Universidade Estadual Paulista "Júlio de Mesquita Filho"
Instituto de Geociências e Ciências Exatas
Câmpus de Rio Claro

Rotações no espaço tridimensional por meio de produtos quaterniônicos

Aline de Freitas Moroni

Dissertação apresentada ao Programa de Pós-Graduação em Matemática como requisito parcial para a obtenção do grau de Mestre

Orientador
Prof. Dr. Thiago de Melo

2016

512 Moroni, Aline de Freitas
M868r Rotações no espaço tridimensional por meio de produtos quaterniônicos/ Aline de Freitas Moroni. - Rio Claro, 2016.
53 f.: il.,figs.

Dissertação (mestrado) - Universidade Estadual Paulista, Instituto de Geociências e Ciências Exatas.
Orientador: Thiago de Melo

1. Álgebra. 2. Matriz de rotação. 3. Números reais. I. Título

Ficha Catalográfica elaborada pela STATI - Biblioteca da UNESP
Câmpus de Rio Claro/SP

TERMO DE APROVAÇÃO

Aline de Freitas Moroni

ROTAÇÕES NO ESPAÇO TRIDIMENSIONAL POR MEIO DE PRODUTOS QUATERNIÔNICOS

Dissertação APROVADA como requisito parcial para a obtenção do grau de Mestre no Curso de Pós-Graduação em Matemática do Instituto de Geociências e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", pela seguinte banca examinadora:

Prof. Dr. Thiago de Melo
Orientador

Profa. Dra. Alice Kimie Miwa Libardi
IGCE–Unesp

Profa. Dra. Érika Capelato
FCLAR–Unesp

Rio Claro, 20 de abril de 2016

Agradecimentos

Inicio meus agradecimentos por Deus, já que ele colocou pessoas especiais a meu lado, que diretamente ou indiretamente me ajudaram no decorrer deste trabalho.

Agradeço ao Prof. Dr. Thiago de Melo, pela sua orientação, paciência, disponibilidade, e saber transmitido. Muito Obrigada!

Agradeço às Profas. Dras. Alice Kimie Miwa Libardi e Érika Capelato, que aceitaram compor minha banca de qualificação e de defesa, pelas sugestões e análises significativas.

Agradeço aos professores e colegas da turma do mestrado, pelos momentos de orientação e de discussão sobre o conteúdo.

Aos companheiros de viagem (Birigui à Rio Preto e Rio Preto à Rio Claro), agradeço todos os momentos de incentivo e auxílio. Em particular, agradeço à Débora e à Marina pelo companheirismo nessa trajetória.

Aos meus pais, José e Valdeci, meu infinito agradecimento. Pelo apoio incondicional e incentivo demonstrados ao longo desta caminhada. A eles dedico este trabalho!

Resumo

Neste trabalho pretendemos descrever o processo de construção da álgebra dos quatérnios, e a interpretação da multiplicação desses objetos via rotações no espaço. Para isto, vimos a necessidade de iniciar com conceitos que formam a base da álgebra, listando axiomas para o sistema de números reais e complexos.

Palavras-chave: Álgebra, Matriz de rotação, Números reais.

Abstract

The aim of this work is to describe the construction of the quaternion algebra and to interpret the multiplication operation via tridimensional rotations. For that we begin with basic algebraic concepts, and we list the axioms for the real and complex number systems.

Keywords: Algebra, Rotation matrix, Real numbers.

Sumário

1 Introdução **11**

2 Preliminares **13**
2.1 Álgebra 13
2.2 Álgebra Linear 15

3 Números Reais e Complexos **19**
3.1 Números Reais 19
3.2 Números Complexos 29

4 Quatérnios e Rotações **33**
4.1 Quatérnios 33
4.2 Rotações por meio de quatérnios 41
4.2.1 Quatérnios e rotações tridimensionais 43

Referências **53**

1 Introdução

A álgebra dos quatérnios teve origem nos números complexos, no início do século XIX, quando William Rowan Hamilton (1805–1865) apresentou o primeiro conceito moderno dos complexos como pares ordenados de reais e tentou generalizar esta ideia para o espaço tridimensional. Após inúmeras tentativas, verificou-se que não era possível a existência de um complexo tridimensional. Com isso, Hamilton descobriu os quatérnios, com uma álgebra de dimensão quatro sobre o corpo dos números reais e que possui todas as propriedades de um corpo, exceto a comutatividade da multiplicação. A álgebra dos quatérnios foi a primeira álgebra não comutativa da história.

Nessa dissertação queremos mostrar que da mesma forma que a multiplicação complexa representa rotações bidimensionais, existe uma estrutura de quatro dimensões que também com uma multiplicação, representa rotações tridimensionais.

Dividiremos este trabalho em quatro capítulos. No segundo capítulo, exibiremos definições da Álgebra que serão essenciais para o desenvolvimento dos capítulos seguintes.

No terceiro capítulo, lidamos com conceitos que formam a base da Álgebra, definindo a estrutura dos sistemas de números reais e complexos, e listamos os axiomas para tais sistemas, bem como derivamos algumas consequências.

No quarto capítulo, inicialmente, daremos uma motivação para uma estrutura tridimensional, e as tentativas do matemático, físico e astrônomo Sir William Rowan Hamilton (1805–1865), em representar vetores não só no $\mathbb{R}^2$. Hamilton havia conseguido operar vetores no plano através de uma identificação dos complexos da forma $a + bi$ com os pares ordenados de reais da forma (a, b) e aspirava, em meados do século XIX, encontrar uma Álgebra que representasse e operasse vetores em dimensões maiores. Generalizando de alguma forma os complexos, ele desejava fornecer a $\mathbb{R}^3$ uma estrutura algébrica que permitisse operar vetores, para que tivessem aplicabilidade no espaço tridimensional.

Como os números complexos possuem duas partes, uma parte que é um número real e outra parte que é "imaginária", o primeiro pensamento de Hamilton foi de que ele precisava de uma componente "imaginária" adicional. Ele lutou durante anos tentando provar a possibilidade de um complexo ser tridimensional, mas não conseguiu. Foi quando, em 1843, Hamilton encontrou a solução do problema, inventando um sistema

com três partes imaginárias que se tornou a álgebra de quatérnios.

Escritas revelam que a brilhante ideia lhe ocorreu quando estava andando em Dublin, ao longo do Canal Real, a caminho de uma reunião da Real Academia Irlandesa. No momento sua emoção foi tanta em encontrar finalmente um resultado, que o medo de entrar em colapso e morrer antes que ele tivesse a chance de dizer a alguém, o impulsionou a registrar a descoberta em uma pedra de Brougham Bridge, ao lado da ponte arqueia sobre o canal ao longo do qual ele estava andando, e então, talhando com uma faca registrou a fórmula fundamental com os símbolos i, j, k; $i^2 = j^2 = k^2 = kij = -1$. Com o tempo, a escultura de Hamilton desapareceu e uma placa foi depois colocada sobre o local.

Definido o sistema dos quatérnios, trabalharemos com a sua álgebra e propriedades. Ainda no capítulo 4, em um segundo momento, examinaremos a aplicação de quatérnios em rotações, e a forma como estes representam rotações tridimensionais, em particular, as vantagens quando trabalhamos com essas rotações que permitem três graus de liberdade, proporcionando rotações suaves quando comparadas às rotações por ângulos de Euler.

2 Preliminares

Neste capítulo apresentaremos alguns conceitos básicos para o entendimento do trabalho. Noções de anéis e grupos bem como algumas propriedades estão presentes na primeira seção. Já na segunda seção, encontramos algumas ferramentas da Álgebra Linear.

2.1 Álgebra

Muitos dos resultados apresentados aqui, bem como conteúdo adicional, podem ser encontrados em [1, 2], por exemplo.

Definição 2.1. *Um anel comutativo K com identidade é um conjunto não vazio com pelo menos dois elementos, munido de uma operação denotada por $+$, chamada adição, e de uma operação denotada por $\cdot$, chamada multiplicação, satisfazendo às seguintes condições:*

1. $\forall a, b, c \in K, (a + b) + c = a + (b + c)$.
2. $\exists 0 \in K$ *tal que* $a + 0 = 0 + a = a$, $\forall a \in K$.
3. $\forall a \in K$, $\exists \alpha \in K$, *tal que* $a + \alpha = \alpha + a = 0$.
4. $\forall a, b \in K$, $a + b = b + a$.
5. $\forall a, b, c \in K$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
6. $\exists 1 \in K$, *tal que* $a \cdot 1 = 1 \cdot a = a$, $\forall a \in K$.
7. $\forall a, b, c \in K$, $a \cdot (b + c) = a \cdot b + a \cdot c$ *e* $(a + b) \cdot c = a \cdot c + b \cdot c$.
8. $\forall a, b \in K$, $a \cdot b = b \cdot a$.

Definição 2.2. *Um anel comutativo com identidade $(K, +, \cdot)$ é chamado corpo se todo elemento de K diferente de zero possuir um inverso com respeito a multiplicação.*

Por exemplo, o conjunto $\mathbb{C}$ dos números complexos munido da soma e produto usuais de números complexos, é um corpo.

Vale observar que $+$ e $\cdot$ são meros símbolos para as operações binárias utilizadas. O fato de escrevermos '$a + b$' não significa que estamos trabalhando com adição de números, pois o conjunto em questão pode ser conjunto defunções, de matrizes, de polinômios, entre outros.

Se o conjunto possui apenas uma operação binária sob certas condições, temos a noção de grupo, que é uma estrutura algébrica mais simples quando comparada com a estrutura de anel. Mais precisamente, temos:

Definição 2.3. *Um grupo G é um conjunto não vazio, munido de uma operação binária $+$ que a cada par de elementos x, y de G associa um elemento de G, indicado por $x+y$ possuindo as seguintes propriedades:*

1. *Quaisquer que sejam x, y, z em G, vale a associatividade, isto é, $(x + y) + z = x + (y + z)$.*
2. *Existe um elemento e de G tal que $e + x = x + e = x$, para todo $x \in G$.*
3. *Se x é um elemento de G, então existe um elemento y de G tal que $x + y = y + x = e$.*

Se além disso a operação $+$ for comutativa, isto é, se

4. *$x + y = y + x$ para quaisquer $x, y \in G$*

o grupo G é chamado grupo abeliano.

Pode-se provar que o elemento y proveniente do item 3 acima é único, para cada x dado. Assim, será denotado por $-x$ e chamado de *elemento oposto de x*.

Definição 2.4. *Sejam G um grupo e H um subconjunto de G. Dizemos que H é um subgrupo de G se quaisquer que sejam $x, y \in H$, então $x + y$ e $-x$ também pertencem a H.*

Nestas condições, se H é subgrupo de G, H é ele próprio um grupo quando a operação em H for a mesma de G.

Definição 2.5. *Sejam G um grupo com a operação $*$ e G' um grupo com a operação $\triangle$. Um homomorfismo de G para G' é uma função $f : G \longrightarrow G'$ tal que para todos $x, y \in G$, $f(x * y) = f(x) \triangle f(y)$. Se f é homomorfismo e bijetora, chamamos de isomorfismo e dizemos que os grupos são isomorfos.*

2.2 Álgebra Linear

Algumas referências para os assuntos abordados nesta seção são [3, 4].

Uma outra estrutura algébrica que pode ser posta em um conjunto é a estrutura de espaço vetorial, apresentada a seguir. Tal estrutura engloba a estrutura de grupos, porém carrega uma nova operação que necessita de um corpo, chamado de corpo de escalares. Mais precisamente, temos:

Definição 2.6. *Um espaço vetorial V sobre um corpo K é um conjunto não vazio de elementos que podem ser somados entre si, e multiplicados por elementos de K, de tal maneira que a soma de dois elementos de V é novamente um elemento de V, o produto de um elemento de V por um elemento de K é um elemento de V, e tal que possui as propriedades:*

1. *Dados u, v, w elementos de V, vale:*

$$(u+v)+w = u+(v+w).$$

2. *Quaisquer que sejam u e v de V, vale*

$$u+v = v+u.$$

3. *Existe um elemento de V, indicado por 0, tal que*

$$0+u = u \ (= u+0 \text{ por } 2\ )$$

   qualquer que seja u em V.

4. *Dado um elemento u de V, o elemento $(-1)u$ é tal que*

$$u+(-1)u = 0,$$

   onde $-1 \in K$ é o oposto de $1 \in K$.

5. *Se $c \in K$, então*

$$c(u+v) = cu+cv.$$

6. *Se a e b são dois elementos de K, então*

$$(a+b)v = av+bv.$$

7. *Se a e b são dois elementos de K, então*

$$(ab)v = a(bv).$$

8. *Qualquer que seja u de V, vale*
$$1u = u,$$
onde $1 \in K$.

Os elementos de um espaço vetorial são chamados de *vetores* e os elementos do corpo são chamados de *escalares*.

Uma das principais noções para espaços vetoriais é a de base. Para isso, precisamos lidar com conjunto de geradores e combinações lineares, ou seja, soma de vetores.

Dizemos que um conjunto de vetores $v_1, v_2, \ldots, v_n$ de V gera o espaço V se todo vetor de V pode ser escrito como combinação linear de $v_1, v_2, \ldots, v_n$, ou seja, para qualquer $v \in V$ existem escalares $a_1, \ldots, a_n \in K$ tais que

$$v = a_1v_1 + a_2v_2 + \cdots + a_nv_n.$$

Outra noção importante é a de (in)dependência linear, definida a seguir.

Definição 2.7. *Dizemos que o conjunto $v_1, v_2, \ldots, v_n$ de vetores de V é linearmente dependente se existem escalares $a_1, a_2, \ldots, a_n$ não todos nulos tais que*

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0.$$

Caso contrário, dizemos que o conjunto é linearmente independente, e neste caso, a única solução para a equação acima é a solução trivial onde todos os escalares $a_1, a_2, \ldots, a_n$ são nulos.

Utilizando conjunto de geradores e independência linear, podemos introduzir *base* para um espaço vetorial.

Definição 2.8. *Definimos uma base de V sobre K como um conjunto linearmente independente $\{v_1, \ldots, v_n\}$ de vetores de V que gera V.*

É possível provar que todo espaço vetorial não nulo possui uma base (ver, por exemplo, [2, III.5]).

Se V é um espaço vetorial tendo uma base constituída por n elementos, dizemos que n é a dimensão de V, ou que V é n-dimensional. Se V for formado unicamente pelo elemento 0, então V não possui base, e neste caso dizemos que V tem dimensão 0.

Definição 2.9. *Seja V um espaço vetorial sobre um corpo K. Um produto escalar sobre V é uma regra que a cada par v, w de elementos de V associa um escalar, indicado por $\langle v, w\rangle$ ou também por $v \cdot w$, verificando as propriedades seguintes:*

1. $\langle v, w\rangle = \langle w, v\rangle$ *para todos* $v, w \in V$.

2. *Se u, v, w são elementos de V, então*
$$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle.$$

3. *Se $x \in K$, então*
$$\langle xu, v\rangle = x\langle u, v\rangle = \langle u, xv\rangle,$$
para quaisquer $u, v \in V$.

Vale observar que um produto escalar é diferente de fazer o produto por um escalar. Como funções, o primeiro é $V \times V \to K$, já o segundo, $K \times V \to V$.

