partir da identidade de Bianchi

$$R_{ijkl;m} + R_{ijlm;k} + R_{ijmk;\ell} = 0,$$

a qual pode ser obtida a partir de (6.68e). Multiplicando esta identidade por $g^{i\ell} g^{jk}$, usando a antissimetria do tensor de curvatura e lembrando que a derivada covariante do tensor de métrica é nula, obtém-se

$$g^{jk} R^{\ell}_{jkl;m} - g^{jk} R^{\ell}_{jml;k} - g^{i\ell} R^k_{imk;\ell} = 0,$$

¹⁷Die Feldgleichungen der Gravitation (As equações de campo da gravitação). *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*, p. 844-847, 25 de novembro de 1915. Acesso: <https://www.biodiversitylibrary.org/item/92536#page/920/mode/lup> (versão em inglês: <https://einsteinpapers.press.princeton.edu/vol6-trans/129>).

¹⁸Die Grundlage der allgemeinen Relativitätstheorie (Os fundamentos da teoria da relatividade geral). *Annalen der Physik*, v. 354, n. 7, p. 769-822, 11 de maio de 1916. Acesso doi: [10.1002/andp.19163540702](https://doi.org/10.1002/andp.19163540702) (versão em inglês: <https://einsteinpapers.press.princeton.edu/vol6-trans/158>).

ou seja,

$$g^{jk} R_{jk;m} - g^{jk} R_{jm;k} - g^{i\ell} R_{im;\ell} = 0,$$

que também pode ser escrita como

$$R^k_{k;m} - 2R^k_{m;k} = 0.$$

Dado agora o escalar de Ricci (6.70),

$$R \equiv g^{ij} R_{ij},$$

observa-se que $R_{,k} = g^{ij} R_{ij;k} = R^j_{j;k}$; ou seja, a identidade anterior pode ser reescrita como

$$R_{,m} - 2R^k_{m;k} = 0,$$

ou como

$$\left(R^k_m - \frac{1}{2} \delta^k_m R \right)_{;k} = 0.$$

A condição suficiente para o cumprimento da identidade acima forma o sistema de equações

$$G^i_j = 0, \tag{6.71a}$$

sendo que o tensor

$$G^i_j \doteq R^i_j - \frac{1}{2} \delta^i_j R \tag{6.71b}$$

é o **tensor de Einstein**.

Em um espaço de dimensão 4, as 10 equações diferenciais parciais em (6.71a) são as **equações de Einstein para o espaço livre**, isto é, as equações de campo da teoria geral da relatividade para um corpo massivo no vácuo (sem a presença de outros campos) e sem a constante cosmológica.

6.15 APLICAÇÕES FÍSICAS

Algumas aplicações físicas do conteúdo desenvolvido neste capítulo serão apresentadas agora.

6.15.1 A TRANSFORMAÇÃO DE LORENTZ, O ESPAÇO-TEMPO DE MINKOWSKI E A FORMULAÇÃO COVARIANTE DO ELETROMAGNETISMO CLÁSSICO

Será realizada aqui uma breve abordagem da transformação de Lorentz e do espaço vetorial de Minkowski, dentro do contexto da relatividade restrita. Não se pretende realizar aqui uma introdução à teoria da relatividade, mas sim somente apresentar um conjunto de expressões relevantes para a próxima seção, onde será discutida a formulação covariante do eletromagnetismo clássico.

As origens da teoria da relatividade restrita estão relacionadas com o desenvolvimento do eletromagnetismo clássico, ocorrido principalmente durante a segunda metade do século XIX, com os trabalhos de James Clerk Maxwell (1831 – 1879). Pode-se afirmar que a unificação da eletricidade e do magnetismo nas equações de Maxwell forçou o desenvolvimento da relatividade.

Os fundamentos da teoria foram construídos a partir das contribuições pioneiras de diversos matemáticos e físicos, destacando-se em particular o trabalho do físico holandês Hendrik Antoon Lorentz (1853 – 1928), a partir de 1890, com contribuições importantes fornecidas pelo físico-matemático Jules Henri Poincaré (1854 – 1912). Contudo, foi o físico alemão Albert Einstein (1879 – 1955) quem contribuiu com as ideias cruciais, principalmente com a publicação em 1905 do artigo:

*Zur Elektrodynamik bewegter Körper.*¹⁹ *Annalen der Physik, v. 322, n. 10, pp. 891 – 921, Juni 1905.*

¹⁹Acerca da eletrodinâmica dos corpos em movimento. Acesso livre no doi: [10.1002/andp.19053221004](https://doi.org/10.1002/andp.19053221004). Versão em inglês: <https://einsteinpapers.press.princeton.edu/vol2-trans/154>.

Na física moderna, a teoria da relatividade restrita é suposta válida para todas as formas de interações, exceto em fenômenos gravitacionais de larga escala, onde é necessário empregar a teoria generalizada da relatividade. Em particular, fenômenos atômicos, nucleares e subatômicos não podem ser adequadamente descritos sem o uso da relatividade.

6.15.1.1 A SITUAÇÃO ANTERIOR A 1900. A TRANSFORMAÇÃO DE GALILEU

Nos 40 anos anteriores a 1900, a eletricidade, o magnetismo e a óptica tiveram as suas relações mutuas devidamente estabelecidas na teoria eletromagnética descrita pelas equações de Maxwell.

Uma vez que a experiência prévia com movimento ondulatório sempre envolvera um meio necessário para a propagação da onda, era natural assumir que a luz também necessitaria de um meio, denominado *éter*, o qual permearia todo o espaço, possuía densidade desprezível e interagía fracamente com a matéria ordinária. Este meio existia somente para permitir a propagação de ondas eletromagnéticas.

A hipótese do éter colocava os fenômenos eletromagnéticos em um plano conceitual distinto da mecânica Newtoniana. Nesta, os conceitos de espaço e tempo são separáveis, sendo o tempo uma quantidade absoluta, independente do referencial e que era determinado por um referencial absoluto, possivelmente em repouso em relação às estrelas distantes. Paralelamente, as leis da mecânica são as mesmas em diferentes referenciais inerciais, isto é, elas são invariantes frente a uma **transformação de Galileu**.

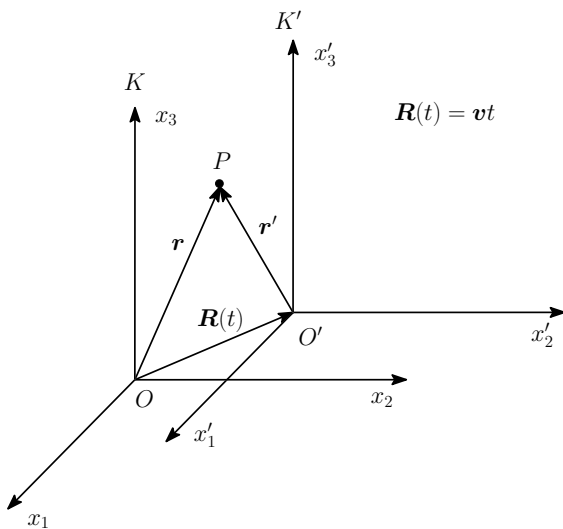


Figura 6.15: Transformação entre os referenciais inerciais K e K' .

Para enfatizar a distinção entre a mecânica Newtoniana e o eletromagnetismo, será considerada brevemente a aplicação da relatividade Galileana a cada uma destas teorias. A figura 6.15 mostra dois sistemas de referência K e K' , com coordenadas $(x_1, x_2, x_3; t)$ e $(x'_1, x'_2, x'_3; t')$, respectivamente, sendo que o referencial K' se move em relação a K com velocidade relativa v (constante) e as origens de ambos os referenciais concordam no instante $t = t' = 0$. Na transformação de Galileu, as coordenadas espaciais e o tempo em cada referencial estão relacionados de acordo com

$$\begin{cases} \mathbf{r}' = \mathbf{r} - \mathbf{v}t \\ t' = t \end{cases} \iff \begin{cases} \mathbf{r} = \mathbf{r}' + \mathbf{v}t' \\ t = t'. \end{cases} \quad (6.72)$$

A lei de transformação (6.72) descreve uma simples translação, para um determinado instante de tempo. No contexto desta transformação, o sistema físico é descrito matematicamente no espaço Euclidiano \mathbb{E}^3 . Verifica-se facilmente que nesta transformação os elementos de comprimento nos

dois referenciais são os mesmos,

$$d\ell^2 = dx_1^2 + dx_2^2 + dx_3^2 = dx_1'^2 + dx_2'^2 + dx_3'^2,$$

pois ambos são mensurados no mesmo instante de tempo, o qual não influi na métrica do espaço. Ou seja, o tensor de métrica é, simplesmente, $g_{ij} = \delta_{ij}$, como se espera em um espaço Euclidiano com coordenadas Cartesianas.

A lei de transformação (6.72) é perfeitamente adequada para a mecânica Newtoniana, pois as suas leis físicas são invariantes frente a uma transformação de Galileu. Para exemplificar, considera-se um sistema de partículas interagindo por meio de potenciais centrais. Supondo que a equação de movimento da i -ésima partícula no sistema K' seja

$$\frac{d\mathbf{p}'_i}{dt'} = -\nabla'_i \left[\sum_{j \neq i} V_{ij}(|\mathbf{r}'_i - \mathbf{r}'_j|) \right],$$

a qual é a expressão para a segunda lei de Newton, sendo $\mathbf{r}'_i(t')$ e $\mathbf{p}'_i(t')$ respectivamente a posição e o momentum linear instantâneos da i -ésima partícula medidos no referencial K' . Nesta

expressão também está indicado o potencial de interação entre as partículas i e j , $V_{ij}(|\mathbf{r}'_i - \mathbf{r}'_j|)$, o qual origina-se da força central conservativa de interação entre as mesmas. Estas interações podem ser gravitacionais ou elétricas, por exemplo. Finalmente, $\nabla'_i = \hat{x}_j \partial / \partial x'_{ij}$ é o operador gradiente, calculado no referencial K' , sendo $x'_{ij}(t')$ a j -ésima coordenada da posição instantânea $\mathbf{r}'_i(t')$.

Denominando $\mathbf{u}(t) = d\mathbf{r}/dt$ a velocidade instantânea de uma partícula, mensurada por um determinado referencial, a lei de transformação (6.72) implica na seguinte fórmula de transformação para as velocidades medidas em cada referencial,

$$\mathbf{u}' = \mathbf{u} - \mathbf{v} \iff \mathbf{u} = \mathbf{u}' + \mathbf{v}. \quad (6.73)$$

Como a métrica do espaço é a mesma em ambos os referenciais, então $\nabla'_i = \nabla_i$, onde ∇_i é o operador gradiente aplicado sobre \mathbf{r}_i . Portanto, a lei (6.72) implica em que

$$\frac{d\mathbf{p}'_i}{dt'} = \frac{d\mathbf{p}_i}{dt} \quad \text{e} \quad \mathbf{r}'_i - \mathbf{r}'_j = \mathbf{r}_i - \mathbf{r}_j,$$

e a equação de movimento para a i -ésima partícula, transformada do referencial K' para K resulta

$$\frac{d\mathbf{p}_i}{dt} = -\nabla_i \left[\sum_{j \neq i} V_{ij}(|\mathbf{r}_i - \mathbf{r}_j|) \right],$$

possuindo a mesma forma matemática apresentada em relação a K' . Ou seja, a mecânica Newtoniana é *invariante frente a uma transformação de Galileu*.

6.15.1.2 EQUAÇÕES DE MAXWELL E A TRANSFORMAÇÃO DE GALILEU

Em contraste com a invariância da mecânica, as leis do eletromagnetismo mudam conforme o referencial, se for assumida uma transformação de Galileu. As equações de Maxwell e expressões relacionadas para distribuições de carga e corrente elétricas no vácuo e no sistema Gaussiano de unidades são:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= 4\pi\rho && \text{(Lei de Gauss: eletricidade)} \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} &= \frac{4\pi}{c} \mathbf{J} && \text{(Lei de Ampère)} \\ \nabla \cdot \mathbf{B} &= 0 && \text{(Lei de Gauss: magnetismo)} \end{aligned} \quad (6.74a)$$

$$\begin{aligned} \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} &= 0 && \text{(Lei de Faraday)} \\ \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} &= 0 && \text{(Equação da continuidade),} \end{aligned} \quad (6.74b)$$

sendo $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$, $\rho(\mathbf{r}, t)$ e $\mathbf{J}(\mathbf{r}, t)$ respectivamente os campos elétrico e de indução magnética e as densidades de carga e corrente elétricas. Os campos também podem ser expressos em termos dos potenciais escalar $\Phi(\mathbf{r}, t)$ e vetor $\mathbf{A}(\mathbf{r}, t)$ como

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t) \quad \text{e} \quad \mathbf{E}(\mathbf{r}, t) = -\nabla\Phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t}. \quad (6.75)$$

Combinando as leis de Ampère e Faraday de forma a desacoplar os campos, obtém-se as equações da onda eletromagnética

$$\begin{aligned} \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} + \nabla \times \nabla \times \mathbf{E} &= -\frac{4\pi}{c^2} \frac{\partial \mathbf{J}}{\partial t} \\ \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} + \nabla \times \nabla \times \mathbf{B} &= \frac{4\pi}{c} \nabla \times \mathbf{J}, \end{aligned}$$

sendo que a densidade de corrente cumpre aqui o papel de fonte. Pode-se também obter equações para os potenciais eletromagnéticos, as quais são

$$\begin{aligned} \nabla^2 \Phi + \frac{1}{c} \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) &= -4\pi\rho \\ \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} + \nabla \times \nabla \times \mathbf{A} + \frac{1}{c} \nabla \left(\frac{\partial \Phi}{\partial t} \right) &= \frac{4\pi}{c} \mathbf{J}. \end{aligned}$$

Nota-se que nas formas gerais acima, as equações para os potenciais são acopladas. Contudo, sobre os potenciais podem ser aplicadas **transformações de calibre (gauge)**

$$\begin{aligned} \mathbf{A}(\mathbf{r}, t) &\longrightarrow \mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) + \nabla\Lambda(\mathbf{r}, t) \\ \Phi(\mathbf{r}, t) &\longrightarrow \Phi'(\mathbf{r}, t) = \Phi(\mathbf{r}, t) - \frac{1}{c} \frac{\partial\Lambda}{\partial t}, \end{aligned}$$

sem que as equações de campo sejam alteradas. Nestas transformações, $\Lambda(\mathbf{r}, t)$ é um campo escalar arbitrário, denominado *função de calibre*. Há infinitas transformações de calibre possíveis, mas uma transformação de particular importância para a relatividade restrita é o *calibre de Lorenz*, proposto pelo físico-matemático dinamarquês Ludvig Valentin Lorenz (1829 – 1891). Neste calibre, o campo $\Lambda(\mathbf{r}, t)$ é escolhido de tal forma que a *condição de Lorenz*,

$$\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial\Phi}{\partial t} = 0, \quad (6.76)$$

é satisfeita. Com esta condição, as equações para os potenciais tornam-se desacopladas e formalmente idênticas às equações para os campos, isto é,

$$\frac{1}{c^2} \frac{\partial^2\Phi}{\partial t^2} - \nabla^2\Phi = 4\pi\rho \quad (6.77a)$$

$$\frac{1}{c^2} \frac{\partial^2\mathbf{A}}{\partial t^2} - \nabla^2\mathbf{A} = \frac{4\pi}{c} \mathbf{J} \quad (6.77b)$$

$$\frac{1}{c^2} \frac{\partial^2\Lambda}{\partial t^2} - \nabla^2\Lambda = 0, \quad (6.77c)$$

sendo que a última equação determina a função de calibre $\Lambda(\mathbf{r}, t)$. Esta equação foi obtida assumindo que a condição de Lorenz é também invariante frente a uma transformação de calibre.

Portanto, observa-se que no calibre de Lorenz tanto os campos quanto os potenciais, bem como a própria função de calibre, satisfazem equações diferenciais que possuem toda a estrutura formal de uma equação de onda. Este fato é relevante para a discussão a seguir, acerca da modificação que uma transformação de Galileu $K \rightarrow K'$ impõe sobre a equação da onda.

Já na descoberta do fenômeno da indução magnética, realizada em meados do ano de 1831 de forma independente pelo físico britânico Michael Faraday (1791 – 1867) e pelo físico americano Joseph Henry (1797 – 1878) e que foi formalizada pela lei de Faraday, que a discussão acerca da transformação dos campos eletromagnéticos frente a uma mudança de referencial é levada a cabo.

Na lei de Faraday, a variação do fluxo de indução magnética \mathbf{B} induz o surgimento do campo elétrico \mathbf{E} . Para a descoberta e determinação desta lei foi empregado um aparato experimental que consistia em um circuito elétrico que delimita uma área que é atravessada pelo fluxo magnético variável, fazendo surgir então uma força eletromotriz (fem) induzida. Esta fem surge porque a variação temporal do fluxo magnético faz surgir um campo elétrico ao longo do circuito. Com base neste tipo de equipamento que a lei de Faraday foi proposta. Contudo, esta lei somente é válida quando está envolvido o campo elétrico medido no *referencial de repouso do circuito*. Se o circuito está em movimento relativo ao equipamento de medida, então a lei de Faraday impõe uma transformação no campo elétrico devido ao movimento relativo entre o circuito e o laboratório.

