

[www.im.ufrj.br/cvga](http://www.im.ufrj.br/cvga)

**CÁLCULO VETORIAL  
&  
GEOMETRIA ANALÍTICA**  
livro 1: o plano

**Felipe Acker**



**Instituto de Matemática  
Universidade Federal do Rio de Janeiro**

agosto de 2019

copyright ©2019 by Felipe Acker

Este trabalho foi contemplado com auxílio financeiro, no âmbito do edital de *Apoio à produção de material didático para atividades de ensino e/ou pesquisa*, 2014, da



# Sumário

<b>Prefácio</b>	<b>i</b>
<b>Índice dos vídeos de 2015</b>	<b>iii</b>
<b>Índice dos vídeos de 2018</b>	<b>vii</b>
<b>1 Sistemas de coordenadas</b>	<b>1</b>
<b>2 Medindo distâncias</b>	<b>5</b>
<b>3 Os objetos geométricos: retas e círculos</b>	<b>9</b>
<b>4 Curvas e equações. Lugares geométricos</b>	<b>15</b>
<b>5 Interseções. Sistemas de equações</b>	<b>19</b>
<b>6 Geometria Analítica, Desenho Geométrico e Computação Gráfica</b>	<b>21</b>
<b>7 Vetores</b>	<b>25</b>
7.1 Flechinhas . . . . .	25
7.2 Norma . . . . .	26
7.3 Produto por escalar . . . . .	26
7.4 Soma de vetores . . . . .	27
7.5 Somando vetores a pontos . . . . .	29
7.6 Vetores e parametrizações . . . . .	30
<b>8 O mistério da Santíssima Trindade</b>	<b>31</b>
8.1 Vetores e coordenadas . . . . .	31
8.2 Pontos e vetores . . . . .	33
8.3 Vetores e pares ordenados . . . . .	33
8.4 A Santíssima Trindade . . . . .	36
<b>9 Equações paramétricas</b>	<b>37</b>
<b>10 Curvas de Bézier</b>	<b>47</b>
<b>11 Transformações e animações</b>	<b>51</b>

<b>12</b>	<b>Translações</b>	<b>55</b>
12.1	Movimento retilíneo uniforme . . . . .	55
12.2	Movimento retilíneo não uniforme . . . . .	58
12.3	Trajetórias não retilíneas . . . . .	60
12.4	Resumindo e Simplificando . . . . .	63
<b>13</b>	<b>Rotações</b>	<b>65</b>
13.1	Rotações em torno da origem . . . . .	65
13.2	Rotação em torno de um ponto qualquer . . . . .	67
13.3	Rotação de vetores . . . . .	68
<b>14</b>	<b>Homotetias</b>	<b>73</b>
<b>15</b>	<b>Reflexões</b>	<b>75</b>
15.1	Reflexão de ponto através de reta passando pela origem . . . . .	75
15.2	Reflexão de vetores . . . . .	76
15.3	Animando reflexões . . . . .	79
<b>16</b>	<b>Deformações</b>	<b>81</b>
16.1	Casos elementares . . . . .	81
16.2	Deformações em outras dimensões . . . . .	83
<b>17</b>	<b>Transformações lineares</b>	<b>85</b>
17.1	Definição . . . . .	85
17.2	Transformações preservando distâncias . . . . .	86
<b>18</b>	<b>Produto interno</b>	<b>89</b>
<b>19</b>	<b>A Decomposição em Valores Singulares</b>	<b>95</b>
19.1	O Teorema . . . . .	95
19.2	A geometria do Teorema . . . . .	95
19.3	Agora com matrizes . . . . .	96
19.4	A SVD em termos de mudanças de base . . . . .	97
19.5	Mudanças de base . . . . .	100
19.6	De novo a Decomposição em Valores Singulares, versão matricial . . . . .	101
<b>20</b>	<b>Otimização</b>	<b>103</b>
20.1	Amendoim torrado . . . . .	103
<b>21</b>	<b>Áreas e determinantes</b>	<b>107</b>
21.1	Orientação . . . . .	108
21.2	Áreas com sinal . . . . .	109
21.3	O determinante de uma transformação linear . . . . .	112
<b>22</b>	<b>O Teorema Espectral</b>	<b>117</b>
22.1	Matrizes simétricas e transformações autoadjuntas . . . . .	117
22.2	O Teorema Espectral . . . . .	118
22.3	Matriz de transformação linear em uma base qualquer . . . . .	120

22.4	Teorema Espectral, agora com matrizes	121
22.5	O polinômio característico	122
22.6	O Teorema Espectral e a SVD	123
22.7	Exercícios Básicos 4: Teorema Espectral & Decomposição em Valores Singulares	123
<b>23</b>	<b>Números complexos e coordenadas polares</b>	<b>125</b>
23.1	Os complexos	125
23.2	Inversões	128
<b>24</b>	<b>Equações Polinomiais</b>	<b>131</b>
24.1	Equações polinomiais no Renascimento	131
24.2	A fórmula de del Ferro-Tartaglia	132
24.3	A fórmula de Ferrari	133
24.4	Problemas inerentes à solução de del Ferro-Tartaglia	134
24.5	A invenção dos números complexos	136
<b>25</b>	<b>O Teorema Fundamental da Álgebra</b>	<b>137</b>
<b>26</b>	<b>O Número de Voltas</b>	<b>143</b>
26.1	O número de Voltas	143
26.2	Curvas e Homotopias	143
26.3	O princípio fundamental	144
26.4	Lema de Brouwer	145
26.5	Lema de Rouché	146
<b>27</b>	<b>Demonstração do Teorema Fundamental da Álgebra</b>	<b>149</b>
<b>28</b>	<b>O Teorema de Brouwer</b>	<b>153</b>
<b>29</b>	<b>Iterações</b>	<b>155</b>
29.1	De Ferrari a Galois	155
29.2	Soluções <i>aproximadas</i>	155
29.3	A iteração de Newton	156
29.4	A Ferradura de Smale	157



# Prefácio

Este pequeno texto foi concebido como uma espécie de abertura. Nele procurei concentrar ideias básicas para um curso de Geometria Analítica no plano. A primeira versão, escrita em janeiro de 1998, destinou-se a um minicurso de aperfeiçoamento, ministrado na UFRJ, para professores de Matemática do ensino médio. De lá para cá, o conteúdo aumentou um bocado. Creio que o caráter de minicurso se mantém: os capítulos iniciais podem ser usados nas primeiras aulas de um curso de Geometria Analítica para apresentar aos alunos, de maneira rápida, as coordenadas, os vetores, as curvas e as transformações do plano. O texto vai, pouco a pouco, ficando mais denso, mas, acredito, ainda guarda uma certa leveza. Dois índices com vídeos das aulas estão incluídos, mas há mais em meu canal no youtube: <https://www.youtube.com/c/FelipeAcker>.

A ideia de ressaltar as relações entre a Geometria Analítica e a Computação Gráfica teve, na versão original (que se chamava Geometria Analítica para Computação Gráfica), e continua tendo agora, um caráter algo oportunista.<sup>1</sup> Computação Gráfica remete aos videogames, que são para boa parte dos estudantes um espaço associado ao prazer. Mas trata-se, também, de enfatizar o papel central assumido pela Geometria Analítica no Desenho, em suas múltiplas facetas, do artístico ao técnico, do estático ao animado. Mesmo para quem se contenta em usar programas de computador prontos, um certo domínio da Matemática envolvida é de grande valia. Afinal, ninguém discutiria a importância, para um pintor, de conhecer um pouco do processo de fabricação das tintas e, mesmo, de ser capaz de produzir e misturar seus próprios pigmentos.

De qualquer forma, em sua versão atual, o livrinho já vai bem além de um texto voltado para computação gráfica.

Vários colegas ajudaram, cada um à sua maneira: Bernardo Costa, Dinamérico Pombo Jr (que leram e sugeriram mudanças, nem todas acatadas), Monique Carmona, Orestes Piermatei Filho, Ricardo Rosa, Umberto Hryniewicz e Waldecir Bianchini (que criou os applets no Geogebra); agradeço a todos.

Felipe Acker  
Santa Teresa, agosto de 2019

---

<sup>1</sup>As figuras, paradoxalmente, foram feitas por mim, na época, com lapiseiras Caran d'Ache, esquadros e compasso; para esta edição, os desenhos foram convertidos em arquivos digitais por João Paulo Pinto Siqueira



# Índice dos vídeos de 2015

Os vídeos abaixo listados foram filmados durante as aulas da disciplina MAE115 Cálculo Vetorial e Geometria Analítica - parte 1, do Departamento de Matemática Aplicada - Instituto de Matemática da Universidade Federal do Rio de Janeiro, no primeiro semestre de 2015, Professor FELIPE ACKER.

Ver também

[https://www.youtube.com/playlist?list=PLCn-wCoLo2pHeG0kD49zV470p8K\\_EH4x1](https://www.youtube.com/playlist?list=PLCn-wCoLo2pHeG0kD49zV470p8K_EH4x1)

para maior comodidade, as aulas foram divididas em 4 partes menores (cerca de 20 minutos, cada, em 4K), conforme o índice abaixo

## **aula 1 :**

parte 1 <https://www.youtube.com/watch?v=PCpjeREAzMw>  
coordenadas  
a reta real  
um número irracional  
distância

parte 2 <https://www.youtube.com/watch?v=hA5bTZqioyY>  
equações cartesianas  
algebrizar a Geometria & geometrizar a Álgebra  
algebrizando a Geometria  
o Teorema de Aproximação de Weierstrass  
geometrizando a Álgebra

parte 3 <https://www.youtube.com/watch?v=7QnPPY-fTE8>  
vetores  
multiplicação de vetor por número  
soma de vetores  
coordenadas  
o mistério da Santíssima Trindade  
equação paramétrica de uma reta  
somando vetor a ponto

parte 4 <https://www.youtube.com/watch?v=RqJ8DopXHlo>  
animações  
transformando uma imagem em outra imagem (morphing)

dimensões mais altas  
animando objetos

parte 5 <https://www.youtube.com/watch?v=7zjiMZc9w7s>  
curvas e animações  
um ponto se movendo

## aula2 :

parte 1 [https://www.youtube.com/watch?v=qkQ\\_9juiepY](https://www.youtube.com/watch?v=qkQ_9juiepY)  
pontos, vetores & pares ordenados  
vetores  
relações de equivalência  
o que é um vetor?  
operações com vetores  
sistemas de coordenadas  
base  
origem  
o que é um sistema de coordenadas  
o mistério da Santíssima Trindade

parte 2 <https://www.youtube.com/watch?v=qaDwMauCRsY>  
a bijeção entre vetores e pares ordenados preserva as operações  
curvas e equações  
curvas e trajetórias

parte 3 <https://www.youtube.com/watch?v=SXGquzdJSgU>  
cicloide  
o professor mente e massacra os alunos  
geometria analítica e desenho técnico  
referencial móvel

parte 4 <https://www.youtube.com/watch?v=FzYlKmfsh9Y>  
transformações  
rotação  
matriz de rotação

parte 5 <https://www.youtube.com/watch?v=VIZ1ID0U0v0>  
transformações lineares  
o descanso da vovó  
matriz de transformação linear  
transformações lineares & sistemas lineares  
reflexão

## aula 3 :

parte 1 <https://www.youtube.com/watch?v=w9Ru4FPptrg>  
área  
área com sinal

propriedades da área com sinal  
a fórmula

parte 2 <https://www.youtube.com/watch?v=XrMSRsCEGcs>  
o determinante  
determinante de transformação linear

#### **aula 4 :**

parte 1 <https://www.youtube.com/watch?v=1CTQbSd16aE>  
produto escalar  
definição geométrica  
propriedades  
expressão do produto escalar em função das coordenadas

parte 2 <https://www.youtube.com/watch?v=CwB30VpyPLI>  
arco capaz de ângulo reto

parte 3 <https://www.youtube.com/watch?v=p8lCbN24hoQ>  
números complexos  
equações do segundo grau  
equações do terceiro grau  
necessidade dos números imaginários  
Bombelli  
uma história italiana

parte 4 [https://www.youtube.com/watch?v=Fim\\_Z1N6y9M](https://www.youtube.com/watch?v=Fim_Z1N6y9M)  
os números complexos  
o plano complexo  
interpretação geométrica da multiplicação  
rotações e as fórmulas para cosseno e seno da soma de dois ângulos  
a conta conjunta dos professores de Matemática

parte 5 <https://www.youtube.com/watch?v=mbmrOJ9oXkY>  
potenciação e radiciação

#### **aula 5 :**

parte 1 [https://www.youtube.com/watch?v=ZYW\\_VOCZfPs](https://www.youtube.com/watch?v=ZYW_VOCZfPs)  
números complexos, definições  
multiplicação de números complexos e interpretação geométrica  
multiplicação por número complexo como composta de rotação com homotetia;  
forma matricial  
coordenadas polares

parte 2 <https://www.youtube.com/watch?v=RxKichfTt0g>  
potências inteiras  
raízes enésimas da unidade  
raízes enésimas de número complexo são vértices de polígono regular

conjugado de número complexo  
o Teorema Fundamental da Álgebra

parte 3 <https://www.youtube.com/watch?v=ntSHjjxeSlI>

polinômios

polinômios reais de grau ímpar têm raiz real

polinômios a coeficientes complexos

polinômios como funções de  $\mathbb{C}$  em  $\mathbb{C}$

estratégia de demonstração do Teorema Fundamental da Álgebra

parte 4 <https://www.youtube.com/watch?v=cMNnsjZEfWQ>

demonstração do Teorema Fundamental da Álgebra

comentários de caráter histórico

parte 5 <https://www.youtube.com/watch?v=iOLkWa5EqB8>

$f(z)=1/z$  e inversões

inversões

propriedades geométricas das inversões

comentário sobre preservação de ângulos e derivabilidade complexa

## aula 6 :

parte 1 <https://www.youtube.com/watch?v=8x-xw5r2GF4>

números complexos como pares ordenados

módulo e conjugado de número complexo

os 3 mosqueteiros: ponto, vetor, par ordenado e número complexo

inversões e números complexos

inversões e a transformação  $f(z) = 1/z$

a transformação  $f(z) = 1/z$

parte 2 <https://www.youtube.com/watch?v=18EU4IpMkU8>

$f(z) = 1/z$  leva retas que não passam por 0 em círculos passando por 0

não leiam demonstrações!

$f(z) = 1/z$  leva círculos que não passam por 0 em círculos que não passam por 0

inversões preservam ângulos

# Índice dos vídeos de 2018

Os vídeos abaixo listados foram filmados durante as aulas da disciplina MAE115 Cálculo Vetorial e Geometria Analítica - parte 1, do Departamento de Matemática Aplicada - Instituto de Matemática da Universidade Federal do Rio de Janeiro, no primeiro semestre de 2018, Professor FELIPE ACKER.

Link para a playlist:

<https://www.youtube.com/playlist?list=PLsVKfJEcnr2d3BALAt0zLaRHA5QREUgz>

CVGA 2018 1 - Vetores e Sistemas de Coordenadas

00:08 Introdução  
01:17 Apresentação do curso  
05:37 Apresentação do Primeiro Minicurso: O plano  
07:05 Novo Testamento: Criação do Sistema de Coordenadas  
Definição Analítica de Distância 10:37  
Definição Analítica de círculo 13:12  
"Algebrizando" a Geometria 14:13  
Geometrizando a Álgebra 15:53  
Generalização para Problemas com muitas variáveis 17:49  
Perguntem! 19:57  
Vetores e Escalares 21:48  
Soma de Vetores 24:19  
Sistemas de Coordenadas (revisitado) 29:09  
Base 32:10  
Bijecção entre Vetores e Pares Ordenados 34:17  
Soma de Vetores coordenada a coordenada 36:09  
Bijecção entre pontos e vetores 42:41  
Santíssima Trindade: ponto, vetor, par ordenado 45:50  
Vetores "alinhados" não formam base 49:23  
Soma de Pontos e Vetores: a operação bastarda 57:48  
Equação Vetorial de segmento e de reta 01:00:34  
Geometria e Álgebra se encontram nos Convexos 01:07:02  
Um pouco de Curvas Parametrizadas 01:08:14  
Desenhando Curvas com o computador 01:13:36  
Movimento dos ponteiros do relógio 01:15:34  
Espiral 01:23:30

Sistema de coordenadas móvel (Vidro) 01:26:15  
Cicloide 01:32:58

## CVGA 2018 2 - Lugares Geométricos e Curvas Parametrizadas

Curvas: Lugares Geométricos e Movimentos 00:00:19  
Círculo como Lugar Geométrico 00:02:33  
Círculo como Movimento de Partícula 00:03:55  
Diferença das Abordagens: Plotando o Gráfico 00:07:01  
Site com exemplos de curvas e suas equações 00:10:18  
A reta 00:11:22  
Por que equações, tipicamente, definem curvas? 00:15:26  
a Cicloide 00:18:54  
A Geometria Analítica, um pouco de história 00:27:44  
Sistema de coordenadas móvel 00:32:00  
a Hipocicloide 00:37:04  
Definição de Ângulo 00:39:51  
Parametrização da Hipocicloide por ângulos 00:46:30  
A Hipocicloide como composição de movimentos 00:53:46  
Curvas de Bézier 01:01:13  
Curva de Bézier dada por  $n+1$  pontos 01:07:55  
Binômio de Newton 01:11:48  
Parametrização das Curvas de Bézier 01:16:28  
Prova da fórmula das Curvas de Bézier por Indução 01:22:07  
Curvas de Bézier e animações 01:33:37

## CVGA 2018 3 - Transformações Lineares e Matrizes; Produto Escalar

Transformações que não cabem em gráficos 00:00:22  
Sobre Transformações Lineares 00:01:39  
 $\mathbb{R}^2$  visto como um plano virtual 00:05:18  
Onde entram as Matrizes 00:06:45  
Transf. Lineares preservam distância? R: Não 00:07:41  
Sobre Transformações e Sistemas Lineares 00:08:39  
Sobre Produto Escalar e Lei dos Cossenos 00:15:57  
Propriedades Lineares: Soma 00:18:33  
De Matrizes para Transformações e de volta para Matrizes 00:20:55  
Visualizando a Mudança do Sistema de Coordenadas 00:23:39  
Toda Transformação dada por Matriz é Linear 00:29:45  
Toda Transformação Linear é dada por Matriz 00:38:32  
Uma Transformação Linear Simples: Rotação 00:42:02  
Cosseno e Seno da soma 00:47:37  
Composição de Transformações e Produto de Matrizes 00:51:45  
Prévia: O determinante 01:02:13  
O Produto Escalar 01:07:30  
Definição e Propriedade do Produto Escalar 01:11:23  
Propriedades específicas no caso canônico 01:14:47  
O Primeiro Produto Escalar da História 01:23:22

Usando os Vetores Linha e Coluna de uma Matriz 01:25:13  
Equação da Reta 01:28:33  
Outra Forma de Mostrar a Equação da Reta 01:31:30  
Propriedades do Produto Escalar em qualquer Sistema Ortonormal 01:33:12  
Prévia: Decomposição em Valores Singulares 01:37:08

## CVGA 2018 4 Decomposição em Valores Singulares; Determinante; Programação Linear

Definição algébrica do Produto Escalar 00:00:58  
Produto escalar em  $\mathbb{R}$ ? 00:03:32  
O Teorema de Pitágoras 00:09:41  
A Lei dos Cossenos 00:14:41  
Definição de ângulo a partir do produto escalar 00:16:31  
Outros produtos escalares 00:19:52  
A "Desigualdade do Cosseno"(Cauchy-Scwarz-Buniacóvsqui) 00:22:30  
Otimização - Programação Linear 00:31:46  
A Decomposição em Valores Singulares (SVD) 00:42:44  
Demonstração da existência da SVD 00:48:13  
Sistemas de Coordenadas e o Produto Escalar 00:59:31  
O Determinante 01:07:24  
Determinante como uma "Área com Sinal"01:14:50  
Propriedades da área com sinal 01:16:58  
Fórmula do determinante 01:25:00  
Determinante de matriz 01:27:00  
Determinante de Transformação Linear 01:27:40  
o Determinante do produto é o produto dos determinantes 01:34:20  
Nova explicação do determinante 01:40:27

## CVGA 2018 5 Números Complexos, Equações Polinomiais, Inversões

Origem dos Números Complexos 00:00:59  
A Equação do Terceiro Grau 00:10:00  
Condição para Existência de Raízes Reais 0:14:38  
Uma Substituição Não-Trivial para Resolver a Cúbica (Vieta) 00:16:43  
Um Pouco Mais da História dos Complexos 00:23:13  
A Representação Gráfica dos Complexos e a Sistematização Hamiltoniana 00:25:52  
Multiplicação de Complexos 00:29:44  
Potenciação e Radiciação - interpretação geométrica 00:39:18  
Todo polinômio de grau ímpar a coeficientes reais tem raiz real 00:49:25  
O Teorema Fundamental da Álgebra 00:50:40  
Comentário sobre a possibilidade de demonstrar a existência de raízes sem fornecer fórmulas para calculá-las 00:54:19  
Comportamento de um polinômio  $p(z)$  quando  $|z|$  é grande 00:55:20  
Inversões em Relação a Um Círculo e o Análogo Complexo 01:00:43  
Propriedades Geométricas da Inversão 01:07:47  
Prova (esboço): Inversão de uma Reta 01:13:03  
Prova (esboço): Inversão de Um Círculo 01:23:32

Prova (geométrica): Conservação de Ângulos 01:27:51  
Problema dos Círculos Inscritos na Coroa Circular 01:30:53  
Prévia da Próxima Aula: TFA e Teorema de Brouwer 01:35:21

## CVGA 2018 6 Teorema Fundamental da Álgebra, Teorema de Brouwer, Número de Voltas

O Teorema Fundamental da Álgebra: enunciado 00:00:50  
O Teorema de Brouwer: enunciado 00:03:10  
TFA: Esquema de Demonstração 00:07:55  
O Número de Voltas de uma Curva 00:17:17  
Distância entre duas curvas 00:24:08  
Princípio Fundamental do Número de Voltas 00:31:16  
O "Lema de Brouwer" 00:36:23  
O Lema de Rouché 00:42:20  
Demonstração do Teorema Fundamental da Álgebra 00:50:15  
Demonstração do Teorema de Brouwer 01:06:45  
Exemplo de Curva Contínua que preenche totalmente uma região do plano 01:18:44

## CVGA 2018 7 Método de Newton; a Ferradura de Smale

Resolução de Equações Polinomiais: De Ferrari a Galois 00:00:27  
O Método de Newton para aproximação de raízes 00:13:09  
Sobre Métodos Iterativos e Órbitas 00:21:18  
A Ferradura de Smale 00:24:58  
Conjunto de Cantor e variantes 00:33:11  
Ferradura de Smale e sua relação com o Conjunto de Cantor 00:57:55  
O Shift e o Itinerário da Aplicação 01:22:16  
Pontos Periódicos na Ferradura de Smale 01:28:00

## CVGA 2018 8 O Método de Newton por Bernardo Freitas Paulo da Costa

Apresentação: o Método de Newton 00:00:08  
O problema: calculando raízes de polinômios 00:02:00  
Implementando o código para clarificar a questão 00:12:50  
Resolução de Equações polinomiais: Fórmula Fechada vs Método Iterativo 00:23:30  
Um pouco mais de programação 00:35:00  
Uma fórmula com o chute da raiz 00:43:00  
Mudando de figura: Quem ganha, os pontos onde dá certo ou onde dá errado?  
00:52:00  
Mudando (mais) de figura: como a variação dos coeficientes afeta o mapa 01:09:55  
Variando o Grau do Polinômio e o efeito na figura 01:14:18  
Perguntas e Respostas 01:27:00

# Capítulo 1

## Sistemas de coordenadas

▶ Sistemas de Coordenadas 1

▶ Sistemas de Coordenadas 2

Do ponto de vista prático, a Geometria Analítica começa pela introdução de um **sistema de coordenadas**: traça-se no plano um par de retas concorrentes, toma-se como **origem** do sistema o ponto **O** de interseção das duas retas (que passaremos a chamar **eixos de coordenadas** e a notar por **Ox** e **Oy**) e marcam-se dois pontos, um sobre cada eixo, que indicarão a unidade de medida e o sentido positivo em cada eixo. Note que não é obrigatória, embora seja usual, a ortogonalidade entre os eixos; tampouco somos forçados ao uso da mesma unidade de medida em  $Ox$  e em  $Oy$ .