Definição 2.10. *Seja V um espaço vetorial sobre $\mathbb{R}$, munido de um produto escalar. Diremos que este produto escalar é positivo definido se $\langle v, v\rangle \geq 0$ para todo $v \in V$, e $\langle v, v\rangle > 0$ se $v \neq 0$.*

Terminamos esta seção com outro importante conceito, justificando o termo Álgebra Linear, que é o de *transformação linear*.

Definição 2.11. *Sejam dois espaços vetoriais V e V' sobre um corpo K. Uma transformação linear $F : V \to V'$ é uma função que satisfaz às seguintes condições:*

1. *Quaisquer que sejam u e v em V,*
$$F(u + v) = F(u) + F(v).$$

2. *Quaisquer que sejam c em K e v em V,*
$$F(cv) = cF(v).$$

Observamos que o item 1 acima é o mesmo que $F : V \to V'$ ser um homomorfismo de grupos.

Definição 2.12. *Seja V um espaço vetorial de dimensão finita sobre $\mathbb{R}$, munido de um produto escalar positivo definido. Seja $A : V \to V$ uma transformação linear. Diremos que A é uma transformação ortogonal se*
$$\langle Av, Aw\rangle = \langle v, w\rangle,$$
para quaisquer $v, w \in V$.

Definição 2.13. *Sejam V e W espaços vetoriais sobre um corpo K. Dizemos que $G : V \times W \to K$ é bilinear se:*

1. *Para quaisquer $v_1, v_2 \in V$ e $w \in W$,*
$$G(v_1 + v_2, w) = G(v_1, w) + G(v_2, w)$$
e para quaisquer $v \in V$ e $w_1, w_2 \in W$, vale
$$G(v, w_1 + w_2) = G(v, w_1) + G(v, w_2).$$

2. *Para quaisquer $c \in K$, $v \in V$ e $w \in W$, vale*
$$G(cv, w) = cG(v, w) = G(v, cw).$$

Nestas condições, podemos dizer que uma aplicação bilinear é tal que, para todo $v \in V$, a aplicação $G_v : w \mapsto G(v, w)$ é linear; e para todo $w \in W$, a aplicação $G_w : v \mapsto G(v, w)$ é linear.

3 Números Reais e Complexos

Neste capítulo, vamos listar alguns axiomas para o sistema de números reais, e tirar as primeiras consequências de que vamos precisar. Existem outras propriedades dos números reais que não são dedutíveis a partir destes. No entanto, não serão abordadas neste texto. A seguir, apresentaremos o corpo dos números complexos e algumas de suas propriedades.

A seção a seguir é um estudo detalhado de [5, Chapter I].

3.1 Números Reais

Assumimos, então, que para o conjunto $\mathbb{R}$ dos números reais, temos:

1. Uma função $+ : (x, y) \to x + y$ de $\mathbb{R} \times \mathbb{R}$ em $\mathbb{R}$;
2. Uma função $\cdot : (x, y) \to x \cdot y$ de $\mathbb{R} \times \mathbb{R}$ em $\mathbb{R}$;
3. Uma relação $\leq$ entre os elementos de $\mathbb{R}$.

Além disso, vamos supor que essas funções possuem as seguintes propriedades:

(P1) Para quaisquer $x, y \in \mathbb{R}$, $x + y = y + x$;

(P2) Para quaisquer $x, y, z \in \mathbb{R}$, $x + (y + z) = (x + y) + z$;

(P3) Existe $\xi \in \mathbb{R}$ tal que, para todo $x \in \mathbb{R}$, $\xi + x = x$;

(P4) Para cada $x \in \mathbb{R}$, existe um elemento $x' \in \mathbb{R}$, tal que $x + x' = \xi$;

(P5) Para quaisquer $x, y \in \mathbb{R}$, $x \cdot y = y \cdot x$;

(P6) Para quaisquer $x, y, z \in \mathbb{R}$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$;

(P7) Existe $\zeta \neq \xi$ em $\mathbb{R}$ tal que, para todo $x \in \mathbb{R}$, $\zeta \cdot x = x$;

(P8) Para cada $x \neq \xi$ em $\mathbb{R}$, existe $x'' \in \mathbb{R}$, tal que $x \cdot x'' = \zeta$;

(P9) Para quaisquer $x, y, z \in \mathbb{R}$, $x \cdot (y + z) = x \cdot y + x \cdot z$;

(P10) As relações $x \leq y$ e $y \leq z$ implicam $x \leq z$;

(P11) As relações $x \leq y$ e $y \leq x$ implicam $x = y$;

(P12) Para quaisquer dois elementos x, y em $\mathbb{R}$, ocorre $x \leq y$ ou $y \leq x$;

(P13) A relação $x \leq y$ implica $x + z \leq y + z$, para qualquer $z \in \mathbb{R}$;

(P14) As relações $\xi \leq x$ e $\xi \leq y$ implicam $\xi \leq x \cdot y$;

(P15) Dado um polinômio $P(x) = \alpha \cdot x^3 + \beta \cdot x^2 + \gamma \cdot x + \delta$ com $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, se acontece que $P(w) \leq \xi$ e $\xi \leq P(t)$, quando w, t são elementos de $\mathbb{R}$ tais que $w \leq t$, então existe z em $\mathbb{R}$, tal que $w \leq z \leq t$ e $P(z) = \xi$.

Observação 3.1. As propriedades (P1–P4) garantem que $\mathbb{R}$ com a operação $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto x + y$ é grupo abeliano, ou seja, um grupo no qual a operação é comutativa. Ao adicionarmos (P5–P9) exceto (P8), obtemos uma estrutura de anel comutativo com identidade. Por fim, incluindo (P8) obtemos um corpo.

Proposição 3.1. *O elemento ξ de $\mathbb{R}$ tal que $\xi + x = x$ para todo $x \in \mathbb{R}$ é único.*

Demonstração. De fato, se ξ' é um segundo elemento de $\mathbb{R}$ tal que $\xi' + x = x$, para todo x em $\mathbb{R}$, em particular, para o elemento ξ garantido por (P3), temos que

$$\xi = \xi' + \xi = \xi + \xi' = \xi'.$$

Sendo assim, é único o elemento ξ de $\mathbb{R}$ que satisfaz (P3), chamado de "zero" e denotado por 0 a partir deste ponto. □

Proposição 3.2. *Em $\mathbb{R}$, se $x + y = x + z$ então $y = z$.*

Demonstração. De fato, para cada x, seja x' o elemento dado por (P4). Pelas propriedades (P1–P4), temos que

$$x' + (x + y) = (x' + x) + y = (x + x') + y = 0 + y = y$$

e

$$x' + (x + z) = (x' + x) + z = (x + x') + z = 0 + z = z.$$

Como $x' + (x + y) = x' + (x + z)$, o resultado segue. □

Em particular, para cada $x \in \mathbb{R}$, o elemento x' tal que $x + x' = 0$ é único. De fato, se existe x'' tal que $x + x'' = 0$, então $x + x' = 0 = x + x''$, e pela Proposição 3.2, $x' = x''$. Já que tal elemento é único, é denotado por $-x$ e chamado *o oposto* de x.

Proposição 3.3. *Em $\mathbb{R}$, $x \cdot y = 0$ se, e somente se, $x = 0$ ou $y = 0$.*

Demonstração. De fato, pelas propriedades (P9), (P1) e (P3), temos

$$x \cdot 0 + x \cdot x = x \cdot (0 + x) = x \cdot x,$$

isto é, $x \cdot 0 + x \cdot x = 0 + x \cdot x$. Pela Proposição 3.2

$$x \cdot 0 = 0.$$

Por (P5)

$$0 \cdot x = 0.$$

Se agora,

$$x \cdot y = 0,$$

com $x \neq 0$, seja x'' o elemento dado por (P8), e em virtude de (P5), temos:

$$x'' \cdot (x \cdot y) = (x'' \cdot x) \cdot y = (x \cdot x'') \cdot y = \zeta \cdot y = y.$$

E por outro lado,

$$x'' \cdot (x \cdot y) = x'' \cdot 0 = 0 \cdot x'' = 0,$$

o que mostra que $y = 0$. □

A Proposição 3.3 diz que $\mathbb{R}$ não tem divisor de zero, ou seja, se $x \neq 0$ e $y \neq 0$ então $x \cdot y \neq 0$.

Vamos denotar o conjunto complementar de $\{0\}$, em $\mathbb{R}$, por $\mathbb{R}^*$, isto é, $\mathbb{R}^* = \mathbb{R} - \{0\}$. Então, seguem da Proposição 3.3 que a restrição à $\mathbb{R}^* \times \mathbb{R}^*$ da função $\cdot : (x, y) \to x \cdot y$ assume valores em $\mathbb{R}^*$ e que não podemos ter $x'' = 0$ em (P8) (isto é, 0 não é o inverso de nenhum elemento). As condições (P5–P8) garantem que $\mathbb{R}^*$ com a operação $\cdot$ é um grupo abeliano, chamado *grupo multiplicativo* de números diferentes de zero.

Pela Definição 2.3 e observando que nas Proposições 3.1 e 3.2 apenas as propriedades (P1–P4) são usadas, podemos demonstrar as duas propriedades seguintes:

Proposição 3.4. *Há um único elemento em $\zeta \in \mathbb{R}$ tal que $\zeta \cdot x = x$ para todo $x \in \mathbb{R}$.*

Demonstração. A existência de ζ é garantida por (P7). Supondo que exista ζ', tal que $\zeta' \cdot x = x$ para todo $x \in \mathbb{R}$, temos que

$$\zeta = \zeta' \cdot \zeta = \zeta \cdot \zeta' = \zeta'.$$ □

Sendo assim, é único o elemento ζ que satisfaz (P7), que portanto será chamado de *identidade* e será escrito 1 a partir deste ponto.

Proposição 3.5. *Se $x \neq 0$ e $x \cdot y = x \cdot z$ então $y = z$.*

Demonstração. De fato, temos pelas propriedades (P5–P8), que

$$x \cdot y = x \cdot z \Rightarrow x'' \cdot (x \cdot y) = x'' \cdot (x \cdot z) \Rightarrow$$

$$(x'' \cdot x) \cdot y = (x'' \cdot x) \cdot z \Rightarrow \zeta \cdot y = \zeta \cdot z \Rightarrow y = z.$$ □

Para cada $x \neq 0$ em $\mathbb{R}$, o elemento x'' tal que $x'' \cdot x = 1$ é único, chamado de *inverso* de x e escrito como x^{-1}. Como $0 \cdot z = 0$, para $z \in \mathbb{R}$, usando a Proposição 3.3 e o fato de que $1 \neq 0$, vemos que o elemento 0 não tem inverso em $\mathbb{R}$. Para mostrarmos a unicidade de x'', vamos supor que exista x''' tal que $x \cdot x''' = 1$, desta forma $x \cdot x'' = 1 = x \cdot x'''$, e pela Proposição 3.5 $x'' = x'''$.

Proposição 3.6. *Se $x \neq 0$ e $y \neq 0$, temos $(x \cdot y)^{-1} = y^{-1} \cdot x^{-1}$.*

Demonstração. De fato, $(y^{-1} \cdot x^{-1}) \cdot (x \cdot y) = y^{-1} \cdot (x^{-1} \cdot x) \cdot y = y^{-1} \cdot 1 \cdot y = y^{-1} \cdot y = 1$. O resultado segue por unicidade do inverso e comutatividade. □

Proposição 3.7. *Para quaisquer $x, y \in \mathbb{R}$, $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$.*

Demonstração. De fato, para mostrarmos $-(x \cdot y) = (-x) \cdot y$, fazemos

$$x \cdot y + (-x) \cdot y = (x + (-x)) \cdot y = 0 \cdot y = 0 = x \cdot y + (-(x \cdot y)),$$

e para mostrarmos $-(x \cdot y) = x \cdot (-y) =$ fazemos

$$x \cdot y + x \cdot (-y) = x \cdot (y + (-y)) = x \cdot 0 = 0 = x \cdot y + (-(x \cdot y)).$$ □

Em particular, $(-1) \cdot y = -(1 \cdot y) = -y$, para cada $y \in \mathbb{R}$.

Definição 3.1. *A relação "$x \leq y$ e $x \neq y$" entre dois elementos x, $y \in \mathbb{R}$ é também escrita como "$x < y$" ou "$y > x$". A relação "$x \leq y$" (respectivamente "$x \geq y$", "$x < y$", "$x > y$") é lida como "y é maior que ou igual a x" (respectivamente "x é maior que ou igual a y", "x é estritamente menor que y", "x é estritamente maior que y").*

Uma relação entre x e y ($x \leq y$) que satisfaz simultaneamente (P10–P12), é chamada de *relação de ordem total.* Se um conjunto E tem uma relação $R \subset E \times E$ que satisfaz (P10–P12) (com $x \leq y$ denotado por $R(x, y)$) dizemos que o conjunto E é *totalmente ordenado* com a relação R. Quaisquer resultados que são provados usando somente (P10–P12) são válidos para conjuntos totalmente ordenados (qualquer que seja a forma particular de $R(x, y)$ que define a relação de ordem).

Proposição 3.8. *A relação $x \leq y$ é equivalente à condição "$x < y$ ou $x = y$".*

Demonstração. De fato, se $x \leq y$ e $x \neq y$, então $x < y$, por definição. Assim, se $x \leq y$, então $x < y$ ou $x = y$. Para a recíproca, temos pela definição que $x < y$ pode ser escrito como $x \leq y$, e da mesma forma, se $x = y$. □

Proposição 3.9. *Para quaisquer dois números reais x, y, uma e somente uma das relações, $x < y$, $x = y$, $x > y$ pode ocorrer.*

Demonstração. De fato, por (P12) $x \leq y$ ou $y \leq x$. Agora, usando a Proposição 3.8

$$x \leq y \Rightarrow x < y \text{ ou } x = y$$

ou

$$x \geq y \Rightarrow x > y \text{ ou } x = y.$$

Assim, $x < y$ ou $x = y$ ou $x > y$. □

Proposição 3.10. *Se $x \leq y$ e $y < z$, então $x < z$. Do mesmo modo, se $x < y$ e $y \leq z$, então $x < z$.*

Demonstração. De fato, pela definição 3.1, $y < z$ garante que $y \neq z$ e assim, a partir de (P10) vemos que se $x \leq y$ e $y \leq z$, temos $x \leq z$. No entanto, $x = z$ é impossível, porque isso significaria que $x \leq y$ e $y < x = z$ aconteceriam simultaneamente, contradizendo a Proposição 3.9.