Realizando uma análise cuidadosa da lei de Faraday, pode-se mostrar que se no referencial de laboratório (referencial K na figura 6.15) o campo elétrico é igual a \mathbf{E} e no referencial K' o campo é \mathbf{E}' , então estes campos relacionam-se por

$$\mathbf{E} = \mathbf{E}' - \frac{1}{c} (\mathbf{v} \times \mathbf{B}') \quad \text{e} \quad \mathbf{E}' = \mathbf{E} + \frac{1}{c} (\mathbf{v} \times \mathbf{B}). \quad (6.78)$$

Estas expressões são as leis de transformação de Galileu para o campo elétrico. Nota-se que não é feita menção sobre a possibilidade de transformação também do campo \mathbf{B} . De fato, neste contexto, $\mathbf{B}' = \mathbf{B}$.

Para se investigar a questão da invariância (ou não) das equações de Maxwell frente a uma transformação de Galileu, é necessário primeiro verificar como os operadores diferenciais ∇ e $\partial/\partial t$ se transformam neste contexto. Para auxiliar nesta derivação, será suposta a existência de um campo escalar $\psi = \psi(\mathbf{r}, t)$, o qual é um invariante de Galileu, ou seja, o valor do campo na posição P , ilustrada na figura 6.15, e no instante t , é o mesmo, quer este seja mensurado a partir

do referencial K , quer seja mensurado a partir do referencial K' . Se ψ for uma componente de um vetor, é necessário considerar separadamente a lei de transformação da mesma, além da transformação dos operadores diferenciais.

Para o operador gradiente, o valor de $\nabla\psi$ no ponto P , mensurado em relação ao referencial K , será tomado da seguinte maneira. Assumindo que o observador em K consiga mensurar simultaneamente os valores do campo no ponto P , na posição \mathbf{r} , e na posição $\mathbf{r}_2 = \mathbf{r} + \Delta\mathbf{r}$, suficientemente próxima ao ponto P , então a quantidade

$$\frac{\psi(x_1 + \Delta x_1, x_2, x_3, t) - \psi(x_1, x_2, x_3, t)}{\Delta x_1} \equiv \frac{\Delta_1 \psi}{\Delta x_1}$$

será igual, no limite $\Delta x_1 \rightarrow 0$, a $\partial\psi/\partial x_1$. O mesmo ocorrendo para as outras coordenadas. Já no referencial K' , também é possível fazer a mesma suposição. Então, pode-se definir a quantidade

$$\frac{\psi(x'_1 + \Delta x'_1, x'_2, x'_3, t) - \psi(x'_1, x'_2, x'_3, t)}{\Delta x'_1} \equiv \frac{\Delta'_1 \psi}{\Delta x'_1}.$$

Mas, de acordo com a transformação (6.72), $\Delta x'_1 = x'_{12} - x'_1 = x_{12} - x_1$, já que $t' = t$. Além disso como $dl' = dl$, resulta então que

$$\nabla' \psi = \nabla \psi \implies \nabla' = \nabla.$$

Por outro lado a derivada temporal é mais complicada. Assumindo que o ponto P está em repouso em relação a K , então realiza-se duas medições do valor do campo neste ponto, nos instantes t_1 e t_2 , definindo a quantidade

$$\frac{\psi(\mathbf{r}, t_2) - \psi(\mathbf{r}, t_1)}{t_2 - t_1} \equiv \frac{\Delta_t \psi}{\Delta t}.$$

Já em relação a K' , o ponto P está em movimento. Portanto, realizando duas medidas de ψ neste ponto nos instantes t'_1 e t'_2 , sendo que em relação a K' o ponto P está, respectivamente, nas posições \mathbf{r}'_1 e \mathbf{r}'_2 , pode-se definir a quantidade

$$\frac{\psi(\mathbf{r}'_2, t'_2) - \psi(\mathbf{r}'_1, t'_1)}{t'_2 - t'_1} \equiv \frac{\Delta'_t \psi}{\Delta t'}.$$

Se $\Delta t' = t'_2 - t'_1$ for pequeno o suficiente, pode-se realizar um desenvolvimento de ψ em série de Taylor, mantendo somente os termos de ordem mais baixa,

$$\psi(\mathbf{r}'_2, t'_2) \approx \psi(\mathbf{r}'_1, t'_1) + (t'_2 - t'_1) \frac{\partial \psi}{\partial t'_1} + (\mathbf{r}'_2 - \mathbf{r}'_1) \cdot \frac{\partial \psi}{\partial \mathbf{r}'_1}.$$

Então,

$$\frac{\Delta'_t \psi}{\Delta t'} \approx \frac{\partial \psi}{\partial t'_1} + \frac{\mathbf{r}'_2 - \mathbf{r}'_1}{t'_2 - t'_1} \cdot \frac{\partial \psi}{\partial \mathbf{r}'_1}.$$

Como ψ , Δt e $\Delta t'$ são invariantes de Galileu, então, chamando $t'_1 = t'$ e $\mathbf{r}'_1 = \mathbf{r}'$, no limite $\Delta t = \Delta t' \rightarrow 0$ obtém-se

$$\frac{\partial \psi}{\partial t} = \frac{\partial \psi}{\partial t'} - (\mathbf{v} \cdot \nabla') \psi.$$

Portanto, as leis de transformação de Galileu dos operadores diferenciais são

$$\begin{cases} \nabla = \nabla' \\ \frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - (\mathbf{v} \cdot \nabla') \end{cases} \iff \begin{cases} \nabla' = \nabla \\ \frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla). \end{cases} \quad (6.79)$$

Cabe mencionar aqui que o mesmo resultado poderia ser obtido caso fosse feita uma antecipação na definição de quadrivetores e se considerasse a quádrupla $\{x^\mu\} = (ct, \mathbf{r})$, ($\mu = 0, \dots, 3$). Neste caso, com a lei de transformação (6.72) e empregando a regra da cadeia, seria possível escrever

$$\frac{\partial}{\partial x^\mu} = \frac{\partial x'^\nu}{\partial x^\mu} \frac{\partial}{\partial x'^\nu} \implies \begin{cases} \frac{\partial}{\partial x^i} = \frac{\partial x'^0}{\partial x^i} \frac{\partial}{\partial x'^0} + \frac{\partial x'^j}{\partial x^i} \frac{\partial}{\partial x'^j} = \delta_i^j \frac{\partial}{\partial x'^j} = \frac{\partial}{\partial x'^i} \\ \frac{\partial}{\partial x^0} = \frac{\partial x'^0}{\partial x^0} \frac{\partial}{\partial x'^0} + \frac{\partial x'^i}{\partial x^0} \frac{\partial}{\partial x'^i} \Rightarrow \frac{1}{c} \frac{\partial}{\partial t} = \frac{1}{c} \frac{\partial}{\partial t'} - \frac{v_i}{c} \frac{\partial}{\partial x'^i}, \end{cases}$$

onde os índices latinos variam de 1 a 3.

Para verificar como as densidades de carga e corrente se transformam, considera-se um conjunto de cargas puntiformes. Neste caso,

$$\rho(\mathbf{r}, t) = \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i(t))$$

$$\mathbf{J}(\mathbf{r}, t) = \sum_i q_i \mathbf{u}_i(t) \delta(\mathbf{r} - \mathbf{r}_i(t)).$$

Então, de acordo com (6.72) e (6.73),

$$\rho'(\mathbf{r}', t') = \sum_i q_i \delta(\mathbf{r}' - \mathbf{r}'_i(t')) = \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i(t)) = \rho(\mathbf{r}, t)$$

$$\mathbf{J}'(\mathbf{r}', t') = \sum_i q_i \mathbf{u}'_i(t') \delta(\mathbf{r}' - \mathbf{r}'_i(t')) = \sum_i q_i (\mathbf{u}_i(t) - \mathbf{v}) \delta(\mathbf{r} - \mathbf{r}_i(t)) = \mathbf{J}(\mathbf{r}, t) - \mathbf{v}\rho(\mathbf{r}, t).$$

Ou seja,

$$\begin{cases} \rho'(\mathbf{r}', t') = \rho(\mathbf{r}, t) \\ \mathbf{J}'(\mathbf{r}', t') = \mathbf{J}(\mathbf{r}, t) - \mathbf{v}\rho(\mathbf{r}, t) \end{cases} \iff \begin{cases} \rho(\mathbf{r}, t) = \rho'(\mathbf{r}', t') \\ \mathbf{J}(\mathbf{r}, t) = \mathbf{J}'(\mathbf{r}', t') + \mathbf{v}\rho'(\mathbf{r}', t') \end{cases} \quad (6.80)$$

Pode-se então verificar a transformação das equações de Maxwell e da equação de continuidade. Supondo que estas equações assumem as suas formas conhecidas no referencial K , emprega-se as transformações (6.78), (6.79) e (6.80) inicialmente na equação da continuidade para verificar a sua forma no referencial K' :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \xrightarrow{K \rightarrow K'} \frac{\partial \rho'}{\partial t'} + \nabla' \cdot \mathbf{J}' + \nabla' \cdot (\mathbf{v}\rho') - (\mathbf{v} \cdot \nabla') \rho' = 0.$$

Como $\nabla' \cdot (\mathbf{v}\rho') = (\mathbf{v} \cdot \nabla') \rho'$, observa-se que a equação da continuidade é invariante frente a uma equação de Galileu. Fisicamente, isto significa que a transformação de Galileu respeita o princípio da conservação da carga elétrica.

Contudo, as equações de Maxwell tomam as seguintes formas frente a uma transformação de Galileu:

$$\begin{array}{ll} \nabla \cdot \mathbf{E} = 4\pi\rho & \xrightarrow{K \rightarrow K'} \nabla' \cdot \mathbf{E}' - \nabla' \cdot (\boldsymbol{\beta} \times \mathbf{B}') = 4\pi\rho' \\ \nabla \cdot \mathbf{B} = 0 & \xrightarrow{K \rightarrow K'} \nabla \cdot \mathbf{B}' = 0 \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J} & \xrightarrow{K \rightarrow K'} \nabla' \times \mathbf{B}' - \frac{1}{c} \frac{\partial}{\partial t'} (\mathbf{E}' - \boldsymbol{\beta} \times \mathbf{B}') \\ & + \frac{1}{c} (\mathbf{v} \cdot \nabla') (\mathbf{E}' - \boldsymbol{\beta} \times \mathbf{B}') = \frac{4\pi}{c} (\mathbf{J}' + \mathbf{v}\rho') \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 & \xrightarrow{K \rightarrow K'} \nabla' \times \mathbf{E}' + \frac{1}{c} \frac{\partial \mathbf{B}'}{\partial t'} \\ & - \nabla' \times (\boldsymbol{\beta} \times \mathbf{B}') - \frac{1}{c} (\mathbf{v} \cdot \nabla') \mathbf{B}' = 0, \end{array}$$

onde $\boldsymbol{\beta} = \mathbf{v}/c$. Na lei de Faraday, os últimos dois termos são nulos, uma vez que $\nabla' \times (\boldsymbol{\beta} \times \mathbf{B}') = -(\boldsymbol{\beta} \cdot \nabla') \mathbf{B}'$. Portanto, somente a lei de Gauss do magnetismo e a lei de Faraday são invariantes em forma. As leis de Gauss da eletricidade e de Ampère são modificadas pela transformação de Galileu. Isto implica também que as equações da onda eletromagnética também se transformam.

A conclusão é que do ponto de vista da transformação de Galileu, as equações de Maxwell somente são válidas em um referencial absoluto. Pode-se dizer que a transformação de Galileu implica na existência do éter.

6.15.1.3 A TRANSFORMAÇÃO DE LORENTZ E OS PRINCÍPIOS DA RELATIVIDADE RESTRITA

Em uma tentativa de contornar o conflito que surgiu entre as leis do eletromagnetismo e as leis de transformação entre referenciais, Hendrik Lorentz propôs, em 1985, uma nova lei de transformação que tratava o tempo não mais como um parâmetro livre nas equações de campo,

mas como uma nova coordenada que também é alterada na transformação de referenciais. Outras pessoas contribuíram na elaboração desta nova transformação, dentre as quais pode-se mencionar o físico-matemático alemão Woldemar Voigt (1850 – 1919), o físico-matemático britânico Joseph Larmor (1857 – 1942), Henri Poincaré, que consolidou o desenvolvimento matemático desta transformação e a denominou **transformação de Lorentz** e, finalmente, Albert Einstein, que em 1905 incorporou a mesma na sua teoria da relatividade restrita. Na literatura atual, pode-se encontrar diversas derivações distintas da transformação de Lorentz. Aqui, a mesma será apresentada sem a sua dedução.

Fazendo novamente referência à figura 6.15, considera-se dois referenciais inerciais K e K' , sendo que este último desloca-se com velocidade v constante em relação ao primeiro. Agora, porém, deve existir uma quarta coordenada, a coordenada temporal $x_0 = ct$, cujo eixo Cartesiano está orientado perpendicularmente ao hiper-plano 3D no qual os eixos espaciais $\{x_i\}$ estão localizados. Se as origens O e O' de ambos os referenciais coincidem em $t = t' = 0$, então a transformação entre as coordenadas de K e K' é dada por:

$$\begin{cases} ct' = \gamma(ct - \boldsymbol{\beta} \cdot \mathbf{r}) \\ \mathbf{r}' = \mathbf{r} + \frac{\gamma - 1}{\beta^2} (\boldsymbol{\beta} \cdot \mathbf{r}) \boldsymbol{\beta} - \gamma \boldsymbol{\beta} ct \end{cases} \iff \begin{cases} ct = \gamma(ct' + \boldsymbol{\beta} \cdot \mathbf{r}') \\ \mathbf{r} = \mathbf{r}' + \frac{\gamma - 1}{\beta^2} (\boldsymbol{\beta} \cdot \mathbf{r}') \boldsymbol{\beta} + \gamma \boldsymbol{\beta} ct', \end{cases} \quad (6.81a)$$

onde foram definidos os parâmetros *beta* ($\boldsymbol{\beta}$) e *gama* (γ) de Lorentz, respectivamente dados por

$$\boldsymbol{\beta} = \frac{\mathbf{v}}{c} \quad \text{e} \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}. \quad (6.81b)$$

Posteriormente, será demonstrado que as equações de Maxwell são invariantes em forma frente a transformação (6.81a).

Definindo as componentes paralela (r_{\parallel}) e perpendicular (r_{\perp}) do vetor posição em relação a $\boldsymbol{\beta}$,

$$\mathbf{r}_{\parallel} = \frac{\boldsymbol{\beta} \cdot \mathbf{r}}{\beta^2} \boldsymbol{\beta} \quad \text{e} \quad \mathbf{r}_{\perp} = \mathbf{r} - \mathbf{r}_{\parallel} = \frac{(\boldsymbol{\beta} \times \mathbf{r}) \times \boldsymbol{\beta}}{\beta^2},$$

a transformação de Lorentz pode também ser expressa de uma forma mais simples e intuitiva como

$$\begin{cases} ct' = \gamma(ct - \boldsymbol{\beta} \cdot \mathbf{r}) \\ r'_{\parallel} = \gamma(r_{\parallel} - \beta ct) \\ \mathbf{r}'_{\perp} = \mathbf{r}_{\perp} \end{cases} \iff \begin{cases} ct = \gamma(ct' + \boldsymbol{\beta} \cdot \mathbf{r}') \\ r_{\parallel} = \gamma(r'_{\parallel} + \beta ct') \\ \mathbf{r}_{\perp} = \mathbf{r}'_{\perp}. \end{cases} \quad (6.81c)$$

Com relação à teoria da relatividade restrita, esta será abordada de uma maneira sucinta. Como foi visto na seção anterior, as equações de Maxwell não são invariantes frente a uma transformação de Galileu. De acordo com estas, uma onda eletromagnética necessariamente requer um meio para se propagar, tal como ocorre com ondas mecânicas. Este meio seria o éter, o qual não havia sido detectado em experimentos específicos, tais como o experimento de Michelson-Morley. Embora acredita-se que Albert Einstein não estava ciente dos resultados destes experimentos em 1905, mesmo assim a hipótese do éter não era aceitável a partir de suas concepções individuais a respeito do espaço-tempo e do comportamento dos campos eletromagnéticos. Confrontado com o fato de que as equações de Maxwell não são invariantes frente a uma transformação de Galileu, Albert Einstein considerou as seguintes possibilidades:

1. As equações de Maxwell estavam incorretas. Uma teoria correta do eletromagnetismo deveria ser invariante frente a uma transformação de Galileu.
2. Existia um referencial privilegiado para o eletromagnetismo, no qual o éter está em repouso.
3. Existe um princípio da relatividade, distinto da relatividade galileana, e válido tanto para a mecânica quanto para o eletromagnetismo. Isto implicaria em alterações nas leis da mecânica.

A escolha realizada por Einstein a favor da possibilidade 3 levou-o a propor 2 postulados fundamentais:

Postulado 1 (Postulado da relatividade). As leis da natureza e os resultados decorrentes de quaisquer experimentos realizados em um dado sistema de referências são independentes do movimento de translação do sistema como um todo.

Postulado 2 (Postulado da constância da velocidade da luz). A velocidade da luz é constante e independente do movimento da fonte emissora.