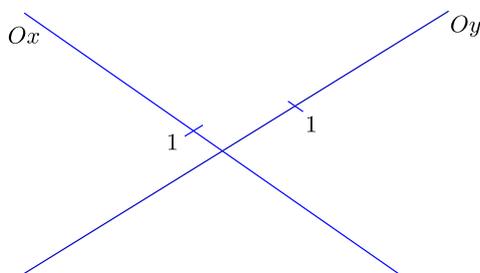


Figura 1.1:

Consideremos agora nosso plano com um sistema de coordenadas e seja **P** um ponto do plano. Vamos definir as **coordenadas** de **P**. Traçando por **P** uma reta **r** paralela a  $Oy$ , tomamos a interseção de **r** com  $Ox$ ; ao ponto assim obtido corresponde um número real **x** (em função de termos definido em  $Ox$  um sentido positivo e uma unidade). Da mesma forma, traçando por **P** uma reta **s** paralela a  $Ox$  e tomando sua interseção com  $Oy$ , temos um ponto sobre  $Oy$  ao qual, pelas mesmas razões, corresponde um número **y**. Os números **x** e **y**, tomados nesta ordem, são ditos as **coordenadas** de **P** (no sistema dado). Dizemos que **P** é representado pelo par ordenado **(x,y)**. Na presente figura, **x** é um pouquinho maior do que 2, enquanto **y** está entre 0,5 e 1.

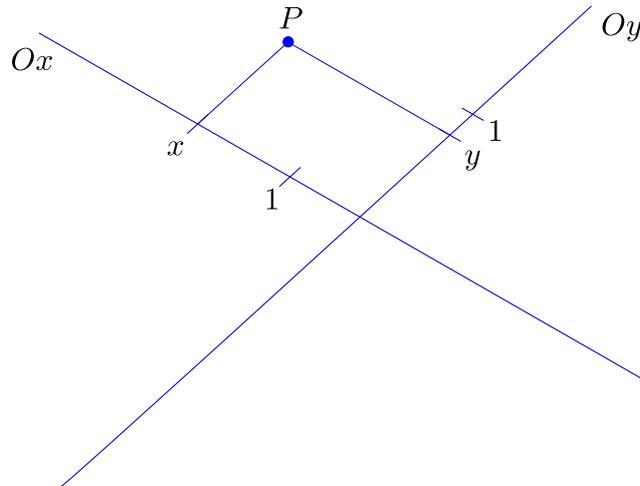


Figura 1.2:

**Exercício 1.1** *Certifique-se de que você é capaz de inverter o processo: fixado o sistema de coordenadas e dado um par ordenado  $(x, y)$ , é sempre possível determinar o ponto  $P$  do plano correspondente a  $(x, y)$ . Descreva o modus operandi a ser adotado.*

Salvo menção em contrário, usaremos o sistema de eixos **canônico**: eixos ortogonais,  $Ox$  *horizontal* com sentido positivo *da esquerda para a direita*,  $Oy$  *vertical* com sentido positivo *de baixo para cima*, e a mesma unidade de medida em ambos os eixos (as expressões em itálico têm aqui o significado que lhes atribuí o senso comum e não serão definidas).

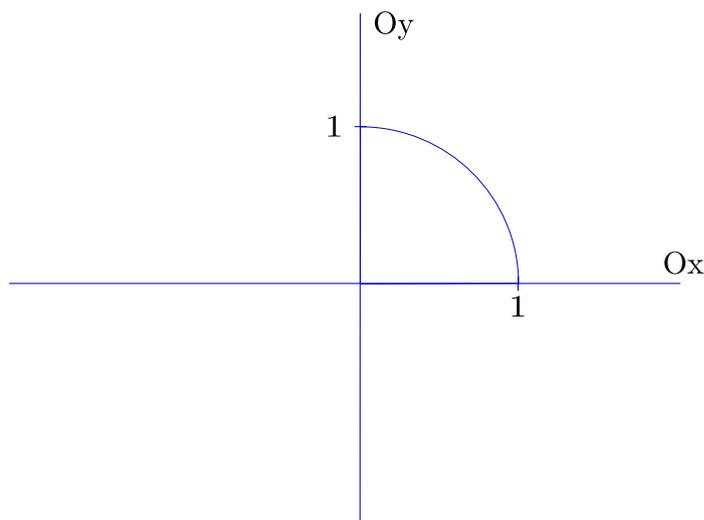


Figura 1.3:

Os procedimentos acima descritos estabelecem uma bijeção entre o plano e o conjunto  $\mathbb{R}^2$  dos pares ordenados. Desta forma, paralelamente ao **plano geométrico**, passamos a ter um **plano virtual**, formado pelos pares ordenados. O plano virtual, pois, nada mais é que

$$\mathbb{R}^2 = \{(x, y) \mid x \in \mathbb{R}, y \in \mathbb{R}\}.$$

Nossa primeira preocupação, como é natural, será obter, para esse plano virtual, *traduções algébricas* dos objetos e procedimentos geométricos a que estamos habituados.

**Exercício 1.2** *Examine o seguinte exemplo de tradução algébrica: se o ponto  $P_0$  é dado, em um sistema canônico, por  $(x_0, y_0)$  e  $r$  é a reta vertical passando por  $P$ , então, no plano virtual,  $r$  é dada pelo conjunto*

$$\{(x, y) \in \mathbb{R}^2 \mid x = x_0\}.$$

Como vamos centrar nossa atenção no plano virtual, escreveremos, usualmente,

$$r = \{(x, y) \in \mathbb{R}^2 \mid x = x_0\},$$

e diremos que  $x = x_0$  é a equação da reta  $r$  (mais corretamente, devemos dizer que  $x = x_0$  é, apenas, uma equação que descreve, ou caracteriza,  $r$ ).

**Exercício 1.3** *Mostre que as equações  $x = x_0$  e  $(x - x_0)^4 = 0$  descrevem a mesma reta.*



# Capítulo 2

## Medindo distâncias

No capítulo anterior mencionamos a possibilidade de se associar a cada ponto de uma reta um número real (fixados uma origem, correspondente ao número 0, e um segundo ponto, correspondente ao número 1). Esta não é, na verdade, uma questão simples: o processo de medição de um segmento é, de fato, um dos mais famosos processos de construção dos números reais.

Analisemos o método que empregamos para medir segmentos, mesmo que sem a preocupação de um extremo rigor geométrico. Consideremos dois pontos **A** e **B** sobre uma reta  $r$  e um segmento,  $l$ , que nos sirva de unidade (pensemos  $l$  como uma espécie de régua, que pode ser transportada, rigidamente, para qualquer lugar do plano).

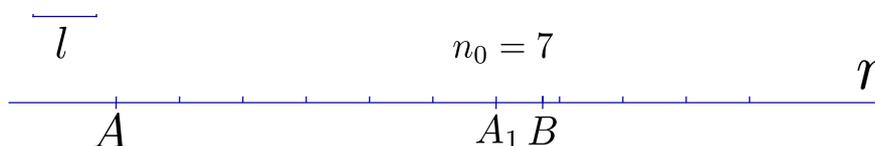


Figura 2.1: medindo o segmento  $AB$ , parte inteira

A partir do ponto  $A$ , e caminhando na direção de  $B$ , podemos alinhar um certo número máximo  $n_0$  de segmentos congruentes a  $l$  de forma a não ultrapassar  $B$ , obtendo o ponto  $A_1$ . Em seguida partimos  $l$  em 10 (obtendo um segmento  $l_1$ ) e repetimos o processo, trocando  $A$  por  $A_1$  e  $l$  por  $l_1$ . O número de segmentos congruentes a  $l_1$  utilizados será  $n_1$ , e obteremos o ponto  $A_2$ .

**Exercício 2.1** Note que  $0 \leq n_1 \leq 9$ .

Agora dividimos  $l_1$  em 10 (obtendo  $l_2 = l_1/10$ ) e trocamos  $A_1$  por  $A_2$  e  $l_1$  por  $l_2$ , repetindo novamente o processo, obtendo o número  $n_2$  e o ponto  $A_3$ , e assim sucessivamente.

**Exercício 2.2** Se  $l$  é a unidade de medida, convença-se de que o número real que expressa a distância entre  $A$  e  $B$  é  $n_0, n_1 n_2 \dots$

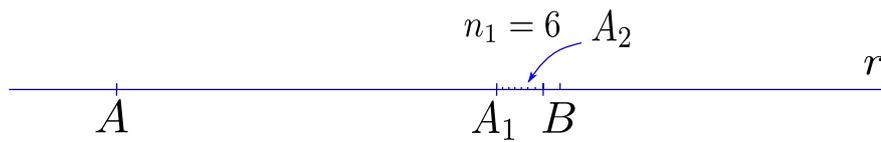


Figura 2.2: medindo o segmento  $AB$ , primeira casa decimal

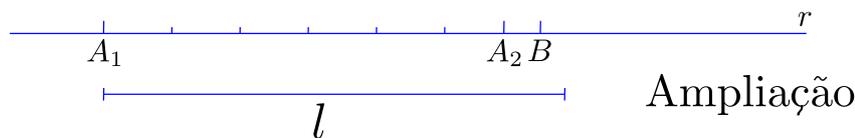


Figura 2.3:

O que acabamos de descrever é o processo de determinação da **distância** entre dois pontos no plano geométrico. Suponhamos agora que estamos lidando com o plano virtual. Isto é, nossos “pontos”  $A$  e  $B$  são agora dois pares ordenados,  $A = (x_1, y_1)$  e  $B = (x_2, y_2)$ . Podemos desenhar os pontos do plano geométrico correspondentes a  $(x_1, y_1)$  e  $(x_2, y_2)$  e, usando nossa régua  $l$ , repetir o procedimento anterior sobre a reta  $r$  determinada por  $A$  e  $B$ .

Alternativamente, podemos, aplicando o **Teorema de Pitágoras**, concluir, sem fazer qualquer medição, que a **distância** entre  $A$  e  $B$  é dada por

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Se a figura acima serve de justificativa para a fórmula que a precede, podemos por outro lado observar que o *plano virtual* pode perfeitamente dispensá-la, assim como qualquer outro recurso gráfico, quando se trata de calcular a distância entre  $(x_1, y_1)$  e  $(x_2, y_2)$ . Os procedimentos envolvidos são os da Álgebra - adições (subtrações) e multiplicações- e os da Análise - radiciações. Aos pares ordenados  $(x_1, y_1)$  e  $(x_2, y_2)$  que descrevem os **pontos virtuais** corresponde um número,  $\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ , sua distância, que se obtém diretamente, “sem desenhos”.

Podemos dizer que a fórmula

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

é a *tradução* para a Geometria Analítica da noção de distância da Geometria Sintética.

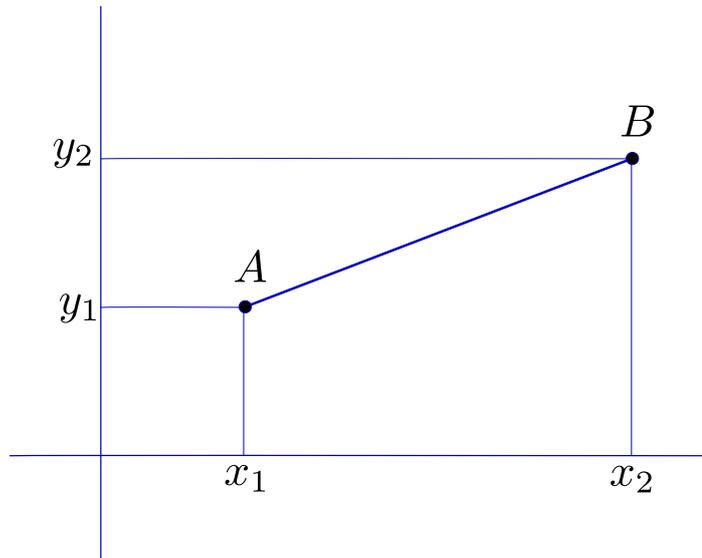


Figura 2.4: distância entre pontos do plano virtual

Enfatizemos, mais uma vez, a diferença. Dados dois pontos,  $A$  e  $B$ , do plano geométrico, sobre o qual suporemos fixado um sistema de coordenadas canônico, podemos, usando milênios de cultura matemática, efetuar os seguintes passos: obtemos, por meio de retas paralelas aos eixos (como visto no capítulo 1) e do processo de medição acima descrito, as coordenadas  $(x_1, y_1)$  de  $A$  e  $(x_2, y_2)$  de  $B$ ; em seguida, usando o Teorema de Pitágoras (fruto de elaborada construção teórica), concluímos que a distância entre  $A$  e  $B$  é dada por

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

Já no plano virtual, o ponto  $A$  é o par ordenado  $(x_1, y_1)$  e o ponto  $B$  é o par ordenado  $(x_2, y_2)$ . Aqui, respaldados por milênios de cultura matemática, simplesmente **definimos a distância** entre os pontos  $(x_1, y_1)$  e  $(x_2, y_2)$  por

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

**Exercício 2.3** *Refleta sobre as ideias acima.*



## Capítulo 3

# Os objetos geométricos: retas e círculos

Continuemos trabalhando sobre a dualidade entre o *plano geométrico*, feito de pontos, e o *plano virtual* (isto é, o  $\mathbb{R}^2$ ), feito de pares ordenados de números.

Duas classes particulares de subconjuntos do plano podem, por sua importância, servir-nos de ponto de partida: **retas** e **círculos**. Começemos com os círculos e partamos da definição tradicional: dados um ponto  $C$  do plano e um número positivo  $r$ , o **círculo de centro  $C$  e raio  $r$**  é o conjunto dos pontos do plano que distam  $r$  de  $C$ .

Façamos a tradução para o plano virtual: o ponto  $C$  será dado por suas coordenadas  $(x_0, y_0)$ , os pontos do círculo serão designados por suas coordenadas  $(x, y)$ , e a distância será calculada pela fórmula vista no capítulo anterior. Teremos então que, dados um par ordenado  $(x_0, y_0)$  e um número positivo  $r$ , o **círculo de centro  $(x_0, y_0)$  e raio  $r$**  é o conjunto dos pares ordenados  $(x, y)$  tais que

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = r.$$

Observando que a fórmula acima é equivalente a  $(x - x_0)^2 + (y - y_0)^2 = r^2$ , podemos dizer que, no plano virtual, o círculo de centro  $(x_0, y_0)$  e raio  $r$  é o conjunto  $c$  dado por

$$c = \{ (x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 = r^2 \}.$$

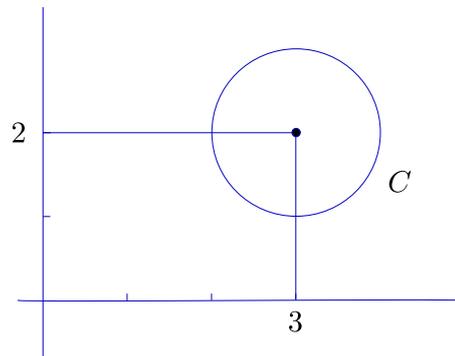


Figura 3.1: círculo "de verdade"

Podemos, é claro, construir um *círculo geométrico* a partir do *círculo virtual* definido acima. Mais concretamente, suponhamos que o centro é dado pelas coordenadas  $x_0 = 3$ ,  $y_0 = 2$  e que  $r=1$ ; suponhamos também dado, no plano geométrico, um sistema de coordenadas. Ao conjunto  $c$  do plano virtual dado por  $c = \{ (x, y) \in \mathbb{R}^2 \mid (x - 3)^2 + (y - 2)^2 = 1 \}$  corresponde o conjunto dos pontos do plano geométrico cujas coordenadas  $(x, y)$  satisfazem a equação  $(x - 3)^2 + (y - 2)^2 = 1$ .

**Observação:** Note que a equação nos fornece um critério, um teste, para decidirmos se um ponto do plano geométrico está ou não no círculo: para cada ponto  $P$  do plano geométrico devemos medir suas coordenadas  $x$  e  $y$  e substituí-las na equação;  $P$  está no círculo se e só se a igualdade é satisfeita. Imagine o que aconteceria se não tivéssemos qualquer experiência anterior com círculos geométricos, ou simplesmente não soubéssemos que àquela equação corresponde um círculo de raio 1 - provavelmente ficaríamos testando às cegas os pontos mais disparatados e levaríamos muito tempo antes de conseguirmos uma figura parecida com um círculo "de verdade".

Passemos agora às retas. Reta é usualmente considerado um conceito primitivo em Geometria Sintética; não podemos, ao contrário do que fizemos com os círculos, partir da definição. Tentemos outra estratégia: vamos desenhar uma reta em um plano dotado de um sistema de coordenadas e ver que relação conseguimos entre as coordenadas de seus pontos.

Fixemos como caso padrão o da reta  $r$  passando por dois pontos dados  $P_1$  e  $P_2$ , digamos  $P_1 = (1, 2)$ ,  $P_2 = (4, 1)$ . Se  $P = (x, y)$  é um ponto de  $r$  à esquerda de  $P_1$ , temos, por semelhança de triângulos,

$$\frac{y - 2}{1 - x} = \frac{2 - 1}{4 - 1} = \frac{1}{3}$$

ou, multiplicando em cruz,

$$3y - 6 = 1 - x.$$

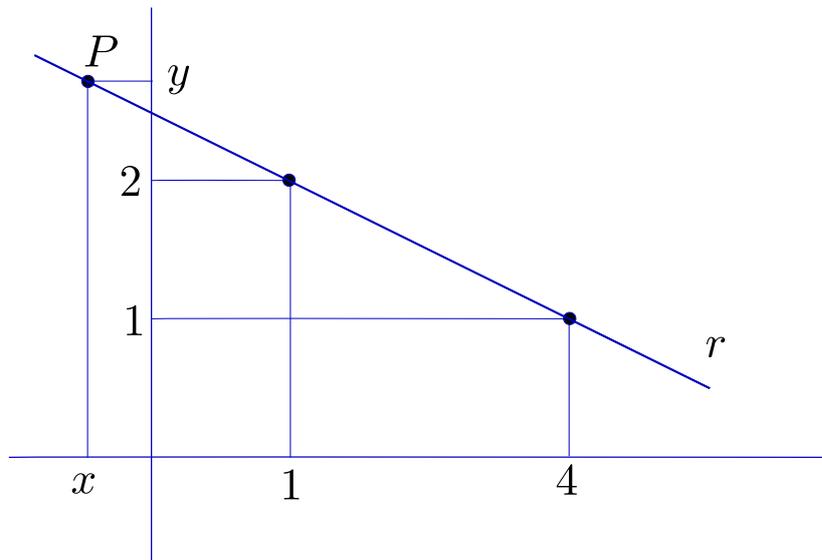


Figura 3.2:

**Exercício 3.1** Verifique que se  $P$  está à direita de  $P_2$  ou entre  $P_1$  e  $P_2$  a mesma relação é válida.

Os pontos (virtuais) de nossa reta (virtual) devem, portanto, satisfazer a equação  $x + 3y - 7 = 0$ . Isto quer dizer que  $r$  é o subconjunto do plano virtual dado por

$$r = \{ (x, y) \mid x + 3y - 7 = 0 \}.$$

Observe que podemos repetir o raciocínio para o caso geral em que  $P_1 = (x_1, y_1)$ ,  $P_2 = (x_2, y_2)$ ,  $x_1 \neq x_2$ . Obteremos então, se  $P = (x, y)$  é um ponto da reta (virtual):

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1},$$

o que nos fornece uma equação do tipo  $ax + by + c = 0$ .

**Exercício 3.2** Verifique isso (note que  $a = (y_2 - y_1)$ ,  $b = (x_1 - x_2)$ ,  $c = (x_2y_1 - y_2x_1)$ ).

Não custa nada observar que a semelhança de triângulos em que baseamos nossas deduções fica comprometida se a reta  $r$  é vertical ou horizontal.

Note que se o ponto  $(h, 0)$  é a interseção de uma reta vertical  $r$  com o eixo dos  $x$ , então um ponto  $(x, y)$  do plano está em  $r$  se, e só se,  $x = h$ , o que corresponde a uma equação do tipo  $ax + by + c = 0$ , com  $a = 1$ ,  $b = 0$ ,  $c = -h$ . Da mesma forma, uma reta horizontal passando por  $(0, k)$  terá equação da forma  $ax + by + c = 0$ , com  $a = 0$ ,  $b = 1$ ,  $c = -k$ .