Vamos mostrar a segunda afirmação "$x < y$ e $y \leq z \Rightarrow x < z$" usando as mesmas propriedades (P's). Temos que $x < y$ implica $x \leq y$, e assim $x < y$ e $y \leq z$ implicam em $x \leq z$. Mas $x = z$ é impossível, pois $z < y$ e $y \leq z$ não podem ocorrer. □

Proposição 3.11. *Se $x_1 \leq y_1$ e $x_2 \leq y_2$, então $x_1 + x_2 \leq y_1 + y_2$. Além disso, se $x_1 < y_1$ e $x_2 < y_2$, então $x_1 + x_2 < y_1 + y_2$.*

Demonstração. Aplicando (P13) (e (P10) implicitamente),

$$x_1 + x_2 \leq y_1 + x_2 \leq y_1 + y_2.$$

Para mostrar $x_1 + x_2 < y_1 + y_2$, por contra-positiva suponhamos que acima valha a igualdade,

$$x_1 + x_2 = y_1 + x_2 = y_1 + y_2,$$

e daí, pela Proposição 3.2, $x_1 = y_1$ e $x_2 = y_2$. □

Proposição 3.12. *A condição $x \leq y$ é equivalente à $x + z \leq y + z$, $\forall\, z$; a condição $x < y$ é equivalente à $x + z < y + z$, $\forall\, z$.*

Demonstração. Por (P13) temos que $x \leq y$ implica $x + z \leq y + z$, e por outro lado,

$$x + z \leq y + z \Rightarrow (x + z) + (-z) \leq (y + z) + (-z)$$

$$\Rightarrow (z + (-z)) + x \leq (z + (-z)) + y \Rightarrow x \leq y.$$

Por fim, a implicação $x + z = y + z \Rightarrow x = y$ é uma consequência da Proposição 3.2, enquanto que $x = y \Rightarrow x + z = y + z$ obviamente. Logo, $x = y \Leftrightarrow x + z = y + z$ e isto fornece a segunda equivalência, $x < y \Leftrightarrow x + z < y + z$. □

Proposição 3.13. *As relações $x \leq y$, $0 \leq y - x$, $x - y \leq 0$, $-y \leq -x$ são todas equivalentes. Também são $x < y$, $0 < y - x$, $x - y < 0$, $-y < -x$.*

Demonstração. Como vimos na Proposição 3.12, a relação $x \leq y$ é equivalente a qualquer uma das relações $x - x \leq y - x$, $x - y \leq y - y$, $x - (x + y) \leq y - (x + y)$, e simplificando $0 \leq y - x$, $x - y \leq 0$, $-y \leq -x$.

A segunda parte provamos de forma análoga. A partir da relação $x < y$, tomamos as relações equivalentes $x - x < y - x$, $x - y < y - y$ e $x - (x + y) < y - (x + y)$, que simplificam-se para $0 < y - x$, $x - y < 0$ e $-y < -x$. □

Os números reais x que satisfazem $x \geq 0$ (resp. $x > 0$) são chamados *positivos* (resp. *estritamente positivos*) e os que satisfazem $x \leq 0$ (resp. $x < 0$) são chamados *negativos* (resp. *estritamente negativos*). Escrevemos $\mathbb{R}_+^*$ (resp. $\mathbb{R}_-^*$) para o conjunto de estritamente positivos (resp. estritamente negativos).

Proposição 3.14. *Se $x \geq 0$ e $y \geq 0$, então $x + y \geq 0$. Nestas condições, ou $x + y > 0$ ou $x = y = 0$. (Este é um caso particular da Proposição 3.11).*

Demonstração. Temos que, se $x \geq 0$ e $y \geq 0$, pela Proposição 3.11, $x+y \geq 0$. A partir daí, $x + y > 0$ ou $x + y = 0$. Analisando todos os possíveis casos, vemos que

- $x > 0$ e $y > 0$, novamente pela Proposição 3.11, $x + y > 0$;
- $x > 0$ e $y = 0$, por (P3), $x + y = x > 0$;
- $x = 0$ e $y > 0$, por (P3), $x + y = y > 0$;
- $x = 0$ e $y = 0$, por (P3), $x + y = 0 + 0 = 0$.

Portanto, $x + y > 0$ ou $x = y = 0$. □

Definição 3.2. *Para cada número real x, definimos $|x| = x$ se $0 \leq x$, $|x| = -x$ se $x \leq 0$. Chamamos $|x|$ de valor absoluto de x.*

Claramente $|x|$ é sempre positivo e, pela Proposição 3.13, $|-x| = |x|$. Caso $-x \geq 0$, então $x \leq 0$ e $|x| = -x = |-x|$. Caso $-x \leq 0$, então $x \geq 0$ e $|x| = x = |-x|$. A igualdade $|x| = 0$ é equivalente a $x = 0$.

Proposição 3.15. *Se $\alpha > 0$ então $|x| \leq \alpha$ se, e somente se, $-\alpha \leq x \leq \alpha$. Da mesma forma, $|x| < \alpha$ se, e somente se, $-\alpha < x < \alpha$.*

Demonstração. No caso em que $x \geq 0$, temos $|x| = x$ e $|x| \leq \alpha$, então, $x \leq \alpha$. Se $x < 0$, temos $|x| = -x$ e $|x| \leq \alpha$, então, $-x \leq \alpha$, ou seja, $x \geq -\alpha$. A relação $x > -\alpha$ sempre se mantém, pois pela hipótese $-\alpha < 0$ ($\alpha > 0$), e daí segue que $-\alpha < 0 \leq x$, ou seja, $x > -\alpha$. Enquanto $|x| \leq \alpha$ (resp. $|x| < \alpha$) é equivalente (neste caso $x \geq 0$) a $x \leq \alpha$ (resp. $x > \alpha$).

No caso em que $x \leq 0$, então $x < \alpha$ se mantém sempre, enquanto $|x| \leq \alpha$ (resp. $|x| < \alpha$) é equivalente a $-x \leq \alpha$ (resp. $-x < \alpha$). Isto é, se e somente se $x \geq -\alpha$ (resp. $x > -\alpha$). □

Proposição 3.16. *Para números reais arbitrários x, y, tem-se $|x + y| \leq |x| + |y|$ e $||x| - |y|| \leq |x - y|$.*

Demonstração. Para o caso $x \geq 0$ e $y \geq 0$, temos que $x + y \geq 0$ e assim $|x + y| = x + y = |x| + |y|$. Da mesma forma se $x \leq 0$ e $y \leq 0$, temos que $x + y \leq 0$ e assim $|x + y| = -(x + y) = -x - y$.

Agora, se $x \leq 0 \leq y$, temos $x + y \leq y \leq y + |x| = |y| + |x|$ e $x + y \geq x \geq x - |y| = -|x| - |y|$. Daí,

$$-(|x| + |y|) \leq x + y \leq |y| + |x|.$$

Pela Proposição 3.15

$$|x + y| \leq |x| + |y|.$$

Da mesma forma, quando $y \leq 0 \leq x$, temos

$$x + y \leq x \leq x + |y| = |x| + |y|$$

e

$$x + y \geq y \geq -|x| + y \geq -|x| - |y|.$$

Daí, novamente pela Proposição 3.15

$$|x + y| \leq |x| + |y|.$$

Para provarmos a segunda afirmação, lembramos que

$$|x| = |y + (x - y)| \leq |y| + |x - y|$$

e

$$|y| = |x + (y - x)| \leq |x| + |y - x|.$$

Assim, da primeira desigualdade $|x| - |y| \leq |x - y|$ e, da segunda, $-|y - x| = -|x - y| \leq |x| - |y|$. Logo

$$-|x - y| \leq |x| - |y| \leq |x - y|.$$

E, pela Proposição 3.15

$$||x| - |y|| \leq |x - y|. \qquad \square$$

Proposição 3.17. *Se $z \geq 0$, então $x \leq y$ implica $x \cdot z \leq y \cdot z$.*

Demonstração. Temos pela Proposição 3.13 que $x \leq y \Rightarrow 0 \leq y - x$ e por (P14) e P(9), $0 \leq (y - x) \cdot z = y \cdot z - x \cdot z$. O resultado segue da Proposição 3.13, i.e., $x \cdot z \leq y \cdot z$. $\square$

Proposição 3.18 (Regra de sinais)**.**

1. *Se $x \leq 0$ e $y \geq 0$, então $x \cdot y \leq 0$.*
2. *Se $x \leq 0$ e $y \leq 0$, então $x \cdot y \geq 0$.*
3. *Se $x > 0$ e $y > 0$, então $x \cdot y > 0$.*
4. *Se $x > 0$ e $y < 0$, então $x \cdot y < 0$.*
5. *Se $x < 0$ e $y < 0$, então $x \cdot y > 0$.*

Demonstração. Para provar 1, lembramos que $(-x) \cdot y = -(x \cdot y)$ e $(-x) \cdot (-y) = -(x \cdot (-y)) = x \cdot y$ (pelas Proposições 3.10 e 3.7).

Para cada número real x diferente de zero, temos pela regra de sinais que $x \cdot x > 0$. Pela Proposição 3.3, $x \cdot x = 0$ se, e somente se, $x = 0$. Portanto, para qualquer x, $x \cdot x \geq 0$. Para os demais itens, a prova segue o mesmo raciocínio. □

Assim, $x \neq 0$ é equivalente a $x \cdot x > 0$. Em particular, como $1 \cdot 1 = 1$ (pois vale P(7)), temos $1 > 0$. Aplicando a Proposição 3.14, temos $2 = 1 + 1 > 0$, $3 = 2 + 1 > 0$, e assim por diante.

Proposição 3.19. *Para todos x, $y \in \mathbb{R}$, $|x| \cdot |y| = |x \cdot y|$.*

Isso segue facilmente da definição 3.2 de valor absoluto, juntamente com a regra de sinais. Em cada caso:

- Se $x \leq 0$ e $y \leq 0$, então $|x \cdot y| = x \cdot y = (-x) \cdot (-y) = |x| \cdot |y|$.
- Se $x \leq 0$ e $y \geq 0$, então $|x \cdot y| = -(x \cdot y) = (-x) \cdot y = |x| \cdot |y|$.
- Se $x \geq 0$ e $y \geq 0$, então $|x \cdot y| = x \cdot y = |x| \cdot |y|$.
- Se $x \geq 0$ e $y \leq 0$, então $|x \cdot y| = -(x \cdot y) = x \cdot (-y) = |x| \cdot |y|$.

Em particular, isto implica que

$$|x|^2 = |x| \cdot |x| = |x \cdot x| = |x^2| = x^2.$$

Proposição 3.20.

1. *Se $x > 0$, então $x^{-1} > 0$.*
2. *Se $z > 0$, então $x \leq y$ se, e somente se, $x \cdot z \leq y \cdot z$ (da mesma forma $x < y$ se, e somente se, $x \cdot z < y \cdot z$).*
3. $0 < x < y \Leftrightarrow 0 < y^{-1} < x^{-1}$.

Demonstração. (1) De fato, como $x > 0$ e $x \cdot x^{-1} = 1 > 0$, pela regra de sinais, obrigatoriamente $x^{-1} > 0$.

(2) De $x \leq y$, temos $0 \leq y - x$. Como $z > 0$,

$$0 \leq (y - x) \cdot z \Rightarrow 0 \leq y \cdot z - x \cdot z \Rightarrow x \cdot z \leq y \cdot z.$$

Reciprocamente, se $x \cdot z \leq y \cdot z$, como $z^{-1} > 0$ pelo item 1, segue que $x \cdot z \cdot z^{-1} \leq y \cdot z \cdot z^{-1}$, ou seja, $x \leq y$.

(3) A partir de $0 < x < y$, temos $0 < x \cdot (x^{-1} \cdot y^{-1}) < y \cdot (y^{-1} \cdot x^{-1})$ e assim $0 < y^{-1} < x^{-1}$. Como $(x^{-1})^{-1} = x$, segue a prova da recíproca. □

Em consequência das Proposições 3.18 e 3.20, vemos que podemos restringir as funções $(x, y) \to x \cdot y$ e $x \to x^{-1}$ para os domínios $\mathbb{R}_+^* \times \mathbb{R}_+^*$ e $\mathbb{R}_+^*$, respectivamente, com valores em $\mathbb{R}_+^*$. Além disso, como $1 > 0$, podemos concluir de (P5–P8) que $\mathbb{R}_+^*$ é um grupo abeliano com a operação produto, um subgrupo de $\mathbb{R}^*$.

Definição 3.3. *Definimos o sinal de um número real x, e escrevemos* $\operatorname{sinal}(x)$*, como sendo* 1*,* -1 *ou* 0*, dependendo se* $x > 0$*,* $x < 0$ *ou* $x = 0$*, respectivamente.*

Usando a regra de sinais, segue facilmente que, para todos x, y em $\mathbb{R}$,

$$\operatorname{sinal}(x \cdot y) = \operatorname{sinal}(x) \cdot \operatorname{sinal}(y). \tag{3.1}$$

Além disso, pela definição de valor absoluto, temos

$$x = |x| \cdot \operatorname{sinal}(x). \tag{3.2}$$

É evidente que o conjunto $\{-1, 1\}$ é um subgrupo de $\mathbb{R}^*$ com a operação $\cdot$. Agora temos que a função $x \mapsto \operatorname{sinal}(x)$ é um homomorfismo (por (3.1)) sobrejetor de $(\mathbb{R}^*, \cdot)$ em $(\{\pm 1\}, \cdot)$.

Além disso, pela Proposição 3.19, a função $x \mapsto |x|$ é um homomorfismo de $\mathbb{R}^*$ em $\mathbb{R}^*$.

Temos que a fórmula (3.2) dá a decomposição única de um número real $x \neq 0$ como um produto $\rho \cdot \zeta$, em que $\rho > 0$ e ζ é igual a 1 ou -1. Em outras palavras, o grupo $\mathbb{R}^*$ é o produto direto dos seus subgrupos $\{-1, 1\}$ e $\mathbb{R}_+^*$,

$$\mathbb{R}^* = \mathbb{R}_+^* \times \{-1, 1\}.$$

Proposição 3.21. *Se* $0 < x_1 \leq y_1$ *e* $0 < x_2 \leq y_2$ *então* $x_1 \cdot x_2 \leq y_1 \cdot y_2$*. Além disso, se* $x_1 < y_1$ *e* $x_2 < y_2$ *então* $x_1 \cdot x_2 < y_1 \cdot y_2$*.*

Demonstração. De fato, de $x_1 \leq y_1$ e $x_2 > 0$ e de $y_1 > 0$ e $x_2 \leq y_2$, temos que

$$x_1 \cdot x_2 \leq y_1 \cdot x_2 \leq y_1 \cdot y_2. \qquad \square$$

Para simplificar, denotemos $x \cdot x$ por x^2, e $x \cdot x \cdot x$ por x^3, para qualquer $x \in \mathbb{R}$.

Proposição 3.22. *A relação* $x^2 \leq y^2$ *é equivalente a* $|x| \leq |y|$*.*

Demonstração. Primeiro, se $x = 0$ o resultado é imediato. Suponhamos então $x \neq 0$ e portanto $0 < x^2$.

Se $0 < x^2 \leq y^2$, então $x^2 = |x^2| = |x|^2$ e $y^2 = |y^2| = |y|^2$, ou seja, $|x|^2 \leq |y|^2$. Como ambos são estritamente positivos, segue que $|x| \leq |y|$, pois caso $|y| < |x|$, pela Proposição 3.21, teríamos $|y|^2 < |x|^2$.

Reciprocamente, se $0 < |x| \leq |y|$, novamente pela Proposição 3.21, $0 < |x|^2 \leq |y|^2$, ou seja, $0 < |x^2| \leq |y^2|$, isto é, $0 < x^2 \leq y^2$, o que conclui a prova. $\square$

Proposição 3.23. *$x^3 \leq y^3$ se, e somente se, $x \leq y$.*

Demonstração. Para provarmos $x^3 \leq y^3 \Rightarrow x \leq y$, observamos que se $x = 0$ o resultado é imediado, pois $0 \leq y^3 \Rightarrow 0 \leq y$, pela regra de sinais da Proposição 3.18. Suponhamos então $x \neq 0$. Se $x < 0$ e $x > y$ então $0 < -x < -y \Rightarrow y^3 < x^3$, o que não ocorre. Logo $x \leq y$. Se $y < 0 < x$, novamente $y^3 < x^3$, que não ocorre. Portanto $y \geq 0$. Mas se $0 \leq y < x$, então $y^3 < x^3$, novamente impossível por hipótese. Resta então a única possibilidade $0 < x \leq y$. Em resumo, $x^3 \leq y^3 \Rightarrow x \leq y$.