Postulado 2' (alternativo). Em todo sistema inercial de referências, existe uma velocidade limite finita universal, denotada por c , para todas as entidades físicas.

Com base nestes postulados e assumindo outros princípios físicos tais como a isotropia do espaço livre, Einstein realizou sua própria derivação da transformação de Lorentz e demonstrou que as equações de Maxwell são invariantes em forma para estas transformações.

Fazendo referência novamente aos referenciais K e K' na figura 6.15, se no instante $t = t' = 0$, quando as origens O e O' coincidem, for emitido um pulso luminoso a partir da origem comum, então, de acordo com o postulado 2, as frentes de onda observadas por observadores situados nas origens de cada referencial irão se propagar na forma de ondas esféricas que se deslocam com a mesma velocidade, igual a c . Do ponto de vista do observador no referencial K , a frente de onda no instante t localiza-se no ponto (x_1, x_2, x_3) determinado pela equação

$$\sum_{i=1}^3 x_i^2 - c^2 t^2 = 0,$$

ao passo que a mesma frente de onda, vista pelo observador em K' , atinge o ponto (x'_1, x'_2, x'_3) no instante t' , determinado por

$$\sum_{i=1}^3 x_i'^2 - c'^2 t'^2 = 0.$$

Definindo uma nova coordenada em cada referencial como

$$x_0 = ict \quad \text{e} \quad x'_0 = ict',$$

os postulados da relatividade e o princípio de isotropia do espaço demandam então que

$$\sum_{\mu=0}^3 x_\mu^2 = \sum_{\mu=0}^3 x_\mu'^2.$$

Com base nesta expressão, a qual é análoga a uma rotação de eixos em um espaço de dimensão 4 que preserva a métrica do espaço,²⁰ Einstein derivou novamente a transformação (6.81).

6.15.1.4 O ESPAÇO-TEMPO DE MINKOWSKI E OS QUADRIVETORES

Com base nos trabalhos de Lorentz, Poincaré e Einstein, em 1907 o matemático alemão Hermann Minkowski (1864 – 1909) mostrou que a teoria da relatividade restrita poderia ser melhor formalizada assumindo que os fenômenos físicos ocorrem não em um espaço de dimensão 3, mas em um espaço de dimensão 4, doravante denominado o **espaço-tempo de Minkowski**, aqui representado por \mathbb{M}^4 . Neste espaço, uma transformação de Lorentz corresponde a uma rotação arbitrária dos eixos em torno da origem.

Einstein, que foi estudante de Minkowski no Instituto Eletrotécnico de Zurique, inicialmente considerou o trabalho de Minkowski como um simples artifício matemático, porém posteriormente percebeu que a interpretação geométrica do \mathbb{M}^4 seria necessária para a compreensão e desenvolvimento de sua posterior teoria da relatividade geral.

Formalmente, o espaço-tempo de Minkowski é um espaço vetorial real de dimensão 4. Este possui, portanto, 4 vetores de base, os quais não necessitam em geral ser considerados. Um vetor do \mathbb{M}^4 , portanto, possui 4 componentes que podem estar na forma contravariante ou covariante. Estes vetores do \mathbb{M}^4 são usualmente denominados **quadrivetores**. Cada quadrivector possui 3 componentes que correspondem às coordenadas espaciais usuais do \mathbb{E}^3 mais uma coordenada temporal igual a ct . Uma coordenada qualquer de um quadrivector, quer seja temporal ou espacial, será identificada por um índice grego (μ, ν, \dots) . Quando for necessário especificar explicitamente uma coordenada espacial, será empregado um índice latino (i, j, \dots) . Finalmente, quando for necessário especificar explicitamente a coordenada temporal, será empregado o índice “0”. O quadrivector posição no \mathbb{M}^4 tem os seus componentes escritos na forma contravariante

²⁰Porém agora em um espaço vetorial complexo.

definidos como uma quádrupla de números que pode ser representada de diferentes maneiras como:

$$(x^0, x^1, x^2, x^3) = (\{x^\mu\}) = (x^0, \{x^i\}) = (ct, \mathbf{r}), \quad \text{onde } \mu = 0, \dots, 3 \text{ e } i = 1, \dots, 3.$$

No espaço \mathbb{M}^4 está suposta a validade de uma transformação $x^\mu \rightarrow x'^\mu$ que leva as coordenadas $\{x^\mu\}$ a novas coordenadas

$$x'^\mu = x'^\mu(\{x^\nu\}), \quad (\mu = 0, \dots, 3),$$

de acordo com a transformação de Lorentz (6.81). Esta transformação pode ser escrita de uma forma compacta como

$$x'^\mu = \Lambda^\mu_\nu x^\nu, \quad (6.82a)$$

sendo que Λ^μ_ν é denominada a **matriz de transformação de Lorentz**. Obviamente, é válida também a transformação inversa

$$x^\mu = (\Lambda^{-1})^\mu_\nu x'^\nu, \quad (6.82b)$$

onde

$$\Lambda^\mu_\alpha (\Lambda^{-1})^\alpha_\nu = \delta^\mu_\nu, \quad \text{sendo } [\delta^\mu_\nu] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

a delta de Kronecker escrita na forma de um tensor de posto 2 misto. Comparando (6.82a,b) com (6.81), conclui-se que

$$\begin{aligned} \Lambda^0_0 &= \gamma & (\Lambda^{-1})^0_0 &= \gamma \\ \Lambda^0_i &= \Lambda^i_0 = -\gamma\beta_i & (\Lambda^{-1})^0_i &= (\Lambda^{-1})^i_0 = \gamma\beta_i \\ \Lambda^i_j &= \delta_{ij} + (\gamma - 1) \frac{\beta_i\beta_j}{\beta^2} & (\Lambda^{-1})^i_j &= \delta_{ij} + (\gamma - 1) \frac{\beta_i\beta_j}{\beta^2}. \end{aligned} \quad (6.82c)$$

Ou, na forma matricial,

$$\Lambda \equiv \begin{pmatrix} \gamma & -\gamma\beta_1 & -\gamma\beta_2 & -\gamma\beta_3 \\ -\gamma\beta_1 & 1 + (\gamma - 1) \frac{\beta_1^2}{\beta^2} & (\gamma - 1) \frac{\beta_1\beta_2}{\beta^2} & (\gamma - 1) \frac{\beta_1\beta_3}{\beta^2} \\ -\gamma\beta_2 & (\gamma - 1) \frac{\beta_1\beta_2}{\beta^2} & 1 + (\gamma - 1) \frac{\beta_2^2}{\beta^2} & (\gamma - 1) \frac{\beta_2\beta_3}{\beta^2} \\ -\gamma\beta_3 & (\gamma - 1) \frac{\beta_1\beta_3}{\beta^2} & (\gamma - 1) \frac{\beta_2\beta_3}{\beta^2} & 1 + (\gamma - 1) \frac{\beta_3^2}{\beta^2} \end{pmatrix} \quad (6.82d)$$

$$\Lambda^{-1} \equiv \begin{pmatrix} \gamma & \gamma\beta_1 & \gamma\beta_2 & \gamma\beta_3 \\ \gamma\beta_1 & 1 + (\gamma - 1) \frac{\beta_1^2}{\beta^2} & (\gamma - 1) \frac{\beta_1\beta_2}{\beta^2} & (\gamma - 1) \frac{\beta_1\beta_3}{\beta^2} \\ \gamma\beta_2 & (\gamma - 1) \frac{\beta_1\beta_2}{\beta^2} & 1 + (\gamma - 1) \frac{\beta_2^2}{\beta^2} & (\gamma - 1) \frac{\beta_2\beta_3}{\beta^2} \\ \gamma\beta_3 & (\gamma - 1) \frac{\beta_1\beta_3}{\beta^2} & (\gamma - 1) \frac{\beta_2\beta_3}{\beta^2} & 1 + (\gamma - 1) \frac{\beta_3^2}{\beta^2} \end{pmatrix}. \quad (6.82e)$$

De acordo com os postulados da relatividade, dados dois referenciais inerciais K e K' , dois eventos quaisquer sempre ocorrem com separações espaço-temporais dadas respectivamente por $\{dx^\mu\}$ e $\{dx'^\mu\}$, de tal forma que a quantidade

$$ds^2 \doteq (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = ds'^2 \equiv (dx'^0)^2 - (dx'^1)^2 - (dx'^2)^2 - (dx'^3)^2 \quad (6.83a)$$

permanece sempre invariante frente a uma transformação de Lorentz.

A quantidade ds^2 definida em (6.83a) é denominada um **intervalo infinitesimal** entre eventos no espaço-tempo de Minkowski, ou, simplesmente, a **métrica de Minkowski**. Nota-se que ds^2 não é positivo-definido, pois em princípio sempre é possível encontrar uma transformação tal que $ds^2 < 0$.²¹

Desta maneira, é possível definir o *tensor de métrica* do \mathbb{M}^4 a partir do intervalo ds^2 , de tal forma que

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \quad (6.83b)$$

²¹Porém, esta transformação, denominada *tipo-espaço*, viola a causalidade implícita nos postulados da relatividade. As outras transformações possíveis são *tipo-tempo*, para a qual $ds^2 > 0$ e *tipo-luz*, para a qual $ds^2 = 0$ (para $x^\mu \neq 0$).

onde

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \text{ sendo que } g = \det(g) = -1. \quad (6.83c)$$

O tensor de métrica (6.83b,c) mostra que o espaço \mathbb{M}^4 é plano; porém, de acordo com a definição 6.9, é um tipo de espaço pseudo-Riemanniano, os quais são assim denominados justamente por apresentarem sua métrica dada por expressões como (6.83a). Em tais espaços, é comum caracterizar-se o tensor de métrica também por sua *assinatura*, a qual é o conjunto de números indicando a quantidade de autovalores positivos, negativos e nulos do tensor de métrica. No caso do \mathbb{M}^4 , a assinatura do tensor de métrica é indicada por $\{+, -, -, -\}$.

Propriedades matemáticas do tensor de métrica do \mathbb{M}^4 :

1. $g_{\mu\nu} = g_{\nu\mu}$ (simétrico).
2. $g^{\mu\nu} = g_{\mu\nu}$ (forma contravariante).
3. $g^{\mu\alpha} g_{\alpha\nu} = g^{\mu}_{\nu} = \delta^{\mu}_{\nu}$ (forma mista).
4. $\text{Tr}(g) = -2$.

Não será realizada aqui uma discussão ampla das consequências físicas das transformações de Lorentz e dos postulados da relatividade. Discussões a respeito do conceito de simultaneidade, contração espacial, adição de velocidades e cone de luz são referidas a textos específicos sobre relatividade. O único conceito relevante para a presente discussão é do **tempo próprio**.

Considera-se uma partícula com velocidade instantânea $u(t)$ em relação ao referencial K . Em um intervalo de tempo dt , sua posição muda por $dr = u dt$. De (6.83a), o elemento de arco no espaço-tempo percorrido pela partícula é

$$ds^2 = c^2 dt^2 - |dr|^2 = c^2 dt^2 (1 - \beta_u^2) \text{ em } K,$$

sendo $\beta_u = u/c$ o fator beta da velocidade instantânea da partícula em K .

Dado agora o referencial K' onde a partícula está instantaneamente em repouso, como $dr' = 0$ em K' , a partícula percorre o elemento de arco ds'^2 somente ao longo da coordenada temporal, ou seja,

$$ds'^2 = c^2 dt'^2 \equiv c^2 d\tau^2 \text{ em } K'.$$

Como este elemento de arco é um invariante frente a transformação de Lorentz, isto é, $ds'^2 = ds^2$, isto implica que

$$d\tau = \sqrt{1 - \beta_u^2(t)} dt = \frac{dt}{\gamma_u(t)}, \text{ sendo } \gamma_u(t) = (1 - \beta_u^2(t))^{-1/2}.$$

A outra implicação é que a quantidade $d\tau$ também é um invariante de Lorentz. Esta quantidade é o elemento de tempo próprio da partícula, ou seja, o intervalo infinitesimal de tempo mensurado no referencial instantaneamente em repouso com a mesma.

Se for possível resolver a equação de movimento da partícula no referencial K entre os instantes t_1 e t_2 , então o intervalo de tempo próprio transcorrido no referencial em repouso com a partícula será dado por

$$\Delta\tau = \int_{t_1}^{t_2} \frac{dt}{\gamma_u(t)}.$$

Este resultado mostra também que $\Delta\tau \leq \Delta t = t_2 - t_1$, uma vez que $\gamma_u \geq 1$ sempre. Ou seja, em qualquer referencial, o intervalo de tempo medido para um determinado processo físico será sempre maior ou igual que o tempo transcorrido no referencial em repouso com a partícula. Este fenômeno é denominado de *dilatação temporal*.

Dado o tensor de métrica na teoria da relatividade restrita, pode-se agora definir os seguintes objetos que compõe o espaço-tempo de Minkowski:

Tensores de posto zero. Também denominados escalares ou **invariantes de Lorentz**.

Tensores de posto um. Também denominados **quadrivetores**. Como os componentes de um quadrivetor podem estar na forma contravariante ou covariante, será empregada uma notação própria que indica explicitamente qual é a forma dos mesmos.

Vetores contravariantes. Tomando como exemplo o quadrivetor posição, um quadrivetor com componentes contravariantes possui uma componente temporal a^0 e três componentes espaciais $\mathbf{a} = (a^1, a^2, a^3)$, sendo este denotado de forma genérica como

$$\tilde{\mathbf{a}} \equiv (a^0, \mathbf{a}) = (a^0, a^1, a^2, a^3).$$

As componentes $a^\mu = a^\mu(\tilde{\mathbf{r}})$ transformam-se de acordo com (6.41) e (6.82a) como

$$a'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} a^\nu = \Lambda^\mu{}_\nu a^\nu, \quad (6.84a)$$

ou, de forma explícita,

$$a'^0 = \Lambda^0{}_\nu a^\nu = \Lambda^0{}_0 a^0 + \Lambda^0{}_i a^i = \gamma a^0 - \gamma \boldsymbol{\beta} \cdot \mathbf{a} \quad (6.84b)$$

$$a'^i = \Lambda^i{}_\nu a^\nu = \Lambda^i{}_0 a^0 + \Lambda^i{}_j a^j = a^i - \gamma a^0 \beta_i + (\gamma - 1) \frac{\boldsymbol{\beta} \cdot \mathbf{a}}{\beta^2} \beta_i. \quad (6.84c)$$

A transformação inversa, de acordo com (6.82b), é

$$a^\mu = \frac{\partial x^\mu}{\partial x'^\nu} a'^\nu = (\Lambda^{-1})^\mu{}_\nu a'^\nu. \quad (6.84d)$$

Vetores covariantes. Um quadrivetor com componentes covariantes possui uma componente temporal a_0 e três componentes espaciais $(\{a_i\})$. Usando a propriedade de rebaiamento de índice do tensor de métrica, percebe-se facilmente, com o uso de (6.83c) que a relação $a_\mu = g_{\mu\nu} a^\nu$ implica que

$$a_0 = a^0 \quad \text{e} \quad a_i = -a^i.$$

Este quadrivetor com componentes covariantes é denotado de forma genérica como

$$\underline{\mathbf{a}} = (a_0, a_1, a_2, a_3) = (a_0, -\mathbf{a}).$$

Suas componentes $a_\mu = a_\mu(\tilde{\mathbf{r}})$ transformam-se, de acordo com (6.41) e (6.82b) como

$$a'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} a_\nu = (\Lambda^{-1})^\nu{}_\mu a_\nu, \quad (6.85a)$$

$$a'_0 = (\Lambda^{-1})^\nu{}_0 a_\nu = \gamma a_0 - \gamma \boldsymbol{\beta} \cdot \mathbf{a}$$

$$a'_i = (\Lambda^{-1})^\nu{}_i a_\nu = a_i + \gamma a_0 \beta_i - (\gamma - 1) \frac{\boldsymbol{\beta} \cdot \mathbf{a}}{\beta^2} \beta_i.$$

A transformação inversa, de acordo com (6.82a), é

$$a_\mu = \frac{\partial x'^\nu}{\partial x^\mu} a'_\nu = (\Lambda^{-1})^\nu{}_\mu a'_\nu. \quad (6.85b)$$

Da mesma forma como ocorre com o quadrivetor covariante, a forma contravariante pode ser obtida a partir da primeira usando o tensor de métrica: $a^\mu = g^{\mu\nu} a_\nu$.

Tensores de posto dois. Podem ter seus componentes em três formas: contravariantes, covariantes ou mistos.