Assim, toda reta do plano virtual é um conjunto  $r$  da forma

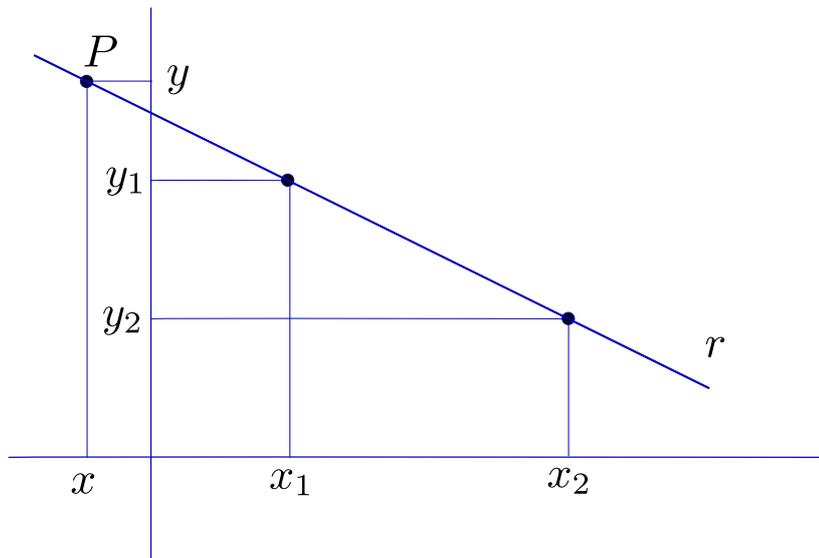


Figura 3.3:

$$r = \{ (x, y) \in \mathbb{R}^2 \mid ax + by + c = 0 \},$$

com  $a$ ,  $b$  e  $c$  fixos.

**Exercício 3.3** *E a recíproca? É verdade que a todo subconjunto  $r$  do plano virtual definido por uma equação (em  $x$  e  $y$ ) do tipo  $ax + by + c = 0$  corresponde uma reta no plano geométrico?*

Pelo que acabamos de ver, aos objetos geométricos **reta** e **círculo** correspondem objetos virtuais **reta** e **círculo**. As retas virtuais são subconjuntos de  $\mathbb{R}^2$  (que é o nome técnico do plano virtual) dados por equações do tipo  $ax + by + c = 0$ . Os círculos virtuais são subconjuntos de  $\mathbb{R}^2$  dados por equações do tipo  $(x - x_0)^2 + (y - y_0)^2 = r^2$ . Também vimos que não existe propriamente *a equação* do círculo: as equações

$$(x - x_0)^2 + (y - y_0)^2 = r^2$$

e

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = r$$

definem o mesmo círculo. Não é difícil ver que a mesma observação vale para retas: as equações  $x + 3y - 7 = 0$ ,  $-2x - 6y + 14 = 0$  e  $\pi x + 3\pi y - 7\pi = 0$  definem a mesma reta.

Poder-se-ia objetar que essas três últimas equações são todas do tipo  $ax + by + c = 0$ , o que é diferente do caso do círculo. Vejamos então o seguinte.

A equação  $x + 3y - 7 = 0$  corresponde à reta  $r$  passando por  $P_1 = (1, 2)$  e  $P_2 = (4, 1)$ , que vem também a ser a mediatriz do segmento de reta de extremidades  $A = (2, 0)$  e  $B = (3, 3)$ . Usando a definição de mediatriz temos que  $P = (x, y)$  está em  $r$  se e só se a distância de  $P$  a  $A$  é igual à de  $P$  a  $B$ , o que nos fornece a seguinte equação para  $r$ :

$$\sqrt{(x - 2)^2 + y^2} = \sqrt{(x - 3)^2 + (y - 3)^2}.$$

**Exercício 3.4** *Mostre que a equação acima é equivalente a  $x + 3y - 7 = 0$ .*

**Exercício 3.5** *Determine uma equação para a reta que passa pelo ponto  $(1, 2)$  e é normal à de equação  $x + 3y - 7 = 0$ .*



## Capítulo 4

# Curvas e equações. Lugares geométricos

A lição a extrair do capítulo anterior não se restringe a retas e círculos: outras curvas do plano geométrico devem poder ser “traduzidas” para o plano virtual por meio de uma equação. Tomemos um novo exemplo. Considere no plano geométrico um ponto  $F$  e uma reta  $d$ . Seja  $p$  o **lugar geométrico** dos pontos do plano que equidistam de  $F$  e de  $d$  (isto é, o ponto  $P$  está em  $p$  se e só se a distância de  $P$  a  $F$  é igual à de  $P$  à reta  $d$ )<sup>1</sup>.

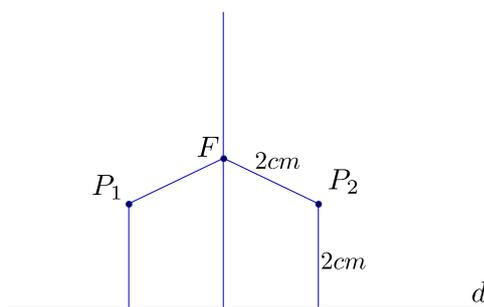


Figura 4.1:

Procuremos traduzir algebricamente a propriedade que define  $p$ . Suponhamos que a reta  $d$  coincide com o eixo horizontal e que o ponto  $F$  está sobre o eixo vertical, digamos  $F = (0, y_0)$ , com  $y_0 \neq 0$ . Se  $P = (x, y)$  é um ponto do plano, então sua distância a  $F$  é

$$\sqrt{x^2 + (y - y_0)^2}.$$

Sua distância à reta  $d$  será dada pelo valor absoluto de  $y$ . Assim,  $P$  pertence a  $p$  se e só se  $P$  satisfaz a equação

$$|y| = \sqrt{x^2 + (y - y_0)^2}.$$

---

<sup>1</sup> $p$  é uma curva bastante famosa, conhecida pelo nome de **parábola**

**Exercício 4.1** Mostre que a equação acima é equivalente a  $y = \frac{1}{2y_0}(x^2 + y_0^2)$ .

Há uma constatação impressionante a ser feita. A equação para  $p$  foi obtida diretamente de sua definição; não foi preciso desenhar  $p$  ou ter qualquer conhecimento geométrico anterior sobre parábolas para obtê-la. Poderíamos repetir esse processo para uma outra curva qualquer, definida arbitrariamente. Só precisamos de uma boa “tradução algébrica” para a definição.

A idéia é animadora, vamos experimentar um exemplo um pouco mais esquisito. Seja  $c$  a curva definida da seguinte forma: o ponto  $P$  está em  $c$  se e só se sua distância ao ponto  $P_0$  de coordenadas  $(4, 5)$  é igual a duas vezes o quadrado de sua distância ao eixo dos  $x$  mais três vezes a quarta potência de sua distância ao eixo dos  $y$ .

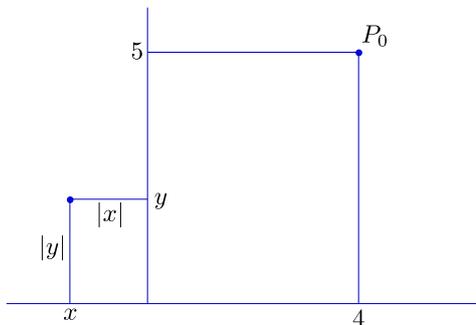


Figura 4.2:

Isto nos dá, se  $P = (x, y)$ ,

$$\sqrt{(x-4)^2 + (y-5)^2} = 2|y|^2 + 3|x|^4.$$

**Exercício 4.2** Mostre que a equação acima é equivalente a

$$9x^8 + 12x^4y^2 + 4y^4 - x^2 - y^2 + 8x + 10y - 41 = 0.$$

Deu certo! Mesmo sem termos a menor idéia de como seja a curva  $c$  (isto é, sem termos jamais visto seu desenho) somos capazes de obter uma equação para sua correspondente no plano virtual. Assim, a introdução de sistemas de coordenadas nos dá acesso a territórios geométricos jamais visitados pelos companheiros de Euclides. Podemos ir mais além, invertendo o processo. Se até agora nos limitamos a traduzir algebricamente objetos previamente definidos no plano geométrico, por que não fazer o contrário? Por que não partir da equação?

Mais concretamente, considere a equação

$$y^4 + x^2 = 1.$$

Seja  $c = \{ (x, y) \in \mathbb{R}^2 \mid y^4 + x^2 = 1 \}$ . Ora,  $c$  é um subconjunto do plano virtual, ao qual corresponde, uma vez fixado um sistema de coordenadas, um subconjunto do plano geométrico. Este conjunto é precisamente o **lugar geométrico** dos pontos do plano tais que a soma da quarta potência de suas coordenadas  $y$  com o quadrado de suas coordenadas  $x$  é igual a 1.

O exemplo acima nos mostra como proceder para criar curvas no plano geométrico a partir de equações. Como já vimos que uma mesma curva tem diversas equações (infinitas, na verdade), como saber se duas equações definem a mesma curva? O que temos não é um critério prático, mas uma simples observação: duas equações definem a mesma curva se e só se o conjunto dos pares ordenados que satisfazem a primeira é igual ao conjunto dos que satisfazem a segunda; ora, esta é precisamente a definição de equivalência algébrica entre equações. Assim, duas equações definem a mesma curva se e somente se são algebricamente equivalentes.

**Exercício 4.3** Como é a curva definida pela equação  $\log(x + 3y - 6) = 0$ ?

Passemos agora a uma questão mais delicada. Andamos insinuando, de forma algo leviana, que toda equação em  $x$  e  $y$  define uma curva no plano (por meio da correspondência entre o plano virtual e o plano geométrico introduzida no capítulo 1). Uma análise rigorosa da questão pode ser (e é) feita no âmbito do Cálculo Diferencial. Podemos porém dar uma indicação das razões por que, em geral (mas nem sempre - pense em  $x^2 + y^2 = 0$ ), isto acontece.

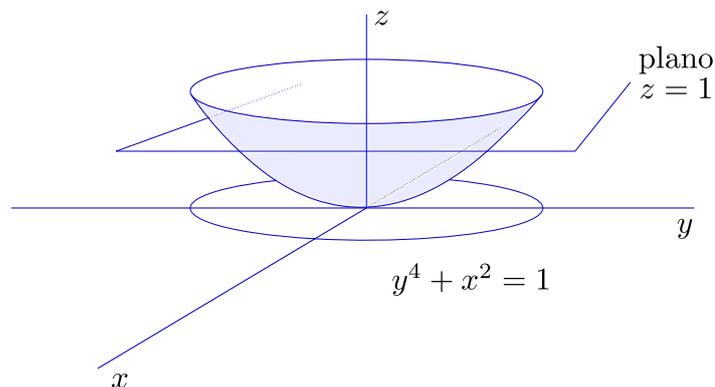


Figura 4.3:

Tomemos como exemplo nossa curva  $c$  dada pela equação  $y^4 + x^2 = 1$ . Para estudar a questão pensemos que nossos pontos vivem em um plano horizontal situado em um espaço tridimensional. Fixado um sistema de coordenadas no plano, consideremos um terceiro eixo, vertical (das coordenadas  $z$ ), passando pela origem. No lugar de nos restringirmos à equação  $y^4 + x^2 = 1$ , consideremos a função

$$z = y^4 + x^2.$$

Isto significa que estamos “levantando” os pontos de coordenadas  $(x, y)$  e marcando, no espaço, os pontos  $(x, y, z)$ , para os quais a “altura”  $z$  é dada por  $z = y^4 + x^2$ . Obtemos assim uma superfície. Os pontos da curva  $y^4 + x^2 = 1$  são aqueles para os quais a altura  $z$  é 1 e correspondem, na superfície, à interseção com o plano horizontal de altura 1 (c é dita uma **curva de nível**).

**Exercício 4.4** Note que, mesmo partindo de equações polinomiais simples, a “curva” nem sempre é, de fato, uma curva (veja  $y^4 + x^2 = 0$ ), ou corresponde a um objeto geométrico palpável (veja  $y^4 + x^2 = -1$ ). Podemos, também, obter “curvas bicudas” ( $y^2 - x^3 = 0$ ), ou com autointerseções ( $y^2 - (x^2 - 1)^2 = 0$ ).

# Capítulo 5

## Interseções. Sistemas de equações

### Álgebra e Geometria

Um procedimento fundamental em Geometria é a tomada de interseções entre curvas. Nas construções geométricas clássicas, as únicas curvas obtidas diretamente são as retas e os círculos; as demais são desenhadas ligando pontos obtidos um a um. No caso de curvas dadas por equações a coisa é um pouco pior: as equações nos fornecem critérios para decidir se os pontos estão ou não na curva, mas, para efeitos de desenho, estes devem ser marcados um a um, mesmo quando se trata de uma reta ou de um círculo.

**Exercício 5.1** *Dê uma pensada sobre o parágrafo acima.*

No caso das interseções entre duas curvas,  $c_1$  e  $c_2$ , por outro lado, a situação é um tanto diferente. Os métodos gráficos clássicos exigem que cheguemos à interseção por meio de retas e círculos. Mas se a curva  $c_1$  é dada por uma equação e a curva  $c_2$  é dada por outra equação, então os pontos de  $c_1$  são os que satisfazem a primeira equação e os de  $c_2$  são os que satisfazem a segunda. Assim, a interseção entre  $c_1$  e  $c_2$  é o conjunto dos pontos que satisfazem simultaneamente ambas as equações.

Vejamos o caso particular de duas retas, digamos  $r_1$ , dada pela equação  $x + 3y - 7 = 0$ , e  $r_2$ , dada por  $2x - y - 1 = 0$ . A interseção de  $r_1$  e  $r_2$  é o ponto de coordenadas  $(x, y)$  que satisfaz o sistema

$$\begin{cases} x + 3y - 7 = 0 \\ 2x - y - 1 = 0 \end{cases}$$

**Exercício 5.2** *Resolva o sistema acima.*

Da mesma forma, se quisermos a interseção da reta de equação  $x + 3y - 7 = 0$  com o círculo de equação  $(x - 2)^2 + (y - 2)^2 = 1$ , devemos resolver o sistema

$$\begin{cases} x + 3y - 7 = 0 \\ (x - 2)^2 + (y - 2)^2 = 1 \end{cases}$$

**Exercício 5.3** *Resolva o sistema acima.*

Consideremos, agora, a interseção de dois círculos, digamos  $c_1$ , de equação  $(x - 2)^2 + (y - 2)^2 = 1$ , e  $c_2$ , de equação  $(x - 3)^2 + y^2 = 5$ . Agora já temos um sisteminha um pouco mais emocionante:

$$\begin{cases} (x - 3)^2 + y^2 = 5 \\ (x - 2)^2 + (y - 2)^2 = 1 \end{cases}$$

**Exercício 5.4** *Resolva o sistema acima. Sugestão: desenvolva as duas equações e em seguida subtraia uma da outra, obtendo uma terceira equação sem termos do segundo grau (a que corresponde essa nova equação?); obtenha nessa última o valor de  $y$  em função de  $x$  e substitua na primeira.*

Bom, já deve estar claro que ao procedimento (gráfico) de achar a interseção de duas curvas no plano geométrico corresponde, no plano virtual, o procedimento (algébrico) de calcular as soluções de um sistema de duas equações a duas incógnitas. O mínimo que se pode dizer é que não é evidente que o segundo seja mais fácil do que o primeiro, ou de que possa ajudar a simplificar a vida.

## Capítulo 6

# Geometria Analítica, Desenho Geométrico e Computação Gráfica

O nascimento da **Geometria Analítica** costuma ser datado de 1637, ano da publicação do livro *A Geometria*, de **René Descartes**. Descartes não se preocupa em explicitar sistemas de coordenadas, nem descreve seus pontos por pares ordenados, como fazemos hoje. Principalmente não se partia, nos primórdios da Geometria Analítica, de um sistema de eixos preexistente - as coordenadas eram apenas grandezas  $x$  e  $y$  a serem relacionadas e eram introduzidas a partir da figura que se considerava (os eixos, em geral, sequer eram desenhados). O que caracteriza o trabalho de Descartes e praticamente tudo que se fez a partir daí, é a intensiva utilização de equações para a descrição das curvas e tratamento das questões geométricas. Com a “algebrização” proposta por Descartes e o desenvolvimento do **Cálculo Infinitesimal** por **Isaac Newton** e seus contemporâneos, poucas décadas depois, a Geometria pôde enfim tomar novos rumos, abordar novas questões (ou dar novas formulações a velhas questões) e ir além dos conhecimentos herdados da Antiguidade Clássica.

Não é nosso propósito tratar aqui os avanços da Geometria nos últimos três séculos e meio. O que queremos ressaltar é que, a partir da introdução da Geometria Analítica, o desenvolvimento da Matemática deixa para trás a Geometria Sintética - régua e compasso são trocados por coordenadas e equações (sem falar em outros instrumentos poderosos, como derivada e integral). Para darmos uma idéia, ainda que pobre, da situação, é como se o “plano geométrico” ao qual nos temos referido nos capítulos anteriores fosse substituído pelo “plano virtual” (isto é, o espaço  $\mathbb{R}^2$ ). A idéia de *plano* é, hoje em dia, praticamente inseparável de  $\mathbb{R}^2$ , uma *curva* é quase que automaticamente associada a uma equação.

Ao longo desses anos, enquanto os matemáticos desbravavam novos espaços, novas Geometrias, a Geometria Euclidiana não foi, como se poderia imaginar, recolhida aos museus. Além de constituir base indispensável para a construção e compreensão das porções mais avançadas da Matemática, manteve-se entrincheirada em um ramo particular da atividade humana: a representação gráfica. Desenho de Arquitetura e de Engenharia, Desenho Industrial, Desenho Técnico em geral, permaneceram inextinguíveis à Geometria Analítica. Suas ferramentas continuaram sendo o velho e bom Desenho Geométrico e sua versão mais moderna, a Geometria Descritiva, que trabalha com coordenadas mas não com equações. A razão é simples: os métodos analíticos são

potentes para a compreensão de propriedades geométricas, as equações são eficazes na descrição das curvas e das superfícies, mas não trouxeram consigo instrumentos mais efetivos de desenho: os cálculos são demorados e a conversão das equações em curvas se faz ponto a ponto. Régua e compasso permaneceram insubstituíveis sempre que se tratou de dar visibilidade às idéias geométricas, e as figuras possíveis, em condições normais de tempo e esforço, estiveram sempre limitadas ao alcance destas ferramentas.

Os avanços da Matemática e da Física a partir do mesmo século XVII em que veio à luz a Geometria Analítica vão abrir caminho para o surgimento, no século XX, do elemento que faltava. Veloz nos cálculos e na conversão de coordenadas em pontos luminosos, o **computador** é, entre outras coisas, máquina de desenhar. Engenhoca essencialmente algébrica, desprovida de visão e de tato, é no entanto capaz de armazenar em sua memória uma “realidade virtual”, feita de coordenadas, na qual as formas de nosso mundo são subconjuntos do espaço  $R^3$ , curvas e superfícies *são* equações. **Sua Geometria é a Geometria Analítica.**

A utilização do computador como “máquina geométrica” exige o uso de Geometria Analítica tanto quando lhe comunicamos os dados e as instruções sobre os procedimentos que desejamos ver cumpridos, como quando queremos receber os resultados obtidos. A tela do monitor é constituída de pequenos pontos luminosos (chamados *pixels*), que podem ser localizados por meio de coordenadas.

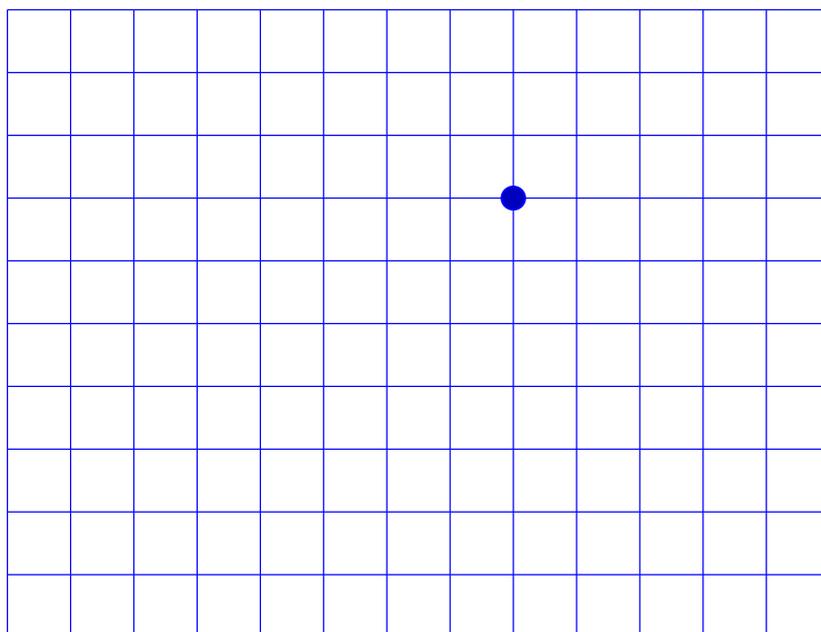


Figura 6.1:

Embora o número de pixels seja finito (milhões, nos monitores de alta definição atuais, com alguns milhares em cada linha horizontal ou vertical), é, ainda assim, suficiente para dar a sensação de continuidade. Para simplificar as coisas, vamos deixar de

considerar este aspecto da situação, trabalhando como se o número de pixels fosse infinito, em bijeção com os pontos da porção do plano representada na tela. O essencial é a compreensão de que, para que o computador marque um determinado ponto na tela, precisamos dizer-lhe onde este se localiza, o que é feito informando suas coordenadas. Mais, o computador não visualiza, como nós podemos fazer de olhos fechados, as imagens dos objetos com que trabalha: os pontos, em sua memória, *são* os pares ordenados.<sup>1</sup>

---

<sup>1</sup>e os objetos geométricos são arquivos em que estão armazenados os pontos que os constituem, estruturas de dados abstratas ou algoritmos que permitam gerar esses pontos



# Capítulo 7

## Vetores

▶ Vetores

▶ Coordenadas e Vetores

No capítulo anterior andamos usando e abusando de *flechinhas*. Estávamos tentando preparar o espírito do leitor para uma nova entidade: os **vetores**, que vão surgir dois séculos depois de Descartes e Fermat <sup>1</sup>.