Provemos a recíproca. Caso $0 \leq x \leq y$, utilizando a Proposição 3.21 duas vezes, temos respectivamente, $x^2 \leq y^2$ e $x^2 \cdot x \leq y^2 \cdot y$, ou seja, $x^3 \leq y^3$. Caso $x \leq y \leq 0$, então $0 \leq -y \leq -x$ e pelo caso anterior, $0 \leq (-y)^3 \leq (-x)^3$, ou seja, $0 \leq -y^3 \leq -x^3$, que equivale a $x^3 \leq y^3 \leq 0$. Caso $x \leq 0 \leq y$, aplicando as regras de sinais, segue imediatamente que $x^3 \leq 0 \leq y^3$. Em resumo, $x \leq y \Rightarrow x^3 \leq y^3$. □

Proposição 3.24. *Para cada número real $\alpha > 0$, existe um, e somente um, número real $\beta > 0$ tal que $\beta^2 = \alpha$. Os números reais β e $-\beta$ são os únicos reais, tais que $\beta^2 = \alpha$.*

Demonstração. Para mostrar a existência e unicidade de β aplicamos (P15) para o polinômio $f(x) = x^2 - \alpha$. Em primeiro lugar, temos que $f(0) = -\alpha$, e então tudo o que precisamos para mostrar a existência de β é mostrar que f assume um valor positivo ou zero.

Se $\alpha \leq 1$, $x = 1$ será suficiente (pois $f(1) \geq 0$), enquanto se $\alpha > 1$ precisamos tomar $x = \alpha$, uma vez que $\alpha^2 > \alpha$ neste caso, e daí $f(\alpha) > 0$. Por (P15), existe β tal que $f(\beta) = 0$, ou seja, $\beta^2 = \alpha$. Note que $\beta \neq 0$ pois $\alpha > 0$. Como $(-\beta)^2 = \beta^2$ então $-\beta \neq 0$ também é solução, e portanto sempre existe uma solução estritamente positiva. Sem perda de generalidade, suponhamos $\beta > 0$. Se $\beta' > 0$ também satisfaz $\beta'^2 = \alpha$, então pela Proposição 3.22 segue que $|\beta| = |\beta'|$ e assim, $\beta = \beta'$ pois ambos são positivos.

Um raciocínio análogo mostra claramente que $-\beta < 0$ é a única solução negativa da equação. □

Como $x^2 \geq 0$, para todo número x, vemos então que a equação $x^2 = \alpha$ tem *duas* soluções se $\alpha > 0$, que são iguais em valor absoluto, mas com sinais opostos; *uma* solução (igual a 0) se $\alpha = 0$ e *nenhuma* solução se $\alpha < 0$.

Quando $\alpha \geq 0$, a única solução $\beta \geq 0$ é chamada *raiz quadrada de* α e escrita como $\sqrt{\alpha}$ ou $\alpha^{\frac{1}{2}}$. Devido à Proposição 3.22, $0 \leq x \leq y$ é equivalente a $\sqrt{x} \leq \sqrt{y}$.

Proposição 3.25. *Uma equação quadrática $\alpha \cdot x^2 + \beta \cdot x + \gamma = 0$ (com $\alpha \neq 0$) tem duas soluções distintas se $\beta^2 - 4 \cdot \alpha \cdot \gamma > 0$, e uma única solução se $\beta^2 - 4 \cdot \alpha \cdot \gamma = 0$.*

Demonstração. Primeiramente, como $\alpha \neq 0$, podemos reescrever a equação na forma

$$\alpha \cdot x^2 + \beta \cdot x + \gamma = \alpha \cdot \left(\left(x + \frac{\beta}{2 \cdot \alpha} \right)^2 - \left(\frac{\beta^2 - 4 \cdot \alpha \cdot \gamma}{4 \cdot \alpha^2} \right) \right).$$

Assim, para aplicarmos a Proposição 3.24, basta definirmos $y = x + \dfrac{\beta}{2 \cdot \alpha}$ e daí buscamos uma solução para a equação $y^2 = \dfrac{\beta^2 - 4 \cdot \alpha \cdot \gamma}{4 \cdot \alpha^2}$, observando que o sinal do lado direito depende simplesmente do numerador $\beta^2 - 4 \cdot \alpha \cdot \gamma$. □

Proposição 3.26. *Cada equação cúbica $\alpha \cdot x^3 + \beta \cdot x^2 + \gamma \cdot x + \delta = 0$ (com $\alpha \neq 0$) tem pelo menos uma solução.*

Demonstração. Desde que $\alpha \neq 0$, podemos dividir por α, se necessário, ou melhor, podemos supor que $\alpha = 1$. Aplicaremos (P15) em $f(x) = x^3 + \beta \cdot x^2 + \gamma \cdot x + \delta$. Para isso, seja $\lambda = 1 + |\beta| + |\gamma| + |\delta| \geq 1$. Então, para $x \neq 0$,

$$f(x) = x^3 \cdot \Big(1 + \frac{\beta}{x} + \frac{\gamma}{x^2} + \frac{\delta}{x^3}\Big)$$

e, tomando x tal que $|x| > \lambda$ (por exemplo, $|x| = \lambda + 1$), temos $|x| > 1$ e

$$\Big|\frac{\beta}{x} + \frac{\gamma}{x^2} + \frac{\delta}{x^3}\Big| \leq \frac{|\beta|}{|x|} + \frac{|\gamma|}{|x^2|} + \frac{|\delta|}{|x^3|} < \frac{1}{|x|} \cdot (|\beta| + |\gamma| + |\delta|) < \frac{\lambda}{|x|} < 1,$$

ou seja,

$$-1 < \frac{\beta}{x} + \frac{\gamma}{x^2} + \frac{\delta}{x^3} < 1.$$

Assim,

$$0 < 1 + \frac{\beta}{x} + \frac{\gamma}{x^2} + \frac{\delta}{x^3}$$

de modo que $f(x)$ tem o mesmo sinal que x^3 para x tal que $|x| > \lambda$. Em particular, para $x = \lambda + 1 > 0$, $f(-\lambda - 1) < 0 < f(\lambda + 1)$, e o resultado segue de (P15). □

3.2 Números Complexos

Nesta seção, $\mathbb{R}$ denotará o corpo ordenado dos números reais, estudado na seção anterior, e $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ será um espaço vetorial real (i.e. sobre o corpo $\mathbb{R}$) com as seguintes operações: para $a, b, c, d, \lambda \in \mathbb{R}$,

$$\begin{aligned} \text{adição}: &\qquad (a,b) + (c,d) = (a + c, b + d), \\ \text{multiplicação por escalar}: &\qquad \lambda \cdot (a,b) = (\lambda \cdot a, \lambda \cdot b). \end{aligned}$$

Dados $(a,b), (c,d), (e,f) \in \mathbb{R}^2$, seguem imediatamente que a operação de adição satisfaz:

$$\begin{aligned} \text{Associativa}: &\qquad (a,b) + [(c,d) + (e,f)] = [(a,b) + (c,d)] + (e,f). \\ \text{Comutativa}: &\qquad (a,b) + (c,d) = (c,d) + (a,b). \\ \text{Existência do neutro}: &\qquad (a,b) + (0,0) = (a,b). \\ \text{Existência do oposto}: &\qquad (a,b) + (-a,-b) = (0,0). \end{aligned}$$

Dados $\lambda, \lambda_1, \lambda_2 \in \mathbb{R}$ e $(a,b),(c,d) \in \mathbb{R}^2$, também é fácil ver que a multiplicação por escalar satisfaz:

$$\begin{aligned}
\lambda_1 \cdot [\lambda_2 \cdot (a,b)] &= (\lambda_1\lambda_2) \cdot (a,b),\\
1 \cdot (a,b) &= (a,b),\\
(\lambda_1 + \lambda_2) \cdot (a,b) &= \lambda_1 \cdot (a,b) + \lambda_2 \cdot (a,b),\\
\lambda \cdot [(a,b) + (c,d)] &= \lambda \cdot (a,b) + \lambda \cdot (c,d).
\end{aligned}$$

Com tais operações $\mathbb{R}^2$ é um espaço vetorial real de dimensão 2, uma vez que o conjunto $\{(1,0),(0,1)\}$ é uma base, comumente chamada de *base canônica.*

Definiremos a seguir uma multiplicação $*$ em $\mathbb{R}^2$ de modo a torná-lo um corpo.

Dados $(a,b),(c,d) \in \mathbb{R}^2$, $*$ é dada por

$$(a,b) * (c,d) = (ac - bd, ad + bc).$$

Proposição 3.27. *O conjunto $\mathbb{R}^2$ munido das operações $+$ e $*$, indicado $(\mathbb{R}^2, +, *)$, é um corpo.*

Demonstração. As propriedades da adição para $(\mathbb{R}^2, +, *)$ decorrem das propriedades da adição para $\mathbb{R}^2$ como um espaço vetorial, apresentadas acima.

Verifiquemos as propriedades abaixo para $*$.

Associativa: $\forall (a,b),(c,d),(e,f) \in \mathbb{R}^2$, $(a,b) * [(c,d) * (e,f)] = (a,b) * (ce - df, cf + de) = (a(ce-df) - b(cf+de), a(cf+de) + b(ce-df)) = ((ace-adf) - (bcf+bde), (acf + ade) + (bce - bdf)) = (ace - adf - bcf - bde, acf + ade + bce - bdf) = (ace - bde - adf - bcf, acf - bdf + ade + bce) = ((ac-bd)e - (ad+bc)f, (ac-bd)f + (ad+bc)e) = [(ac - bd, ad + bc)] * (e,f) = [(a,b) * (c,d)] * (e,f)$.

Comutativa: $\forall (a,b),(c,d) \in \mathbb{R}^2$, $(a,b)*(c,d) = (ac-bd, ad+bc) = (ca-db, cb+dc) = (c,d) * (a,b)$.

Existência de identidade: $\forall (a,b) \in \mathbb{R}^2$, $(a,b)*(c,d) = (a,b) \Leftrightarrow (ac-bd, ad+bc) = (a,b) \Leftrightarrow a - bd = a$ e $ad + bc = b \Leftrightarrow c = 1$ e $d = 0$. Assim, o elemento neutro para $*$, elemento identidade, é $(1,0)$.

Existência de elemento inverso: $\forall (a,b) \neq (0,0)$, $(a,b) * (u,v) = (1,0) \Leftrightarrow (au - bv, av + bu) = (1,0) \Leftrightarrow au - bv = 1$ e $av + bu = 0$. Logo $av = -bu$.

Distributiva: $\forall (a,b),(c,d),(e,f) \in \mathbb{R}^2$, $(a,b)*[(c,d)+(e,f)] = (a,b)*(c+e, d+f) = (a(c+e) - b(d+f), a(d+f) + b(c+e)) = ((ac-bd) + (ae-bf), (ad+bc) + (af+be)) = (ac - bd, ad + bc) + (ae - bf, af + be) = (a,b) * (c,d) + (a,b) * (e,f)$. $\square$

Definição 3.4. *O conjunto $\mathbb{R}^2$ com a estrutura $(\mathbb{R}^2, +, *)$ descrita acima é chamado corpo dos números complexos, denotado por $\mathbb{C}$.*

Mostraremos a seguir que $\mathbb{C}$ contém um subcorpo isomorfo ao corpo $\mathbb{R}$, justificando assim a inclusão $\mathbb{R} \subset \mathbb{C}$.

Consideremos a função, evidentemente injetora,

$$j : \mathbb{R} \to \mathbb{C}, \quad a \mapsto (a, 0).$$

É claro que j preserva as operações de adição e multiplicação, isto é, $\forall a, b \in \mathbb{R}$,

$$j(a + b) = (a + b, 0) = (a, 0) + (b, 0) = j(a) + j(b), \tag{3.3}$$

$$j(ab) = (ab, 0) = (a, 0) * (b, 0) = j(a) * j(b). \tag{3.4}$$

Assim, j é um isomorfismo entre os corpos $\mathbb{R}$ e $\mathrm{Im}(j) = \{(a, 0) : a \in \mathbb{R}\}$, este último subcorpo de $\mathbb{C}$ (isomorfo a $\mathbb{R}$).

Por tal isomorfismo, não há diferença algébrica entre $\mathbb{R}$ e $\mathrm{Im}(j)$ de modo que passamos a identificá-los, não distinguindo entre um número real a e sua imagem isomorfa $j(a) = (a, 0)$. Quando necessário, reforçamos tal identificação escrevendo $a \equiv (a, 0)$ ao invés de $a = (a, 0)$.

A multiplicação por um escalar real, proveniente da estrutura de espaço vetorial de $\mathbb{R}^2$, não conflita com $*$. Mais precisamente, se $\lambda \in \mathbb{R}$ e $(a, b) \in \mathbb{R}^2$ temos,

$$\lambda \cdot (a, b) = (\lambda a, \lambda b) \quad \text{e} \tag{3.5}$$

$$\lambda * (a, b) \equiv (\lambda, 0) * (a, b) = (\lambda a + 0b, \lambda b + 0a) = (\lambda a, \lambda b). \tag{3.6}$$

O corpo $\mathbb{C}$ tem três elementos distinguidos a saber, $(0, 0)$, $(1, 0)$ e $(0, 1)$. Os elementos $0 \equiv (0, 0)$ e $1 \equiv (1, 0)$ são, respectivamente, os elementos neutros da adição $+$ e da multiplicação $*$ em $\mathbb{C}$. O elemento $(0, 1)$ satisfaz à identidade $(0, 1) * (0, 1) = (-1, 0) \equiv -1$, via isomorfismo j, e é indicado por i, ou seja, $i = (0, 1) \in \mathbb{C}$.

Logo, temos a famosa expressão

$$i^2 = -1$$

(mais precisamente, $i^2 \equiv -1$) e portanto $\mathbb{C}$ é uma extensão do corpo $\mathbb{R}$, ou seja, é um corpo que contém um subcorpo isomorfo a $\mathbb{R}$, na qual o elemento $-1 \equiv (-1, 0)$ tem raiz quadrada $i \in \mathbb{C}$, ou seja, $i = \sqrt{-1}$.

Segue que todo número real a admite raiz quadrada complexa. De fato, se $a \geq 0$ as propriedades dos números reais garantem a existência de $\sqrt{a}$ (ver Proposição 3.24) e se $a < 0$ temos $(i * \sqrt{|a|})^2 = -|a| = a$.

Pelo exposto acima é conveniente denotar o complexo $(a, b) \in \mathbb{C}$ por $a + b \cdot i$, uma vez que

$$(a, b) = (a, 0) + (0, b) = (a, 0) + b \cdot (0, 1) = a + b \cdot i.$$

Com esta notação temos,

$$\begin{aligned}(a + b \cdot i) * (c + d \cdot i) &= ac + ad \cdot i + bc \cdot i + bdi^2 \\ &= (ac - bd) + (ad + bc) \cdot i.\end{aligned}$$

Se $z = (a, b) = a + b \cdot i \in \mathbb{C}$, a é chamado *parte real* e b *parte imaginária* de z, denotadas por $\mathrm{Re}(z)$ e $\mathrm{Im}(z)$, respectivamente. Assim

$$z = \mathrm{Re}(z) + \mathrm{Im}(z) \cdot i, \ \forall z \in \mathbb{C}.$$

Também o complexo $\overline{z} = a - b \cdot i$ é chamado *conjugado* de z, e segue facilmente que

$$z * \overline{z} = a^2 + b^2.$$

4 Quatérnios e Rotações

Neste último capítulo, estudamos [5, Appendix IV], que apresenta a construção da álgebra dos quatérnios. Por fim, seguimos [6, Cap. 6] para relacionar o produto de quatérnios com rotações tridimensionais.