Tensor contravariante. Trata-se de um tensor $F^{\mu\nu}$ com 16 componentes que se transforma de acordo com as regras dadas por (6.41) e (6.82a):

$$F'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} F^{\alpha\beta} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta}. \quad (6.86a)$$

Tensor covariante. Trata-se de um tensor $F_{\mu\nu}$ com 16 componentes que se transforma de acordo com (6.41) e (6.82b):

$$F'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} F_{\alpha\beta} = (\Lambda^{-1})^\alpha{}_\mu (\Lambda^{-1})^\beta{}_\nu F_{\alpha\beta}. \quad (6.86b)$$

Tensor misto. Uma das possíveis formas é o tensor F^μ_ν com 16 componentes que se transforma como:

$$F'^\mu_\nu = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x'^\nu} F^\alpha_\beta = \Lambda^\mu_\alpha (\Lambda^{-1})^\beta_\nu F^\alpha_\beta. \quad (6.86c)$$

As diferentes formas do tensor \mathcal{F} podem ser obtidas usando a propriedade de elevação ou rebaixamento de índices do tensor de métrica. Por exemplo,

$$F^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta}, \quad F_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} F^{\alpha\beta}, \quad F^\mu_\nu = g^{\mu\alpha} F_{\alpha\nu}.$$

O produto interno no espaço-tempo de Minkowski possui uma definição ligeiramente distinta do mesmo produto em um espaço Euclidiano. Dados os quadrivetores pertencentes ao \mathbb{M}^4 , o produto interno entre quaisquer dois vetores corresponde ao mapeamento $\mathbb{M}^4 \times \mathbb{M}^4 \mapsto \mathbb{R}$ que satisfaz as seguintes propriedades:

1. **Simetria.** Dados os quadrivetores \tilde{a} e \tilde{b} , o seu produto interno é

$$a^\mu b_\mu = a_\mu b^\mu = a^0 b_0 - \mathbf{a} \cdot \mathbf{b}.$$

2. **Bilinearidade.** Dados \tilde{a} , \tilde{b} e $\tilde{v} \in \mathbb{M}^4$ e $d \in \mathbb{R}$, então

$$(da^\mu + b^\mu) v_\mu = da^\mu v_\mu + b^\mu v_\mu.$$

3. Dados \tilde{a} e $\tilde{b} \in \mathbb{M}^4$, se $a^\mu b_\mu = 0$ para qualquer $\tilde{b} \in \mathbb{M}^4$, então, necessariamente, $\tilde{a} = 0$.

Dado o quadrivetor $\tilde{a} \in \mathbb{M}^4$, a sua *magnitude* é definida por

$$\|\tilde{a}\|^2 = \|\underline{a}\|^2 = a^\mu a_\mu = a^0 a_0 - a^2,$$

sendo a^2 a norma Euclidiana de seus componentes espaciais. Nota-se que, ao contrário do que ocorre em um espaço Euclidiano, a magnitude de um quadrivetor não é necessariamente positivo-definida, uma vez que o espaço de Mikowski é pseudo-Euclidiano.²²

Como era esperado, o produto interno de dois quadrivetores é um invariante de Lorentz, isto é, dados \tilde{a} e \tilde{b} , de acordo com (6.84) e (6.85),

$$a'^\mu b'_\mu = a'_\mu b'^\mu = \Lambda^\mu_\alpha (\Lambda^{-1})^\beta_\mu a^\alpha b_\beta = \delta^\beta_\alpha a^\alpha b_\beta = a^\alpha b_\alpha.$$

Se os vetores \tilde{a} e \tilde{b} forem ambos não nulos, mas $a^\mu b_\mu = a_\mu b^\mu = 0$, então estes são ditos serem (algebricamente) *ortogonais*.

Cabe também discutir a maneira como o operador diferencial $\partial/\partial x^\mu$ se transforma. Dadas as transformações $x'^\mu = x'^\mu(\tilde{r})$ e $x^\mu = x^\mu(\tilde{r}')$, pode-se usar a regra da cadeia e escrever

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu_\mu \frac{\partial}{\partial x^\nu}.$$

Comparando a expressão acima com (6.85a), percebe-se que a diferenciação frente a uma coordenada *contravariante* transforma-se como um quadrivetor *covariante*. Pode-se então usar a seguinte notação para o operador quadrivetorial gradiente na forma covariante:

$$\underline{\partial} \equiv (\{\partial_\mu\}) = \left(\frac{\partial}{\partial x^0}, \left\{ \frac{\partial}{\partial x^i} \right\} \right) = \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right) = \left(\frac{\partial}{\partial x^0}, \nabla \right), \quad (6.87a)$$

onde ∇ é o operador gradiente usual do \mathbb{E}^3 .

Agora, como $x^\mu = g^{\mu\nu} x_\nu$, pode-se escrever

$$\frac{\partial}{\partial x_\mu} = \frac{\partial x^\nu}{\partial x_\mu} \frac{\partial}{\partial x^\nu} = g^{\nu\alpha} \delta^\mu_\alpha \frac{\partial}{\partial x^\nu} = g^{\mu\nu} \frac{\partial}{\partial x^\nu} = g^{\mu\nu} \partial_\nu,$$

o que mostra que a diferenciação frente a uma coordenada *covariante* transforma-se como um quadrivetor *contravariante*. Especificamente,

$$\frac{\partial}{\partial x_0} = g^{0\nu} \frac{\partial}{\partial x^\nu} = \frac{\partial}{\partial x^0} = \partial_0 \quad \text{e} \quad \frac{\partial}{\partial x_i} = g^{i\nu} \frac{\partial}{\partial x^\nu} = -\frac{\partial}{\partial x^i} = -\partial_i.$$

²²Ver discussão na seção 4.8.1.

Portanto, definindo o operador quadrivetorial gradiente na forma contravariante, este resulta igual a

$$\tilde{\partial} \equiv (\{\partial^\mu\}) = (\partial^0, \{\partial^i\}) = \left(\frac{\partial}{\partial x_0}, \left\{ \frac{\partial}{\partial x_i} \right\} \right) = \left(\frac{\partial}{\partial x_0}, -\nabla \right). \quad (6.87b)$$

A *quadrividivergência* de um quadrivetor \tilde{a} é o invariante de Lorentz

$$\partial^\mu a_\mu = \partial_\mu a^\mu = \frac{\partial a^0}{\partial x^0} + \nabla \cdot \mathbf{a}.$$

O operador laplaciano no \mathbb{M}^4 é a contração invariante

$$\square \equiv \partial^\mu \partial_\mu = \partial_\mu \partial^\mu = \frac{\partial^2}{\partial (x^0)^2} - \nabla^2,$$

sendo que o operador “ \square ” é denominado o *d’Alembertiano*.

6.15.1.5 FORMULAÇÃO COVARIANTE DO ELETROMAGNETISMO CLÁSSICO

A invariância de forma ou *covariância* das equações de Maxwell, bem como das leis físicas relacionadas, tais como a equação de continuidade e a força de Lorentz, foi demonstrada por Lorentz e por Poincaré. A covariância implica que as grandezas $\rho(\mathbf{r}, t)$, $\mathbf{J}(\mathbf{r}, t)$, $\mathbf{E}(\mathbf{r}, t)$ e $\mathbf{B}(\mathbf{r}, t)$ transformam-se de uma maneira bem definida pelas transformações de Lorentz.

Inicia-se a derivação das formas covariantes das leis do eletromagnetismo pela equação da continuidade (6.74b), a qual é assumida válida em um determinado referencial K . Inicialmente, é necessário postular uma forma para o quadrivetor densidade de corrente elétrica $\tilde{\mathbf{J}}$. Escrevendo-se os operadores diferenciais em termos dos componentes do quadrivetor posição contravariante $\tilde{\mathbf{r}}$,

$$\frac{\partial (c\rho)}{\partial x^0} + \sum_{i=1}^3 \frac{\partial J_i}{\partial x^i} = 0.$$

Percebe-se então que se for definido o quadrivetor

$$\tilde{\mathbf{J}} \equiv (J^0, \{J^i\}) = (c\rho, \mathbf{J}),$$

pode-se escrever a equação da continuidade, de acordo com (6.87a), como

$$\partial_\mu J^\mu = 0 \quad (\text{em } K).$$

Mostra-se facilmente que a equação acima é invariante em forma frente a uma transformação de Lorentz, ou seja, na mudança $K \rightarrow K'$, pode-se mostrar que esta equação se torna igual a

$$\partial'_\mu J'^\mu = 0 \quad (\text{em } K').$$

É interessante também mostrar explicitamente a transformação das componentes de $\tilde{\mathbf{J}}$ para o referencial K' . De (6.84b,c), obtém-se

$$\begin{aligned} \rho' &= \gamma \left(\rho - \frac{1}{c} \boldsymbol{\beta} \cdot \mathbf{J} \right) \\ \mathbf{J}' &= \mathbf{J} - \gamma c \rho \boldsymbol{\beta} + (\gamma - 1) \frac{\boldsymbol{\beta} \cdot \mathbf{J}}{\beta^2} \boldsymbol{\beta}, \end{aligned}$$

a qual pode ser comparada diretamente com a transformação de Galileu (6.80) para estas quantidades. É digno de menção que a densidade de carga é mensurada de forma distinta entre os referenciais K e K' .

Agora, para se derivar a forma covariante das equações de Maxwell (6.74a), a abordagem tradicionalmente adotada parte dos potenciais eletrodinâmicos. É também conveniente adotar-se o calibre de Lorenz nesta derivação, dentro do qual os potenciais $\Phi(\mathbf{r}, t)$ e $\mathbf{A}(\mathbf{r}, t)$ são determinados pela condição de Lorenz (6.76) e pelas equações (6.77a,b). Como antes, assume-se que estas equações são respeitadas em um determinado referencial K .

Observa-se então ser possível escrever as equações para os potenciais como

$$\frac{\partial^2 \Phi}{\partial (x^0)^2} - \sum_{i=1}^3 \frac{\partial^2 \Phi}{\partial (x^i)^2} = \frac{4\pi}{c} (c\rho)$$

$$\frac{\partial^2 \mathbf{A}}{\partial (x^0)^2} - \sum_{i=1}^3 \frac{\partial^2 \mathbf{A}}{\partial (x^i)^2} = \frac{4\pi}{c} \mathbf{J}$$

$$\frac{\partial \Phi}{\partial x^0} + \sum_{i=1}^3 \frac{\partial A_i}{\partial x^i} = 0.$$

Dada a definição acima para $\tilde{\mathbf{J}}$, se for definido o quadrivetor potencial

$$\tilde{\mathbf{A}} \equiv (A^0, \{A^i\}) = (\Phi, \mathbf{A}),$$

as 3 equações acima são escritas nas suas formas covariantes como

$$\square A^\mu = \frac{4\pi}{c} J^\mu, \quad (\mu = 0, \dots, 3)$$

$$\partial_\mu A^\mu = 0.$$

Uma vez definidos o quadrivetor potencial e as equações que o mesmo obedece, considera-se agora a relação dos mesmos com o campo. Esta relação é fornecida pelas expressões (6.75). Escrevendo as mesmas para a i -ésima componente dos campos, ainda empregando a notação do \mathbb{E}^3 ,

$$E_i = -\frac{\partial \Phi}{\partial x_i} - \frac{1}{c} \frac{\partial A_i}{\partial t} \quad \text{e} \quad B_i = \sum_{j,k=1}^3 \epsilon_{ijk} \frac{\partial A_k}{\partial x_j}.$$

Reescrevendo agora estas expressões em termos de derivações das componentes *contravariantes* do quadrivetor potencial $\tilde{\mathbf{A}}$ em relação aos componentes do quadrivetor posição *covariante* $\underline{r} = (x_0, -\mathbf{r})$, obtém-se

$$E_i = \frac{\partial A^0}{\partial x_i} - \frac{\partial A^i}{\partial x_0} = -(\partial^0 A^i - \partial^i A^0)$$

$$B_i = -\epsilon_{ijk} \frac{\partial A^k}{\partial x_j} = -\epsilon_{ijk} \partial^j A^k \implies B_i = -(\partial^j A^k - \partial^k A^j),$$

sendo que na última expressão deve-se sempre realizar uma permutação cíclica de $\{i, j, k\} = \{1, 2, 3\}$. Ainda na mesma expressão,

$$B_i = -\epsilon_{ijk} \partial^j A^k \implies \partial^i A^j - \partial^j A^i = -\epsilon^{ijk} B_k.$$

Percebe-se então que não é possível estas expressões em um único quadrivetor. A maneira correta de agrupar os campos é na forma de um tensor de posto dois denominado **tensor do campo eletromagnético**, o qual, na forma contravariante $F^{\mu\nu}$ é definido como

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu.$$

Identificando os componentes de $F^{\mu\nu}$:

$$F^{\mu\mu} \stackrel{\text{NS}}{=} 0$$

$$F^{0i} = \partial^0 A^i - \partial^i A^0 = -E_i$$

$$F^{i0} = \partial^i A^0 - \partial^0 A^i = E_i$$

$$F^{ij} = \partial^i A^j - \partial^j A^i = -\epsilon^{ijk} B_k.$$

Ou seja, o tensor de campo é antissimétrico. Pode-se escrever o mesmo forma matricial como

$$[F^{\mu\nu}] \equiv \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_3 & -B_2 & B_1 & 0 \end{pmatrix}.$$

O tensor de campo eletromagnético na forma covariante é obtido por

$$F_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} F^{\alpha\beta},$$

resultando na forma matricial

$$[F_{\mu\nu}] \equiv \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}.$$

Uma outra forma útil para o tensor de campo é a sua forma dual, a qual é obtida de acordo com a definição de um tensor dual, apresentada na seção 6.6.2. Para tanto, inicialmente define-se o símbolo de Levi-Civita de quarta ordem $\epsilon^{\alpha\beta\gamma\delta}$ como

$$\begin{aligned} \epsilon^{\alpha\beta\gamma\delta} &= \begin{cases} +1; & \text{permutações pares de } \{0, 1, 2, 3\} \\ -1; & \text{permutações ímpares de } \{0, 1, 2, 3\} \\ 0; & \text{outras combinações} \end{cases} \\ &= \frac{1}{12} [(\alpha - \beta)(\alpha - \gamma)(\alpha - \delta)(\beta - \gamma)(\beta - \delta)(\gamma - \delta)]. \end{aligned}$$

Da mesma forma como ocorre com o símbolo de terceira ordem, o qual está relacionado com o cálculo do determinante de uma matriz 3×3 , o símbolo $\epsilon^{\alpha\beta\gamma\delta}$ pode ser empregado no cálculo do correspondente determinante de uma matriz 4×4 ,

$$\epsilon^{\alpha\beta\gamma\delta} A = \epsilon^{\mu\nu\kappa\theta} A_{\mu}^{\alpha} A_{\nu}^{\beta} A_{\kappa}^{\gamma} A_{\theta}^{\delta},$$

como ocorre nas expressões (6.33). Também como ocorre com o símbolo de terceira ordem, $\epsilon^{\alpha\beta\gamma\delta}$ é um tensor relativo de peso $w = +1$. Contudo, como no caso das transformações de Lorentz, $\Lambda = \det(\Lambda) = +1$, o símbolo de Levi-Civita se transforma como um tensor de posto quatro na transformação entre referenciais.

Portanto, o tensor dual do campo eletromagnético, definido como

$$\mathcal{F}^{\mu\nu} = \frac{1}{2} \epsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}$$

age como um tensor frente a uma transformação de Lorentz. Na representação matricial, este tensor é dado por

$$[\mathcal{F}^{\mu\nu}] = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & E_3 & -E_2 \\ B_2 & -E_3 & 0 & E_1 \\ B_3 & E_2 & -E_1 & 0 \end{pmatrix}.$$

Finalmente, em termos dos tensores de campo $F^{\mu\nu}$, $F_{\mu\nu}$ e $\mathcal{F}^{\mu\nu}$ e do quadrivetor \tilde{J} , as equações de Maxwell (6.74a) podem ser escritas de uma forma manifestamente covariante. Principiando pelas equações inhomogêneas (lei de Gauss da eletricidade e lei de Ampère), a forma covariante das mesmas é

$$\partial_{\mu} F^{\mu\nu} = \frac{4\pi}{c} J^{\nu}.$$

Por sua vez, as equações homogêneas (lei de Gauss do magnetismo e a lei de Faraday) têm a sua forma covariante de duas maneiras equivalentes,

$$\partial_{\mu} \mathcal{F}^{\mu\nu} = 0,$$

ou

$$\partial_{\mu} F_{\nu\sigma} + \partial_{\nu} F_{\sigma\mu} + \partial_{\sigma} F_{\mu\nu} = 0,$$

onde μ, ν e σ são quaisquer combinações dos índices $\{0, 1, 2, 3\}$.

Um ponto importante consiste em mostrar a lei de transformação dos campos. Dado o tensor de campo $F^{\mu\nu}$ no referencial K , um observador no referencial K' irá observar estes campos de acordo com (6.86a),

$$F'^{\mu\nu} = \Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} F^{\alpha\beta}.$$

Realizando a transformação, os campos observados no referencial K' são

$$\begin{aligned} \mathbf{E}' &= \gamma(\mathbf{E} + \boldsymbol{\beta} \times \mathbf{B}) - \frac{\gamma^2}{\gamma + 1} (\boldsymbol{\beta} \cdot \mathbf{E}) \boldsymbol{\beta} \\ \mathbf{B}' &= \gamma(\mathbf{B} - \boldsymbol{\beta} \times \mathbf{E}) - \frac{\gamma^2}{\gamma + 1} (\boldsymbol{\beta} \cdot \mathbf{B}) \boldsymbol{\beta}. \end{aligned}$$

Exercício 6.21. Obtenha os campos transformados para o referencial K' .