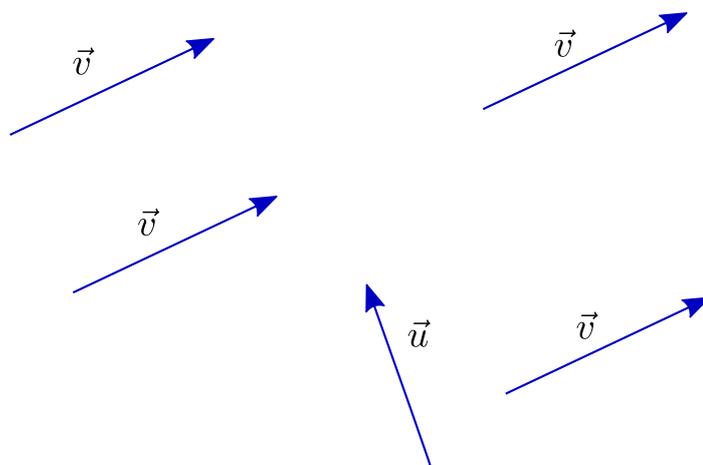


Figura 7.1:

### 7.1 Flechinhas

---

<sup>1</sup>Como geralmente ocorre no processo histórico, diversos matemáticos contribuíram para a construção do conceito de vetor, que vai, no século XX, tomar amplitude bem maior do que a que lhe daremos aqui. Nossas *flechinhas* podem ser vistas como um subproduto dos **quatérnions**, criados em 1843 pelo irlandês **William Rowan Hamilton**

De maneira informal, um **vetor** é uma flechinha que pode ser fixada em qualquer ponto do plano, por simples translação. O vetor correspondente à flecha que liga o ponto  $A$  ao ponto  $B$ , apontando de  $A$  para  $B$ , é usualmente designado por  $\overrightarrow{AB}$ . Na figura anterior, todas as flechinhas designadas por  $\vec{v}$  representam o mesmo vetor. Já a flechinha designada por  $\vec{u}$ , embora do mesmo tamanho (e, poderíamos quase dizer, igualzinha), representa outro vetor, distinto de  $\vec{v}$ .<sup>2</sup>

## 7.2 Norma

O comprimento de um vetor  $\vec{v}$  é chamado de **norma** de  $\vec{v}$  e é notado por  $|\vec{v}|$ . Um vetor de norma 1 é dito **unitário**.

## 7.3 Produto por escalar

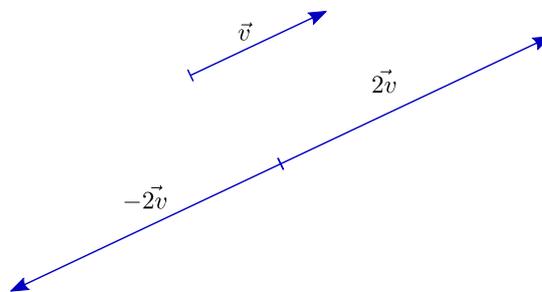


Figura 7.2:

Vetores podem ser multiplicados por números<sup>3</sup>: se  $t$  é um número real e  $\vec{v}$  é um vetor, então  $t\vec{v}$  é o vetor que se obtém esticando (ou comprimindo)  $\vec{v}$  de forma que seu comprimento fique multiplicado por  $t$  (se  $t$  é negativo, então, além disso, trocamos a orientação de  $\vec{v}$ , isto é,  $t\vec{v}$  aponta no sentido contrário ao de  $\vec{v}$ ).

**Exercício 7.1** Seja  $\vec{v}$  o vetor  $\overrightarrow{OP}$ , sendo  $O$  a origem do sistema de coordenadas e  $P$  o ponto de coordenadas  $(a, b)$ . Pense até chegar à conclusão de que  $t\vec{v}$  é o vetor  $\overrightarrow{OQ}$ , sendo  $Q$  o ponto de coordenadas  $(ta, tb)$ . Não prossiga enquanto não conseguir.

**Exercício 7.2** Conclua que se  $\vec{v} = \overrightarrow{OP}$ , então a reta passando pela origem e por  $P$  é o conjunto dos pontos  $Q$  tais que o vetor  $\vec{u} = \overrightarrow{OQ}$  é da forma  $t\vec{v}$  para algum número real  $t$ .

**Exercício 7.3** Seja  $\vec{v}$  um vetor. Mostre que  $|t\vec{v}| = |t||\vec{v}| \forall t \in \mathbb{R}$ .

<sup>2</sup>Para não dizerem que não definimos corretamente o conceito de vetor, aqui vai: considere o conjunto de pares ordenados  $(A, B)$  de pontos do plano (ou mesmo do espaço), o que equivale a considerar segmentos orientados  $AB$ ; diremos que  $(A, B)$  é equivalente a  $(C, D)$  se os segmentos  $AB$  e  $CD$  têm o mesmo comprimento, são paralelos e se também são paralelos e têm o mesmo comprimento os segmentos  $AC$  e  $BD$  (note que assim as flechas  $\overrightarrow{AB}$  e  $\overrightarrow{CD}$  representam de fato o mesmo vetor); um **vetor** é uma classe de equivalência definida por essa relação (isto é, um vetor é um conjunto de segmentos orientados equivalentes).

<sup>3</sup>Neste contexto, também chamados **escalares**

## 7.4 Soma de vetores

Vetores também podem ser somados. A soma  $\vec{u} + \vec{v}$  é ilustrada no paralelogramo abaixo (note que a congruência entre os triângulos garante a comutatividade da operação).

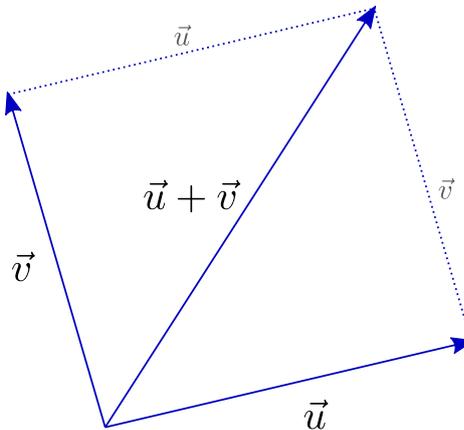


Figura 7.3:

**Exercício 7.4** Desenhe e pense até concluir que se  $O$  é a origem,  $P$  é o ponto de coordenadas  $(a, b)$ ,  $Q$  o de coordenadas  $(c, d)$ ,  $\vec{u} = \vec{OP}$  e  $\vec{v} = \vec{OQ}$ , então  $\vec{u} + \vec{v} = \vec{OR}$ , onde  $R$  é o ponto de coordenadas  $(a + c, b + d)$ . Não prossiga enquanto não conseguir.

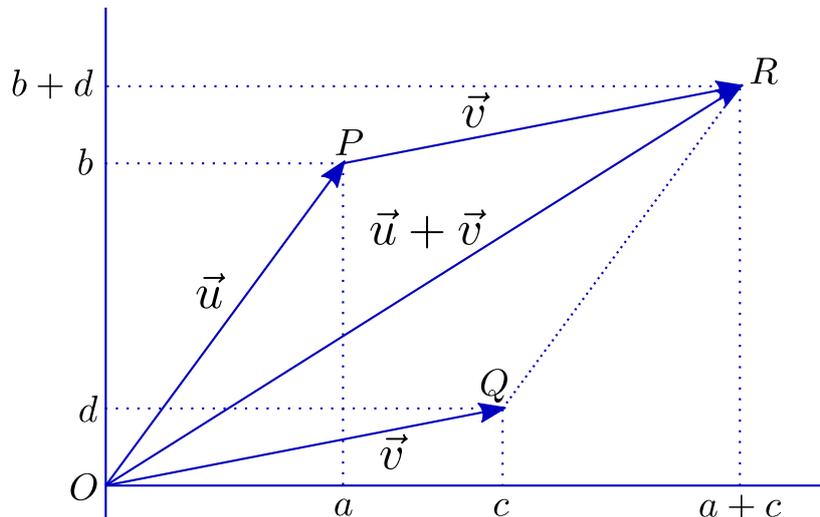


Figura 7.4:

Usando coordenadas, ou simplesmente olhando para a figura a seguir, conclua que a adição de vetores é uma operação associativa, isto é: para quaisquer vetores  $\vec{u}$ ,  $\vec{v}$  e  $\vec{w}$ , é verdade que  $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ .

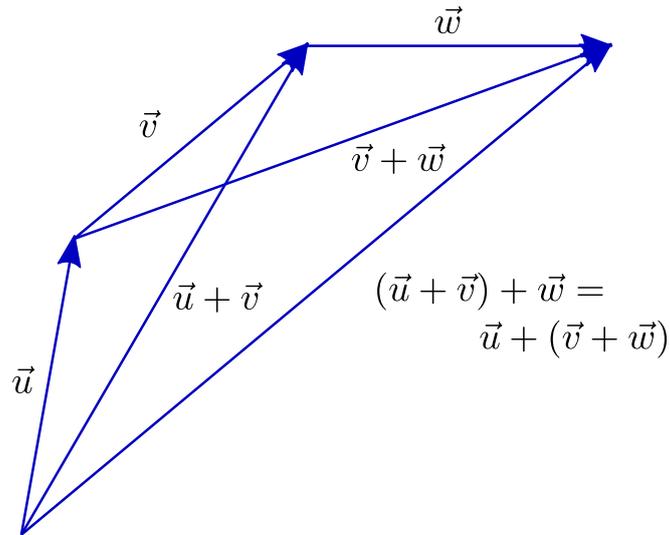


Figura 7.5:

Convença-se, também, de que valem as seguintes propriedades, quaisquer que sejam os escalares  $s$  e  $t$  e quaisquer que sejam os vetores  $\vec{u}$  e  $\vec{v}$ :

$$s(t\vec{u}) = (st)\vec{u};$$

$$t(\vec{u} + \vec{v}) = t\vec{u} + t\vec{v}$$

$$(s + t)\vec{u} = s\vec{u} + t\vec{u}.$$

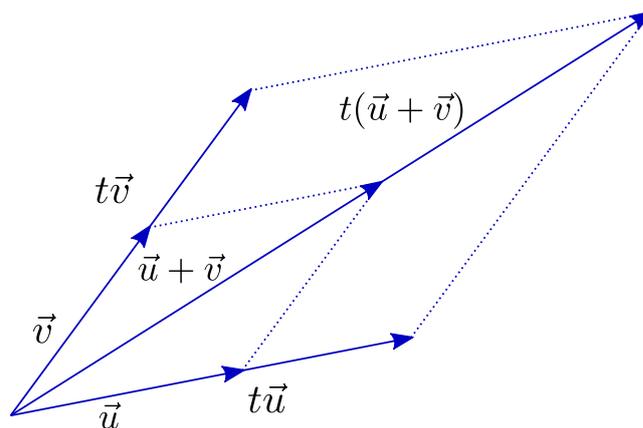


Figura 7.6:

**Exercício 7.5** Considere a reta  $r$  que passa pelo ponto  $P_0$  e é paralela ao vetor  $\vec{v}$ . Se  $A$  é um ponto qualquer do plano, seja  $\vec{u}_0 = \vec{AP}_0$ . Entenda que o ponto  $P$  do plano está em  $r$  se e somente se o vetor  $\vec{u} = \vec{AP}$  é da forma  $\vec{u}_0 + t\vec{v}$  para algum escalar  $t$ .

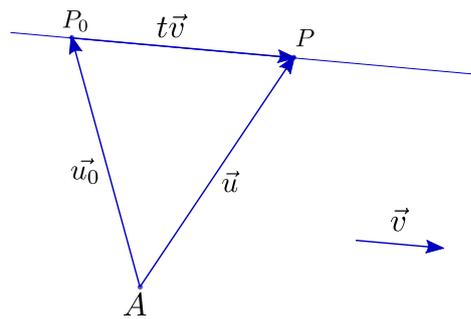


Figura 7.7:

**Exercício 7.6** Sejam  $\vec{u}$  e  $\vec{v}$  dois vetores. Mostre que  $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$ .

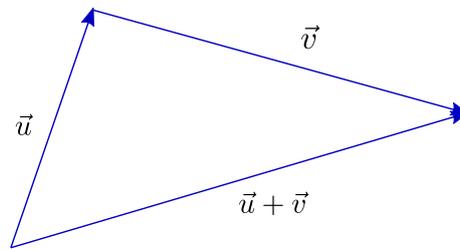


Figura 7.8:

## 7.5 Somando vetores a pontos

Podemos, ainda, definir uma operação "bastarda", somando o vetor  $\vec{v}$  ao ponto  $P$ . Neste caso,  $P + \vec{v}$  é um novo ponto,  $Q$ , definido por:  $P + \vec{v} = Q$  se  $\vec{PQ} = \vec{v}$ . Às vezes dizemos que o ponto  $Q$  é obtido *aplicando* o vetor  $\vec{v}$  ao ponto  $P$ .

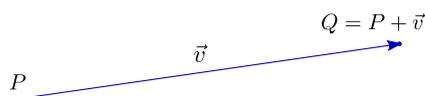


Figura 7.9:

Observe que essa operação também é associativa: para qualquer ponto  $P$  e quaisquer vetores  $\vec{u}$  e  $\vec{v}$ , vale  $(P + \vec{u}) + \vec{v} = P + (\vec{u} + \vec{v})$ .

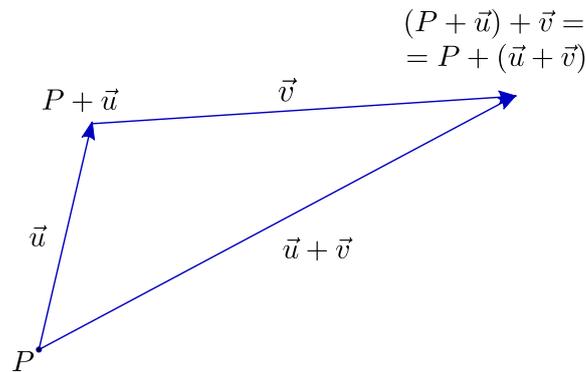


Figura 7.10:

## 7.6 Vetores e parametrizações

Consideremos o seguinte desafio: dados dois pontos  $P_1$  e  $P_2$ , parametrizar o segmento  $P_1P_2$  de forma que partamos de  $P_1$  no tempo  $t = 0$  e cheguemos a  $P_2$  no tempo  $t = 1$ .

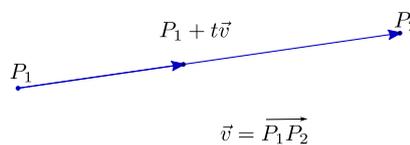


Figura 7.11:

Vamos trabalhar de forma *intrínseca*, isto é, sem fazer uso de coordenadas. Seja  $\vec{v}$  o vetor  $P_1P_2$ , de forma que podemos escrever  $P_1 + \vec{v} = P_2$ . Se considerarmos os vetores  $t\vec{v}$ , com  $0 \leq t \leq 1$ , teremos os pontos do segmento  $P_1P_2$  dados por  $P_1 + t\vec{v}$ .

É claro então que o ponto  $P_1 + t\vec{v}$  percorre o segmento  $P_1P_2$  quando  $t$  varia de 0 a 1, começando em  $P_1$  e terminando em  $P_2$ .

**Exercício 7.7** E se quisermos partir de  $P_1$  no tempo  $t = t_1$  e chegar a  $P_2$  no tempo  $t = t_2$ ?

# Capítulo 8

## O mistério da Santíssima Trindade

### O Mistério da Santíssima Trindade

#### 8.1 Vetores e coordenadas

Vejam agora uma versão vetorial do que fizemos no capítulo 1, com a introdução de sistemas de coordenadas. Começemos com as coordenadas canônicas.

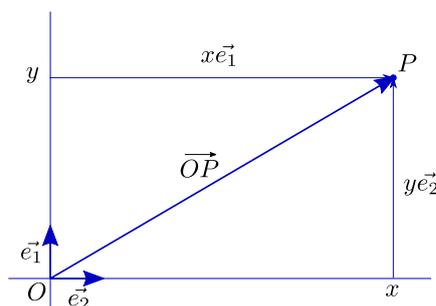


Figura 8.1:

Sejam  $\vec{e}_1$  o vetor unitário horizontal e  $\vec{e}_2$  o unitário vertical <sup>1</sup>. Se  $P = (x, y)$  é um ponto do plano, o correspondente vetor posição  $\vec{OP}$  pode ser expresso por

$$\vec{OP} = x\vec{e}_1 + y\vec{e}_2.$$

(dizemos que  $\vec{OP}$  está expresso como **combinação linear** de  $\vec{e}_1$  e  $\vec{e}_2$ ). Assim, expressar o ponto  $P$  por meio de suas coordenadas é essencialmente a mesma coisa que escrever seu vetor posição  $\vec{OP}$  como combinação linear de  $\vec{e}_1$  e  $\vec{e}_2$ .

Vamos, mais à frente, examinar com um pouco mais de cuidado esse processo em que um mesmo par ordenado pode fornecer coordenadas de um ponto ou de um vetor, conforme o caso. Mas vejamos, ainda, o caso de um sistema de coordenadas qualquer.

<sup>1</sup>Por razões históricas, também conhecidos como  $\mathbf{i}$  e  $\mathbf{j}$ . Nos quatérnions de Hamilton,  $\mathbf{i}$  é o mesmo

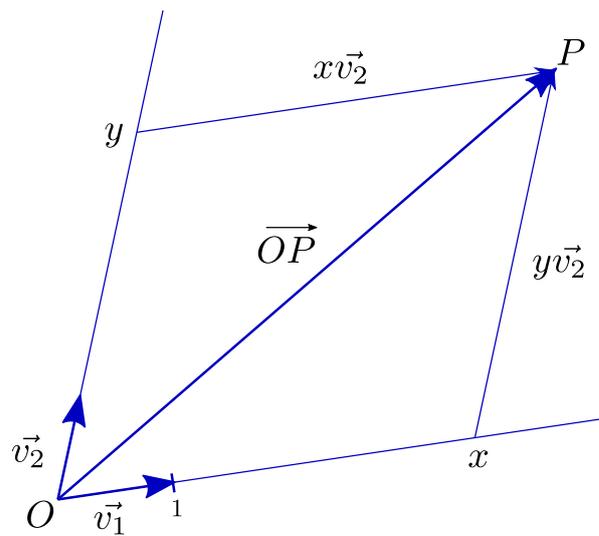


Figura 8.2:

Tomemos os vetores  $\vec{v}_1$  e  $\vec{v}_2$  definidos da seguinte forma:  $\vec{v}_1$  é o vetor posição do ponto que marca a unidade no primeiro eixo,  $\vec{v}_2$  o correspondente no segundo eixo. Então, se o ponto  $P$  tem, nesse sistema,  $(x, y)$  por coordenadas, o vetor  $\vec{OP}$  se expressa como combinação linear de  $\vec{v}_1$  e  $\vec{v}_2$  da seguinte forma:

$$\vec{OP} = x\vec{v}_1 + y\vec{v}_2.$$

Por extensão, os números  $x$  e  $y$  são chamados **coordenadas** do vetor  $\vec{OP}$  na **base**  $\{\vec{v}_1, \vec{v}_2\}$ <sup>2</sup>.

**Exercício 8.1** Considere fixado um sistema canônico de coordenadas, por meio do qual vamos expressar os dados do problema. Sejam  $\vec{v}_1 = (2, -1)$ ,  $\vec{v}_2 = (1, 3)$  e  $\vec{v} = (1, 1)$ . Escreva  $\vec{v}$  como combinação linear de  $\vec{v}_1$  e  $\vec{v}_2$ . A resposta depende mesmo de ser canônico o sistema de coordenadas?

**Exercício 8.2** Sejam  $\vec{v}_1$  e  $\vec{v}_2$  como acima. Calcule as coordenadas  $(a_{11}, a_{21})$  de  $\vec{e}_1$  e  $(a_{12}, a_{22})$  de  $\vec{e}_2$  na base  $\{\vec{v}_1, \vec{v}_2\}$ . Mostre que se o vetor  $\vec{v}$  tem coordenadas  $(x_1, x_2)$  na base  $\{\vec{e}_1, \vec{e}_2\}$  e  $(y_1, y_2)$  na base  $\{\vec{v}_1, \vec{v}_2\}$ , então:

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 \\ y_2 = a_{21}x_1 + a_{22}x_2 \end{cases} ,$$

ou, na forma matricial,

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} .$$

dos complexos;  $j$  e  $k$  foram concebidos como novos números, com o propósito de ir além do conjunto dos complexos

<sup>2</sup>Uma base (no plano) é qualquer par de vetores  $\{\vec{u}_1, \vec{u}_2\}$  tal que todo vetor  $\vec{u}$  (do plano) se expressa de forma única como combinação linear de  $\vec{u}_1$  e  $\vec{u}_2$ . A base composta por  $\vec{e}_1$  e  $\vec{e}_2$  é dita a **base canônica** do plano

## 8.2 Pontos e vetores

Para melhor compreensão desta seção e da próxima, é conveniente que o leitor apague de sua memória tudo que sabe de sistemas de coordenadas e volte a pensar no plano "puro", isto é, sem eixos ou coordenadas. Voltemos, pois, à Geometria Sintética, mas sem eliminar o conceito de vetor.

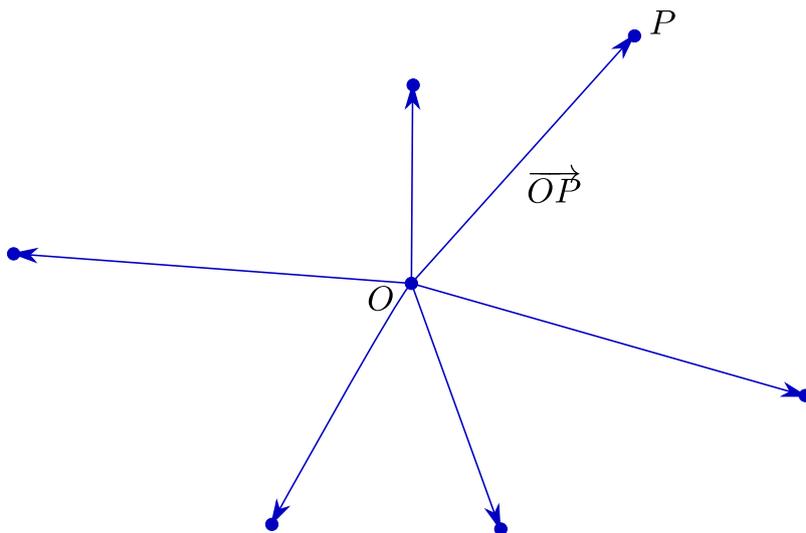


Figura 8.3:

Sabemos que a cada par ordenado de pontos,  $(A, B)$ , podemos associar um único vetor, usualmente denotado por  $\vec{AB}$ . No entanto, a cada vetor  $\vec{u}$  corresponde uma infinidade de pares ordenados  $(A, B)$  de pontos, tais que  $\vec{u} = \vec{AB}$ . A coisa muda de figura, porém, se fixarmos um ponto de origem,  $O$ , do qual partirão as flechas que representam nossos vetores: a cada vetor  $\vec{u}$  corresponde um único ponto  $P$  tal que  $\vec{OP} = \vec{u}$ ; reciprocamente, a cada ponto  $P$  corresponde um único vetor  $\vec{u}$  tal que  $\vec{u} = \vec{OP}$ . Assim, estabelece-se uma bijeção entre os pontos e os vetores do plano. A chave que permite tal identificação é a fixação de uma origem  $O$ .