4.1 Quatérnios

O fato de que é possível definir uma multiplicação no conjunto $\mathbb{R} \times \mathbb{R}$ que, com a adição usual, transforma $\mathbb{R} \times \mathbb{R}$ em um corpo comutativo, levanta interesse sobre a questão se é possível generalizar este processo para espaços vetoriais sobre $\mathbb{R}$ de dimensões maiores que 2. Primeiro mostraremos que é impossível obter alguma estrutura análoga em um espaço tridimensional.

Mais precisamente, seja E um espaço vetorial real tridimensional, e suponha que exista uma função $\cdot : E \times E$, $(x, y) \mapsto x \cdot y$ tal que, juntamente com a adição $+ : (x, y) \mapsto x + y$, satisfaz os axiomas (P1–P9), e também para cada escalar $\alpha \in \mathbb{R}$, e para cada par de vetores $x, y \in E$,

$$\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y). \tag{4.1}$$

Seja $e_0 \neq 0$ a identidade para a multiplicação definida em E. Em seguida, para cada vetor $x \neq e_0 \in E$, os quatro vetores e_0, x, x^2, x^3 são necessariamente linearmente dependentes (pois E tem dimensão 3), e assim existem quatro escalares $\alpha, \beta, \gamma, \delta$ não todos nulos tais que

$$\alpha x^3 + \beta x^2 + \gamma x + \delta e_0 = 0.$$

Se x não é múltiplo de e_0, afirmamos que $\alpha \neq 0$ ou $\beta \neq 0$. De fato, se $\alpha = \beta = 0$, a equação acima torna-se $\gamma x + \delta e_0 = 0$, ou seja, $\gamma x = -\delta e_0$, e neste caso $\gamma \neq 0$ pois caso contrário teríamos $\delta \neq 0$ (pois os escalares não são todos nulos) e $0 = \gamma x = -\delta e_0 \neq 0$, o que é um absurdo. Porém, se $\gamma \neq 0$, temos $x = -\gamma^{-1}\delta e_0 = \lambda e_0$, ou seja, um múltiplo de e_0, o que não ocorre por hipótese. Assim, de fato, $\alpha \neq 0$ ou $\beta \neq 0$.

Se $\alpha \neq 0$, sabemos pela Proposição 3.26 que podemos escrever, para escalares adequados,

$$(x - \lambda e_0) \cdot (x^2 + \mu x + \nu e_0) = 0.$$

Como a Proposição 3.3 é válida para E, temos $x = \lambda e_0$ ou $x^2 + \mu x + \nu e_0 = 0$, mas como a primeira é descartada por hipótese, então

$$x^2 + \mu x + \nu e_0 = 0. \tag{4.2}$$

Além disso, devemos ter $\mu^2 - 4\nu < 0$, porque se não teríamos uma fatoração

$$(x - \eta e_0) \cdot (x - \theta e_0),$$

para η, θ adequados, e novamente teríamos uma contradição quando aplicássemos a Proposição 3.3, ou seja, teríamos x múltiplo de e_0.

Observe que (4.2) pode ser reescrita como

$$(x + \frac{\mu}{2} e_0)^2 = (\frac{\mu^2}{4} - \nu) e_0,$$

e, colocando $x' = \sigma^{-1}(x + \frac{\mu}{2} e_0)$, onde $\sigma^2 = \nu - \frac{\mu^2}{4}$, obtemos

$$x'^2 = -e_0. \tag{4.3}$$

Antes de prosseguirmos, vale reforçar que o processo descrito acima nos garante que é possível, para x linearmente independente com e_0, obter uma combinação linear x' de x e e_0 cujo quadrado é $-e_0$.

Suponha agora que $\{e_0, e_1, e_2\}$ é uma base de E. Como e_1 e e_2 não são múltiplos de e_0, aplicando o argumento acima, podemos substitui-los, se necessário, por dois vetores da forma $\alpha_1 e_1 + \beta_1 e_0$, $\alpha_2 e_2 + \beta_2 e_0$, com $\alpha_1 \neq 0$ e $\alpha_2 \neq 0$ (para manter a propriedade de base), satisfazendo (4.3). Portanto, podemos assumir que

$$e_1^2 = -e_0, \qquad e_2^2 = -e_0. \tag{4.4}$$

Assim sendo, existem escalares λ, μ e ν tais que

$$e_1 \cdot e_2 = \lambda e_0 + \mu e_1 + \nu e_2. \tag{4.5}$$

Multiplicando ambos os lados à direita por e_2, deduzimos que

$$e_1 \cdot e_2^2 = e_1 \cdot (-e_0) = -e_1 = \lambda e_2 + \mu e_1 \cdot e_2 - \nu e_0, \tag{4.6}$$

assim,

$$e_1 \cdot e_2 = \frac{-e_1}{\mu} - \frac{\lambda}{\mu} + \frac{v e_0}{\mu}. \tag{4.7}$$

por (4.1) e (4.4), e por assumirmos que E satisfaz (P6) e (P7).

Comparando com (4.5) e lembrando que $\{e_0, e_1, e_2\}$ é uma base, segue em particular que $\mu = \frac{-1}{\mu}$, o que é absurdo, pois $\mu^2 = -1 \notin \mathbb{R}$. O mesmo argumento se aplica se $\beta \neq 0$.

Assim, nenhuma estrutura de corpo do tipo citado é possível em espaços tridimensionais.

Notamos que a *comutatividade* da multiplicação (P5) não é realmente usada no argumento acima; apenas os axiomas (P1–P9), e os axiomas "simétricos" a (P7–P9), a saber:

(P7') $\xi \cdot \epsilon = \xi$,

(P8') $\xi'' \cdot \xi = \epsilon$,

(P9') $(\eta + \zeta) \cdot \xi = \eta \cdot \xi + \zeta \cdot \xi$.

Estes axiomas decorrem facilmente de (P5) com os demais, mas quando (P5) não é assumido, eles devem ser exigidos separadamente.

Definição 4.1. *Um* anel de divisão*, ou* corpo não-comutativo *é um conjunto no qual duas operações são dadas satisfazendo (P1–P4), (P6–P9), (P7'), (P8'), (P9'), mas* ***não*** *(P5).*

Veremos agora que é possível dar a um espaço vetorial de dimensão quatro sobre $\mathbb{R}$ uma estrutura de anel de divisão que também satisfaz (4.1). As propriedades de distributividade (P9) e (P9') (juntamente com a relação (4.1)) expressam o fato de que a multiplicação $(x, y) \mapsto x \cdot y$ deve ser uma função bilinear de $E \times E$ em E. Ela é, portanto, totalmente determinada pelos seus valores no conjunto de todos os pares de vetores em alguma base de E.

Seja $\{e_0, e_1, e_2, e_3\}$ uma base de E e defina

$$\begin{gathered} e_0^2 = e_0, \quad e_1^2 = -e_0, \quad e_2^2 = -e_0, \quad e_3^2 = -e_0, \\ e_0 \cdot e_1 = e_1 \cdot e_0 = e_1, \quad e_0 \cdot e_2 = e_2 \cdot e_0 = e_2, \quad e_0 \cdot e_3 = e_3 \cdot e_0 = e_3, \\ e_1 \cdot e_2 = -e_2 \cdot e_1 = e_3, \quad e_2 \cdot e_3 = -e_3 \cdot e_2 = e_1, \quad e_3 \cdot e_1 = -e_1 \cdot e_3 = e_2. \end{gathered} \tag{4.8}$$

Daí, para $x = \xi_0 e_0 + \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3$, $y = \eta_0 e_0 + \eta_1 e_1 + \eta_2 e_2 + \eta_3 e_3$, temos

$$\begin{aligned} x \cdot y = {} & (\xi_0\eta_0 - \xi_1\eta_1 - \xi_2\eta_2 - \xi_3\eta_3)e_0 + (\xi_0\eta_1 + \xi_1\eta_0 + \xi_2\eta_3 - \xi_3\eta_2)e_1 \\ & + (\xi_0\eta_2 + \xi_2\eta_0 + \xi_3\eta_1 - \xi_1\eta_3)e_2 + (\xi_0\eta_3 + \xi_3\eta_0 + \xi_1\eta_2 - \xi_2\eta_1)e_3. \end{aligned} \tag{4.9}$$

Queremos mostrar que (P1–P4), (P6–P9) do capítulo 3, (P7'–P9') acima e (4.1) são satisfeitas.

A condição (4.1) é trivial (a partir de (4.9)), enquanto que (P1–P4) são verdadeiras por hipótese, pois estamos supondo E um espaço vetorial.

O elemento $e_0 \neq 0$ satisfaz (P7) e (P7') e fornece a identidade multiplicativa. (P9) e (P9') são satisfeitas simplesmente pela forma como a multiplicação foi definida (bilinear), a partir dos vetores da base $\{e_0, e_1, e_2, e_3\}$.

Resta verificar (P6), (P8) e (P8'). Para provarmos que $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, para $x, y, z \in E$, é suficiente mostrarmos que $(e_i \cdot e_j) \cdot e_k = e_i \cdot (e_j \cdot e_k)$ para $i, j, k \in \{0, 1, 2, 3\}$, pois daí usamos o fato de que, para $x, y, z \in E$,

$$x = \sum \alpha_i e_i, \qquad y = \sum \beta_i e_i, \qquad z = \sum \gamma_i e_i,$$

e o fato de que o produto se distribui sobre a soma.

Como e_0 é o elemento identidade, eliminamos todos os casos em que um dos elementos é e_0, e portanto podemos supor que $i, j, k \in \{1, 2, 3\}$. Isso significa $3^3 = 27$ possíveis triplas.

No entanto, perceba que o sistema (4.8) mantém-se inalterado por qualquer permutação cíclica de e_1, e_2 e e_3, e assim podemos reduzir os cálculos, bastando checar $3^2 = 9$ relações:

$$\begin{aligned}
(e_1^2) \cdot e_1 &= e_1 \cdot (e_1^2), & (e_1^2) \cdot e_2 &= e_1 \cdot (e_1 \cdot e_2), & (e_1^2) \cdot e_3 &= e_1 \cdot (e_1 \cdot e_3),\\
(e_1 \cdot e_2) \cdot e_1 &= e_1 \cdot (e_2 \cdot e_1), & (e_1 \cdot e_2) \cdot e_2 &= e_1 \cdot (e_2^2), & (e_1 \cdot e_3) \cdot e_1 &= e_1 \cdot (e_3 \cdot e_1),\\
(e_1 \cdot e_3) \cdot e_3 &= e_1 \cdot (e_3^2), & (e_1 \cdot e_2) \cdot e_3 &= e_1 \cdot (e_2 \cdot e_3), & (e_1 \cdot e_3) \cdot e_2 &= e_1 \cdot (e_3 \cdot e_2).
\end{aligned}$$

Estes podem ser individualmente verificados, utilizando (4.8). Por exemplo,

$$(e_1 \cdot e_2) \cdot e_1 = e_3 \cdot e_1 = e_2 \quad \text{e} \quad e_1 \cdot (e_2 \cdot e_1) = e_1 \cdot (-e_3) = -(e_1 \cdot e_3) = e_2.$$

As demais são verificadas analogamente.

Para provar (P8) e (P8'), primeiro associamos a cada elemento $x = \xi_0 e_0 + \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3$, o elemento

$$\overline{x} = \xi_0 e_0 - \xi_1 e_1 - \xi_2 e_2 - \xi_3 e_3 \tag{4.10}$$

chamado *conjugado* de x.

Claramente a função $x \to \overline{x}$ é um automorfismo do espaço vetorial E. Além disso, uma verificação imediata mostra que

$$\overline{x \cdot y} = \overline{y} \cdot \overline{x}. \tag{4.11}$$

De fato, por um lado

$$\begin{aligned}
\overline{x \cdot y} = {} & (\xi_0\eta_0 - \xi_1\eta_1 - \xi_2\eta_2 - \xi_3\eta_3)e_0 - (\xi_0\eta_1 + \xi_1\eta_0 + \xi_2\eta_3 - \xi_3\eta_2)e_1\\
& - (\xi_0\eta_2 + \xi_2\eta_0 + \xi_3\eta_1 - \xi_1\eta_3)e_2 - (\xi_0\eta_3 + \xi_3\eta_0 + \xi_1\eta_2 - \xi_2\eta_1)e_3,
\end{aligned}$$

e por outro lado,

$$\begin{aligned}
\overline{y} \cdot \overline{x} &= (\eta_0 e_0 - \eta_1 e_1 - \eta_2 e_2 - \eta_3 e_3) \cdot (\xi_0 e_0 - \xi_1 e_1 - \xi_2 e_2 - \xi_3 e_3)\\
&= (\xi_0\eta_0 - \xi_1\eta_1 - \xi_2\eta_2 - \xi_3\eta_3)e_0 + (-\xi_0\eta_1 - \xi_1\eta_0 - \xi_2\eta_3 + \xi_3\eta_2)e_1\\
&\quad + (-\xi_0\eta_2 - \xi_3\eta_1 + \xi_1\eta_3 - \xi_2\eta_0)e_2 + (-\xi_0\eta_3 - \xi_1\eta_2 + \xi_2\eta_1 - \xi_3\eta_0)e_3\\
&= (\xi_0\eta_0 - \xi_1\eta_1 - \xi_2\eta_2 - \xi_3\eta_3)e_0 - (\xi_0\eta_1 + \xi_1\eta_0 + \xi_2\eta_3 - \xi_3\eta_2)e_1\\
&\quad - (\xi_0\eta_2 + \xi_2\eta_0 + \xi_3\eta_1 - \xi_1\eta_3)e_2 - (\xi_0\eta_3 + \xi_3\eta_0 + \xi_1\eta_2 - \xi_2\eta_1)e_3.
\end{aligned}$$

Além disso, por (4.9) e (4.10), temos

$$x \cdot \overline{x} = \overline{x} \cdot x = (\xi_0^2 + \xi_1^2 + \xi_2^2 + \xi_3^2)e_0. \tag{4.12}$$

Portanto, se $x \neq 0$, $x \cdot \overline{x} \neq 0$, e $x \cdot \overline{x} = \overline{x} \cdot x$ comuta (assim como e_0) com todos os elementos de E.

Assim, (P8) e (P8') podem ser verificados imediatamente, pois se $x \neq 0$, o elemento $x' = (\xi_0^2 + \xi_1^2 + \xi_2^2 + \xi_3^2)^{-1}\overline{x}$ satisfaz $x \cdot x' = x' \cdot x = e_0$. Por isso escreveremos x^{-1} ao invés de x'.

Definição 4.2. *O anel de divisão que acabamos de definir é chamado de* Corpo não comutativo dos Quatérnios *e seus elementos são chamados de quatérnios.*

A função $\xi \mapsto \xi e_0$ de $\mathbb{R}$ em E é um isomorfismo de $\mathbb{R}$ sobre um subcorpo $Z = \{\xi e_0;\ \xi \in \mathbb{R}\}$ de E. Denotaremos $Z \equiv \mathbb{R}e_0$. Antes de mostrarmos o próximo resultado, lembramos que, em um grupo G, seu centro $Z(G)$ é o conjunto dos elementos de G que comutam com todos os elementos de G.