Solução. Realiza-se a transformação inicialmente para o campo elétrico. A i -ésima componente em K' é $E'^i = F'^{i0}$. Então,

$$\begin{aligned} F'^{i0} &= \Lambda^i_{\alpha} \Lambda^0_{\beta} F^{\alpha\beta} \\ &= \Lambda^i_{\alpha} \Lambda^0_{\beta} F^{\alpha\beta} \\ &= \Lambda^i_0 \Lambda^0_{\beta} F^{0\beta} + \Lambda^i_j \Lambda^0_{\beta} F^{j\beta} \\ &= \Lambda^i_0 \Lambda^0_j F^{0j} + \Lambda^i_j \Lambda^0_0 F^{j0} + \Lambda^i_j \Lambda^0_k F^{jk} \\ &= \gamma E_i + \gamma (\boldsymbol{\beta} \times \mathbf{B})_i - \frac{\gamma^2}{\gamma + 1} (\boldsymbol{\beta} \cdot \mathbf{E}) \beta_i. \end{aligned}$$

Já os componentes de \mathbf{B}' são dados a partir de

$$F^{ij} = -\epsilon^{ij\ell} B_{\ell} \implies \epsilon_{ijm} F^{ij} = -\epsilon_{ijm} \epsilon^{ij\ell} B_{\ell} = -2\delta_m^{\ell} B_{\ell} = -2B_m \implies B_i = -\frac{1}{2} \epsilon_{ijk} F^{jk}.$$

Portanto, no referencial K' ,

$$\begin{aligned} B'_i &= -\frac{1}{2} \epsilon_{ijk} F'^{jk} \\ &= -\frac{1}{2} \epsilon_{ijk} \Lambda^j_{\alpha} \Lambda^k_{\beta} F^{\alpha\beta} \\ &= -\frac{1}{2} \epsilon_{ijk} \left[\Lambda^j_0 \Lambda^k_m F^{0m} + \Lambda^j_{\ell} \Lambda^k_0 F^{\ell 0} + \Lambda^j_{\ell} \Lambda^k_m F^{\ell m} \right] \\ &= \gamma B_i - (\gamma - 1) \frac{\boldsymbol{\beta} \cdot \mathbf{B}}{\beta^2} \beta_i - \gamma (\boldsymbol{\beta} \times \mathbf{E})_i \\ &= \gamma B_i - \gamma (\boldsymbol{\beta} \times \mathbf{E})_i - \frac{\gamma^2}{\gamma + 1} (\boldsymbol{\beta} \cdot \mathbf{B}) \beta_i. \end{aligned}$$

Uma expressão interessante também pode ser obtida para a equação de movimento de uma partícula carregada sob a ação dos campos eletromagnéticos. Na mecânica Newtoniana, esta equação é

$$\frac{d\mathbf{p}}{dt} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right),$$

sendo \mathbf{p} o momentum linear (Newtoniano) da partícula e q a sua carga. Para derivar a forma covariante desta equação, inicialmente define-se o quadrivetor velocidade a partir de $\tilde{\mathbf{r}}$ e do tempo próprio τ como

$$\tilde{\mathbf{U}} = \frac{d\tilde{\mathbf{r}}}{d\tau} = (\gamma_u c, \gamma_u \mathbf{v}),$$

sendo $\tilde{\mathbf{U}}$ a quadrivelocidade. A partir desta, pode-se definir o quadrivetor momentum linear como $\tilde{\mathbf{p}} = m_0 \tilde{\mathbf{U}} = (E/c, \gamma_u \mathbf{p})$, sendo m_0 a *massa de repouso* da partícula, ou seja, a massa que seria mensurada em um referencial sempre em repouso com a mesma e

$$E = \gamma_u m_0 c^2 = \sqrt{(m_0 c^2)^2 + p^2 c^2}$$

a sua *energia relativística*. Esta quantidade também é um invariante de Lorentz. Portanto, a forma covariante da equação de movimento da carga q é

$$\frac{dp^{\mu}}{d\tau} = m_0 \frac{dU^{\mu}}{d\tau} = \frac{q}{c} F^{\mu\nu} U_{\nu}.$$

É interessante separar as componentes temporal e espacial desta equação. Estas são, respectivamente,

$$\begin{aligned} \frac{dp^0}{d\tau} &= \frac{1}{c} \frac{dE}{d\tau} = \frac{q}{c} \gamma_u (\mathbf{E} \cdot \mathbf{u}) \\ \frac{dp^i}{d\tau} &= \frac{d}{d\tau} (\gamma_u p_i) = q \gamma_u \left[E_i + \frac{1}{c} (\mathbf{u} \times \mathbf{B})_i \right]. \end{aligned}$$

Das equações acima, a última corresponde à forma relativística da força de Lorentz, enquanto que a primeira descreve a taxa de transferência de energia entre a carga e os campos.

Por fim, quando há um sistema de partículas carregadas (caracterizadas pelo quadrivetor \tilde{J}) interagindo com os campos, observa-se a conservação da energia, do momentum linear e do momentum angular totais deste sistema. A primeira lei de conservação em particular é denominada *teorema de Poynting*. A lei de conservação geral possui dois termos, um termo correspondente às ondas e outro contendo a interação destas com as partículas. O primeiro termo é descrito em função do *tensor energia-momentum* $\Theta^{\mu\nu}$, dado por

$$\Theta^{\mu\nu} = \frac{1}{4\pi} \left(g^{\mu\sigma} F_{\sigma\alpha} F^{\alpha\nu} + \frac{1}{4} g^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right).$$

Pode-se perceber que $\Theta^{\mu\nu}$ é simétrico.

A lei de conservação para o sistema cargas + campos é

$$\partial_\mu \Theta^{\mu\nu} = -\frac{1}{c} F^{\nu\sigma} J_\sigma.$$

6.15.2 A MÉTRICA DE SCHWARZSCHILD NA RELATIVIDADE GERAL

Uma solução exata para as equações de Einstein (6.71) foi encontrada por Karl Schwarzschild (1873 - 1916) ainda em 1915, menos de dois meses após a primeira publicação da teoria da relatividade geral. Schwarzschild obteve sua solução enquanto estava servindo no exército alemão durante a I Guerra Mundial, publicou a mesma no início de 1916²³ e pouco tempo depois acabou falecendo em virtude de uma doença contrída no campo de batalha.

A solução de Schwarzschild descreve o campo gravitacional (ou seja, a métrica do espaço-tempo) gerado no espaço que rodeia um corpo de massa M esférico homogêneo, sem carga elétrica, com momentum angular nulo e fixa na origem do sistema de referência.

6.15.2.1 DERIVAÇÃO DO TENSOR DE MÉTRICA

Considera-se um contínuo espaço-tempo de 4 dimensões, o qual se reduz assintoticamente ao espaço de Minkowski introduzido na seção 6.15.1.4 a uma distância infinita do corpo de massa M . Neste espaço, o quadrivetor posição continua sendo determinado pela 4-upla (x^0, x^1, x^2, x^3) . Assumindo que as equações de Einstein (6.71a) podem ser escritas como

$$G^\alpha_\beta = 0,$$

contraíndo o tensor de Einstein (6.71b) resulta

$$G^\alpha_\alpha = R^\alpha_\alpha - \frac{1}{2} \delta^\alpha_\alpha R = R - \frac{1}{2} 4R = 0,$$

o que implica que $R = 0$. Portanto, as equações de Einstein se reduzem a

$$R_{\alpha\beta} = 0.$$

Lembrando que o tensor de Ricci é dado por (6.69), no espaço R^4 em questão as equações de Einstein ficam

$$R_{\alpha\beta} = \frac{\partial \Gamma^\gamma_{\alpha\beta}}{\partial x^\gamma} - \frac{\partial \Gamma^\gamma_{\alpha\gamma}}{\partial x^\beta} + \Gamma^\delta_{\alpha\beta} \Gamma^\gamma_{\delta\gamma} - \Gamma^\delta_{\alpha\gamma} \Gamma^\gamma_{\delta\beta} = 0, \quad (\alpha, \beta, \gamma, \delta = 0, 1, 2, 3), \quad (6.88)$$

na forma (6.69b). Alternativamente, pode-se empregar a forma (6.69d).

Busca-se uma solução do sistema acima para um espaço esfericamente simétrico. Por isso, as coordenadas espaciais do quadrivetor serão escritas em termos das coordenadas polares esféricas $x^1 = r$, $x^2 = \theta$ e $x^3 = \varphi$. Além disso, como já mencionado, a distâncias infinitas da

²³Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie (A respeito do campo gravitacional de uma partícula massiva na teoria de Einstein). *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, v. 7, p. 189–196, 1916. Acesso: <https://www.biodiversitylibrary.org/item/93032#page/215/mode/lup> (Versão em inglês: <https://arxiv.org/abs/physics/9905030v1>).

massa que gera o campo, a métrica do espaço-tempo deve se reduzir à métrica do espaço-tempo de Minkowski M^4 , dada por

$$ds^2 = (cdt)^2 - dr^2 - r^2 d\theta^2 - r^2 \sin^2 \theta d\varphi^2.$$

Por estas razões, Schwarzschild propôs a seguinte métrica para o espaço-tempo em torno da massa M :

$$ds^2 = e^{2\nu(r)} c^2 dt^2 - e^{2\lambda(r)} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2), \quad (6.89)$$

onde $\nu(r)$ e $\lambda(r)$ são funções que devem ser derivadas pelas equações de campo e que estão sujeitas à condição de contorno

$$\lim_{r \rightarrow \infty} \nu(r) = \lim_{r \rightarrow \infty} \lambda(r) = 0.$$

Comparando a métrica proposta acima com a métrica geral no R^n definida em (6.26), conclui-se que

$$\mathbf{g} \equiv [g_{\alpha\beta}] = \begin{pmatrix} e^{2\nu} & 0 & 0 & 0 \\ 0 & -e^{2\lambda} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2 \theta \end{pmatrix}.$$

Ou seja, o tensor de métrica é diagonal. O determinante de $g_{\alpha\beta}$ é

$$g = g_{00}g_{11}g_{22}g_{33} = -e^{2(\nu+\lambda)} r^4 \sin^2 \theta$$

e sua forma contravariante é

$$[g^{\alpha\beta}] = \begin{pmatrix} e^{-2\nu} & 0 & 0 & 0 \\ 0 & -e^{-2\lambda} & 0 & 0 \\ 0 & 0 & -\frac{1}{r^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{r^2 \sin^2 \theta} \end{pmatrix}.$$

De acordo com a expressão (6.46), os símbolos de Christoffel são escritos

$$\Gamma^\gamma_{\alpha\beta} = \frac{1}{2} g^{\gamma\delta} \left(\frac{\partial g_{\beta\delta}}{\partial x^\alpha} + \frac{\partial g_{\delta\alpha}}{\partial x^\beta} - \frac{\partial g_{\alpha\beta}}{\partial x^\delta} \right).$$

Para a derivação destes símbolos, escreve-se inicialmente

$$g_{\alpha\beta} \stackrel{\text{NS}}{=} f_\alpha^\delta \delta_{\alpha\beta}, \quad g^{\alpha\beta} \stackrel{\text{NS}}{=} f_\alpha^{-2} \delta^{\alpha\beta},$$

onde

$$f_0 = e^\nu, \quad f_1 = e^\lambda, \quad f_2 = r, \quad f_3 = r \sin \theta,$$

o que resulta em

$$\begin{aligned} \Gamma^\gamma_{\alpha\beta} &= \frac{1}{2} f_\gamma^{-2} \delta^{\gamma\delta} \left(\frac{\partial f_\beta^2}{\partial x^\alpha} \delta_{\beta\delta} + \frac{\partial f_\alpha^2}{\partial x^\beta} \delta_{\alpha\delta} - \frac{\partial f_\alpha^2}{\partial x^\delta} \delta_{\alpha\beta} \right) \\ &\stackrel{\text{NS}}{=} f_\gamma^{-2} \left(f_\beta \delta_\beta^\gamma \frac{\partial f_\beta}{\partial x^\alpha} + f_\alpha \delta_\alpha^\gamma \frac{\partial f_\alpha}{\partial x^\beta} - f_\alpha \delta_{\alpha\beta} \frac{\partial f_\alpha}{\partial x^\gamma} \right). \end{aligned}$$

Após uma certa álgebra, as matrizes de Christoffel ficam então:

$$\begin{aligned} \Gamma^0 &= \begin{pmatrix} 0 & \nu' & 0 & 0 \\ \nu' & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \Gamma^1 &= \begin{pmatrix} e^{2(\nu-\lambda)} \nu' & 0 & 0 & 0 \\ 0 & \lambda' & 0 & 0 \\ 0 & 0 & -e^{-2\lambda} r & 0 \\ 0 & 0 & 0 & -e^{-2\lambda} r \sin^2 \theta \end{pmatrix} \\ \Gamma^2 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & r^{-1} & 0 \\ 0 & r^{-1} & 0 & 0 \\ 0 & 0 & 0 & -\cos \theta \sin \theta \end{pmatrix} & \Gamma^3 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r^{-1} \\ 0 & 0 & 0 & \cotan \theta \\ 0 & r^{-1} \cotan \theta & 0 & 0 \end{pmatrix}. \end{aligned} \quad (6.90)$$

Destes resultados pode-se concluir que

$$\Gamma^0_{ij} = 0, \quad \Gamma^i_{0j} = \Gamma^i_{j0} = 0.$$

Inserindo os símbolos de Christoffel dados por (6.90) e suas derivadas nas equações de campo (6.88), obtém-se, após uma certa quantidade de álgebra,

$$\begin{aligned} R_{\alpha\beta} &= 0, \quad (\alpha \neq \beta) \\ R_{00} &= [\nu'' + \nu'^2 - \lambda'\nu' + 2r^{-1}\nu'] e^{2(\nu-\lambda)} = 0 \\ R_{11} &= -\nu'' + \lambda'\nu' - \nu'^2 + 2r^{-1}\lambda' = 0 \\ R_{22} &= 1 - [1 + r(\nu' - \lambda')] e^{-2\lambda} = 0 \\ R_{33} &= [1 - (1 + \nu'r - \lambda'r) e^{-2\lambda}] \text{sen}^2 \theta = 0. \end{aligned}$$

Portanto, observa-se que $R_{\alpha\beta} = 0$ para $\alpha \neq \beta$ e a partir dos elementos da diagonal resulta o seguinte sistema de equações para as funções $\nu(r)$ e $\lambda(r)$:

$$\nu'' - \nu'\lambda' + \nu'^2 + 2r^{-1}\nu' = 0 \quad (6.91a)$$

$$\nu'' - \lambda'\nu' + \nu'^2 - 2\lambda'r^{-1} = 0 \quad (6.91b)$$

$$[r(\lambda' - \nu') - 1] e^{-2\lambda} + 1 = 0 \quad (6.91c)$$

$$\{[r(\lambda' - \nu') - 1] e^{-2\lambda} + 1\} \text{sen}^2 \theta = 0. \quad (6.91d)$$

A equação (6.91d) é uma mera repetição de (6.91c). De (6.91a) e (6.91b) conclui-se que

$$\lambda' = -\nu' \implies \lambda(r) = -\nu(r) + \text{cte.}$$

Porém, pelo limite $r \rightarrow \infty$ que ambas têm que satisfazer, resulta que, simplesmente, $\lambda(r) = -\nu(r)$. Com isso, (6.91c) resulta

$$-(2r\nu' + 1) e^{2\nu} + 1 = 0.$$

Chamando $\mu(r) = e^{2\nu}$,

$$2\nu' = \frac{\mu'}{\mu} \implies r\mu' + \mu = 1.$$

A solução desta EDO é

$$\mu(r) = e^{2\nu(r)} = 1 - \frac{r_G}{r},$$

onde r_G é uma constante de integração denominada o *raio gravitacional* ou o *raio de Schwarzschild do corpo*, a qual será em breve identificada com a massa do mesmo.

Assim, inserindo esta solução em (6.89), obteve-se a métrica de Schwarzschild

$$\begin{aligned} ds^2 &= \left(1 - \frac{r_G}{r}\right) c^2 dt^2 - \left(1 - \frac{r_G}{r}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2 \theta d\varphi^2). \\ [g_{\alpha\beta}] &= \begin{pmatrix} \mu & 0 & 0 & 0 \\ 0 & -\mu^{-1} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \text{sen}^2 \theta \end{pmatrix}. \end{aligned} \quad (6.92)$$

6.15.2.2 CONSEQUÊNCIAS E APLICAÇÕES DA MÉTRICA DE SCHWARZSCHILD

Serão discutidos agora de forma breve algumas das principais consequências advindas da métrica de Schwarzschild.

TRAJETÓRIA DE PARTÍCULAS NO CAMPO GRAVITACIONAL E O LIMITE NEWTONIANO

Na teoria da relatividade geral, as equações de movimento de uma partícula de massa m sob a ação do campo gravitacional determinado pela métrica $g_{\alpha\beta}$ do espaço-tempo são exatamente as equações (6.61), as quais resultam nas curvas geodésicas entre dois pontos quaisquer do espaço-tempo determinado pela métrica. Ou seja, na relatividade geral as equações de movimento são

$$\frac{d^2 x^\alpha}{d\ell^2} + \Gamma^\alpha_{\beta\gamma} \frac{dx^\beta}{d\ell} \frac{dx^\gamma}{d\ell} = 0 \quad (\alpha, \beta, \gamma = 0, 1, 2, 3),$$

sendo ℓ a extensão do arco da trajetória da partícula no espaço-tempo.