## 8.3 Vetores e pares ordenados

Fixemos no plano dois vetores,  $\vec{e}_1$  e  $\vec{e}_2$ , **linearmente independentes** (isto significa: não é possível escrever  $\vec{e}_1 = t\vec{e}_2$ , nem  $\vec{e}_2 = s\vec{e}_1$ , quaisquer que sejam os números  $s$  e  $t$ ).

**Exercício 8.3** Mostre que, se  $\vec{e}_1$  e  $\vec{e}_2$ , são linearmente independentes, então  $a\vec{e}_1 + b\vec{e}_2 = \vec{0} \Rightarrow a = b = 0$ . Mostre que vale a recíproca: se a única maneira de obter  $a\vec{e}_1 + b\vec{e}_2 = \vec{0}$  é fazendo  $a = b = 0$ , então  $\vec{e}_1$

e  $\vec{e}_2$  são linearmente independentes.

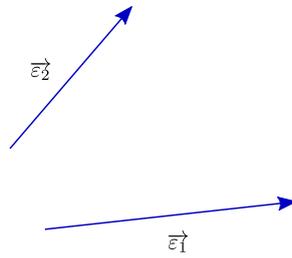


Figura 8.4:

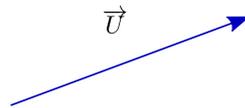


Figura 8.5:

Seja agora  $\vec{u}$  um vetor qualquer do plano. Podemos colocar as flechinhas que representam  $\vec{u}$ ,  $\vec{e}_1$  e  $\vec{e}_2$  partindo de um mesmo ponto. Traçando, pela ponta de  $\vec{u}$ , retas paralelas a  $\vec{e}_1$  e  $\vec{e}_2$ , respectivamente, obtemos vetores  $\vec{v}_2$ , múltiplo de  $\vec{e}_2$ , e  $\vec{v}_1$ , múltiplo de  $\vec{e}_1$ , tais que

$$\vec{u} = \vec{v}_1 + \vec{v}_2.$$

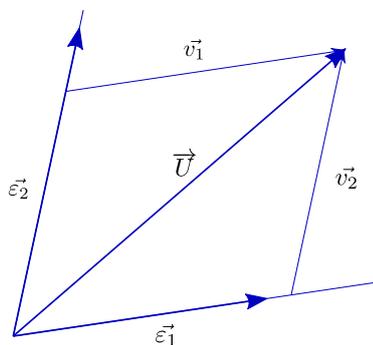


Figura 8.6:

Note que esse procedimento determina perfeitamente  $\vec{v}_1$  e  $\vec{v}_2$  (isto é: não existe outro par de vetores,  $\vec{w}_1$  e  $\vec{w}_2$ , respectivamente múltiplos de  $\vec{e}_1$  e  $\vec{e}_2$ , tais que  $\vec{u} = \vec{w}_1 + \vec{w}_2$ ).

Mais interessante ainda, como podemos, para certos reais  $x_1$  e  $x_2$ , escrever  $\vec{v}_1 = x_1\vec{e}_1$  e  $\vec{v}_2 = x_2\vec{e}_2$ , fica determinado um único par ordenado  $(x_1, x_2)$  de números reais tal que

$$\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2.$$

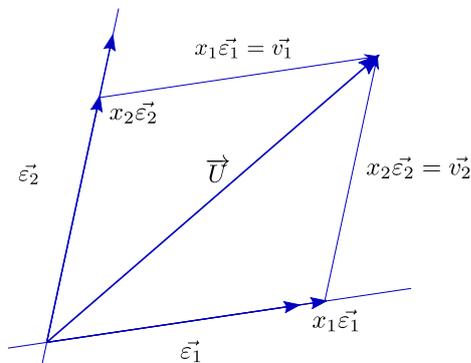


Figura 8.7:

Como, reciprocamente, dado um par ordenado  $(x_1, x_2)$  de números reais, podemos construir um único vetor  $\vec{u}$  tal que  $\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2$ , o que acabamos de estabelecer é uma bijeção entre o conjunto dos vetores do plano e o conjunto  $\mathbb{R}^2$  de pares ordenados de números reais. Mais ainda, usando as propriedades algébricas da adição e da multiplicação por escalar de vetores, temos: se  $\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2$  e  $t$  é um número real, então

$$t\vec{u} = t(x_1\vec{e}_1 + x_2\vec{e}_2) = t(x_1\vec{e}_1) + t(x_2\vec{e}_2) = (tx_1)\vec{e}_1 + (tx_2)\vec{e}_2;$$

se  $\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2$  e  $\vec{v} = y_1\vec{e}_1 + y_2\vec{e}_2$ , então

$$\vec{u} + \vec{v} = (x_1\vec{e}_1 + x_2\vec{e}_2) + (y_1\vec{e}_1 + y_2\vec{e}_2) = (x_1 + y_1)\vec{e}_1 + (x_2 + y_2)\vec{e}_2.$$

Isto significa que a bijeção que acabamos de construir preserva as operações: o par ordenado que corresponde ao vetor obtido pela multiplicação do vetor  $\vec{u}$  pelo escalar  $t$  é obtido multiplicando por  $t$  o par ordenado que corresponde a  $\vec{u}$ ; o par ordenado que corresponde ao vetor obtido pela soma de dois vetores é obtido somando os correspondentes pares ordenados. Esquemáticamente:

$$\begin{aligned} \vec{u} &\longleftrightarrow (x_1, x_2) \\ t\vec{u} &\longleftrightarrow (tx_1, tx_2) \end{aligned}$$

$$\begin{aligned} \vec{u} &\longleftrightarrow (x_1, x_2) \\ \vec{v} &\longleftrightarrow (y_1, y_2) \\ \vec{u} + \vec{v} &\longleftrightarrow (x_1 + y_1, x_2 + y_2) \end{aligned}$$

Podemos, assim, dizer que o par de vetores  $\vec{e}_1$  e  $\vec{e}_2$  é uma espécie de chave que nos permite codificar cada vetor do plano como um par ordenado de números reais, preservando as operações. O termo erudito é **base**.

**Definição:** Um par ordenado de vetores,  $(\vec{e}_1, \vec{e}_2)$ , é dito uma **base** para o conjunto de vetores do plano se, para todo vetor  $\vec{u}$  do plano, existe um único par ordenado  $(x_1, x_2)$  em  $\mathbb{R}^2$  tal que

$$\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2.$$

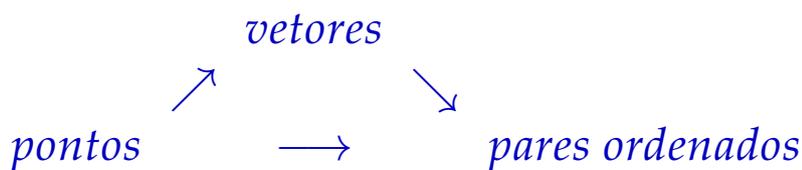
**Observação:** Não vamos, aqui e agora, discutir duas questões gravíssimas: não seria possível, procedendo de forma análoga, escolher adequadamente três ou mais vetores do plano e estabelecer uma bijeção entre o espaço dos vetores do plano e  $\mathbb{R}^3$  ou mesmo um outro  $\mathbb{R}^n$ ? ser linearmente independentes é, de fato, condição necessária e suficiente para que  $\vec{e}_1$  e  $\vec{e}_2$  constituam uma base para o conjunto de vetores do plano?

## 8.4 A Santíssima Trindade

Se a fixação de uma origem  $O$  estabelece uma bijeção entre o plano e o conjunto dos vetores e a fixação de uma base  $(\vec{e}_1, \vec{e}_2)$  estabelece uma bijeção entre conjunto dos vetores do plano e  $\mathbb{R}^2$ , então a fixação simultânea de uma origem  $O$  e de uma base  $(\vec{e}_1, \vec{e}_2)$  estabelece uma bijeção entre o plano e  $\mathbb{R}^2$ .

**Definição:** Um terno ordenado  $(O, \vec{e}_1, \vec{e}_2)$ , sendo  $O$  um ponto do plano (denominado origem) e  $(\vec{e}_1, \vec{e}_2)$  uma base para o conjunto dos vetores do plano, é dito um **sistema de coordenadas** para o plano.

Neste momento, solene, é crucial observar que, se cada sistema de coordenadas estabelece uma bijeção entre o plano e  $\mathbb{R}^2$ , tal bijeção é, de fato, mediada por duas outras: a que a origem cria entre pontos e vetores e a que a base gera entre vetores e pares ordenados. A introdução de um sistema de coordenadas, portanto, leva a uma identificação entre pontos, vetores e pares ordenados, que passam constituir uma espécie de realização matemática do mistério da Santíssima Trindade.



A aceitação do mistério da Santíssima Trindade, aqui, não é uma questão religiosa. Como, em Geometria Analítica, trabalhamos sempre com coordenadas, é crucial ser capaz de, ao lidar com pares ordenados, saber distinguir, apenas pelo contexto, se estes representam pontos ou vetores, para que as ideias geométricas possa ser adequadamente traduzidas algebricamente e para que as manipulações algébricas possam ter sentido geométrico.

# Capítulo 9

## Equações paramétricas

▶ Movimentos

▶ Curvas parametrizadas 1

▶ Curvas parametrizadas 2

Consideremos o problema geral de desenhar curvas definidas por equações. Mais especificamente, consideremos o círculo de equação  $x^2 + y^2 = 1$ . A primeira dificuldade que vamos enfrentar é que a equação nos fornece apenas um teste para decidirmos, para cada ponto  $(x, y)$ , se este pertence ou não à nossa curva. Ao contrário do compasso, que nos indica, preciso e decidido, apenas os pontos que interessam, a equação acima nos obriga a sairmos testando todos e cada um dos pontos do plano, escolhendo os que servem e deixando de lado os que não satisfazem a equação <sup>1</sup>.

Na realidade podemos manipular a equação  $x^2 + y^2 = 1$ , obtendo

$$y^2 = 1 - x^2,$$

ou seja,

$$y = \pm \sqrt{1 - x^2}.$$

Assim, fazendo variar  $x$  de  $-1$  a  $+1$ , obtemos, para cada  $x$ , um par de coordenadas  $y$  correspondentes. Na prática, teremos que estabelecer um número finito de pontos a serem calculados e marcados <sup>2</sup>.

É natural que façamos variar  $x$  a intervalos regularmente espaçados. Uma primeira aproximação pode ser feita, por exemplo, com os valores

$$x = -1, -\frac{3}{4}, -\frac{1}{2}, -\frac{1}{4}, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1.$$

Note que, embora os valores de  $x$  estejam regularmente espaçados, o mesmo não acontece com os pontos do círculo obtidos a partir deles.

<sup>1</sup>Note que essa ideia não é de todo absurda, no caso de estarmos desenhando na tela de um computador, já que o número de pontos na tela é finito e a máquina, veloz

<sup>2</sup>Se estamos desenhando um círculo na tela do computador, é inútil que este número seja superior ao número de pixels na horizontal entre os pontos  $(-1, 0)$  e  $(1, 0)$

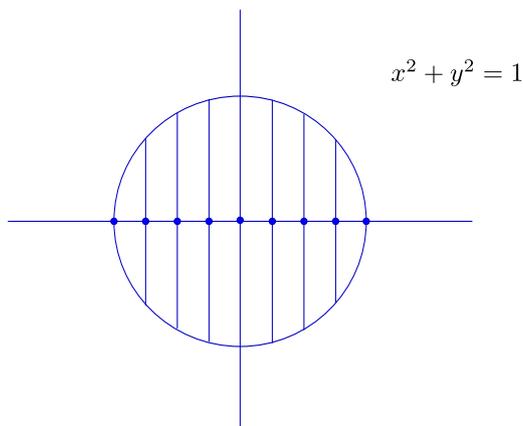


Figura 9.1:

**Exercício 9.1** *Refleta a respeito antes de prosseguir a leitura. Qual o comprimento do arco que vai do ponto  $(1, 0)$  ao ponto  $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ ? Qual o do arco que vai de  $(\frac{1}{2}, \frac{\sqrt{3}}{2})$  até  $(0, 1)$ ? Encontre nove novos valores de  $x$  de forma a obter sobre o círculo pontos regularmente espaçados.*

Um pouco de reflexão nos leva a concluir que pontos regularmente espaçados sobre o círculo podem ser obtidos mais facilmente se trocarmos de parâmetro, usando o ângulo  $\theta$  naturalmente associado a cada ponto no lugar de sua coordenada  $x$ . Chamando de  $\theta$  o ângulo<sup>3</sup> correspondente ao arco que vai de  $(1, 0)$  ao ponto considerado (no sentido trigonométrico), as coordenadas correspondentes serão

$$(x(\theta), y(\theta)) = (\cos\theta, \sin\theta).$$

Como estávamos trabalhando com um total de dezesseis pontos do círculo, devemos variar  $\theta$  de 0 a  $2\pi$ , a intervalos de tamanho  $\frac{\pi}{8}$ :

$$\theta = 0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}, \pi, \frac{9\pi}{8}, \frac{5\pi}{4}, \frac{11\pi}{8}, \frac{3\pi}{2}, \frac{13\pi}{8}, \frac{7\pi}{4}, \frac{15\pi}{8}.$$

É claro que podemos alterar o número de pontos: se queremos  $n$  pontos, basta criar  $\alpha = \frac{2\pi}{n}$  e fazer variar  $\theta$  de  $\alpha$  em  $\alpha$ , começando em  $\theta = 0$  e terminando em  $\theta = (n - 1)\frac{2\pi}{n}$ .

**Exercício 9.2** *Refleta a respeito.*

O fato é que temos aqui uma grande novidade: estamos não mais fornecendo um critério para se testar quais pontos  $(x, y)$  estão sobre o círculo e quais estão fora dele, mas, com a introdução de um parâmetro novo (o ângulo  $\theta$ ), indicando uma fórmula (precisa e decidida como um compasso) para a obtenção direta de pontos do círculo. O

<sup>3</sup>Procuraremos sempre trabalhar com ângulos em radianos, que simplificam o cálculo de derivadas - estas vão nos interessar, pelo menos, por nos facilitarem a determinação das tangentes às curvas

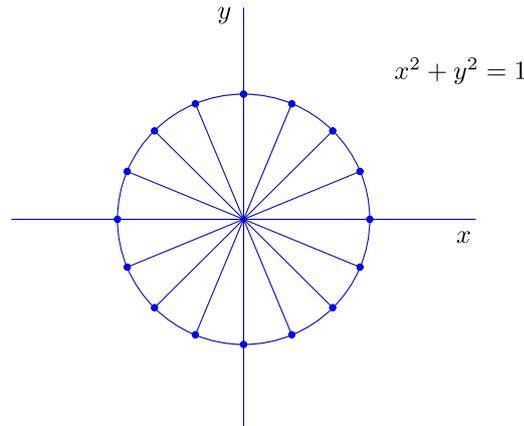


Figura 9.2:

que criamos é uma função que a cada valor do parâmetro  $\theta$  associa um par ordenado  $(x(\theta), y(\theta))$ , dado pelas **equações paramétricas**

$$\begin{cases} x(\theta) = \cos\theta \\ y(\theta) = \sin\theta \end{cases}$$

As equações paramétricas nos fornecem não só a curva, mas um modo de percorrê-la. Quando  $\theta$  varia de 0 a  $2\pi$ , o ponto correspondente  $(x(\theta), y(\theta))$  percorre o círculo, no sentido trigonométrico, a partir do ponto  $(1,0)$ .

Se quisermos o círculo de raio  $r$ , de equação

$$x^2 + y^2 = r^2,$$

basta que façamos

$$\begin{cases} x(\theta) = r\cos\theta \\ y(\theta) = r\sin\theta \end{cases}$$

**Exercício 9.3** Note que se trocarmos as equações para

$$\begin{cases} x(\theta) = \sin\theta \\ y(\theta) = \cos\theta \end{cases}$$

nosso círculo será percorrido no sentido horário a partir do ponto  $(0,1)$ .

**Exercício 9.4** Suponha que o ponto de coordenadas  $(x(t), y(t))$  representa a extremidade móvel do ponteiro dos segundos de um relógio (suponha que o comprimento do ponteiro é 5). Se a origem do sistema de coordenadas está no centro do relógio, encontre as equações paramétricas  $x(t)$  e  $y(t)$ . Faça o mesmo para o ponteiro dos minutos (também de comprimento 5) e para o das horas (suposto de comprimento 3). Nos três casos o parâmetro  $t$  é o tempo (medido em segundos e a partir de 00 : 00).

Examinemos as idéias que acabamos de desenvolver de um ponto de vista mais geral. O que estamos fazendo é, de certa forma, olhar para uma curva não mais como um conjunto mas como a trajetória de uma partícula. Algo assim como acompanhar o movimento da ponta de um lápis que esteja traçando nossa curva. O traçado se faz durante um certo intervalo de tempo, começando, digamos, em  $t = t_0$  e terminando em  $t = t_1$ . Em cada instante  $t$  do intervalo  $[t_0, t_1]$  a ponta do lápis está sobre um ponto  $P(t)$  da curva, sendo as coordenadas de  $P(t)$  dadas por  $(x(t), y(t))$ . As expressões de  $x(t)$  e  $y(t)$  são chamadas de **equações paramétricas** da curva ( $t$  é o **parâmetro**). No caso do ponteiro das horas do exercício acima, as expressões de  $x(t)$  e  $y(t)$  são dadas por

$$\begin{cases} x(t) = 3\text{sen}\theta(t) \\ y(t) = 3\text{cos}\theta(t) \end{cases} ,$$

onde  $\theta(t)$  é o ângulo varrido pelo ponteiro das horas no tempo  $t$  (expresso em segundos). Portanto, visto que o ponteiro varre um ângulo de  $30^\circ$  (o mesmo que  $\pi/6$ , usando a convenção  $^\circ = \pi/180$ ) em uma hora (3600 segundos), temos

$$\theta(t) = \frac{\pi}{21600}t,$$

e, logo, as equações paramétricas são

$$\begin{cases} x(t) = 3\text{sen}\left(\frac{\pi}{21600}t\right) \\ y(t) = 3\text{cos}\left(\frac{\pi}{21600}t\right) \end{cases} .$$

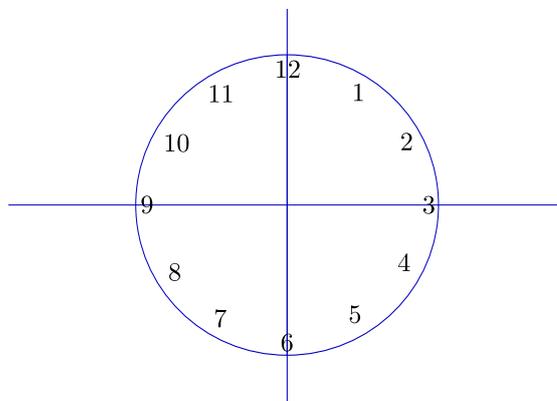


Figura 9.3:

**Exercício 9.5** *Pense bem, refaça o exercício. Tenha certeza de que não vai se confundir com a situação em que a curva representa o gráfico de uma função (nesse caso  $x$  é considerado a variável e temos apenas uma equação, que expressa  $y$  como função de  $x$ ). Aqui a variável  $t$  é invisível como o tempo, o que vemos é o ponto que se move “à medida em que o tempo passa”.*

Vejam mais um exemplo. A espiral abaixo não pode, certamente, representar o gráfico de uma função  $y = f(x)$ . Mas podemos conceber que seja traçada a partir

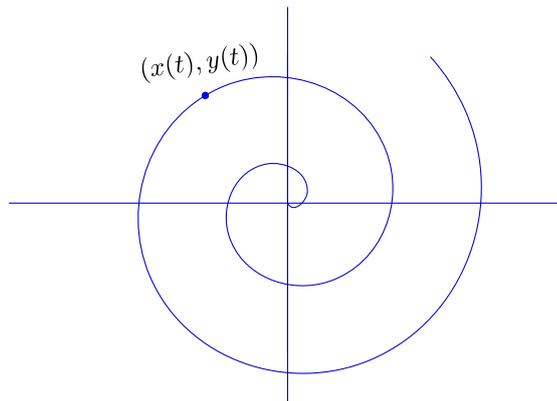


Figura 9.4:

do instante  $t_0 = 0$ , começando da origem. Em cada instante  $t \geq 0$  teremos um ponto da espiral, de coordenadas  $(x(t), y(t))$ .

**Exercício 9.6** Desenhe a curva  $(x(t), y(t))$ ,  $t \geq 0$ , dada por

$$\begin{cases} x(\theta) = t \cos t \\ y(\theta) = t \sin t \end{cases}$$

(note que é como se tentássemos traçar um círculo cujo raio fosse aumentando).

**Exercício 9.7** Observe que o caso em que a curva representa o gráfico de uma função  $y = f(x)$  pode ser visto sob o prisma das equações paramétricas, seja fazendo

$$\begin{cases} x(t) = t \\ y(t) = f(t) \end{cases} ,$$

seja simplesmente considerando que, neste caso, o parâmetro é a própria variável  $x$ .

**Exercício 9.8** Note que uma curva não precisa ser percorrida a velocidade constante. Nada nos impede, por exemplo, de mudar a parametrização do círculo unitário dada anteriormente para

$$\begin{cases} x(t) = \cos t^2 \\ y(t) = \sin t^2 \end{cases}$$

Neste caso, se começarmos de  $t = 0$ , a primeira volta será percorrida no intervalo  $[0, \sqrt{2\pi}]$ , que é maior do que o intervalo  $[\sqrt{2\pi}, 2\sqrt{\pi}]$ , correspondente à segunda volta.

**Exercício 9.9** No exercício anterior, para que valor tende o tempo decorrido entre a  $n$ -ésima e a  $(n+1)$ -ésima primeira passagem de  $(x(t), y(t))$  por  $(1, 0)$ , quando  $n$  tende a infinito?

Passemos agora ao caso da reta. Consideremos a reta passando pela origem e pelo ponto  $(3, 2)$ .