Lema 4.1. *O subcorpo $Z = \mathbb{R}e_0$ é o centro de E, com relação ao produto do corpo.*

Demonstração. Se um elemento $x = \xi_0 e_0 + \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3$ pertence ao centro, então deve comutar com e_1, e_2 e e_3. Escrevendo as condições para que isso aconteça, e usando (4.8), obtemos $\xi_1 = \xi_2 = \xi_3 = 0$. De fato, se $x \in E$ comuta com e_1, obtemos:

$$(\xi_0 e_0 + \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3) \cdot e_1 = e_1 \cdot (\xi_0 e_0 + \xi_1 e_1 + \xi_2 e_2 + \xi_3 e_3) \Rightarrow$$

$$\xi_0 e_0 \cdot e_1 + \xi_1 e_1^2 + \xi_2 e_2 \cdot e_1 + \xi_3 e_3 \cdot e_1 = e_1 \cdot \xi_0 e_0 + e_1 \cdot \xi_1 e_1 + e_1 \cdot \xi_2 e_2 + e_1 \cdot \xi_3 e_3 \Rightarrow$$

$$\xi_0 e_1 - \xi_1 e_0 - \xi_2 e_3 + \xi_3 e_2 = \xi_0 e_1 - \xi_1 e_0 + \xi_2 e_3 - \xi_3 e_2 \Rightarrow$$

$$\xi_2 = -\xi_2 \Rightarrow \xi_2 = 0,$$

$$\xi_3 = -\xi_3 \Rightarrow \xi_3 = 0.$$

De modo análogo, se x comuta com e_2 e e_3, concluímos que $\xi_1 = 0$, e portanto $x = \xi_0 e_0 \in \mathbb{R}e_0$. Como é imediato que $\mathbb{R}e_0$ está contido no centro de E, segue o resultado. □

Para cada $x \in E$, um cálculo rápido mostra que

$$x + \overline{x} = 2\xi_0 e_0, \tag{4.13}$$

e como $x^2 - (x + \overline{x}) \cdot x + \overline{x} \cdot x = 0$, vemos que cada elemento de E satisfaz uma equação quadrática

$$x^2 - \alpha x + \beta e_0 = 0 \tag{4.14}$$

onde $\alpha = 2\xi_0$ e $\beta = \xi_0^2 + \xi_1^2 + \xi_2^2 + \xi_3^2$.

Observação 4.1. Em geral, tal equação pode ter infinita raízes em E. De fato, se x a satisfaz então, como todo elemento da forma $s \cdot x \cdot s^{-1}$ também a satisfaz, basta tomar $s \in E$ não nulo de modo que $s \cdot x \cdot s^{-1} \neq x$ (por exemplo, $s \notin Z$).

Esses resultados mostram que, para cada $x \neq 0$ em E, o conjunto de elementos $\lambda e_0 + \mu x$, onde λ e μ variam em $\mathbb{R}$, é um subcorpo comutativo $\mathbb{R}(x)$ de E. Se $x \in Z$, este subcorpo é Z; se $x \notin Z$, $\mathbb{R}(x)$ é isomorfo a $\mathbb{C}$, pois também mostramos neste caso que $\alpha^2 - 4\beta < 0$ e que existe $x' \in \mathbb{R}(x)$ tal que $x'^2 = -e_0$. Tal elemento x' é o análogo do complexo i.

Definição 4.3. *Os números reais* $\mathrm{Tr}(x) = 2\xi_0$ *e* $\mathrm{N}(x) = \xi_0^2 + \xi_1^2 + \xi_2^2 + \xi_3^2$ *são, respectivamente, chamados de* Traço *e* Norma *de um quatérnio* x.

Podemos ver que $\mathrm{Tr} : E \to \mathbb{R}$, $x \mapsto \mathrm{Tr}(x)$, é um funcional linear em E, pois

$$\mathrm{Tr}(x + y) = \mathrm{Tr}(x) + \mathrm{Tr}(y) \quad \text{e} \quad \mathrm{Tr}(\alpha x) = \alpha\,\mathrm{Tr}(x),$$

para $x, y \in E$ e $\alpha \in \mathbb{R}$. Também, como $x, \overline{x}$ possuem a mesma parte real, é imediato que

$$\mathrm{Tr}(\overline{x}) = \mathrm{Tr}(x). \tag{4.15}$$

Além disso, pelo produto calculado em (4.9), segue que

$$\mathrm{Tr}(y \cdot x) = \mathrm{Tr}(x \cdot y) \tag{4.16}$$

e assim a igualdade

$$\mathrm{Tr}(s \cdot x \cdot s^{-1}) = \mathrm{Tr}(x) \tag{4.17}$$

é válida para todo quatérnio $s \neq 0$, pois $x = (x \cdot s^{-1}) \cdot s$.

Proposição 4.1. *A função norma satisfaz*

$$\mathrm{N}(x \cdot y) = \mathrm{N}(x)\,\mathrm{N}(y). \tag{4.18}$$

Demonstração. De fato, basta notar que $\mathrm{N}(x \cdot y)$ é a parte real do produto

$$(x \cdot y) \cdot (\overline{x \cdot y}) = (x \cdot y) \cdot (\overline{y} \cdot \overline{x}) = (x \cdot \overline{x}) \cdot (y \cdot \overline{y}),$$

no qual a última igualdade segue de $y \cdot \overline{y} \in Z$, e o último termo da equação acima é exatamente o produto $\mathrm{N}(x)\,\mathrm{N}(y)$. □

Os quatérnios x tais que $\overline{x} = -x$ (ou equivalentemente $\mathrm{Tr}(x) = 0$) são chamados *quatérnios puros*, e formam um conjunto F, que é o hiperplano de E tendo $\{e_1, e_2, e_3\}$ como base. De (4.17), para cada quatérnio $s \neq 0$ em E, segue que

$$s \cdot F \cdot s^{-1} = F. \tag{4.19}$$

O espaço E é a soma direta de $Z = \mathbb{R}e_0$ com F, isto é, $E = \mathbb{R}e_0 \oplus F$. De fato, cada quatérnio z pode ser escrito de forma única como

$$z = \frac{1}{2}\,\mathrm{Tr}(z)e_0 + \frac{1}{2}(z - \overline{z}),$$

onde $\frac{1}{2}\,\mathrm{Tr}(z)e_0 = \frac{1}{2}(z + \overline{z})$ é chamada *parte escalar* de z e $\frac{1}{2}(z - \overline{z})$ é chamada *parte pura* de z.

Como Tr é linear, segue de (4.15) que a função $\phi : E \times E \to \mathbb{R}$, $(x, y) \mapsto \phi(x, y) = \mathrm{Tr}(x \cdot \overline{y})$ é uma forma bilinear simétrica, já que

$$\phi(x, y) = \mathrm{Tr}(x \cdot \overline{y}) = \mathrm{Tr}(y \cdot \overline{x}) = \phi(y, x).$$

Além disso, é positiva não-degenerada, pois

$$\phi(x,x) = \mathrm{Tr}(x \cdot \overline{x}) = \mathrm{N}(x) \geq 0 \quad \text{e} \quad \phi(x,x) = 0 \Leftrightarrow x = 0.$$

Consideramos a estrutura de espaço euclidiano em E proveniente de ϕ, e escrevemos $(x \mid y)$ em vez de $\mathrm{Tr}(x \cdot \overline{y})$ e $\|x\|^2$ em vez de $\mathrm{N}(x)$, esta última fornecendo $\|x\| = \sqrt{\mathrm{N}(x)}$ (raiz positiva). A relação (4.18), portanto, garante que

$$\|x \cdot y\| = \|x\|\|y\|, \tag{4.20}$$

e como $\mathrm{N}(e_0) = 1$, obtemos

$$\|x^{-1}\| = \|x\|^{-1}. \tag{4.21}$$

De fato, como $x \cdot x^{-1} = e_0$, segue que $1 = \mathrm{N}(x \cdot x^{-1}) = \mathrm{N}(x)\,\mathrm{N}(x^{-1})$, que por sua vez fornece $\|x\|\|x^{-1}\| = 1$.

As equações (4.20)–(4.21) mostram que o conjunto dos quatérnios com norma 1 é um subgrupo (denotado por S^3) do grupo multiplicativo E^* dos quatérnios não-nulos. Além disso, uma vez que cada quatérnio pode ser escrito de forma única como $x = \rho z$ com $\rho = \|x\|$ e $\|z\| = 1$, vê-se facilmente que E^* é o produto direto de S^3 com o grupo multiplicativo $\mathbb{R}_+^*$ (este identificado com o grupo de quatérnios ρe_0 tais que $\rho > 0$)

$$E^* = \mathbb{R}_+^* \times S^3.$$

A relação (4.20) mostra que se a e b são dois quatérnios com norma 1, então a função

$$x \mapsto a \cdot x \cdot b \tag{4.22}$$

é uma transformação ortogonal, ou seja, preserva a norma:

$$\|a \cdot x \cdot b\| = \|a \cdot x\|\|b\| = \|a\|\|x\|\|b\| = \|x\|.$$

O nosso objetivo é demonstrar que todas as rotações em E podem ser obtidas como em (4.22), para a e b adequados.

Levando em conta (4.19), a observação acima a respeito da transformação (4.22) mostra que F é invariante pela transformação ortogonal $x \mapsto s \cdot x \cdot s^{-1}$ de E.

Proposição 4.2. *Toda rotação no espaço F de quatérnios puros pode ser escrita sob a forma*

$$\mu_s : F \to F,\ x \mapsto s \cdot x \cdot s^{-1}, \tag{4.23}$$

onde s é um quatérnio não nulo. Além disso, $\mu_s = \mu_{s'}$ se, e somente se, $s' = \lambda s$ onde λ é um escalar não nulo.

Demonstração. De fato, a segunda parte é óbvia, pois se $s \cdot x \cdot s^{-1} = s' \cdot x \cdot s'^{-1}$ para cada quatérnio puro x, então multiplicando à direita por s e à esquerda por s'^{-1} temos, $(s'^{-1} \cdot s) \cdot x = x \cdot (s'^{-1} \cdot s)$, ou seja, $s'^{-1} \cdot s$ comuta com todo quatérnio puro. Como

também comuta com os elementos do centro $Z = \mathbb{R}e_0$, pertence necessariamente ao centro de E (lembre que $E = \mathbb{R}e_0 \oplus F$), ou seja, é um real puro.

Para provar a primeira parte, note que

$$\mu_{s\cdot t} = \mu_s \circ \mu_t. \tag{4.24}$$

Também, por [5, Apêndice II, Proposição 2], se E tem dimensão n, então toda transformação ortogonal u é o produto de no máximo n simetrias sobre um hiperplano.

Mostremos então que tais simetrias são da forma (4.23) e daí usamos o resultado citado e a equação (4.24). Em outras palavras, lidaremos com uma simetria v sobre um hiperplano, comprovando a existência de $s \in E$, tal que $v = -\mu_s$.

Como v é simetria através de um hiperplano de F, então, v é da forma

$$\begin{aligned} v(x) &= x - \frac{2(x \mid a)}{(a \mid a)}\, a \\ &= x + 2(x \mid a)\frac{-a}{N(a)} \\ &= x + 2(x \mid a)\, a^{-1} \end{aligned} \tag{4.25}$$

para todo $x \in F$, com $a \neq 0$ em F.

Agora, se $y, z \in F$ são dois quatérnios puros então

$$\begin{aligned} 2(y \mid z)\, e_0 &= ((y+z \mid y+z) - (y \mid y) - (z \mid z))\, e_0 \\ &= -((y+z)^2 - y^2 - z^2)\, e_0 \\ &= -(y \cdot z + z \cdot y). \end{aligned} \tag{4.26}$$

A equação (4.25), portanto, mostra que no anel de divisão E,

$$v(x) = x + 2(x \mid a)\, a^{-1} = x - (x \cdot a + a \cdot x) \cdot a^{-1} = -a \cdot x \cdot a^{-1} = -\mu_a(x).$$

Isto prova o resultado. □

Antes de prosseguir, vamos tirar algumas inferências sobre espaços euclidianos tridimensionais a partir do que acabamos de dizer (lembre-se, quaisquer dois espaços de mesma dimensão são isomorfos e, em particular, todos os espaços tridimensionais são isomorfos a F).

A Proposição 4.2 mostra que sempre podemos escrever uma rotação em F na forma $x \mapsto s \cdot x \cdot s^{-1}$, onde s é um quatérnio de norma 1, multiplicando por um escalar se necessário.

Levando em conta (4.24), temos definido um homomorfismo sobrejetor $s \mapsto \mu_s$ do grupo S^3 de quatérnios unitários para o grupo $O_3^+(\mathbb{R})$ das rotações tridimensionais, cujo núcleo é o subgrupo $\{1, -1\}$ de ordem 2 em S^3 (aqui estamos identificando ± 1 com $\pm e_0$, de onde vemos facilmente que $\mu_{\pm e_0} = \mathrm{id}$ é a identidade).

Isto significa que para cada rotação $v \in O_3^+(\mathbb{R})$ correspondem dois quatérnios opostos $\pm s$, de norma 1, tais que $\mu_{\pm s} = v$. Não há nenhuma maneira "razoável" de "selecionar" um desses dois quatérnios (para $O_3^+(\mathbb{R})$ como um todo).

No entanto, s e μ_s estão relacionados "geometricamente" como mostra a proposição seguinte:

Proposição 4.3. *Seja $s = \alpha e_0 + t$ um quatérnio, onde t é um quatérnio puro não-nulo e α um escalar. Então o eixo de rotação de μ_s é a reta D_{ot} pela origem e determinada por t, e o ângulo de rotação tomado em relação a uma orientação apropriada do complemento ortogonal de D_{ot} em F (plano perpendicular a t dentro de F) é reto quando $\alpha = 0$, e quando $\alpha \neq 0$ é $\theta > 0$ satisfazendo $\operatorname{tg}(\theta/2) = \|t\|/\alpha$.*

Demonstração. D_{ot} ser o eixo de μ_s é uma consequência do fato do quatérnio λt ($\lambda \in \mathbb{R}$) comutar com t e com os reais puros, e portanto comutar com s. Assim,

$$\mu_s(\lambda t) = s \cdot \lambda t \cdot s^{-1} = \lambda t,$$

ou seja, múltiplos de t são fixados pela rotação e então a reta determinada por t é o eixo de rotação.

Para provarmos a afirmação sobre o ângulo de rotação, observamos que, dados dois quatérnios puros x e y de mesma norma, sempre existe um quatérnio z tal que $y = z \cdot x \cdot z^{-1}$. De fato, basta considerarmos uma rotação em F que aplica x em y e usar a Proposição 4.2 que garante que esta rotação é da forma μ_z, para algum z.

Em particular, dados t e ρe_1 ($\rho > 0$) puros, existe z tal que $\rho e_1 = z \cdot t \cdot z^{-1}$, e desde que $\mu_{z \cdot s \cdot z^{-1}} = \mu_z \circ \mu_s \circ \mu_z^{-1}$, podemos limitar ao caso em que $t = \rho e_1$.