Será feito uso agora do conceito de *tempo próprio*, discutido na seção 6.15.1.4 e que consiste no intervalo de tempo registrado no referencial sempre em repouso com o observador. Dada a métrica

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta,$$

no referencial em repouso ao observador, dois eventos infinitesimalmente separados no tempo e que ocorrem no mesmo ponto do espaço ocorrem no intervalo $d\tau$, onde τ é o *tempo próprio*. Então, neste referencial, $dx^i = 0$ ($i = 1, 2, 3$) e

$$ds^2 \doteq c^2 d\tau^2 = g_{00} (dx^0)^2.$$

Portanto,

$$\frac{d\tau}{d\ell} = c^{-1}, \quad d\tau = \sqrt{g_{00}} dt.$$

Então, nas equações de movimento será realizada a mudança de variável $\ell \rightarrow \tau$ via

$$\frac{dx^\alpha}{d\ell} = \frac{dx^\alpha}{d\tau} \frac{d\tau}{d\ell} = c^{-1} \frac{dx^\alpha}{d\tau}.$$

Dessa forma, as equações de movimento da partícula no campo gravitacional ficam

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\gamma} \frac{dx^\beta}{d\tau} \frac{dx^\gamma}{d\tau} = 0 \quad (\alpha, \beta, \gamma = 0, 1, 2, 3).$$

Na métrica de Schwarzschild, os símbolos de Christoffel são dados por (6.90) e

$$d\tau = \sqrt{\mu} dt.$$

Neste caso, as equações de movimento tornam-se

$$\begin{aligned} \frac{d^2 x^0}{d\tau^2} + 2\nu' \frac{dr}{d\tau} \frac{dx^0}{d\tau} &= 0 \\ \frac{d^2 x^1}{d\tau^2} + \mu^2 \nu' \left(\frac{dx^0}{d\tau} \right)^2 - \nu' \left(\frac{dx^1}{d\tau} \right)^2 - \mu r \left(\frac{dx^2}{d\tau} \right)^2 - \mu r \sin^2 \theta \left(\frac{dx^3}{d\tau} \right)^2 &= 0 \\ \frac{d^2 x^2}{d\tau^2} + 2r^{-1} \frac{dx^1}{d\tau} \frac{dx^2}{d\tau} - \sin \theta \cos \theta \left(\frac{dx^3}{d\tau} \right)^2 &= 0 \\ \frac{d^2 x^3}{d\tau^2} + 2r^{-1} \frac{dx^1}{d\tau} \frac{dx^3}{d\tau} + 2 \cotan \theta \frac{dx^2}{d\tau} \frac{dx^3}{d\tau} &= 0. \end{aligned}$$

Mas, como $\nu' = \mu'/2\mu$, pode-se escrever

$$\frac{d}{d\tau} \left(\mu \frac{dx^0}{d\tau} \right) = 0 \quad (6.93a)$$

$$\frac{d^2 r}{d\tau^2} - \mu r \left(\frac{d\theta}{d\tau} \right)^2 - \mu r \sin^2 \theta \left(\frac{d\varphi}{d\tau} \right)^2 + \frac{1}{2} \mu \mu' \left(\frac{dx^0}{d\tau} \right)^2 - \frac{1}{2} \frac{\mu'}{\mu} \left(\frac{dr}{d\tau} \right)^2 = 0 \quad (6.93b)$$

$$\frac{d}{d\tau} \left(r^2 \frac{d\theta}{d\tau} \right) - r^2 \sin \theta \cos \theta \left(\frac{d\varphi}{d\tau} \right)^2 = 0 \quad (6.93c)$$

$$\frac{d}{d\tau} \left(r^2 \sin^2 \theta \frac{d\varphi}{d\tau} \right) = 0. \quad (6.93d)$$

Observa-se que a primeira e a última equações de movimento pode ser imediatamente integradas, resultando em

$$\begin{aligned} \mu \frac{dx^0}{d\tau} &= \alpha \\ r^2 \sin^2 \theta \frac{d\varphi}{d\tau} &= h, \end{aligned}$$

onde α e h são constantes de movimento, as quais foram escolhidas de maneira a facilitar a identificação do limite não relativístico no movimento da partícula. Adicionalmente, será determinada agora a orientação do referencial. Como a métrica é esfericamente simétrica, os eixos do

referencial podem ter qualquer orientação arbitrária, sem que isso comprometa a generalidade da equação de movimento. Observando que a terceira equação de movimento é

$$\frac{d^2\theta}{d\tau^2} + \frac{2}{r} \frac{dr}{d\tau} \frac{d\theta}{d\tau} - \operatorname{sen} \theta \cos \theta \left(\frac{d\varphi}{d\tau} \right)^2 = 0,$$

se a partícula estiver no plano $\theta = \pi/2$ em um determinado instante próprio τ_0 , pode-se desenvolver $\theta(\tau)$ em torno de τ_0 como

$$\theta(\tau) = \frac{\pi}{2} + \left. \frac{d\theta}{d\tau} \right|_{\tau_0} (\tau - \tau_0) + \frac{1}{2!} \left. \frac{d^2\theta}{d\tau^2} \right|_{\tau_0} (\tau - \tau_0)^2 + \frac{1}{3!} \left. \frac{d^3\theta}{d\tau^3} \right|_{\tau_0} (\tau - \tau_0)^3 + \dots$$

Ou seja,

$$\frac{d\theta}{d\tau} = \left. \frac{d\theta}{d\tau} \right|_{\tau_0} + \left. \frac{d^2\theta}{d\tau^2} \right|_{\tau_0} (\tau - \tau_0) + \frac{1}{2!} \left. \frac{d^3\theta}{d\tau^3} \right|_{\tau_0} (\tau - \tau_0)^2 + \dots$$

Se, adicionalmente, $d\theta/d\tau|_{\tau_0} = 0$, então a equação de movimento mostra que $d^2\theta/d\tau^2|_{\tau_0} = 0$. Para se obter $d^3\theta/d\tau^3|_{\tau_0}$, deriva-se a equação para $\theta(\tau)$, resultando então que esta derivada também é nula. Assim, pode-se mostrar que todos os termos na série de Taylor para $\theta(\tau)$ são nulos exceto o primeiro. Portanto, as equações de movimento na métrica de Schwarzschild admitem uma solução $\theta(\tau) = \pi/2$, exatamente como na teoria Newtoniana da gravitação.

As equações de movimento reduzem-se então para

$$\begin{aligned} \mu \frac{dx^0}{d\tau} &= \alpha \\ r^2 \frac{d\varphi}{d\tau} &= h, \\ \frac{d^2r}{d\tau^2} - \mu r \left(\frac{h}{r^2} \right)^2 + \frac{1}{2} \mu \mu' \left(\frac{\alpha}{\mu} \right)^2 - \frac{1}{2} \frac{\mu'}{\mu} \left(\frac{dr}{d\tau} \right)^2 &= 0. \end{aligned}$$

Agora, o termo $(dr/d\tau)^2$ na equação será modificado a partir da métrica. De (6.92), resulta que

$$\begin{aligned} \mu^{-1} \left(\frac{dr}{d\tau} \right)^2 &= \mu \left(\frac{dx^0}{d\tau} \right)^2 - r^2 \left[\left(\frac{d\theta}{d\tau} \right)^2 + \operatorname{sen}^2 \theta \left(\frac{d\varphi}{d\tau} \right)^2 \right] - c^2 \\ &= \mu^{-1} \alpha^2 - \frac{h^2}{r^2} - c^2, \end{aligned}$$

onde também foi empregada a identidade $dl = cd\tau$. Portanto,

$$\begin{aligned} \frac{d^2r}{d\tau^2} - \mu \frac{h^2}{r^3} + \frac{1}{2} \mu' \frac{h^2}{r^2} + \frac{1}{2} \mu' c^2 &= 0, \\ \frac{d^2r}{d\tau^2} - \left(1 - \frac{3}{2} \frac{r_G}{r} \right) \frac{h^2}{r^3} + \frac{r_G c^2}{2r^2} &= 0, \end{aligned} \tag{6.94}$$

uma vez que $\mu = 1 - r_G/r$.

Para identificar finalmente a expressão para o raio gravitacional r_G , considera-se primeiro o limite clássico das equações de movimento. Observa-se que

$$d\tau = \sqrt{\mu} dt = \sqrt{1 - \frac{r_G}{r}} dt.$$

O limite clássico é obtido para $r \gg r_G$, de onde

$$d\tau \approx \left(1 - \frac{r_G}{2r} \right) dt \approx dt.$$

Neste limite, a equação (6.94) fica

$$\frac{d^2r}{dt^2} - \frac{h^2}{r^3} + \frac{r_G c^2}{2r^2} = 0.$$

Esta equação será comparada com as equações de movimento de uma partícula de massa m sob a influência da força gravitacional de uma massa M , de acordo com a lei da gravitação universal de Newton. As equações são as seguintes:

$$\begin{aligned} \ddot{r} - r(\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2) + \frac{1}{m}U' &= 0 \\ \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin \theta \cos \theta \dot{\varphi}^2 &= 0 \\ \frac{d}{dt}(r^2 \sin^2 \theta \dot{\varphi}) &= 0 \implies r^2 \sin^2 \theta \dot{\varphi} = h \end{aligned}$$

onde

$$U = -\frac{GMm}{r}$$

é a energia potencial gravitacional do sistema composto pelas massas M e m e h é proporcional ao momentum angular (constante). Estas equações admitem uma solução $\theta = \pi/2 = \text{cte.}$, em cuja situação a equação radial pode ser escrita

$$\ddot{r} - \frac{h^2}{r^3} + \frac{GM}{r^2} = 0.$$

Comparando esta equação com (6.94) no limite clássico, obtém-se

$$r_G = \frac{2GM}{c^2}.$$

É interessante comparar os raios de Schwarzschild de alguns objetos astronômicos conhecidos. A tabela 6.3 mostra o valores de r_G e também da densidade dos objetos. Observa-se que a Terra tem um raio gravitacional de pouco menos do que 1 cm, ao passo que para o Sol, $r_G \approx 3$ km. Ou seja, os seus raios gravitacionais são muito menores que os seus tamanhos. Contudo, um objeto suficientemente denso pode ser um raio de Schwarzschild maior que o seu tamanho. Se o raio do objeto é

$$r = \sqrt[3]{\frac{3M}{4\pi\rho}},$$

sendo ρ sua densidade, então $r_G > r$ se

$$\rho > \frac{3c^6}{32\pi G^3 M^2}.$$

Isto é o que ocorre com um Buraco Negro.

DISTORÇÕES DO ESPAÇO-TEMPO NA MÉTRICA DE SCHWARZSCHILD

Alguns dos resultados mais conhecidos da métrica de Schwarzschild serão apresentados agora. O espaço-tempo descrito pela métrica (6.92) não é “plano” como o espaço Euclidiano, tacitamente assumido na gravitação Newtoniana. Se forem considerados dois pontos infinitesimalmente próximos na direção radial a partir da massa M , de tal forma que $dt = d\theta = d\varphi$, resulta que a extensão do elemento de arco na direção radial resulta

$$dR = \left(1 - \frac{r_G}{r}\right)^{-1/2} dr > dr.$$

Ou seja, a distância entre os pontos r_1 e r_2 ($> r_1$) é obtida pela integração

$$R_{21} = \int_{r_1}^{r_2} \left(1 - \frac{r_G}{r}\right)^{-1/2} dr = \left[\sqrt{r(r-r_G)} + r_G \ln(\sqrt{r} + \sqrt{r-r_G})\right]_{r_1}^{r_2} > r_2 - r_1.$$

Tabela 6.3: Raios de Schwarzschild de alguns objetos astronômicos.

	r_G (m)	Densidade (g/cm ³)
Terra	$8,83 \times 10^{-3}$	$2,04 \times 10^{27}$
Sol	$2,95 \times 10^3$	$1,84 \times 10^{16}$
Via Láctea	$2,08 \times 10^{15}$ ($\sim 0,2$ a.l.)	$3,72 \times 10^{-8}$

Obviamente, se $M \rightarrow 0$, $R_{21} \rightarrow r_2 - r_1$ e o espaço torna-se plano.

Considerando agora a relação entre o tempo próprio τ e o tempo t medido por um observador distante da massa M , parte-se da relação

$$d\tau = \sqrt{\mu} dt = \sqrt{1 - \frac{r_G}{r}} dt,$$

o que mostra que $d\tau < dt$. O tempo próprio entre dois eventos quaisquer ocorrendo no mesmo ponto do espaço é então dado por

$$\tau = \int_{t_1}^{t_2} \sqrt{1 - \frac{r_G}{r}} dt < t_2 - t_1.$$

Observa-se que $\tau \rightarrow 0$ à medida que $r \rightarrow r_G$. Este resultado caracteriza o denominado *horizonte de eventos*, pois para um observador situado exatamente no ponto $r = r_G$ quaisquer dois eventos que levam um intervalo de tempo τ finito irão ocorrer ao longo de um intervalo de tempo infinito para um outro observador situado distante do raio gravitacional de M . A dilatação temporal prevista na métrica de Schwarzschild foi corroborada pelo experimento de Pound-Rebka em 1959 (POUND; REBKA, 1959, 1960).²⁴

REFERÊNCIAS

- ARFKEN, G. B.; WEBER, H. J. **Mathematical Methods for Physicists**. Sixth. New York: Elsevier, 2005. 1182 + xii pp. ISBN 0-12-059876-0.
- BELLAC, M.; LÉVY-LEBLOND, J.M. Galilean electromagnetism. English. **Il Nuovo Cimento**, Società Italiana di Fisica, v. 14, n. 2, p. 217–234, abr. 1973. ISSN 0369-3554. DOI: [10.1007/BF02895715](https://doi.org/10.1007/BF02895715).
- BISHOP, R. L.; GOLDBERG, S. I. **Tensor Analysis on Manifolds**. New York: Dover, 2012. (Dover Books on Mathematics). ISBN 9780486139234.
- BLOCH, A. M. **Nonholonomic Mechanics and Control**. New York: Springer, 2003. (Interdisciplinary Applied Mathematics). ISBN 9780387955353.
- BORISENKO, A.A.I.; TARAPOV, I.E.; SILVERMAN, R.A. **Vector and Tensor Analysis With Applications**. New York: Dover, 1968. (Dover Books on Mathematics). 257 + x pp. ISBN 9780486638331.
- DALARSSON, M.; DALARSSON, N. **Tensor Calculus, Relativity, and Cosmology: A First Course**. New York: Elsevier, 2005. 280 + xii pp. ISBN 9780122006814.
- DEBENEDICTIS, Andrew. Integration in General Relativity, p. 7, fev. 1998. Disponível em: <http://arxiv.org/abs/physics/9802027>.
- EINSTEIN, Albert. A Generalized Theory of Gravitation. **Reviews of Modern Physics**, American Physical Society, v. 20, n. 1, p. 35–39, 1 jan. 1948. DOI: [10.1103/RevModPhys.20.35](https://doi.org/10.1103/RevModPhys.20.35).
- FITZPATRICK, Richard. **Fluid Mechanics**. [S.l.: s.n.], 2012. Acesso em 10/03/2013. Disponível em: <http://farside.ph.utexas.edu/teaching/3361/fluidhtml/Fluidhtml.html>.
- GARWIN, Richard L.; LEDERMAN, Leon M.; WEINRICH, Marcel. Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon. **Physical Review**, American Physical Society, v. 105, n. 4, p. 1415–1417, 4 fev. 1957. DOI: [10.1103/PhysRev.105.1415](https://doi.org/10.1103/PhysRev.105.1415).
- JACKSON, J. D. **Classical Electrodynamics**. Third. New York: John Wiley & Sons, 1999. 808 + xxi pp. ISBN 0-471-30932-X.
- JOSHI, A. W. **Matrices And Tensors In Physics**. New Delhi: John Wiley & Sons, 1995. 342 + xvi pp. ISBN 9788122405637.
- KAY, D. **Schaums Outline of Tensor Calculus**. New York: McGraw-Hill, 2011. (Schaum's Outline Series). 228 + ix pp. ISBN 9780071756037.
- LEE, T. D.; YANG, C. N. Question of Parity Conservation in Weak Interactions. **Physical Review**, American Physical Society, v. 104, n. 1, p. 254–258, 1 out. 1956. DOI: [10.1103/PhysRev.104.254](https://doi.org/10.1103/PhysRev.104.254).

²⁴Ver também: [Physics Focus: The Weight of Light](#).