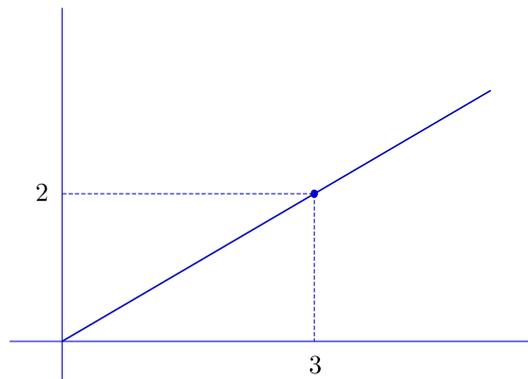


Figura 9.5:

Os pontos  $(x, y)$  da reta são tais que  $x$  e  $y$  são catetos de um triângulo retângulo semelhante ao de catetos 3 e 2. Vale, portanto,

$$\frac{x}{3} = \frac{y}{2} = k.$$

Isto nos conduz a usar a razão de semelhança  $k$  como parâmetro e escrever:

$$\begin{cases} x(k) = 3k \\ y(k) = 2k \end{cases}$$

**Exercício 9.10** Observe que podemos ter  $x$  e  $y$  negativos, o que daria aos catetos correspondentes os valores  $-x$  e  $-y$ , mas isso pode ser facilmente arranjado fazendo  $k$  negativo e mantendo as mesmas equações obtidas acima.

Vejam o que acontece se mudarmos o nome do parâmetro de  $k$  para  $t$ ,  $t$  de tempo. Fazendo variar  $t$  de  $-\infty$  a  $+\infty$ , nossa reta é totalmente percorrida, num certo sentido e com uma certa velocidade: o sentido é da origem para o ponto  $(3, 2)$  e a velocidade é tal que a cada unidade de tempo percorremos distância igual ao comprimento do segmento que vai da origem ao ponto  $(3, 2)$ .

**Exercício 9.11** Note que podemos mudar a velocidade e o sentido do percurso: se  $(a, b)$  é um ponto da reta (outro que a origem), podemos tomar como equações paramétricas:

$$\begin{cases} x(t) = at \\ y(t) = bt \end{cases}$$

Verifique que, neste caso, o sentido é o mesmo se  $a$  e  $b$  forem positivos e muda se forem negativos. A velocidade é dada pela distância de  $(a, b)$  à origem ( $\sqrt{a^2 + b^2}$  por unidade de tempo).

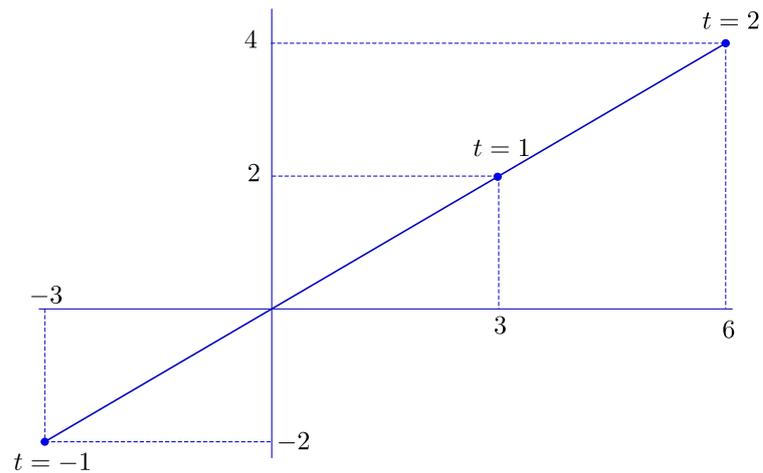


Figura 9.6:

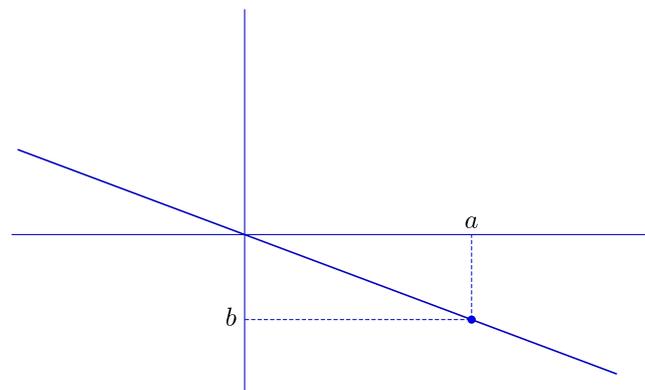


Figura 9.7:

**Exercício 9.12** Note que qualquer reta passando pela origem pode ser parametrizada da mesma forma: escolhamos um ponto  $(a, b)$  outro que a origem e fazemos

$$\begin{cases} x(t) = at \\ y(t) = bt \end{cases}$$

Note que o sentido de percurso é sempre da origem para  $(a, b)$  e que a velocidade é dada pelo comprimento  $\sqrt{a^2 + b^2}$ . Pense nisto cuidadosamente, examine diversos casos.

Vejam agora o que acontece quando a reta considerada não passa pela origem. Tentemos aproveitar o que já foi feito, considerando uma reta paralela à que passa pela origem e por  $(3, 2)$ . Suponhamos que nossa reta passa por um ponto conhecido, digamos  $(1, 2)$ .

Podemos então operar um deslocamento paralelo de todos os pontos da reta que passa pela origem de forma a obtermos a reta desejada (isto é fazemos uma **translação**). Em termos de coordenadas isso se faz somando sempre os mesmos valores às coordenadas  $(3t, 2t)$  do ponto original:

$$\begin{cases} x(t) = 3t + 1 \\ y(t) = 2t + 2 \end{cases}$$

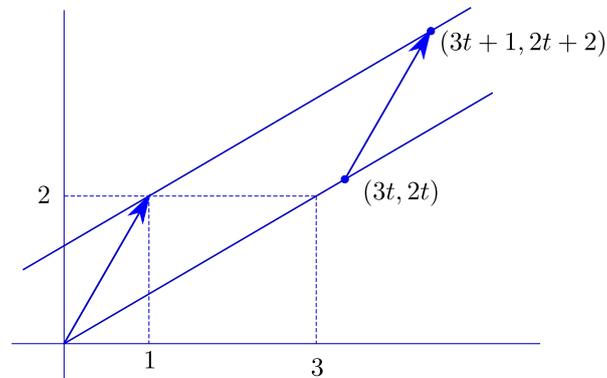


Figura 9.8:

Ora, este é um procedimento geral: se queremos uma reta que passa pelo ponto  $(c, d)$  e é paralela à que passa pela origem e por  $(a, b)$ , basta operarmos um deslocamento análogo

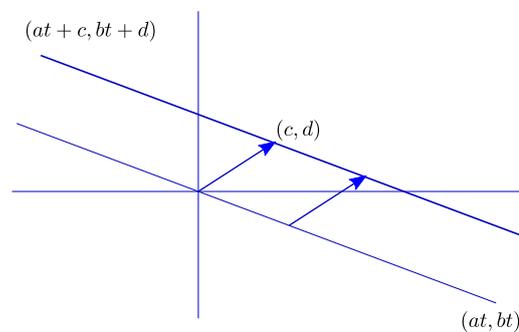


Figura 9.9:

e obteremos

$$\begin{cases} x(t) = at + c \\ y(t) = bt + d \end{cases}$$

**Exercício 9.13** Entenda isto direitinho.

Vejamos o que acontece quando a reta é dada por dois pontos, digamos  $(1, 2)$  e  $(3, 4)$ . Note que se tomarmos uma paralela passando pela origem, esta incluirá o ponto de coordenadas  $(3 - 1, 4 - 2) = (2, 2)$ . A reta passando por  $(2, 2)$  e pela origem será

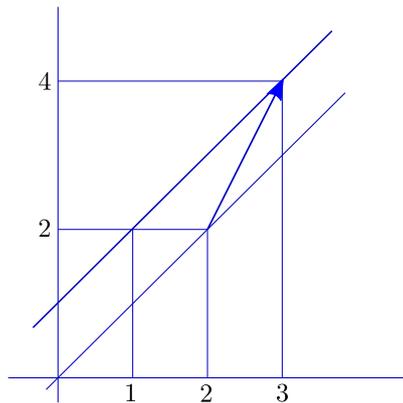


Figura 9.10:

$$\begin{cases} x(t) = 2t \\ y(t) = 2t \end{cases}$$

e a reta desejada pode ser obtida deslocando-se a origem até o ponto  $(1,2)$  (e todos os demais pontos de forma paralela):

$$\begin{cases} x(t) = 2t + 1 \\ y(t) = 2t + 2 \end{cases}$$

Note que mais uma vez temos um procedimento geral: se a reta passa por  $(a_1, b_1)$  e  $(a_2, b_2)$ , podemos escrevê-la na forma paramétrica por

$$\begin{cases} x(t) = (a_2 - a_1)t + a_1 \\ y(t) = (b_2 - b_1)t + b_1 \end{cases}$$

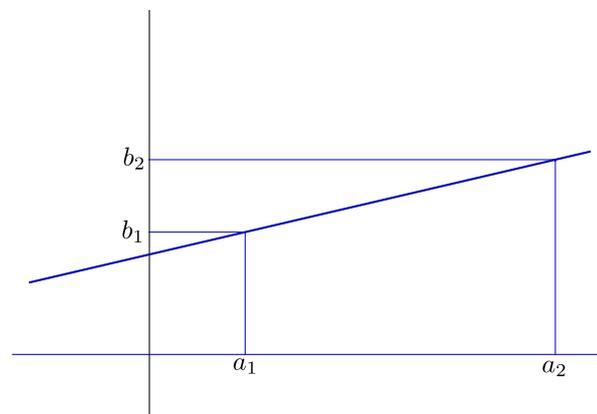


Figura 9.11:

**Exercício 9.14** Escolha dois pontos e ache equações paramétricas para a reta passando por eles.

**Exercício 9.15** Determine equações paramétricas para a reta que passa por  $(1, -3)$  e é normal à reta que passa pela origem e por  $(2, 1)$ .

Solução:

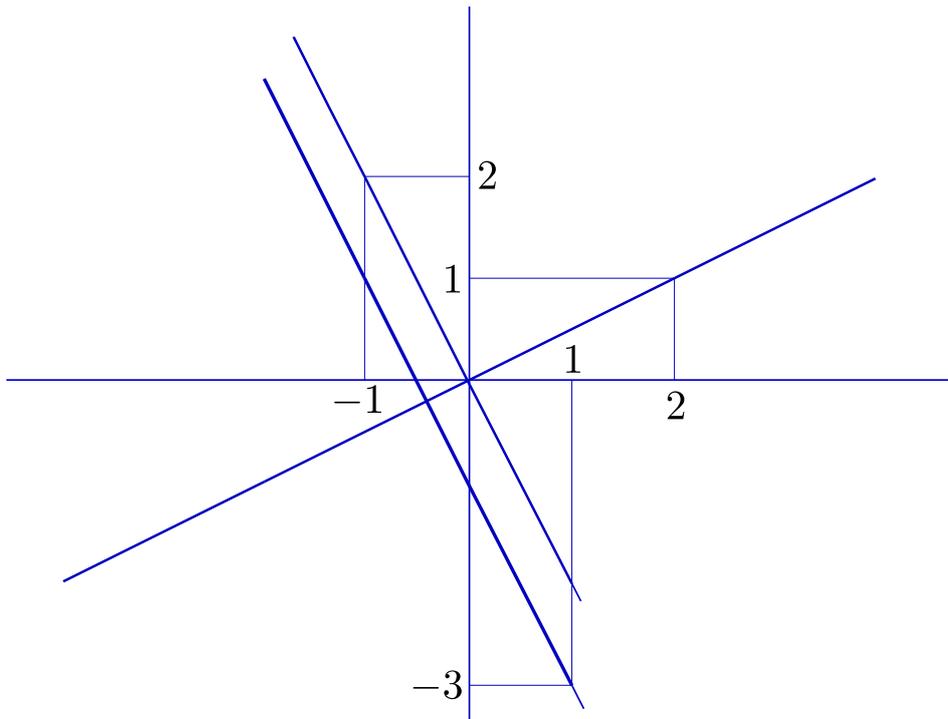


Figura 9.12:

A reta passando pela origem e normal à que passa por  $(2, 1)$  passará pelo ponto  $(-1, 2)$  (veja a semelhança de triângulos na figura). Pode, portanto, ser parametrizada por

$$\begin{cases} x(t) = -1t = -t \\ y(t) = 2t \end{cases} .$$

A reta que queremos é a paralela a esta última passando por  $(1, -3)$ , que parametrizamos por

$$\begin{cases} x(t) = -t + 1 \\ y(t) = 2t - 3 \end{cases} .$$

**SITES:** Existem sites interessantes mostrando variadas curvas. Um (francês) muito bom é

<http://www.mathcurve.com/>

Outro (italiano), com máquinas que geram curvas:

<http://archivioweb.unimore.it/theatrum/inizio.htm>

# Capítulo 10

## Curvas de Bézier

### Curvas de Bézier

Dados, no plano, os pontos  $A_0, \dots, A_n$  ( $n$  é um inteiro, com  $n \geq 0$ ), vamos definir, recursivamente, a curva de Bézier associada a  $A_0, \dots, A_n$ . Nossa curva, que notaremos por  $c$ , é parametrizada:  $c : [0, 1] \rightarrow \mathbb{R}^2$ . Se  $n = 0$ ,  $c$  é constante, dada por  $c(t) = A_0$ , para todo  $t$  em  $[0, 1]$ . Se  $n = 1$ ,  $c$  percorre o segmento de reta que vai de  $A_0$  a  $A_1$ , isto é:

$$c(t) = (1 - t)A_0 + tA_1, \quad t \in [0, 1].$$

De maneira geral, usaremos a notação  $B(A_0, \dots, A_n)$  para designar a **curva de Bézier associada** a  $A_0, \dots, A_n$  (chamaremos os pontos  $A_0, \dots, A_n$  de **pontos de controle**). A definição por recorrência é: se  $B(A_0, \dots, A_{n-1}) = c_0 : [0, 1] \rightarrow \mathbb{R}^2$  e  $B(A_1, \dots, A_n) = c_1 : [0, 1] \rightarrow \mathbb{R}^2$  são, respectivamente, as curvas de Bézier associadas aos  $n$  primeiros e aos  $n$  últimos pontos, então a curva de Bézier associada a  $A_0, \dots, A_n$  é  $B(A_0, A_1, \dots, A_n) = c : [0, 1] \rightarrow \mathbb{R}^2$ , dada por

$$c(t) = (1 - t)c_0(t) + tc_1(t).$$

**Exercício 10.1** Use a definição para provar que, dados os pontos  $A_0, A_1, A_2, A_3$ , temos:

$$B(A_0, A_1, A_2) = (1 - t)^2 A_0 + 2t(1 - t)A_1 + t^2 A_2; \quad B(A_0, A_1, A_2, A_3) = (1 - t)^3 A_0 + 3t(1 - t)^2 A_1 + 3t^2(1 - t)A_2 + t^3 A_3.$$

**Exercício 10.2** Recorde a fórmula do **Binômio de Newton**:

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k.$$

Dos dois exercícios acima, temos razões para especular que,

**Proposição 10.3** : Dados os pontos  $A_0, \dots, A_n$ , a correspondente curva de Bézier,  $B(A_0, \dots, A_n)$  é dada por

$$B(A_0, \dots, A_n)(t) = \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n-k} A_k.$$

**Demonstração:** Provemos por indução sobre  $n$ . O caso  $n = 1$  segue da definição. Suponhamos, agora, o resultado válido para um certo  $n$  e examinemos o caso  $n + 1$ . Da definição e da hipótese de indução, temos

$$\begin{aligned} B(A_0, \dots, A_{n+1}) &= (1-t) \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n-k} A_k + t \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n-k} A_{k+1} = \\ &= \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n+1-k} A_k + \sum_{k=0}^n \binom{n}{k} t^{k+1} (1-t)^{n+1-(k+1)} A_{k+1} = \\ &= \sum_{k=0}^n \binom{n}{k} t^k (1-t)^{n+1-k} A_k + \sum_{k=1}^{n+1} \binom{n}{k-1} t^k (1-t)^{n+1-k} A_k = \\ &= (1-t)^{n+1} A_0 + \sum_{k=1}^n \left( \binom{n}{k} + \binom{n}{k-1} \right) t^k (1-t)^{n+1-k} A_k + t^{n+1} A_{n+1}. \end{aligned}$$

Mas (faça a conta, ou prove por um argumento combinatório)

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

o que nos dá

$$B(A_0, \dots, A_{n+1}) = \sum_{k=0}^{n+1} \binom{n+1}{k} t^k (1-t)^{n+1-k} A_k.$$

■

**Observação:** Trabalhamos, até agora, supondo que os pontos de controle são fixos.

Evidentemente, podemos fazer com que variem, em função de um certo parâmetro  $s$ , de forma a *dar movimento* a nossa curva (isto é, para cada  $s$ , desenhamos a correspondente curva de Bézier, que se moverá com a variação de  $s$  em um certo conjunto  $S$ . Mas podemos, também, permitir que os pontos de controle variem com  $t$ , o próprio parâmetro da curva; evidentemente,  $t$  estará em  $[0, 1]$ . Assim, podemos supor que cada ponto de controle,  $A_i(s, t)$ , é dado por uma função

$$A_i : S \times [0, 1] \longrightarrow \mathbb{R}^2.$$

Note que as definições e a Proposição acima permanecem inalteradas, se trocamos os pontos de controle fixos,  $A_i$ , por  $A_i(s, t)$ . Agora, porém, podemos enunciar com mais simplicidade coisas do seguinte tipo: a curva de Bézier tendo por pontos de controle  $A_0, \dots, A_n$  é igual à que tem por pontos de controle  $B(A_0, A_1), \dots, B(A_{n-1}, A_n)$ , ou seja:

**Lema 10.4** : Para quaisquer  $A_0, \dots, A_n$  (que podem ser pontos fixos ou funções de  $S \times [0, 1]$  em  $E$ ), vale

$$B(A_0, \dots, A_n) = B(B(A_0, A_1), B(A_1, A_2), \dots, B(A_{n-1}, A_n)).$$

**Demonstração:** Provemos por indução. Os casos  $n = 0$  e  $n = 1$  são apenas lógica; para  $n = 2$ , ainda estamos cobertos pela definição. Suponhamos, então, que o resultado vale para um certo  $n$ ; provemos que vale para  $n + 1$ . Da definição, vem

$$B(A_0, \dots, A_{n+1}) = (1 - t)B(A_0, \dots, A_n) + tB(A_1, \dots, A_{n+1}).$$

A hipótese de indução nos dá, então,

$$B(A_0, \dots, A_{n+1}) = (1 - t)B(B(A_0, A_1), \dots, B(A_{n-1}, A_n)) + tB(B(A_1, A_2), \dots, B(A_n, A_{n+1})).$$

Aplicando a fórmula "do Binômio de Newton", temos

$$\begin{aligned} B(A_0, \dots, A_{n+1}) &= (1 - t) \sum_{k=0}^{n-1} \binom{n-1}{k} t^k (1 - t)^{n-1-k} B(A_k, A_{k+1}) + \\ &\quad + t \sum_{k=0}^{n-1} \binom{n-1}{k} t^k (1 - t)^{n-1-k} B(A_{k+1}, A_{k+2}) = \\ &= (1 - t)^n B(A_0, A_1) + \sum_{k=1}^{n-1} \binom{n-1}{k} t^k (1 - t)^{n-k} B(A_k, A_{k+1}) + \\ &\quad + \sum_{k=1}^{n-1} \binom{n-1}{k-1} t^k (1 - t)^{n-k} B(A_k, A_{k+1}) + t^n B(A_n, A_{n+1}) = \\ &= (1 - t) \sum_{k=0}^n \binom{n}{k} t^k (1 - t)^{n-k} B(A_k, A_{k+1}) = B(B(A_0, A_1), B(A_1, A_2), \dots, B(A_n, A_{n+1})). \end{aligned}$$

■

**Proposição 10.5** : Para quaisquer  $A_0, \dots, A_n$  (que podem ser pontos fixos ou funções de  $S \times [0, 1]$  em  $E$ ) e qualquer  $k$  inteiro,  $n > k \geq 1$ , vale

$$B(A_0, \dots, A_n) = B(B(A_0, \dots, A_k), B(A_1, \dots, A_{k+1}), \dots, B(A_{n-k}, A_n)).$$

**Demonstração:** Vamos provar por indução sobre  $k$ . O Lema cuidou do caso  $k = 1$ . Suponhamos agora o resultado provado para um certo  $k$  e tomemos  $n > k + 1$ . Temos, então, pelo Lema,

$$B(A_0, \dots, A_n) = B(B(A_0, A_1), B(A_1, A_2), \dots, B(A_{n-1}, A_n)).$$

Usando a hipótese de indução, segue

$$B(A_0, \dots, A_n) = B(B(B(A_0, A_1), \dots, B(A_k, A_{k+1})), \dots, B(B(A_{n-k-1}, A_{n-k}), \dots, B(A_{n-1}, A_n))).$$

Voltando a aplicar o Lema, temos

$$B(A_0, \dots, A_n) = B(B(A_0, \dots, A_{k+1}), B(A_1, \dots, A_{k+1}), \dots, B(A_{n-(k+1)}, A_n)),$$

o que completa a prova. ■

# Capítulo 11

## Transformações e animações

### ▶ Translações e Deformações

Consideremos agora uma outra possibilidade que o uso de coordenadas nos oferece: transformar figuras no plano em novas figuras. Para melhor visualização do processo, vamos usar dois planos, colocados lado a lado. À esquerda ficarão as figuras originais, cujas coordenadas notaremos por  $(x, y)$ ; à direita as transformadas, de coordenadas  $(u, v)$ .

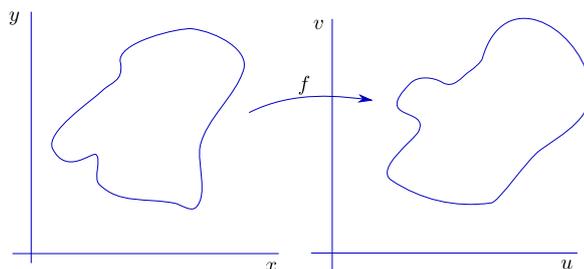


Figura 11.1:

Podemos inventar transformações a nosso bel prazer, a idéia é simples: basta criarmos duas fórmulas que nos dêem as coordenadas  $(u, v)$  em função de  $(x, y)$ . Se  $f_1$  e  $f_2$  são funções das variáveis  $x$  e  $y$ , fazemos:

$$\begin{cases} u = f_1(x, y) \\ v = f_2(x, y) \end{cases} .$$

É mais erudito juntar o par de funções  $f_1, f_2$  em uma só, escrevendo

$$(u, v) = f(x, y),$$

entendido que  $f(x, y)$  tem duas coordenadas, dadas por

$$f(x, y) = (f_1(x, y), f_2(x, y)).$$

Mais chique ainda é escrever

$$f : \begin{array}{l} \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \\ (x, y) \longrightarrow (f_1(x, y), f_2(x, y)) \end{array} '$$

que se lê:

$f$  é a função de  $\mathbb{R}^2$  em  $\mathbb{R}^2$  que associa ao par  $(x, y)$  o par  $(f_1(x, y), f_2(x, y))$ .