Desde que $s^{-1} = \overline{s}/\operatorname{N}(s)$, concluímos que

$$\mu_s(e_2) = \frac{1}{\alpha^2 + \rho^2}(\alpha e_0 + \rho e_1) \cdot e_2 \cdot (\alpha e_0 - \rho e_1) = \frac{\alpha^2 - \rho^2}{\alpha^2 + \rho^2} e_2 + \frac{2\alpha\rho}{\alpha^2 + \rho^2} e_3.$$

Assim, orientando o plano $P_{oe_2e_3}$ gerado por e_2, e_3 de modo que o par (e_2, e_3) seja positivo, vemos que se θ é o ângulo de rotação de μ_s ao redor de $t = \rho e_1$, então

$$\cos\theta = \frac{\alpha^2 - \rho^2}{\alpha^2 + \rho^2}, \qquad \operatorname{sen}\theta = \frac{2\alpha\rho}{\alpha^2 + \rho^2}, \tag{4.27}$$

uma vez que $R_\theta(e_2) = (\cos\theta)e_2 + (\operatorname{sen}\theta)e_3$, para R_θ tal rotação.

Finalmente, usando a identidade trigonométrica $\operatorname{tg}(\theta/2) = \dfrac{1 - \cos\theta}{\operatorname{sen}\theta}$ e (4.27), concluímos que $\operatorname{tg}(\theta/2) = \rho/\alpha$ e como $\rho = \|t\|$, segue o resultado. □

4.2 Rotações por meio de quatérnios

Quatérnios estão relacionados fundamentalmente com rotações tridimensionais, que podem representar orientações de uma base e podem agir para produzir mudanças na orientação desta base. Nesta seção, apresentamos algumas relações entre rotações bidimensionais e suas matrizes com números complexos. Por fim, passamos para matrizes de rotações tridimensionais, que estão relacionadas com os quatérnios.

Rotações bidimensionais

Rotações em um espaço bidimensional são implementadas pela ação de matrizes ortogonais com determinante 1, de ordem 2×2, denotadas a partir daqui por $R_2(\theta)$. Assim, denotando um vetor por uma matriz coluna, sua imagem pela rotação é obtida simplesmente pelo produto matricial, ou seja,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = R_2(\theta) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \end{bmatrix}. \tag{4.28}$$

Escrevendo $A = \cos\theta$ e $B = \operatorname{sen}\theta$, a matriz de rotação torna-se

$$R_2(\theta) = \begin{bmatrix} A & -B \\ B & A \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \tag{4.29}$$

e satisfaz

$$\det R_2(\theta) = A^2 + B^2 = \cos^2(\theta) + \operatorname{sen}^2(\theta) = 1$$

e

$$R_2(\theta)R_2(\theta)^T = \begin{bmatrix} A^2 + B^2 & 0 \\ 0 & A^2 + B^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2. \tag{4.30}$$

Relação com números complexos

Como apresentado na Seção 3.2, o produto $*$ de números complexos é dado por

$$(x_1 + y_1 \cdot i) * (x_2 + y_2 \cdot i) = (x_1x_2 - y_1y_2) + (x_1y_2 + x_2y_1) \cdot i.$$

As rotações bidimensionais podem ser representadas facilmente usando números complexos. Se ao vetor de coordenadas x, y associamos o complexo $z = x + y \cdot i$ então

$$z' = e^{\theta \cdot i} * (x + y \cdot i) = (x\cos\theta - y\operatorname{sen}\theta) + (x\operatorname{sen}\theta + y\cos\theta) \cdot i \tag{4.31}$$

reproduz exatamente a matriz coluna resultante do produto, como em (4.28). Ou seja, por meio da forma polar $e^{\theta \cdot i} = \cos\theta + \operatorname{sen}\theta \cdot i$, a rotação torna-se simplesmente um produto de números complexos.

A forma meio-ângulo

Examinaremos agora uma forma aparentemente simples de reescrever rotações bidimensionais. Embora ainda não seja possível o uso de quatérnios com rotações bidimensionais, uma rotação de $\frac{\theta}{2}$ pode fornecer algo útil.

Sejam $a = \cos\frac{\theta}{2}$ e $b = \operatorname{sen}\frac{\theta}{2}$. Por meio das relações trigonométricas para cosseno/seno de somas, obtemos

$$\begin{aligned} A &= \cos\theta = \cos\Big(\frac{\theta}{2} + \frac{\theta}{2}\Big) = \cos^2\Big(\frac{\theta}{2}\Big) - \operatorname{sen}^2\Big(\frac{\theta}{2}\Big) = a^2 - b^2, \\ B &= \operatorname{sen}\theta = \operatorname{sen}\Big(\frac{\theta}{2} + \frac{\theta}{2}\Big) = 2\cos\Big(\frac{\theta}{2}\Big)\operatorname{sen}\Big(\frac{\theta}{2}\Big) = 2ab. \end{aligned}$$

Assim, a matriz de rotação também pode ser escrita como

$$R_2(a,b) = \begin{bmatrix} a^2 - b^2 & -2ab \\ 2ab & a^2 - b^2 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}. \tag{4.32}$$

Notamos também que, quando calculamos o determinante usando as variáveis a, b, encontramos um resultado compatível, isto é,

$$\det R_2(a,b) = (a^2 - b^2)^2 + (2ab)^2 = a^4 + 2a^2b^2 + b^4 = (a^2 + b^2)^2 = 1. \tag{4.33}$$

Veremos que o par (a, b) pode ser tratado como um caso especial de um quatérnio.

Versão exponencial complexa

Usando a, b como acima, ao escrevermos o complexo $a+b\cdot i$ na forma polar, obtemos uma equação para o meio-ângulo,

$$e^{\frac{\theta}{2}\cdot i} = \cos\frac{\theta}{2} + \operatorname{sen}\frac{\theta}{2}\cdot i = a + b\cdot i, \tag{4.34}$$

onde $e^{\frac{\theta}{2}\cdot i} * e^{\frac{\theta}{2}\cdot i} = e^{\theta\cdot i}$. Assim, por (4.31) vemos que $e^{\frac{\theta}{2}\cdot i}$ é a raiz quadrada do número complexo original $e^{\theta\cdot i}$ que representa a rotação bidimensional.

4.2.1 Quatérnios e rotações tridimensionais

Assim como as matrizes $R_2(\theta)$, uma matriz de rotação tridimensional $R_3(\theta)$ é também ortogonal, ou seja, satisfaz $\det R_3(\theta) = 1$ e

$$R_3(\theta)R_3(\theta)^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3. \tag{4.35}$$

Tais matrizes podem ser descritas de várias maneiras, por exemplo, como em (4.39) ou (4.44). No entanto, aparecem dois novos fatos para rotações tridimensionais:

1. *Ordem de dependência:* O produto de duas matrizes de rotação S e T pode depender da ordem em que a multiplicação ocorre. Isto é, exceto em casos especiais, tem-se
$$ST \neq TS.$$

2. *Único vetor diretor:* $R_3(\theta)$ tem um único vetor diretor (Teorema de Rotação de Euler), e assim toda matriz de rotação, e seus produtos, podem ser escritas da seguinte forma em termos de uma única matriz rotação $R_3(\theta, \vec{n})$ que deixa uma particular direção fixada:
$$R_3(\theta, \vec{n})\vec{n} = \vec{n}.$$
Além de deixar $\vec{n}$ fixo, $R_3(\theta, \vec{n})$ expressa todas as possíveis rotações tridimensionais de um ângulo θ ao redor de uma direção $\vec{n}$.

Para simplificar, escreveremos apenas R e denotaremos as três matrizes de rotação tridimensional, que são simplesmente rotações no plano bidimensional determinado por cada par de eixos ortogonais, por

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\operatorname{sen}\theta \\ 0 & \operatorname{sen}\theta & \cos\theta \end{bmatrix}, \tag{4.36}$$

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \operatorname{sen}\theta \\ 0 & 1 & 0 \\ -\operatorname{sen}\theta & 0 & \cos\theta \end{bmatrix}, \tag{4.37}$$

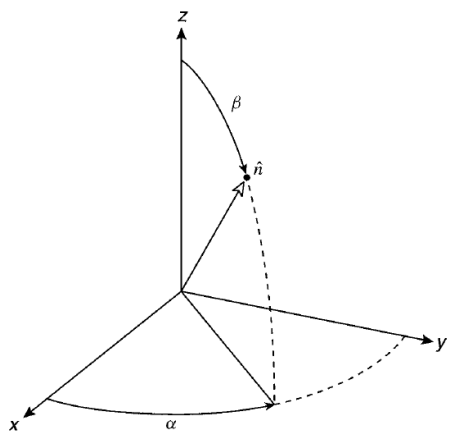
$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.38}$$

Estas produzem, pela regra da mão direita, rotações ao redor dos vetores da base $\vec{x} = (1,0,0)$, $\vec{y} = (0,1,0)$, $\vec{z} = (0,0,1)$, respectivamente. R_x rotaciona no plano yz ao redor de sua origem, R_y no plano zx e R_z no plano xy.

Construção

Podemos explicitamente construir $R(\theta, \vec{n})$ do seguinte modo:

1. Escrevendo $\vec{n}$ com coordenadas esféricas, temos $\vec{n} = (\cos\alpha \operatorname{sen}\beta, \operatorname{sen}\alpha \operatorname{sen}\beta, \cos\beta)$, com $0 \leq \alpha < 2\pi$ e $0 \leq \beta \leq \pi$, que denota o vetor diretor do eixo sobre o qual queremos rotacionar, e portanto deve ser um autovetor da rotação.



2. Se escrevermos $\vec{z}$ como o vetor coluna $(0,0,1)^T$, a figura acima no ajuda a ver que $\vec{n}$ pode ser obtido de $\vec{z}$ por meio de rotações, isto é,

$$\vec{n} = (R_z(\alpha)R_y(\beta))\vec{z}.$$

3. Para construirmos a matriz de rotação ao redor de $\vec{n}$, aplicamos $\vec{n}$ para $\vec{z}$ invertendo a operação do item acima, rotacionamos ao redor de $\vec{z}$ por θ usando a

matriz $R_z(\theta)$, e aplicamos $\vec{z}$ de volta para $\vec{n}$ (de onde começamos). Tal processo é descrito pela composição de rotações

$$R(\theta, \vec{n}) = R_z(\alpha) \circ R_y(\beta) \circ R_z(\theta) \circ R_y^T(\beta) \circ R_z^T(\alpha).$$

Escrevendo todos os componentes deste produto de matrizes, e escrevendo todas ocorrências de α e β em termos de

$$\vec{n} = (\cos\alpha \operatorname{sen}\beta, \operatorname{sen}\alpha \operatorname{sen}\beta, \cos\beta) = (n_1, n_2, n_3),$$

a matriz de rotação desejada que fixa a direção $\vec{n}$ é

$$R(\theta, \vec{n}) = \begin{bmatrix} c + n_1^2(1-c) & n_1n_2(1-c) - sn_3 & n_1n_3(1-c) + sn_2 \\ n_2n_1(1-c) + sn_3 & c + n_2^2(1-c) & n_2n_3(1-c) - sn_1 \\ n_3n_1(1-c) - sn_2 & n_3n_2(1-c) + sn_1 & c + n_3^2(1-c) \end{bmatrix}, \tag{4.39}$$

onde, por simplicidade, escrevemos $c = \cos\theta$ e $s = \operatorname{sen}\theta$.

Podemos ver que de fato $\vec{n}$ é fixado.

$$\begin{aligned} R(\theta, \vec{n})\vec{n} &= \begin{bmatrix} n_1c + n_1n_1^2(1-c) + n_1n_2^2(1-c) + n_1n_3^2(1-c) \\ n_2n_1^2(1-c) + n_2c + n_2n_2^2(1-c) + n_2n_3^2(1-c) \\ n_3n_1^2(1-c) + n_3n_2^2(1-c) + n_3c + n_3n_3^2(1-c) \end{bmatrix} \\ &= \begin{bmatrix} n_1(c+1-c) \\ n_2(c+1-c) \\ n_3(c+1-c) \end{bmatrix} = \vec{n}. \end{aligned} \tag{4.40}$$

Quatérnios e meio-ângulos

Voltemos à ideia de reescrever rotações bidimensionais na forma meio-ângulo. Para isso, consideremos a correspondência entre um par (a, b) de um número complexo e um quatérnio especial. Daí, observemos que em cada subespaço bidimensional pode-se relacionar uma rotação tridimensional de meio-ângulo.

Por exemplo, se $n_1 = n_2 = 0$ devemos ter $n_3 = 1$ e assim a rotação tridimensional ao redor de $\vec{n} = \vec{z}$ pode ser escrita na forma meio-ângulo, ou seja,

$$R_z = \begin{bmatrix} \cos\theta & -n_3 \operatorname{sen}\theta & 0 \\ n_3 \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a^2 - b^2 & -2ab & 0 \\ 2ab & a^2 - b^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.41}$$

onde $a = \cos\theta$, $b = \operatorname{sen}\theta$.

Analogamente, escrevemos as outras duas rotações ao redor de $\vec{x}$ e $\vec{y}$ utilizando meio-ângulo.

Parece altamente provável que qualquer matriz de rotação ao redor de um vetor $\vec{n}$ arbitrário também possa ser expressa em termos do meio-ângulo, e não apenas as três matrizes particulares acima.

Assim, tentamos reescrever toda a matriz em termos de $\frac{\theta}{2}$ usando $a = \cos\frac{\theta}{2}$ e $b = \operatorname{sen}\frac{\theta}{2}$ para simplificar a notação. Porém o fator $1 - c$ em (4.39) não se modifica imediatamente, como ocorre no caso bidimensional em (4.41), até que notemos que $a^2 + b^2 = 1$.