- MATHEWS, Jon; WALKER, R. L. **Mathematical methods of physics**. Second. [S.l.]: W. A. Benjamin, 1970. (World Student Series). 501 + xiii pp. ISBN 9780805370027.
- MISNER, C.W.; THORNE, K.S.; WHEELER, J.A. **Gravitation**. San Francisco: Freeman, 1973. 1279 + xxvii pp. ISBN 9780716703440.
- NEARING, James. **Mathematical Tools for Physics**. New York: Dover Publications, out. 2010. (Dover Books on Physics). 496 pp. ISBN 978-0486482125.
- POUND, R. V.; REBKA, G. A. Apparent Weight of Photons. **Physical Review Letters**, American Physical Society, v. 4, n. 7, p. 337–341, 7 abr. 1960. DOI: [10.1103/PhysRevLett.4.337](https://doi.org/10.1103/PhysRevLett.4.337).
- _____. Gravitational Red-Shift in Nuclear Resonance. **Physical Review Letters**, American Physical Society, v. 3, n. 9, p. 439–441, 9 nov. 1959. DOI: [10.1103/PhysRevLett.3.439](https://doi.org/10.1103/PhysRevLett.3.439).
- PRICE, Richard H. General relativity primer. **American Journal of Physics**, v. 50, n. 4, p. 300–329, abr. 1982. DOI: [10.1119/1.12853](https://doi.org/10.1119/1.12853).
- RINDLER, Wolfgang. **Relativity: Special, General, and Cosmological**. New York: Oxford University Press, 2006. 430 + xvi pp. ISBN 9780198567325.
- SHARIPOV, R. A. **Quick introduction to tensor analysis**. [S.l.], mar. 2004. 47 pp. eprint: [arXiv:math/0403252](https://arxiv.org/abs/math/0403252). Disponível em: <http://arxiv.org/abs/math/0403252>.
- SILVEIRA, V. M. G. da. **On CP violation and relativistic non-inertial effects**. Abr. 2019. 106 f. Diss. (Mestrado) – IF-UFRGS, Porto Alegre. Disponível em: https://sabi.ufrgs.br/F/YM9MTR51JMXFYNG3R1N7BQ4G9TM56454YMGVT6UQKTSIX3M6EI-05313?func=find-b&request=Vinicius+Medeiros+Gomes+da+Silveira&find_code=WAU&adjacent=N&x=0&y=0&filter_code_2=WLN&filter_request_2=&filter_code_3=WYR&filter_request_3=&filter_code_4=WYR&filter_request_4.
- SIMMONDS, James D. **A Brief on Tensor Analysis**. Second. New York: Springer-Verlag, jul. 1997. (Undergraduate Texts in Mathematics). 112 + xiv pp. ISBN 9780387940885.
- SOKOLNIKOFF, I. S. **Tensor Analysis. Theory and Applications**. New York: John Wiley & Sons, 1951. (Applied Mathematics Series). 335 + x pp.
- SYNGE, J.J.L.; SCHILD, A. **Tensor calculus**. New York: Dover, 1969. (Dover books on advanced mathematics). 329 + ix pp. ISBN 9780486636122.
- SZEKERES, Peter. **A Course in Modern Mathematical Physics: Groups, Hilbert Space, and Differential Geometry**. Cambridge: Cambridge University Press, 2004. ISBN 9780521536455. Disponível em: <http://books.google.com.br/books?id=pRgjkgEACAAJ>.
- WU, C. S.; AMBLER, E.; HAYWARD, R. W.; HOPPES, D. D.; HUDSON, R. P. Experimental Test of Parity Conservation in Beta Decay. **Physical Review**, American Physical Society, v. 105, n. 4, p. 1413–1415, 4 fev. 1957. DOI: [10.1103/PhysRev.105.1413](https://doi.org/10.1103/PhysRev.105.1413).

DISTRIBUIÇÕES E A “FUNÇÃO” DELTA DE DIRAC

DELTA DE DIRAC é um artifício matemático utilizado em diversas áreas da física. Por exemplo, no cálculo da Função de Green, em problemas de contorno não homogêneos, na determinação do espectro contínuo de autovalores em mecânica quântica, na determinação das flutuações térmicas de partículas pontuais em mecânica estatística e outras aplicações importantes.

Este roteiro tem por objetivo definir a Delta de Dirac como tipo particular de uma *distribuição* e apresentar algumas de suas propriedades.

A.1 DEFINIÇÃO DE DISTRIBUIÇÕES

Denotando por $\varphi(x_1, x_2, \dots, x_n)$ uma função de n variáveis contínuas x_1, x_2, \dots, x_n , cujos valores não nulos estão todos contidos dentro do domínio da função φ e que possui derivadas de todas as ordens em relação a estas variáveis.

Definição. Uma distribuição $T[\varphi]$ é um funcional linear e contínuo da função φ .

Por linearidade, entende-se que para qualquer combinação linear $\lambda_1\varphi_1 + \lambda_2\varphi_2$, onde $\lambda_1, \lambda_2 \in \mathbb{C}$:

$$T[\lambda_1\varphi_1 + \lambda_2\varphi_2] = \lambda_1T[\varphi_1] + \lambda_2T[\varphi_2].$$

Por continuidade, entende-se que para qualquer seqüência $\varphi_1, \varphi_2, \dots, \varphi_j, \dots$ de funções tais que

$$\lim_{j \rightarrow \infty} \varphi_j = \varphi,$$

tem-se

$$\lim_{j \rightarrow \infty} T[\varphi_j] = T[\varphi].$$

A.1.1 DEFINIÇÃO OPERACIONAL DE DISTRIBUIÇÃO

A uma função *localmente integrável* f qualquer, isto é, qualquer função cuja integral¹ sobre qualquer intervalo finito contido no domínio exista, corresponde uma distribuição \hat{f} , também denominada *funcional* definida pelo produto escalar

$$\hat{f}[\varphi] = \int f(x)\varphi(x)dx = \langle \varphi^*, f \rangle.$$

A.1.1.1 EXEMPLOS

A função $1/x$ não define uma distribuição porque não é integrável no ponto $x = 0$. Contudo, pode-se definir a distribuição

$$\text{PP} \frac{1}{x} [\varphi] \equiv \text{PP} \int \frac{\varphi(x)}{x} dx,$$

¹Integral na definição de *Integral de Lebesgue*. Uma integral de Lebesgue reduz-se a uma integral de Riemann (integral usual) sempre que a última puder ser definida. Contudo, a integral de Lebesgue existe mesmo em casos onde a integral de Riemann não pode ser definida.

onde PP denote a parte principal de Cauchy da integral, isto é,

$$\text{PP} \int_{-\infty}^{\infty} = \lim_{\epsilon \rightarrow 0} \left(\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty} \right).$$

A “função de Dirac” $\delta(x)$ define a distribuição

$$\delta[\varphi] = \varphi(0),$$

assim como a “função” $\delta(x - x_0)$ define a distribuição

$$\delta_{x_0}[\varphi] = \varphi(x_0).$$

A.1.2 PROPRIEDADES DE DISTRIBUIÇÕES

A.1.2.1 COMBINAÇÃO LINEAR DE DISTRIBUIÇÕES

Seja $T = \lambda_1 T_1 + \lambda_2 T_2$, onde $\lambda_1, \lambda_2 \in \mathbb{C}$, tem-se

$$T[\varphi] = \lambda_1 T_1[\varphi] + \lambda_2 T_2[\varphi].$$

A.1.2.2 PRODUTO DE DUAS DISTRIBUIÇÕES

Se \hat{f} uma distribuição associada com uma função localmente integrável f e T uma distribuição arbitrária, a distribuição

$$P = \hat{f}T$$

é bem definida se T é linear, é um funcional contínuo da função $f\varphi$ e tem-se, por definição,

$$P[\varphi] = T[f\varphi].$$

O produto de duas distribuições nem sempre existe. Se f possui derivadas de todas as ordens, $\hat{f}T$ existe para todas distribuições T . Se f é contínua no ponto x_0 ,

$$\left(\hat{f}\delta_{x_0} \right) [\varphi] = f(x_0)\varphi(x_0). \quad (\text{A.1})$$

Se f e g são funções quadraticamente integráveis,² o produto $\hat{f}\hat{g}$ está bem definido. Por outro lado, $[\delta(x)]^2$ não tem sentido, assim como $\left(1/\sqrt{|x|}\right)^2$.

Como um caso especial da equação (A.1), tem-se

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

Por conseguinte, se $xT = 0$, T é um múltiplo de $\delta(x)$: $T = c\delta(x)$, onde c é uma constante.

Portanto, se $f(x)$ e $g(x)$ estão relacionadas pela relação

$$xf(x) = g(x),$$

tem-se, necessariamente,

$$f(x) = \text{PP} \frac{g(x)}{x} + c\delta(x)$$

onde c é uma constante a ser determinada.

A.1.2.3 SÉRIES E INTEGRAIS DE DISTRIBUIÇÕES

Se um conjunto de distribuições $T_1, T_2, \dots, T_j, \dots$ é tal que quando $j \rightarrow \infty$, $T_j[\varphi]$ possui um limite para qualquer φ , este limite é também uma distribuição:

$$T = \lim_{j \rightarrow \infty} T_j.$$

²Isto é, se $\int |f|^2 dx$ e $\int |g|^2 dx$ existirem.

Definição equivalente: Se a série

$$\sum_i T_i [\varphi]$$

for definida para qualquer φ , seu resultado define uma distribuição; neste caso, diz-se que a série de distribuições $\{T_i\}$ é realizável:

$$T [\varphi] = \sum_i T_i [\varphi].$$

Se $T(\lambda)$ é uma distribuição que depende de um parâmetro $\lambda \in \mathbb{C}$, o qual pode variar continuamente em um domínio Λ e se a integral

$$I [\varphi] = \int_{\Lambda} T(\lambda) [\varphi] d\lambda$$

converge para qualquer φ , o objeto $I = \int_{\Lambda} T(\lambda) d\lambda$ define uma distribuição. Uma definição análoga vale para integrais múltiplas.

Em particular, se $f(x, \lambda)$ é uma função localmente integrável de x e λ , a distribuição $\hat{f}(\lambda)$ é integrável em λ e sua integral é a distribuição \hat{g} associada com a função:

$$\hat{g} [1] = \int_{\Lambda} \hat{f}(\lambda) [1] d\lambda = \int_{\Lambda} f(x, \lambda) d\lambda.$$

Se a função $a(k)$ permanece menor que uma potência positiva de $|k|$ quando $|k| \rightarrow \infty$:

$$|a(k)| \leq A |k|^{-\alpha} \quad (A \text{ e } \alpha \text{ constantes positivas}),$$

a integral

$$\int_{-\infty}^{\infty} e^{ikx} a(k) dk$$

é uma distribuição. Em particular,

$$\int_{-\infty}^{\infty} e^{ikx} dk = 2\pi\delta(x).$$

A.1.2.4 DERIVADAS DE DISTRIBUIÇÕES

Por definição, a derivada parcial $\partial T / \partial x_i$ da distribuição T é:

$$\frac{\partial T}{\partial x_i} [\varphi] = -T \left[\frac{\partial \varphi}{\partial x_i} \right].$$

Em particular, se uma função localmente integrável é diferenciável, a derivada da distribuição correspondente é a distribuição correspondente a sua derivada. Isto pode ser visto via integração por partes:

$$\hat{f}' [\varphi] = \int f'(x) \varphi(x) dx = - \int f(x) \varphi'(x) dx = -\hat{f} [\varphi'].$$

Todas as propriedades das derivadas de funções aplicam-se a distribuições. Por exemplo, a derivada do produto $P = \hat{f}T$ é:

$$P' = \hat{f}'T + \hat{f}T'.$$

Além disso, certas propriedades que pertencem a uma classe restrita de funções aplicam-se a todas as distribuições sem restrições. São as seguintes:

1. As distribuições são diferenciáveis em todas as ordens.
2. Derivação é uma operação linear e contínua no espaço das distribuições. Se

$$\lim_{j \rightarrow \infty} T_j = T, \quad \lim_{j \rightarrow \infty} T_j' = T'.$$

Em conseqüência, se a série existe, ela é diferenciável termo a termo sob o símbolo de soma \sum . Da mesma forma, se $T(\lambda)$ é integrável sob o parâmetro λ , $\partial T(\lambda) / \partial x_i$ também é integrável e

$$\frac{\partial I}{\partial x_i} = \int_{\Lambda} \frac{\partial}{\partial x_i} T(\lambda) [\varphi] d\lambda.$$

A.2 PROPRIEDADES DA “FUNÇÃO” δ

A Delta de Dirac é um exemplo de distribuição definida sob o símbolo de integração, isto é, ela somente tem sentido matemático quando aparece em uma integral. Na física, costuma-se usar a notação $\delta(x - x_0)$, no lugar da notação mais correta $\delta_{x_0}[\varphi]$. Esta notação, contudo, é bastante conveniente no uso prático. Também neste contexto, a $\delta(x - x_0)$ é tratada como uma função governada por regras peculiares; contudo, estas regras estão justificadas pela teoria das distribuições.

A.2.1 DEFINIÇÃO DA δ

Seja $f(x)$ uma função definida no domínio Ω e $x_0 \in \mathbb{R}$,

$$\int_{\Omega} f(x)\delta(x - x_0)dx \equiv \delta_{x_0}[f(x)] = \begin{cases} f(x_0), & x_0 \in \Omega \\ 0, & x_0 \notin \Omega. \end{cases} \quad (\text{A.2a})$$

Formalmente, escreve-se então

$$\delta(x - x_0) = \begin{cases} 0, & x \neq x_0 \\ +\infty, & x = x_0, \end{cases}$$

onde

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = 1. \quad (\text{A.2b})$$

Neste caso, a δ pode ser pensada como a generalização da delta de Kronecker

$$\delta_{nm} = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases}$$

para o caso contínuo.

A.2.2 REPRESENTAÇÕES DA $\delta(x - x_0)$ COMO O LIMITE DO KERNEL DE UM OPERADOR INTEGRAL

A $\delta(x - x_0)$ pode ser considerada como o limite de uma função que possui um máximo estreito e alto em torno de x_0 , e cuja integral sobre todo o espaço permanece constante e igual a 1. Assim, existem as seguintes representações:

$$\delta(x - x_0) = \frac{1}{\pi} \lim_{L \rightarrow \infty} \frac{\text{sen}[L(x - x_0)]}{x - x_0} = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0} \frac{\text{sen}[(x - x_0)/\epsilon]}{x - x_0} \quad (\text{A.3a})$$

$$= \frac{1}{\pi} \lim_{\chi \rightarrow \infty} \frac{1 - \cos \chi(x - x_0)}{\chi(x - x_0)^2} \quad (\text{A.3b})$$

$$= \frac{1}{\pi} \lim_{\epsilon \rightarrow +0} \frac{\epsilon}{(x - x_0)^2 + \epsilon^2} \quad (\text{A.3c})$$

$$= \lim_{\eta \rightarrow 0} \frac{H(x - x_0 + \eta) - H(x - x_0)}{\eta} \quad (\text{A.3d})$$

$$= \lim_{\epsilon \rightarrow 0} (2\pi\epsilon)^{-1/2} e^{-(x-x_0)^2/2\epsilon} \quad (\text{A.3e})$$

onde $H(x)$ é a função de Heaviside ou também denominada função degrau:

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0. \end{cases} \quad (\text{A.4})$$

A $\delta(x - x_0)$ pode ser interpretada como a derivada da função $H(x)$:

$$\delta(x - x_0) = \frac{d}{dx} H(x - x_0). \quad (\text{A.5})$$

A figura A.1 mostra outras representações da $\delta(x - x_0)$ juntamente com gráficos ilustrando a tendência das respectivas funções à medida que o parâmetro $\epsilon \rightarrow 0$.

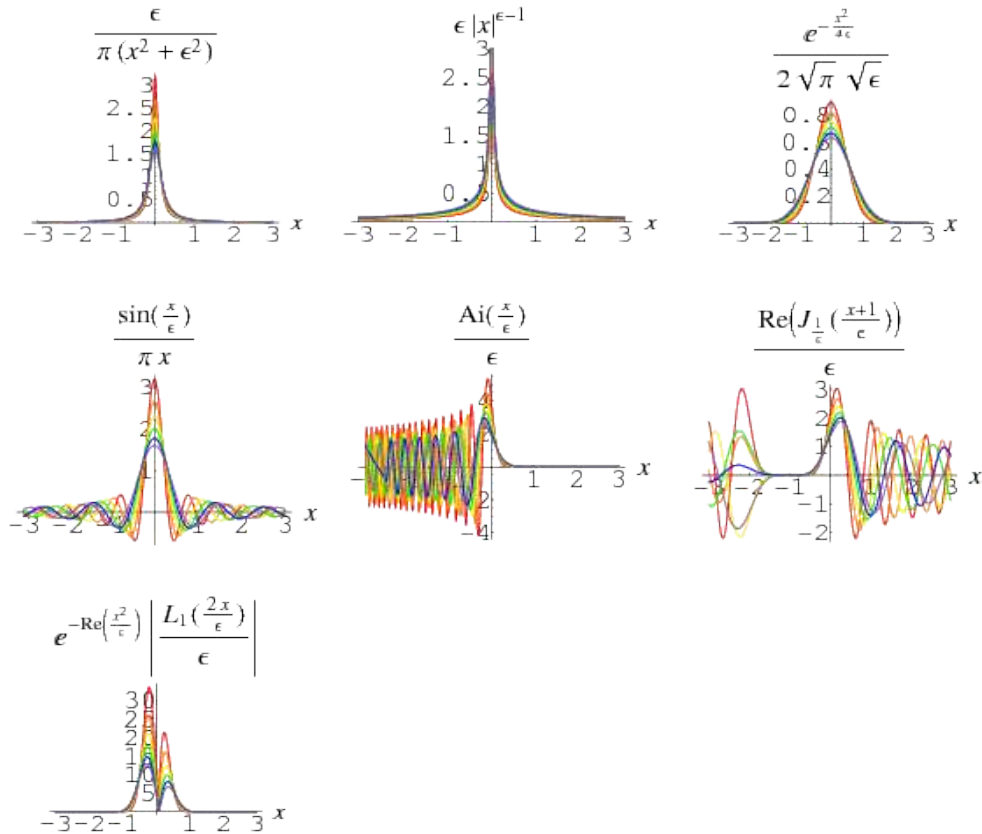


Figura A.1: Outras representações da $\delta(x - x_0)$ junto com gráficos ilustrando a tendência das funções para $\epsilon \rightarrow 0$.