Podemos chutar à vontade, por exemplo:

$$\begin{cases} u = \text{sen}(xy) \\ v = \text{cos}(xy) \end{cases} \quad \begin{cases} u = x^2 - y^2 \\ v = 2xy \end{cases} \quad \begin{cases} u = x + y \\ v = x - y \end{cases} \quad \begin{cases} u = e^x \cos y \\ v = e^x \text{sen } y \end{cases}$$

Uma idéia, para começar, é programarmos o computador para que desenhe, a partir de cada fórmula por nós fornecida, as imagens de diversas figuras (que podem ser dadas por equações, escaneadas, ou mesmo criadas a mão livre). O desafio é termos algum controle prévio sobre os resultados que nossas fórmulas vão produzir, a ponto de podermos criar transformações que resultem em efeitos previamente definidos. Uma aplicação interessante, à qual daremos algum destaque nos próximos capítulos, é o uso de transformações para gerar os quadros que compõem uma animação.

Para darmos brevemente uma idéia de como utilizar transformações para gerar animações, comecemos observando que uma animação é composta por uma sequência de quadros (ou seja, de um conjunto de imagens que se sucedem na tela) e que criar a animação equivale a criar os quadros que a compõem. Uma sequência (da animação) representa a evolução no tempo de um certo número de objetos a partir de uma posição inicial.

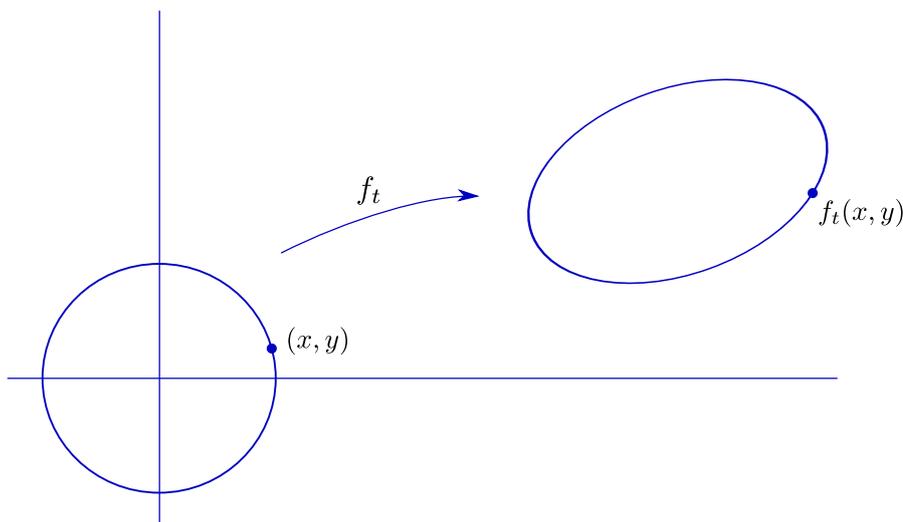


Figura 11.2:

Ora, podemos conceber que uma sequência possa ser construída obtendo-se cada um de seus quadros a partir do primeiro através de uma transformação. Assim, se nossa

sequência descreve a trajetória de uma bola, podemos criar antes de tudo o quadro inicial, que nos descreve a situação no tempo  $t=0$ . Digamos que nossa bola seja descrita por um círculo de raio 1 e centro na origem. Se não houver deformações durante a trajetória, basta-nos dizer para gerar cada quadro da sequência, onde estará o centro do círculo em cada instante; se houver deformações, teremos que, a cada instante, fornecer a transformação que leva o círculo original em uma nova curva que represente, naquele instante, o contorno de nossa “bola”(agora não tão redonda).

De qualquer forma, o processo consiste em fornecer, a cada instante  $t$ , a transformação  $f_t(x, y)$  que leve cada ponto  $(x, y)$  do quadro inicial ( $t = 0$ ) no correspondente ponto  $f_t(x, y)$  do quadro que retrata a situação no instante  $t$ .

**Exercício 11.1** Pegue um computador e brinque com as fórmulas acima e/ou com suas próprias fórmulas.



# Capítulo 12

## Translações

As transformações mais simples (ao menos para quem conhece vetores) são as **translações**: fixa-se um vetor  $\vec{w}$ , que define a translação, e leva-se cada ponto  $P$  no ponto  $P'$  tal que  $\overrightarrow{PP'} = \vec{w}$ .

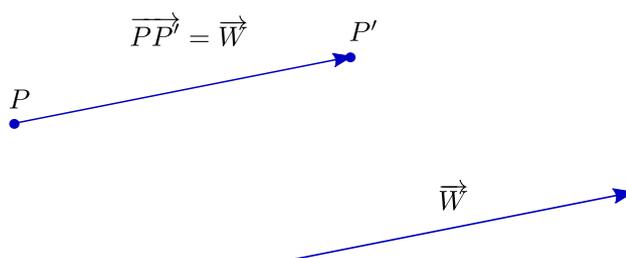


Figura 12.1:

Em termos de coordenadas, se  $\vec{w}=(a,b)$ , nossa transformação levará o ponto  $(x,y)$  no ponto  $(x + a, y + b)$ .

### 12.1 Movimento retilíneo uniforme

Vejam agora algo mais emocionante: animação. Como fazer o ponto  $P$  se mover até o ponto  $P'$  de forma que possamos ver seu deslocamento? Como no cinema, precisamos de uma sucessão de imagens de um ponto (*ponto*, aqui, quer dizer marquinha de tinta, algo visível) ocupando as posições intermediárias entre  $P$  e  $P'$ . O cinema costuma utilizar 24 imagens por segundo; se o tempo do percurso de  $P$  a  $P'$  é de  $n$  segundos, precisaremos de  $24n$  imagens ( $24n + 1$ , contando com  $P$ ). Suporemos que o movimento

se dá em linha reta, com velocidade constante, isto é: que o movimento é **retilíneo e uniforme**.

Para mais simplicidade no raciocínio, comecemos observando que cada uma de nossas imagens retratará um ponto do segmento  $PP'$ . Se  $N$  é o número total de intervalos entre os pontos que queremos, podemos chamar nossos pontos de

$$P = P_0, P_1, P_2, \dots, P_{N-1}, P_N = P'.$$

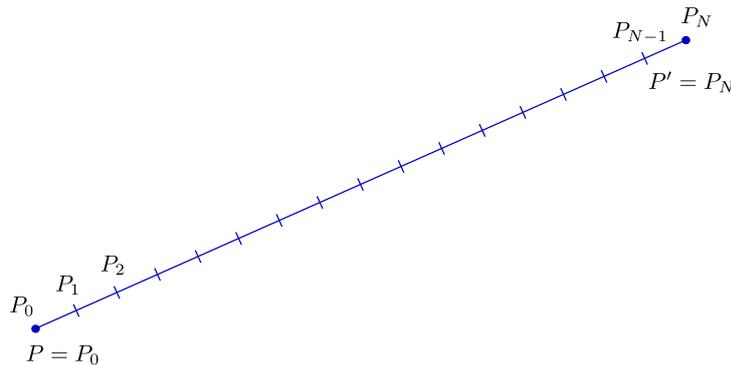


Figura 12.2:

O vetor  $\vec{PP}_i$  ( $i$  é um dos números  $1, 2, 3, \dots, N$ ) é como  $\vec{PP}'$ , só que menorzinho (a menos que  $i = N$ ). Mais precisamente, seu comprimento é  $\frac{i}{N}$  vezes o de  $\vec{PP}'$ . Assim,

$$\vec{PP}_i = \frac{i}{N} \vec{PP}'.$$

Em termos de coordenadas, sendo  $\vec{PP}' = \vec{w} = (a, b)$ , temos

$$\vec{PP}_i = \frac{i}{N} (a, b) = \left( \frac{ia}{N}, \frac{ib}{N} \right),$$

ou seja, escrevendo  $P_i = P + \vec{PP}_i$ <sup>1</sup>, as coordenadas de  $P_i$  serão dadas por

$$\left( x + \frac{ia}{N}, y + \frac{ib}{N} \right),$$

onde  $(x, y)$  representa o ponto  $P$ .

Conseguimos, assim, gerar todos os quadros necessários a nossa animação (cada um deles pode agora ser transformado em um fotograma, como nos desenhos animados do cinema, ou podemos lançá-los sucessivamente na tela do computador, criando diretamente o efeito de animação). Cada quadro será obtido marcando na tela o ponto

<sup>1</sup>Note que, ao somarmos o ponto  $P$  com o vetor  $\vec{PP}_i$ , já estamos identificando livremente pontos e vetores

$P_i$  com as coordenadas obtidas acima. Podemos dizer que cada  $P_i$  é obtido aplicando-se a  $P$  a translação  $f_i$ , com  $f_i$  dada pela fórmula

$$f_i(x, y) = \left(x + \frac{ia}{N}, y + \frac{ib}{N}\right).$$

Note que a fórmula acima pode ser aplicada a outros pontos que não  $P$ . Assim, se tivermos uma figura  $F$  (que é um conjunto de pontos do plano e, para efeitos computacionais, é um arquivo com os pares ordenados correspondentes<sup>2</sup>), podemos aplicar a transformação  $f_i$  aos pontos de  $F$ , gerando um quadro da animação que translada  $F$  até  $F' = F + \vec{w}$ .<sup>3</sup>

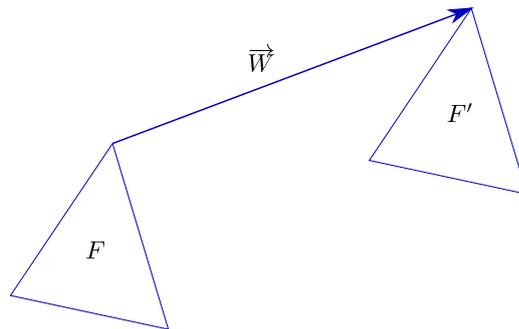


Figura 12.3:

Podemos ainda fazer uma observação interessante: no caso que estamos examinando, basta uma transformação. De fato, a translação de  $P$  a  $P'$  pode ser decomposta em uma sucessão de pequenas translações: de  $P$  a  $P_1$ , de  $P_1$  a  $P_2$ , de  $P_2$  a  $P_3$  e assim sucessivamente, até chegarmos a  $P'$ . Como estamos supondo que o movimento é uniforme, todas essas translações são iguais, já que, em cada caso, o vetor de deslocamento é

$$P_{i-1} \vec{P}_i = \frac{1}{N} \vec{w}.$$

Basta-nos, então, a transformação  $f$  dada por

$$f(x, y) = \left(x + \frac{a}{N}, y + \frac{b}{N}\right).$$

Aplicando  $f$  aos pontos da figura  $F$ , obtemos a figura  $F_1$ ; aplicando  $f$  aos pontos de  $F_1$  obtemos  $F_2$  e assim sucessivamente, até chegarmos a  $F_N = F'$ .<sup>4</sup>

<sup>2</sup>Nos HDs, pendrives ou quaisquer outros meios de armazenamento de dados, mesmo  $F$  sendo, idealmente, um conjunto infinito, só podemos guardar uma quantidade finita de pontos. Essa quantidade está limitada superiormente pela capacidade de armazenamento do meio e pela velocidade de processamento da máquina (já que um número muito grande pode levar a um tempo desmesurado de execução). Por outro lado, temos que zelar para que o número de pontos em nosso arquivo não seja pequeno a ponto de tornar a figura  $F$  irreconhecível

<sup>3</sup> $F + \vec{w}$  é o conjunto formado pelos pontos da forma  $P + \vec{w}$ , com  $P$  em  $F$

<sup>4</sup>Na realidade não precisamos parar em  $F_N$ : enquanto estivermos iterando  $f$ , nossa figura estará andando

As duas seções a seguir são meio chatas. Uma forma de tomar coragem para lê-las é dar primeiro uma olhada na seção **Resumindo e Simplificando**, no final do capítulo. De qualquer forma, é possível viver sem elas.

## 12.2 Movimento retilíneo não uniforme

Voltemos ao movimento (retilíneo) do ponto  $P$  até o ponto  $P'$ . Não é razoável, no mundo físico ou em realidades virtuais, nos limitarmos a movimentos uniformes. Se, por exemplo, uma partícula cai de uma certa altura sob a ação da gravidade, sua velocidade vai aumentando. Se filmarmos seu movimento, obteremos  $N$  quadros, fotografados a intervalos regulares; em cada um deles nossa partícula estará em um ponto  $P_i$ . No entanto, o espaçamento entre os pontos não será regular: o segmento  $P_2P_3$  é maior que  $P_1P_2$ ,  $P_3P_4$  é maior que  $P_2P_3$ , e assim sucessivamente.

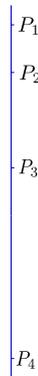


Figura 12.4:

Podemos recorrer a rudimentos de Mecânica para termos uma fórmula descrevendo o movimento de nossa partícula <sup>5</sup>. Se temos queda livre com velocidade inicial nula, a distância percorrida a partir do instante inicial é  $\frac{gt^2}{2}$ , onde  $g$  é a aceleração da gravidade. Para simplificar mais ainda, escolhemos um sistema de unidades em que  $g = 2$ . Assim, o espaço percorrido, na vertical e de cima para baixo, é dado por  $t^2$ , para  $t \geq 0$ .

O vetor unitário vertical é  $\vec{e}_2$ , dado em coordenadas por  $\vec{e}_2 = (0, 1)$ . Se tomarmos o vetor  $t^2\vec{e}_2$  teremos o comprimento certo, mas andaremos para cima. Temos, pois, que somar  $-t^2\vec{e}_2$  a nossa posição inicial  $P$ . Em coordenadas, se  $P = (x, y)$ , nossa posição no instante  $t$  será dada por

$$P + t^2\vec{e}_2 = (x, y) - t^2(0, 1) = (x, y - t^2).$$

Assim, do tempo  $t = 0$  ao tempo  $t = 1$ , nossa partícula cai do ponto  $P = (x, y)$  para o ponto  $P' = (x, y - 1)$ . Para uma boa animação, com  $N$  quadros, o que temos a fazer é

<sup>5</sup> Nosso interesse aqui não é propriamente a Física envolvida; a fórmula que usaremos é apenas uma aproximação, supondo que não há resistência do ar, que a aceleração da gravidade é constante, etc.

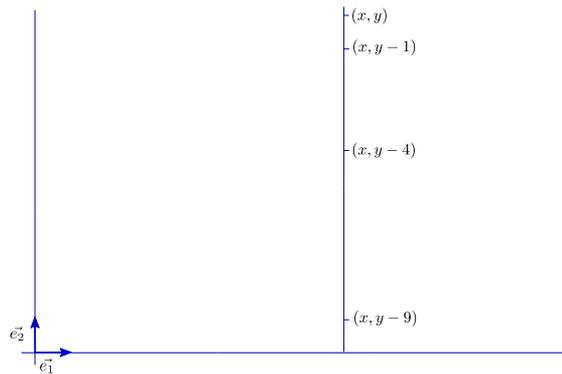


Figura 12.5:

dividir em  $N$  intervalos iguais o *tempo* do percurso, não o *espaço* percorrido!

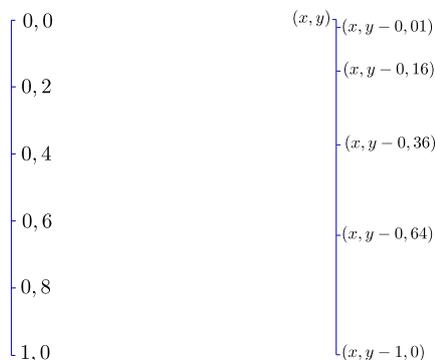


Figura 12.6:

Como temos a posição em função do tempo, dada por  $(x, y - t^2)$ , basta que calculemos as posições correspondentes a cada um dos instantes

$$t_0 = 0, t_1 = \frac{1}{N}, t_2 = \frac{2}{N}, \dots, t_{N-1} = \frac{N-1}{N}, t_N = 1,$$

que dividem o intervalo  $[0, 1]$  em  $N$  subintervalos iguais. Obtemos, para cada  $t_i$ , o ponto  $(x, y - t_i^2)$ .

**Exercício 12.1** Observe que o espaço percorrido a partir do tempo  $t = 0$  não depende da posição inicial, mas só do tempo decorrido. Conclua que se, em vez de uma partícula, deixarmos cair um sólido (que aqui será representado por uma figura plana), cada ponto, no instante  $t$  terá sofrido a mesma translação de  $(0, -t^2)$ .

A observação crucial sobre o exemplo precedente é a seguinte: existem infinitas maneiras de realizar um mesmo trajeto, mesmo retilíneo; a descrição do movimento implica em obter a correspondente **parametrização**, isto é, conhecer o intervalo  $[t_I, t_F]$  do percurso e as equações paramétricas que fornecem as coordenadas do ponto (ou dos pontos) que realiza(m) o movimento.

**Exercício 12.2** Como é o movimento dado por  $(x(t), y(t)) = (\sin t, \sin t)$ ? Sugestão:  $(\sin t, \sin t) = \sin t(1, 1)$ .

### 12.3 Trajetórias não retilíneas

Pelo que acabamos de ver, não há diferenças fundamentais entre fazer animações com trajetórias retilíneas ou curvilíneas, desde que tenhamos as correspondentes parametrizações. Vejamos um exemplo:

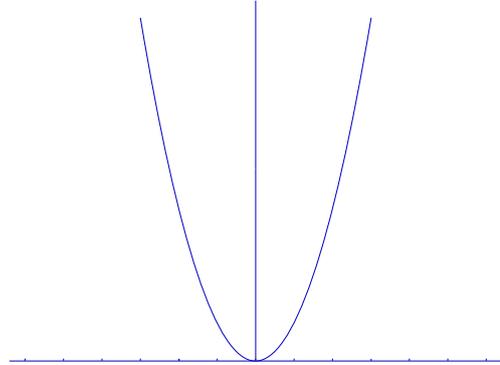


Figura 12.7:

temos um ponto, ocupando a origem em  $t = 0$ , que se desloca sobre a curva descrita por  $(x(t), y(t)) = (t, t^2)$ . Uma animação para este caso, de  $t = 0$  até  $t = T$ , começa pela escolha do número de quadros que vamos utilizar. Digamos que nosso intervalo  $[0, T]$  vai ser subdividido em 20. Geramos então os pontos correspondentes aos tempos

$$t_1 = \frac{T}{20}, t_2 = \frac{2T}{20}, t_3 = \frac{3T}{20}, \dots, t_{20} = T.$$

Temos então, para  $i$  de 0 a 20, os pontos  $P_i = (x(t_i), y(t_i)) = (t_i, t_i^2)$ .

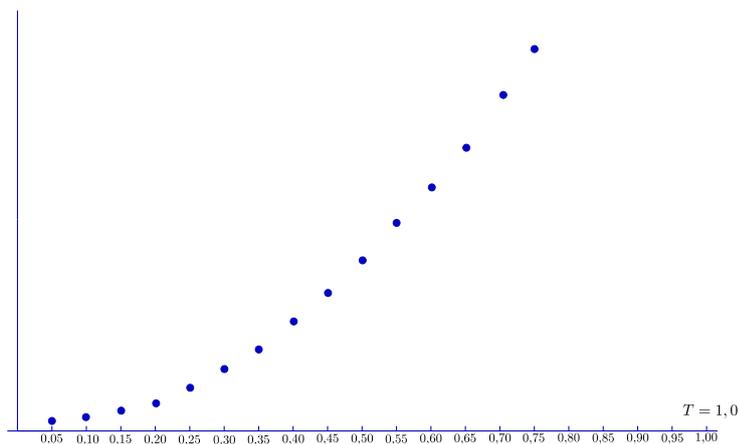


Figura 12.8:

Note que se um outro ponto descreve trajetória igual, mas tendo em  $t = 0$  as coordenadas  $(x_0, y_0)$ , sua posição no tempo  $t$  será dada por

$$(x_0, y_0) + (t, t^2).$$

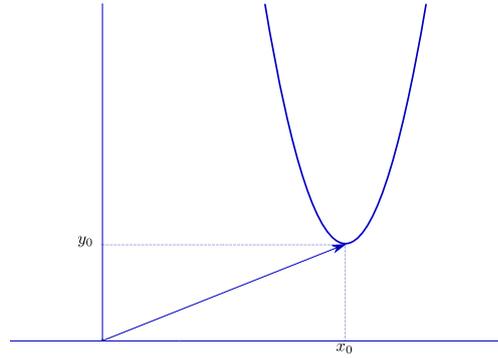


Figura 12.9:

Assim, se toda uma figura do plano descreve a mesma trajetória acima, sua posição no tempo  $t$  será obtida aplicando-se a cada um de seus pontos uma translação de  $(t, t^2)$ . A partir daí podemos proceder como acima para criar uma animação para o movimento da figura.

Tudo que fizemos no exemplo acima pode ser imitado em outras situações, para outras trajetórias: o importante é conseguirmos as equações paramétricas adequadas.

**Exercício 12.3** Estude o movimento do triângulo  $T$  de vértices  $A = (1, 1)$ ,  $B = (2, 0)$ ,  $C = (2, 1)$ , descrito na figura a seguir. Suporemos tratar-se de movimento uniforme.

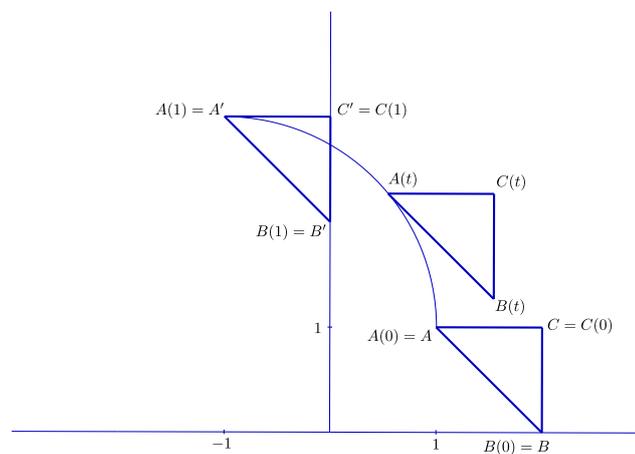


Figura 12.10:

Solução: Observemos, inicialmente, que não se trata de uma rotação - no presente caso, embora cada ponto de  $T$  descreva um círculo, não giram todos em torno de um mesmo centro; em particular note que a horizontal  $AC$  permanece horizontal, e que a vertical  $BC$  permanece vertical. Todos os pontos de  $T$  descrevem círculos de mesmo raio.