Então, fazendo a substituição $1 - c = (a^2 + b^2) - (a^2 - b^2) = 2b^2$ encontramos

$$R(\theta, \vec{n}) = \begin{bmatrix} a^2 - b^2 + n_1^2 2b^2 & 2b^2 n_1 n_2 - 2abn_3 & 2b^2 n_3 n_1 + 2abn_2 \\ 2b^2 n_1 n_2 + 2abn_3 & a^2 - b^2 + n_2^2 2b^2 & 2b^2 n_2 n_3 - 2abn_1 \\ 2b^2 n_3 n_1 - 2abn_2 & 2b^2 n_2 n_3 + 2abn_1 & a^2 - b^2 + n_3^2 2b^2 \end{bmatrix}. \tag{4.42}$$

Fazendo uso da relação $n_1^2 + n_2^2 + n_3^2 = 1$ (pois $\vec{n}$ é unitário), tornamos a expressão ainda mais simétrica por meio das substituições

$$\begin{aligned} -b^2 + n_1^2 2b^2 &= -(n_1^2 + n_2^2 + n_3^2)b^2 + 2b^2 n_1^2 \\ &= b^2(n_1^2 - n_2^2 - n_3^2), \\ -b^2 + n_2^2 2b^2 &= -(n_1^2 + n_2^2 + n_3^2)b^2 + 2b^2 n_2^2 \\ &= b^2(-n_1^2 + n_2^2 - n_3^2), \\ -b^2 + n_3^2 2b^2 &= -(n_1^2 + n_2^2 + n_3^2)b^2 + 2b^2 n_3^2 \\ &= b^2(-n_1^2 - n_2^2 + n_3^2). \end{aligned}$$

Então,

$$R(\theta, n) = \begin{bmatrix} a^2 + b^2(n_1^2 - n_2^2 - n_3^2) & 2b^2 n_1 n_2 - 2ab & 2b^2 n_3 n_1 + 2abn_2 \\ 2b^2 n_1 n_2 + 2abn_3 & a^2 + b^2(n_2^2 - n_3^2 - n_1^2) & 2b^2 n_2 n_3 - 2abn_1 \\ 2b^2 n_3 n_1 - 2abn_2 & 2b^2 n_2 n_3 + 2abn_1 & a^2 + b^2(n_3^2 - n_1^2 - n_2^2) \end{bmatrix}.$$

Esta pode agora ser escrita unicamente em termos de pontos na esfera S^3, e assim em termos de quatérnios unitários. Mais precisamente, escolhemos a seguinte parametrização para pontos de S^3, que garante que $q = (q_0, q_1, q_2, q_3)$ seja unitário:

$$\begin{aligned} q_0 &= a = \cos\frac{\theta}{2}, \\ q_1 &= n_1 b = n_1 \operatorname{sen}\frac{\theta}{2}, \\ q_2 &= n_2 b = n_2 \operatorname{sen}\frac{\theta}{2}, \\ q_3 &= n_3 b = n_3 \operatorname{sen}\frac{\theta}{2}. \end{aligned} \tag{4.43}$$

Substituindo na matriz acima, encontramos o resultado final

$$R(q) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2q_1 q_2 - 2q_0 q_3 & 2q_1 q_3 + 2q_0 q_2 \\ 2q_1 q_2 + 2q_0 q_3 & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2q_2 q_3 - 2q_0 q_1 \\ 2q_1 q_3 - 2q_0 q_2 & 2q_2 q_3 + 2q_0 q_1 & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix}. \tag{4.44}$$

O quatérnio $q(\theta, \vec{n}) = (\cos\frac{\theta}{2}, \operatorname{sen}\frac{\theta}{2}\vec{n})$ de (4.43), quando substituído em (4.44), produz a matriz de rotação $R(\theta, \vec{n})$ de (4.39) para uma rotação de θ no plano perpendicular a $\vec{n}$, um vetor tridimensional unitário.

Agora, podemos verificar várias propriedades da matriz em (4.44) que são importantes na Álgebra Linear. Por exemplo, enquanto o determinante de R_2 em (4.32) é o quadrado de uma soma de quadrados, para R_3 será um cubo porque existe uma potência adicional para o determinante de uma matriz que possui uma dimensão a mais,

$$\det R(q) = \langle q, q\rangle^3 = 1,$$

onde $\langle -, -\rangle$ denota o produto escalar usual.

Além disso, cada linha (ou coluna) da matriz em (4.44) tem norma um. Por exemplo, somando o quadrado das entradas da primeira linha temos

$$(q_0^2 + q_1^2 - q_2^2 - q_3^2)^2 + (2q_1q_2 - 2q_0q_3)^2 + (2q_1q_3 + 2q_0q_2)^2 = \langle q, q\rangle^2 = 1.$$

Também, cada linha (ou coluna) é ortogonal às demais. Por exemplo, para as colunas 1 e 2,

$$\sum_{i=1}^{3} R_{i1}R_{i2} = 0.$$

Tais propriedades confirmam que $R(q)$ é uma matriz ortogonal.

Observação 4.2 (Valores duplos)**.** O fato das entradas da matriz em (4.44) terem grau 2 com relação às coordenadas do quatérnio q, ou seja, somente valores quadráticos aparecem na matriz, facilita vermos que

$$R(q) = R(-q).$$

Assim, os quatérnios q e $-q$ geram a mesma matriz de rotação tridimensional, e portanto temos uma cobertura dupla (função dois-a-um) para as rotações tridimensionais.

Recuperando θ e $\vec{n}$

Dada uma matriz de rotação arbitrária M, podemos descobrir o autovetor $\vec{n}$ e o valor do ângulo θ de rotação usando um cálculo direto e eficiente, evitando abordagens complexas da Álgebra Linear. Procedemos do seguinte modo:

Determinação do eixo: Se $\operatorname{sen}\theta$ for não nulo e grande o suficiente para evitar erros numéricos, $\vec{n}$ pode ser calculado subtraindo-se da matriz R sua transposta e usando a equação (4.39) para encontrar:

$$R - R^T = \begin{bmatrix} 0 & -2n_3\operatorname{sen}\theta & 2n_2\operatorname{sen}\theta \\ 2n_3\operatorname{sen}\theta & 0 & -2n_1\operatorname{sen}\theta \\ -2n_2\operatorname{sen}\theta & 2n_1\operatorname{sen}\theta & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{bmatrix}. \quad (4.45)$$

Daí, simplesmente calculamos $\delta = \sqrt{\alpha^2 + \beta^2 + \gamma^2}$ e normalizamos para produzir o resultado

$$\vec{n} = \left(\frac{\alpha}{\delta}, \frac{\beta}{\delta}, \frac{\gamma}{\delta}\right).$$

Determinação do ângulo: De (4.39), calculamos o traço e obtemos

$$t = \mathrm{Tr}(R) = 1 + 2\cos\theta,$$

de onde segue que

$$\cos\theta = \frac{1}{2}(t-1),$$
$$\operatorname{sen}\theta = +\sqrt{1-\cos^2\theta}.$$

Pelo fato de que $R(\theta, \vec{n}) = R(-\theta, -\vec{n})$, o sinal de $\operatorname{sen}\theta$ pode ser tomado como positivo na raiz quadrada sem perda de generalidade.

Ângulos de Euler e quatérnios

Há muitas maneiras de parametrizarmos uma matriz de rotação tridimensional. Até agora, focamos na versão eixo-ângulo de $R(\theta, \vec{n})$ com o quatérnio correspondente

$$q(\theta, \vec{n}) = \left(\cos\frac{\theta}{2}, \operatorname{sen}\frac{\theta}{2}\vec{n}\right),$$

onde o autovetor de $R(\theta, \vec{n})$ pode ser parametrizado por

$$\vec{n} = (\cos\alpha\operatorname{sen}\beta, \operatorname{sen}\alpha\operatorname{sen}\beta, \cos\beta).$$

Esta forma, por fazer uso explícito do seu autovetor, evita o *gimbal lock*[1] descrito a seguir. No entanto, muitas aplicações tratam rotações como sequências de rotações da forma eixo-ângulo, podendo assim levar ao surgimento de *gimbal lock*.

Efeito *gimbal lock*

Para entendermos como uma sequência de três rotações pode fornecer um resultado indesejado, conhecido como 'efeito *gimbal lock*', precisamos primeiramente entender o que vem a ser um *gimbal*.

Na figura 4.1 podemos ver um aparato, conhecido por *giroscópio*, formado por três circunferências (chamadas de *gimbal*), que estão posicionadas em planos dois a dois ortogonais. Pelo centro de cada *gimbal* existe o eixo de rotação, ao redor do qual ele pode rotacionar. Se pensarmos em termos de coordenadas, no sistema cartesiano ortogonal, teremos os três planos coordenados e os eixos de rotação tornam-se os eixos gerados pelos vetores $\vec{x}, \vec{y}$ e $\vec{z}$, descritos acima.

[1] Ver `https://en.wikipedia.org/wiki/Gimbal_lock`

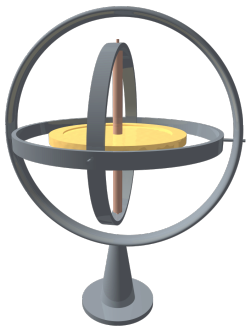


Figura 4.1: Giroscópio

Dizemos que o efeito *gimbal lock* ocorre quando um dos *gimbals* torna-se coplanar com algum outro, após uma rotação sobre seu eixo. Esse alinhamento faz com que dois eixos coincidam, reduzindo assim um grau de liberdade para o movimento de rotação.

Por exemplo, vejamos o que ocorre ao fazermos a rotação R descrita pela composição $R_x(\alpha)R_y(\pi/2)R_z(\gamma)$. Neste caso, calculamos

$$\begin{aligned}
R &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\gamma & -\operatorname{sen}\gamma & 0 \\ \operatorname{sen}\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 1 \\ \operatorname{sen}\alpha & \cos\alpha & 0 \\ -\cos\alpha & \operatorname{sen}\alpha & 0 \end{bmatrix} \begin{bmatrix} \cos\gamma & -\operatorname{sen}\gamma & 0 \\ \operatorname{sen}\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 1 \\ \operatorname{sen}(\alpha+\gamma) & \cos(\alpha+\gamma) & 0 \\ -\cos(\alpha+\gamma) & \operatorname{sen}(\alpha+\gamma) & 0 \end{bmatrix}.
\end{aligned}$$

Assim, notemos que a rotação final depende dos ângulos α, γ somente, sendo sempre uma rotação ao redor do eixo-z. Dizemos então que houve um travamento do mecanismo.

A Figura 4.2 mostra uma situação de *gimbal lock*, não necessariamente para os valores tomados acima.

Para evitarmos o efeito *gimbal lock*, basta descrevermos as rotações por meio de quatérnios, obtendo assim um ângulo e um único eixo de rotação.

Sequências triplas de rotações da forma eixo-ângulo são geralmente chamadas de representações por ângulos de Euler. Estas também possuem uma representação quaterniônica, a mais comum sendo conhecida por representação XYZ.

Representação XYZ. Recebe este nome por ser uma sequência de três rotações ao redor dos três eixos coordenados. Por isso, às vezes nos leva a crer que não permite *gimbal lock*, mas na verdade não é bem assim.

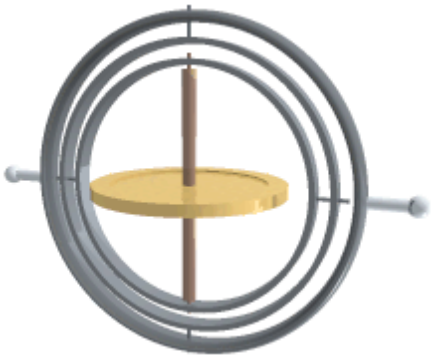


Figura 4.2: Efeito *gimbal lock*

Nesta forma, a matriz é escrita como produto das três matrizes elementares que representam rotações ao redor dos eixos, não necessariamente por um mesmo ângulo. Por isso, é denotada e definida por

$$\begin{aligned} R_{xyz}(\alpha,\beta,\gamma) &= R_x(\alpha)R_y(\beta)R_z(\gamma) \\ &= \begin{bmatrix} \cos\beta\cos\gamma & -\cos\beta\operatorname{sen}\gamma & \operatorname{sen}\beta \\ \operatorname{sen}\alpha\operatorname{sen}\beta\cos\gamma+\cos\alpha\operatorname{sen}\gamma & \cos\alpha\cos\gamma-\operatorname{sen}\alpha\operatorname{sen}\beta\operatorname{sen}\gamma & -\operatorname{sen}\alpha\cos\beta \\ -\cos\alpha\operatorname{sen}\beta\cos\gamma+\operatorname{sen}\alpha\operatorname{sen}\gamma & \operatorname{sen}\alpha\cos\gamma+\cos\alpha\operatorname{sen}\beta\operatorname{sen}\gamma & \cos\alpha\cos\beta \end{bmatrix}. \end{aligned} \tag{4.46}$$

O quatérnio q correspondente é dado por

$$\begin{aligned} q_0 &= \cos\frac{\alpha}{2}\cos\frac{\beta}{2}\cos\frac{\gamma}{2} - \operatorname{sen}\frac{\alpha}{2}\operatorname{sen}\frac{\beta}{2}\operatorname{sen}\frac{\gamma}{2}, \\ q_1 &= \operatorname{sen}\frac{\alpha}{2}\cos\frac{\beta}{2}\cos\frac{\gamma}{2} + \cos\frac{\alpha}{2}\operatorname{sen}\frac{\beta}{2}\operatorname{sen}\frac{\gamma}{2}, \\ q_2 &= \cos\frac{\alpha}{2}\operatorname{sen}\frac{\beta}{2}\cos\frac{\gamma}{2} - \operatorname{sen}\frac{\alpha}{2}\cos\frac{\beta}{2}\operatorname{sen}\frac{\gamma}{2}, \\ q_3 &= \cos\frac{\alpha}{2}\cos\frac{\beta}{2}\operatorname{sen}\frac{\gamma}{2} + \operatorname{sen}\frac{\alpha}{2}\operatorname{sen}\frac{\beta}{2}\cos\frac{\gamma}{2}. \end{aligned}$$

Podemos ver a partir das coordenadas de q que, para $\beta = \pm\frac{\pi}{2}$, temos uma função apenas nas variáveis $\alpha+\gamma$ ou $\alpha-\gamma$, respectivamente, pois

$$q\Big(\alpha, \frac{\pm\pi}{2}, \gamma\Big) = \frac{1}{\sqrt{2}}\Big(\cos\frac{\alpha\pm\gamma}{2}, \operatorname{sen}\frac{\alpha\pm\gamma}{2}, \pm\cos\frac{\alpha\pm\gamma}{2}, \pm\operatorname{sen}\frac{\alpha\pm\gamma}{2}\Big). \tag{4.47}$$

Representação ZYZ. Por fim, apenas para mencionar, existe a representação ZYZ, que é tradicionalmente usada na Física Clássica e Quântica, para parametrizar o movimento de um objeto girando. Neste caso, a matriz é dada por

$$R_{zyz}(\alpha,\beta,\gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma) = R(\gamma,\vec{c})R(\beta,\vec{b})R(\alpha,\vec{z}) \tag{4.48}$$

onde $\vec{c} = (\cos\alpha\operatorname{sen}\beta, \operatorname{sen}\alpha\operatorname{sen}\beta, \cos\beta)$ e $\vec{b} = (-\operatorname{sen}\alpha, \cos\alpha, 0)$.

O quatérnio correspondente é dado por

$$\begin{aligned} q_0 &= \cos\frac{\beta}{2}\cos\frac{1}{2}(\alpha+\gamma), \\ q_1 &= \operatorname{sen}\frac{\beta}{2}\operatorname{sen}\frac{1}{2}(\gamma-\alpha), \\ q_2 &= \operatorname{sen}\frac{\beta}{2}\cos\frac{1}{2}(\gamma-\alpha), \\ q_3 &= \cos\frac{\beta}{2}\operatorname{sen}\frac{1}{2}(\alpha+\gamma). \end{aligned}$$

Podemos ver que, para $\beta = 0$ temos uma função apenas de $(\alpha+\gamma)$, e para $\beta = \pi$, apenas $(\alpha-\gamma)$.

Conclusão. A partir do exposto acima e fazendo uso da Proposição 4.2, é possível então escrevermos explicitamente a matriz de rotação associada a qualquer quatérnio $s = \alpha e_0 + t$, com t quatérnio puro.

Referências

[1] GONÇALVES, A. *Introdução à Álgebra*. 5. ed. Rio de Janeiro: IMPA, 2013.

[2] LANG, S. *Algebra*. 3. ed. New York: Springer-Verlag New York, 2002.

[3] LANG, S. *Linear Algebra*. 3. ed. New York: Springer-Verlag New York, 2004.

[4] LIMA, E. L. *Álgebra Linear*. 9. ed. Rio de Janeiro: IMPA, 2016.

[5] DIEUDONNÉ, J. A. *Linear Algebra and Geometry*. 1. ed. Paris: HERMANN, 1969.

[6] HANSON, A. J. *Visualizing Quaternions*. 1. ed. San Francisco: Morgan Kaufmann, 2006.