A.2.3 PRINCIPAIS PROPRIEDADES

As principais propriedades da δ são:

$$\delta(x) = \delta(-x) \tag{A.6a}$$

$$\delta(ax) = \frac{1}{|a|} \delta(x), \quad (a = \text{cte.} \neq 0) \tag{A.6b}$$

$$\delta(x^2 - a^2) = \frac{1}{|a|} [\delta(x - a) + \delta(x + a)], \quad (a = \text{cte.} \neq 0) \tag{A.6c}$$

$$\delta[g(x)] = \sum_n \frac{\delta(x - x_n)}{|g'(x_n)|}, \quad \left\{ \begin{array}{l} g(x_n) = 0, \\ g'(x) \neq 0 \end{array} \right\} \tag{A.6d}$$

$$x\delta(x) = 0 \tag{A.6e}$$

$$f(x)\delta(x - a) = f(a)\delta(x - a) \tag{A.6f}$$

$$\int \delta(x - y)\delta(y - a)dy = \delta(x - a) \tag{A.6g}$$

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk. \tag{A.6h}$$

Uma outra propriedade importante, que com freqüência é utilizada no tratamento de funções complexas é a fórmula de Plemelj:

$$\lim_{\epsilon \rightarrow +0} \frac{1}{x - x_0 \pm i\epsilon} = \text{PP} \frac{1}{x - x_0} \mp i\pi\delta(x - x_0). \tag{A.6i}$$

Todas as igualdades apresentadas acima indicam que um lado da equação pode ser substituído pelo outro lado quando a δ for multiplicada por uma função regular e o produto integrado sobre a variável x . Um exemplo de aplicação da propriedade (A.6d) é apresentada abaixo:

$$\delta(x^2 + x - 2) \Rightarrow g(x) = x^2 + x - 2 = (x - 1)(x + 2) \Rightarrow g'(x) = 2x + 1$$

$$\begin{aligned} \Rightarrow x_n &= \{1, -2\}, \quad g'(1) = 3 \text{ e } g'(-2) = -3. \\ \Rightarrow \delta(x^2 + x - 2) &= \frac{1}{3} [\delta(x - 1) + \delta(x + 2)]. \end{aligned}$$

A.2.4 DERIVADAS DA $\delta(x)$

A “função” $\delta(x)$ é diferenciável em todas as ordens. A sua m -ésima derivada $\delta^{(m)}(x)$ é definida pela propriedade

$$\int_{-\infty}^{\infty} \delta^{(m)}(x - x_0) f(x) dx \equiv \delta_{x_0}^{(m)} [f(x)] = \begin{cases} (-1)^m f^{(m)}(x_0), & x_0 \in \Omega \\ 0, & x_0 \notin \Omega. \end{cases}$$

definição esta válida para qualquer função $f(x)$ diferenciável m vezes no ponto $x = x_0 \in \Omega$. A $\delta^{(m)}(x - x_0)$ pode ser considerada como o limite da derivada de ordem m de qualquer das funções dadas em (A.3a-d). As propriedades das derivadas da δ são as seguintes:

$$\delta^{(m)}(x) = (-1)^m \delta^{(m)}(-x) \quad (\text{A.7a})$$

$$\int \delta^{(m)}(x - y) \delta^{(n)}(y - a) dy = \delta^{(m+n)}(x - a) \quad (\text{A.7b})$$

$$x^{m+1} \delta^{(m)}(x) = 0. \quad (\text{A.7c})$$

Em particular a derivada primeira tem as propriedades:

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

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

$$\int \delta'(x - y) \delta(y - a) dy = \delta'(x - a)$$

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

$$x^2 \delta'(x) = 0$$

$$\delta'(x) = \frac{i}{2\pi} \int_{-\infty}^{\infty} k e^{ikx} dk.$$

Como exemplo, a propriedade (A.7a) pode ser obtida integrando-se por partes m vezes o funcional

$$\int_{-\infty}^{\infty} \delta^{(m)}(x) f(x) dx = - \int_{-\infty}^{\infty} \delta^{(m-1)}(x) f'(x) dx = \dots = (-1)^m \int_{-\infty}^{\infty} \delta(x) f^{(m)}(x) dx.$$

A.3 DELTAS DE DIRAC EM MAIS DE UMA DIMENSÃO

Quando o problema envolve duas ou mais dimensões, a delta de Dirac é dada pelo produto de deltas, cada uma delas unidimensional. Representa-se então a delta de Dirac multidimensional, no sistema Cartesiano de coordenadas, da seguinte forma:

$$\delta(\mathbf{r} - \mathbf{R}) = \delta(x - X) \delta(y - Y) \delta(z - Z), \quad (\text{A.8})$$

sendo $\mathbf{R} = (X, Y, Z)$ o vetor posição do ponto no espaço \mathbb{R}^3 que é selecionado pela distribuição $\delta(\mathbf{r} - \mathbf{R})$.

De uma forma genérica, a extensão das definições (A.2a,b) para 3 dimensões é determinada pelas condições:

$$\int_{\mathcal{V}} \delta(\mathbf{r} - \mathbf{R}) d^3r = \begin{cases} 1, & \mathbf{R} \in \mathcal{V} \\ 0, & \mathbf{R} \notin \mathcal{V} \end{cases} \quad (\text{A.9a})$$

$$\int_{\mathcal{V}} F(\mathbf{r}) \delta(\mathbf{r} - \mathbf{R}) d^3r = \begin{cases} F(\mathbf{R}), & \mathbf{R} \in \mathcal{V} \\ 0, & \mathbf{R} \notin \mathcal{V}, \end{cases} \quad (\text{A.9b})$$

sendo \mathcal{V} um subespaço vetorial de \mathbb{R}^3 .

Generalizações para espaços com um número arbitrário de dimensões seguem diretamente destas condições.

A.3.1 COORDENADAS CURVILÍNEAS ORTOGONAIS

Quando se está trabalhando em um sistema de coordenadas curvilíneas ortogonais qualquer, a forma simples dada por (A.8) em geral está incorreta. Ou seja, sendo (q_1, q_2, q_3) o conjunto de três coordenadas curvilíneas ortogonais, as quais se relacionam com o sistema Cartesiano pelas leis de transformação (1.1a-c) e (1.2a-c),

$$q_i = q_i(x_1, x_2, x_3) \iff x_i = x_i(q_1, q_2, q_3), \quad (i = 1, 2, 3),$$

não está correto, em geral, escrever

$$\delta(\mathbf{r} - \mathbf{r}') = \delta(q_1 - q'_1) \delta(q_2 - q'_2) \delta(q_3 - q'_3),$$

pois esta expressão não respeita necessariamente a propriedade (A.9a).

Por outro lado, a expressão

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{\delta(q_1 - q'_1)}{h_1} \frac{\delta(q_2 - q'_2)}{h_2} \frac{\delta(q_3 - q'_3)}{h_3}, \quad (\text{A.10})$$

sendo h_i o fator de escala da coordenada i ,³ dado por

$$h_i = \sqrt{\sum_{j=1}^3 \left(\frac{\partial x_j}{\partial q_i} \right)^2},$$

irá respeitar as propriedades (A.9a,b). O produto $h_1 h_2 h_3$ é também denominado o *Jacobiano* da transformação de coordenadas. Portanto, (A.10) é a forma correta para a delta de Dirac em coordenadas curvilíneas ortogonais. Os exemplos a seguir ilustram situações comuns.

Exemplo A.1 (Coordenadas polares cilíndricas). Nas coordenadas cilíndricas,⁴ $q_1 = \rho$, $q_2 = \varphi$ e $q_3 = z$. Por isso, $h_1 = h_3 = 1$ e $h_2 = \rho$. Assim,

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{\rho} \delta(\rho - \rho') \delta(\varphi - \varphi') \delta(z - z').$$

Exemplo A.2 (Coordenadas polares esféricas). Nas coordenadas esféricas,⁵ $q_1 = r$, $q_2 = \theta$ e $q_3 = \varphi$. Os fatores de escala são $h_1 = r^2$, $h_2 = \sin \theta$ e $h_3 = 1$. Portanto,

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{r^2 \sin \theta} \delta(r - r') \delta(\theta - \theta') \delta(\varphi - \varphi').$$

Assim, sendo \mathbf{r}' um ponto contido em \mathcal{V} , (A.9a) será respeitada, pois

$$\int_{\mathcal{V}} \delta(\mathbf{r} - \mathbf{r}') d^3 r = \int_{\mathcal{V}} \frac{1}{r^2 \sin \theta} \delta(r - r') \delta(\theta - \theta') \delta(\varphi - \varphi') r^2 \sin \theta dr d\theta d\varphi = 1.$$

Uma outra forma utilizada com frequência é

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{r^2} \delta(r - r') \delta(\cos \theta - \cos \theta') \delta(\varphi - \varphi').$$

A.3.2 PONTOS DEGENERADOS

A expressão (A.10) assume que \mathbf{r}' não é um ponto degenerado, isto é, não é caracterizado por uma variação contínua em uma ou mais coordenadas. Em algumas situações, o “ponto” degenerado pode ser uma curva ou uma superfície em 3D. Exemplos de pontos degenerados são: a origem num sistema plano-polar (caracterizado por $r = 0$ e qualquer valor de $0 \leq \theta \leq 2\pi$), a origem em um sistema curvilíneo em 3D ($r = 0$, $0 \leq \theta \leq \pi$, $0 \leq \varphi \leq 2\pi$ em coordenadas esféricas), o eixo z em coordenadas cilíndricas ($\rho = 0$, $0 \leq \varphi \leq 2\pi$).

Suponha, então, que a coordenada q_1 assume todos os valores no intervalo $q_{11} < q_1 < q_{12}$. Neste caso, a representação correta de $\delta(\mathbf{r} - \mathbf{r}')$ não é (A.10), uma vez que a coordenada q_1 não

³Vide seção 1.1.

⁴Vide seção 1.5.1.

⁵Vide seção 1.5.2.

mais possui um único valor para q'_1 (sua multiplicidade é coberta pela variação de q_1 dentro de um intervalo). Neste caso, a propriedade (A.9a) é novamente respeitada se

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{\delta(q_2 - q'_2) \delta(q_3 - q'_3)}{\left(\int_{q_{11}}^{q_{12}} h_1 dq_1\right) h_2 h_3} H(q_1 - q_{11}) H(q_{12} - q_1), \quad (\text{A.11a})$$

sendo $H(x)$ a função de Heaviside (A.4), pois (A.9a) fica

$$\int_{\mathcal{V}} \delta(\mathbf{r} - \mathbf{r}') d^3r = \int_{q_{11}}^{q_{12}} h_1 dq_1 \iint h_2 h_3 dq_2 dq_3 \frac{\delta(q_2 - q'_2) \delta(q_3 - q'_3)}{\left(\int_{q_{11}}^{q_{12}} h_1 dq_1\right) h_2 h_3} = 1,$$

desde que o intervalo $[q_{11}, q_{12}]$ seja um subespaço de \mathcal{V} .

Por exemplo, considerando um problema em coordenadas esféricas que possua simetria azimutal, então a coordenada φ deve ser eliminada, pois a mesma será multiplamente definida. Neste caso,

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{r^2 \sin \theta \int_0^{2\pi} d\varphi} \delta(r - r') \delta(\theta - \theta') = \frac{1}{2\pi r^2 \sin \theta} \delta(r - r') \delta(\theta - \theta').$$

Da mesma maneira, se os pontos multiplamente definidos estiverem sobre a superfície $q_1 \times q_2$, estando a coordenada q_2 no intervalo $q_{21} < q_2 < q_{22}$, então

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{\delta(q_3 - q'_3)}{\left(\int_{q_{11}}^{q_{12}} \int_{q_{21}}^{q_{22}} h_1 h_2 dq_1 dq_2\right) h_3} H(q_1 - q_{11}) H(q_{12} - q_1) H(q_2 - q_{21}) H(q_{22} - q_2). \quad (\text{A.11b})$$

Entretanto, as representações (A.11a) e (A.11b) supõe que a degenerescência é uniformemente distribuída nos intervalos considerados. Esta suposição também não é válida em geral e a representação correta deve ser obtida, neste caso, por uma análise cuidadosa acerca da propriedade físico-matemática que está sendo descrita em termos da delta de Dirac. O exercício a seguir ilustra tal situação.

Exercício A.1 (Densidade linear de carga elétrica). Considere uma distribuição contínua de cargas elétricas ao longo de um anel circular de raio a . A distribuição de cargas não é uniforme, de modo que a densidade linear de cargas é descrita por uma função $\lambda = \lambda(\phi)$ C/m ($0 \leq \phi \leq 2\pi$), sendo ϕ a coordenada angular azimutal dos sistemas cilíndrico (seção 1.5.1) ou esférico (seção 1.5.2). Calcule a carga total depositada ao longo do anel e determine a forma adequada para a função $\rho = \rho(r)$, a qual descreve a densidade volumétrica de carga.

Resolução. Posicionando-se o sistema de coordenadas de tal forma que a origem está no centro geométrico do anel e o plano $x - y$ é o plano do anel, obtém-se o arranjo ilustrado na figura A.2. Como $\rho = a$, se a carga estivesse uniformemente distribuída ao longo do anel, então $\lambda = \text{cte.}$ e $Q = 2\pi a \lambda$. Neste caso, a coordenada ϕ é uniformemente degenerada na coordenada $0 \leq \phi \leq 2\pi$ e a densidade volumétrica de carga é dada diretamente de (A.11a) como

$$\rho(\mathbf{r}) = Q \delta(\mathbf{r} - \mathbf{r}_{\text{anel}}) = Q \frac{\delta(r_{\perp} - a) \delta(z)}{\int_0^{2\pi} r d\phi} = \frac{Q}{2\pi r} \delta(r_{\perp} - a) \delta(z) = \frac{Q}{2\pi a} \delta(r_{\perp} - a) \delta(z)$$

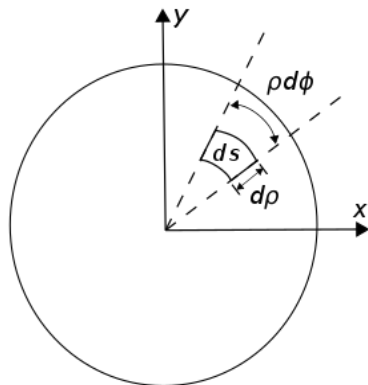


Figura A.2: Anel de raio a contendo uma distribuição não uniforme de cargas elétricas.

em coordenadas cilíndricas, ou

$$\rho(r) = Q \frac{\delta(r-a) \delta(\theta - \pi/2)}{\left(\int_0^{2\pi} r \sin \theta d\phi\right) r} = \frac{Q \delta(r-a) \delta(\theta - \pi/2)}{2\pi r^2 \sin \theta} = \frac{Q \delta(r-a) \delta(\theta - \pi/2)}{2\pi a^2 \sin \theta}$$

em coordenadas esféricas.

Porém, como a distribuição $\lambda = \lambda(\phi)$ não é uniforme, a degenerescência na coordenada ϕ também não o é, o que invalida o método recém empregado. Neste caso, para calcular corretamente a carga total sobre o anel, considera-se inicialmente o elemento de carga $dq(\phi)$ contido no elemento de arco $ad\phi$, situado entre ϕ e $\phi + d\phi$, o qual será dado por $dq = a\lambda(\phi) d\phi$. Portanto, a carga total depositada ao longo do anel será obtida por

$$Q = \int_0^{2\pi} dq(\phi) = a \int_0^{2\pi} \lambda(\phi) d\phi.$$

Para determinar agora a forma correta para a densidade volumétrica de carga $\rho(r)$ em coordenadas cilíndricas, observa-se que o resultado acima será reproduzido se

$$\rho(r_{\perp}, \phi, z) = \lambda(\phi) \delta(r_{\perp} - a) \delta(z) \implies \int \rho(r_{\perp}, \phi, z) d^3r = a \int_0^{2\pi} \lambda(\phi) d\phi = Q.$$

Em coordenadas esféricas, o mesmo resultado será obtido se

$$\rho(r, \theta, \phi) = \lambda(\phi) \frac{\delta(r-a) \delta(\theta - \pi/2)}{r \sin \theta} = \lambda(\phi) \frac{\delta(r-a) \delta(\theta - \pi/2)}{a \sin \theta}.$$

Nesta última expressão, nota-se que foi necessário introduzir o fator de escala da coordenada ϕ para que $\rho(r, \theta, \phi)$ descreva corretamente a distribuição das cargas elétricas.

REFERÊNCIAS

MESSIAH, Albert. **Quantum Mechanics: Two Volumes Bound as One**. Mineola: Dover, 1999. 1136 + xxii pp. ISBN 0-486-40924-4.