Já que a figura nos informa claramente a trajetória de  $A$ , comecemos por ela. Chamaremos de  $A(t)$  a posição de  $A$  no tempo  $t$  e vamos tratar de encontrar as

equações paramétricas correspondentes. Trata-se, sem dúvida, de um círculo de centro  $(-1, 1)$  e raio 2. Se o centro fosse a origem e não tivéssemos qualquer informação sobre o tempo do trajeto, escolheríamos a solução mais simples:

$$(x(t), y(t)) = 2(\cos t, \sin t).$$

A segunda tentativa é transladar o círculo para a posição certa, somando às coordenadas acima o vetor  $(-1, 1)$ :

$$(x(t), y(t)) = (-1, 1) + 2(\cos t, \sin t).$$

Está quase bom, mas assim daremos um quarto de volta em um tempo de  $\frac{\pi}{2}$  e não 1 como mostra a figura. Devemos então fazer uma correção na velocidade angular, obtendo:

$$(x(t), y(t)) = (-1, 1) + 2\left(\cos \frac{\pi}{2}t, \sin \frac{\pi}{2}t\right),$$

que é a forma correta (atenção: certifique-se de que você de fato entendeu esta última passagem).

Cuidemos agora dos demais pontos. Como todos se deslocam da mesma forma, basta determinarmos a translação sofrida por  $A$  no tempo  $t$  e aplicarmos a mesma aos demais para termos suas respectivas posições. Ora, a translação sofrida por  $A$  no tempo  $t$  é dada pelo vetor  $\overrightarrow{AA}(t) = A(t) - A$ , ou seja,

$$\overrightarrow{AA}(t) = (-1, 1) + 2\left(\cos \frac{\pi}{2}t, \sin \frac{\pi}{2}t\right) - (1, 1) = (-2 + 2\cos \frac{\pi}{2}t, 2\sin \frac{\pi}{2}t)$$

(note que, para  $t = 0$ , temos  $\overrightarrow{AA}(t) = \overrightarrow{AA}(0) = \vec{0}$ , já que  $A(0) = A$ ).

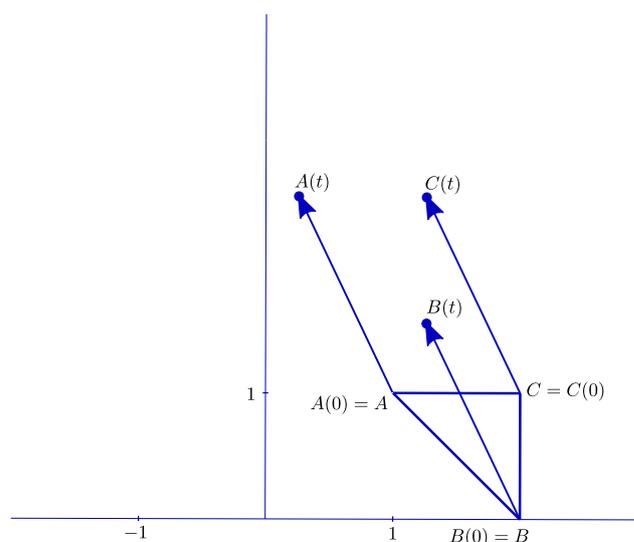


Figura 12.11:

Podemos então obter as coordenadas de qualquer ponto de  $T$  no tempo  $t$  somando o vetor acima a suas coordenadas no instante inicial. Assim, por exemplo, a posição de  $B$  é dada por

$$B(t) = (2, 0) + (-2 + 2\cos \frac{\pi}{2}t, 2\sin \frac{\pi}{2}t) = (2\cos \frac{\pi}{2}t, 2\sin \frac{\pi}{2}t).$$

Uma alternativa é observarmos o seguinte: se  $P$  é um ponto qualquer do triângulo, em qualquer instante  $t$  o vetor  $\vec{A(t)P(t)}$  é igual a  $\vec{AP}$ , ou seja,

$$P(t) = A(t) + \vec{AP} = A(t) + P - A.$$

Em particular, podemos obter de novo  $B(t)$ :

$$B(t) = A(t) + B - A = (-1, 1) + 2(\cos \frac{\pi}{2}t, \sin \frac{\pi}{2}t) + (2, 0) - (1, 1) = (2\cos \frac{\pi}{2}t, 2\sin \frac{\pi}{2}t).$$

**Exercício 12.4** Determine as equações paramétricas de  $C(t)$ .

## 12.4 Resumindo e Simplificando

Uma translação é definida por um único vetor  $\vec{w}$  e leva cada ponto  $P$  no ponto  $P + \vec{w}$ . Em coordenadas, se  $\vec{w} = (a, b)$  e  $P = (x, y)$ , teremos

$$T(x, y) = (x + a, y + b).$$

Para gerarmos uma animação através de translações, basta fornecermos uma figura de referência e, para cada instante  $t$  do intervalo em que o movimento vai ocorrer, um vetor  $\vec{w}(t) = (a(t), b(t))$  que “transporte” os pontos da posição de referência (que pode ou não ser a posição inicial) para a posição no tempo  $t$ .

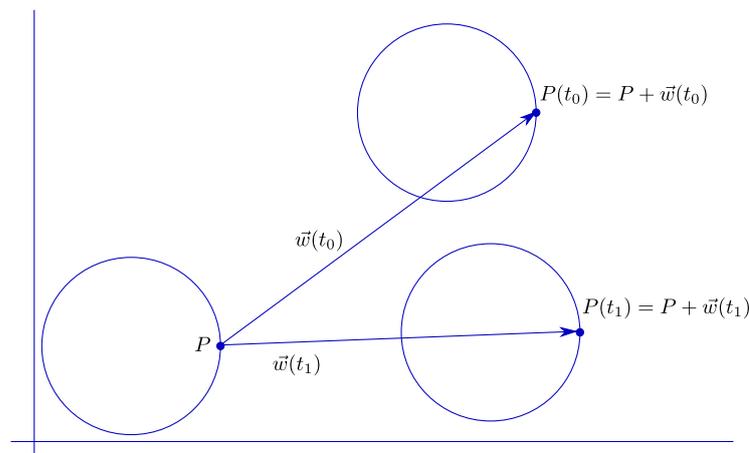


Figura 12.12:

**Exercício 12.5** Refaça os exemplos e exercícios deste capítulo à luz dos esclarecimentos acima.



# Capítulo 13

## Rotações

Uma segunda classe de transformações elementares é a das **rotações**.

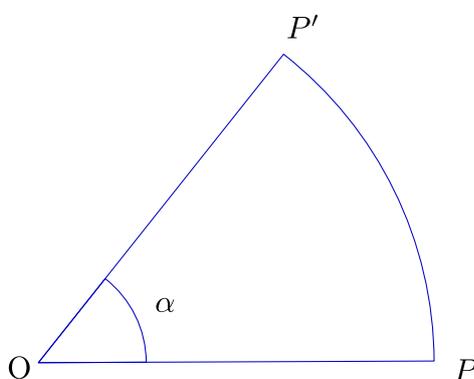


Figura 13.1:

### 13.1 Rotações em torno da origem

Para chegarmos a uma fórmula que expresse as coordenadas do ponto rodado em relação às originais, vamos começar supondo que nossa rotação tem centro na origem do sistema de coordenadas (canônicas) e que o ângulo  $\alpha$  é medido no sentido trigonométrico (o sentido horário será representado pelo sinal negativo).

Seja pois  $P$  um ponto de coordenadas  $(x, y)$  e procuremos obter as correspondentes coordenadas  $(x', y')$  do ponto  $P'$  obtido quando o submetemos a uma rotação de  $\alpha$ . Para auxiliar os cálculos, vamos associar a  $P$  duas novas grandezas: sua distância à origem,  $r$ , e o ângulo de seu vetor posição com o semieixo horizontal positivo,  $\theta$ .<sup>1</sup>

**Exercício 13.1** Verifique que  $x = r \cos \theta$  e  $y = r \sin \theta$ , onde  $\theta$  é medido do semieixo horizontal para  $\vec{OP}$ , considerado positivo o sentido trigonométrico.

<sup>1</sup> $r$  e  $\theta$  são chamados de **coordenadas polares** de  $P$ ; note que o ângulo  $\theta$  não é único: se  $\theta$  serve e  $k$  é um número inteiro, então  $\theta + 2k\pi$  também serve

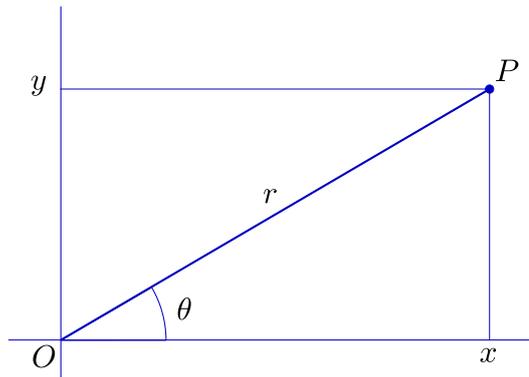


Figura 13.2:

A distância de  $P'$  à origem será, é claro,  $r' = r$ , e seu ângulo com a horizontal,  $\theta'$ , dado por  $\theta' = \theta + \alpha$ . Os valores de  $x'$  e  $y'$  serão dados por:

$$x' = r' \cos \theta' = r \cos(\theta + \alpha) = r(\cos \theta \cos \alpha - \sin \theta \sin \alpha),$$

$$y' = r' \sin \theta' = r \sin(\theta + \alpha) = r(\cos \theta \sin \alpha + \sin \theta \cos \alpha),$$

lançando mão de famosas fórmulas trigonométricas das quais daremos mais tarde demonstrações independentes. Como  $r \cos \theta = x$  e  $r \sin \theta = y$ , temos

$$x' = x \cos \alpha - y \sin \alpha,$$

$$y' = x \sin \alpha + y \cos \alpha.$$

As expressões acima podem, ainda, ser colocadas na forma matricial:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

A matriz

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

é chamada **matriz de rotação** (correspondente ao ângulo  $\alpha$ ).<sup>2</sup>

<sup>2</sup>Lembramos que o produto da matriz  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  pelo vetor  $\begin{pmatrix} x \\ y \end{pmatrix}$  é definido por

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

Assim, a cada rotação associamos a matriz correspondente, com a cara acima. A obtenção de animações se faz como aplicação direta.

**Exercício 13.2** Dado o ponto  $P = (2, 1)$ , gere os quadros para uma animação em que  $P$  roda em torno da origem, percorrendo um ângulo reto em dez segundos.

Solução: vamos trabalhar com o padrão de 24 imagens por segundo. Teremos, então, que gerar 240 imagens, o que implica em dividir o ângulo reto ( $\frac{\pi}{2}$ ) em 240. No entanto, se o movimento é uniforme (o que vamos supor), não é preciso trabalhar com 240 ângulos de rotação diferentes: basta que rodemos nosso ponto, passo a passo, de um ângulo de  $\frac{\pi}{480}$  a cada passo. Isto é, vamos, a partir do ponto  $P_0 = P$ , gerar os 240 pontos  $P_1, P_2, P_3, P_4, \dots, P_{240}$ , de forma que cada um seja obtido do anterior por uma rotação de  $\frac{\pi}{480}$ . Assim, se  $P_i = (x_i, y_i)$ , teremos:

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} \cos \frac{\pi}{480} & -\operatorname{sen} \frac{\pi}{480} \\ \operatorname{sen} \frac{\pi}{480} & \cos \frac{\pi}{480} \end{pmatrix} \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix}, \quad i = 1, 2, 3, \dots, 240.$$

Num caso mais geral de rotação em torno da origem, podemos proceder como no caso das translações: trabalhamos com uma figura de referência  $F$ , um intervalo de tempo  $[T_0, T_1]$  e uma função  $\alpha$  que a cada  $t$  em  $[T_0, T_1]$  associa o ângulo  $\alpha(t)$  de que serão rodados os pontos de  $F$  no tempo  $t$ . Assim, se  $P = (x, y)$  é um ponto de  $F$ , sua posição no tempo  $t$ ,  $(x(t), y(t))$ , será dada por

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos \alpha(t) & -\operatorname{sen} \alpha(t) \\ \operatorname{sen} \alpha(t) & \cos \alpha(t) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Para gerarmos a animação correspondente, basta agora dividir o intervalo de tempo  $[T_0, T_1]$  no número de subintervalos desejado,  $N$ , por meio dos pontos

$$t_0 = T_0, t_1, t_2, t_3, \dots, t_N = T_1,$$

substituir os valores  $t_i$  na expressão matricial acima e computar os correspondentes pontos  $(x(t_i), y(t_i))$ .

## 13.2 Rotação em torno de um ponto qualquer

A maneira mais simples de obtermos uma fórmula para a rotação de um ângulo  $\alpha$  em torno de um ponto  $C = (a, b)$  é “trazermos tudo para a origem”. Isto é, se  $P = (x, y)$  é o ponto a ser rodado em torno de  $C$ , olhamos para  $C$  como se fosse a origem e para  $P$  como se suas coordenadas fossem  $(x - a, y - b)$ . Após efetuarmos a rotação de  $\bar{P} = (x - a, y - b)$  em torno da origem, “levamos de volta”.

Em termos mais eruditos, começamos aplicando a  $P$  a translação de  $(-a, -b)$ , obtendo o ponto  $\bar{P}$ . A  $\bar{P}$  aplicamos a rotação de  $\alpha$  em torno da origem, obtendo o ponto  $\bar{P}'$ . Finalmente, aplicamos a  $\bar{P}'$  a translação de  $(a, b)$ , o que nos dá o ponto  $P'$ , que é  $P$  rodado de  $\alpha$  em torno de  $C$ .

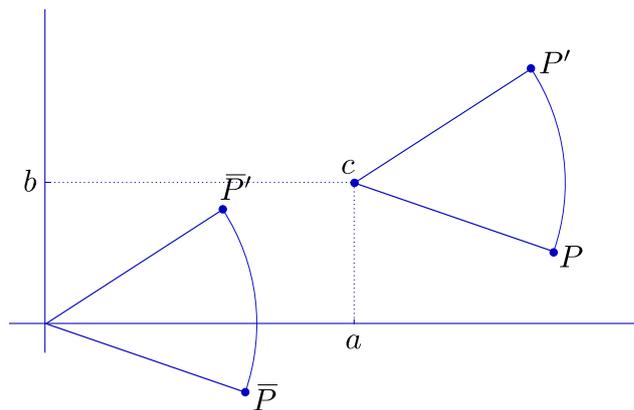


Figura 13.3:

Assim, temos  $P = (x, y)$ ,  $\bar{P} = (x-a, y-b)$ ; as coordenadas de  $\bar{P}'$  serão obtidas aplicando às de  $\bar{P}$  a fórmula de rotação em torno da origem:

$$\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x-a \\ y-b \end{pmatrix};$$

$P'$  terá, então, coordenadas  $(x', y')$ , dadas por

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

### 13.3 Rotação de vetores

Vamos retomar as rotações em torno da origem, mas de um novo ponto de vista. Consideraremos a rotação de  $\alpha$  como uma transformação aplicada a vetores<sup>3</sup>.

Usaremos a notação  $R_\alpha$  para designar a rotação de  $\alpha$  em torno da origem. Se  $\vec{u} = \overrightarrow{OP}$ , definimos  $R_\alpha \vec{u}$  por  $R_\alpha \vec{u} = \overrightarrow{OP'}$ , sendo  $P'$  o ponto obtido pela rotação de  $P$  em torno da origem. Podemos observar<sup>4</sup> que  $R_\alpha$  goza de duas propriedades notáveis:

$$\begin{aligned} (i) \quad R_\alpha(\vec{u} + \vec{v}) &= R_\alpha(\vec{u}) + R_\alpha(\vec{v}) \quad \forall \vec{u}, \vec{v} \in \mathbf{R}^2; \\ (ii) \quad R_\alpha(t\vec{u}) &= tR_\alpha(\vec{u}) \quad \forall \vec{u} \in \mathbf{R}^2, t \in \mathbf{R}. \end{aligned}$$

Se agora utilizarmos o fato de que um vetor  $\vec{u}$  de coordenadas  $(x, y)$  pode ser escrito

$$\vec{u} = xe_1 + ye_2,$$

teremos

<sup>3</sup>Lembramos que estamos identificando pontos, vetores e pares ordenados, de maneira que não estamos operando qualquer alteração formal: ao fim e ao cabo, uma rotação é uma transformação de  $\mathbf{R}^2$  em  $\mathbf{R}^2$

<sup>4</sup>O termo "observar", aqui, varre para debaixo do tapete o fato de que as duas propriedades (particularmente a primeira) merecem demonstração geométrica

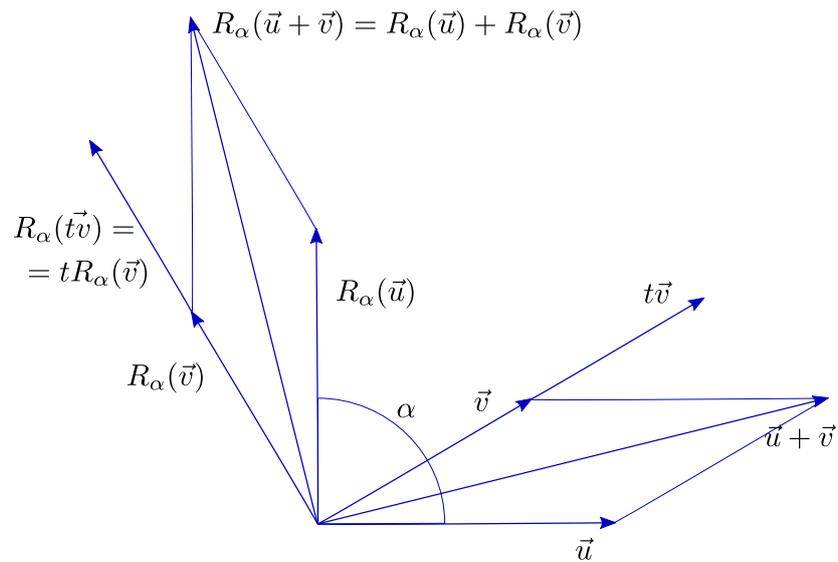


Figura 13.4:

$$R_\alpha(\vec{u}) = R_\alpha(x\vec{e}_1 + y\vec{e}_2) = R_\alpha(x\vec{e}_1) + R_\alpha(y\vec{e}_2) = xR_\alpha(\vec{e}_1) + yR_\alpha(\vec{e}_2).$$

Assim, para obtermos a expressão para  $R_\alpha(x, y)$ , basta conhecermos  $R_\alpha(\vec{e}_1)$  e  $R_\alpha(\vec{e}_2)$ , o que é relativamente fácil:

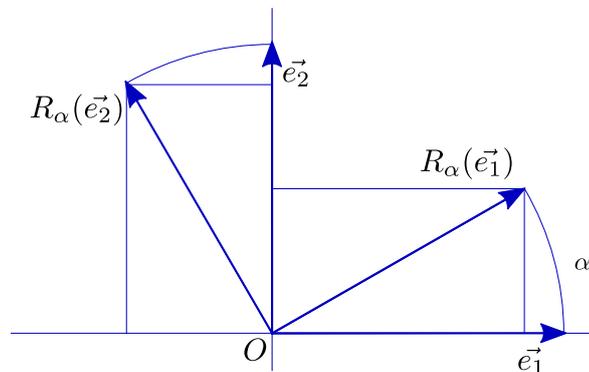


Figura 13.5:

$$R_\alpha(\vec{e}_1) = (\cos\alpha, \sin\alpha), \quad R_\alpha(\vec{e}_2) = (-\sin\alpha, \cos\alpha).$$

Concluimos então que

$$R_\alpha(\vec{u}) = x(\cos\alpha, \sin\alpha) + y(-\sin\alpha, \cos\alpha) = (x\cos\alpha - y\sin\alpha, x\sin\alpha + y\cos\alpha),$$

ou, em notação matricial,

$$R_\alpha(\vec{u}) = \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Note o leitor que acabamos de obter novamente a fórmula para a rotação em torno da origem, mas agora sem fazer uso das famosas fórmulas trigonométricas que prometemos demonstrar. Vamos então a elas.

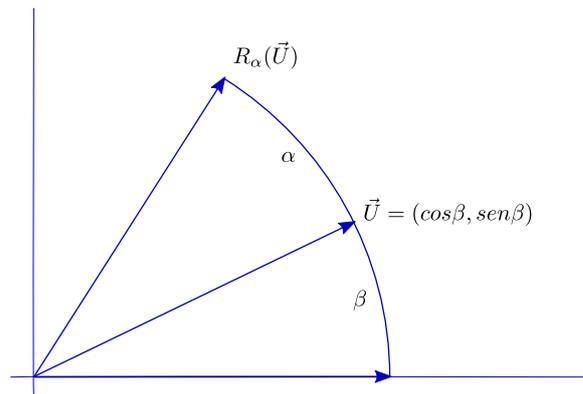


Figura 13.6:

Consideremos o vetor  $\vec{u} = (\cos\beta, \operatorname{sen}\beta)$ . A fórmula que acabamos de obter nos dá

$$R_\alpha(\vec{u}) = (\cos\alpha \cos\beta - \operatorname{sen}\alpha \operatorname{sen}\beta, \operatorname{sen}\alpha \cos\beta + \cos\alpha \operatorname{sen}\beta).$$

Por outro lado, é imediato que

$$R_\alpha(\vec{u}) = (\cos(\alpha + \beta), \operatorname{sen}(\alpha + \beta)).$$

Igualando as duas expressões, obtemos as famosas fórmulas:

$$\cos(\alpha + \beta) = \cos\alpha \cos\beta - \operatorname{sen}\alpha \operatorname{sen}\beta;$$

$$\operatorname{sen}(\alpha + \beta) = \operatorname{sen}\alpha \cos\beta + \cos\alpha \operatorname{sen}\beta.$$

**Exercício 13.3** Verifique que a rotação de  $\alpha$  no sentido horário é dada pela matriz

$$\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$$

**Exercício 13.4** Prove (ou, pelo menos, convença-se de) a veracidade das duas propriedades de  $R_\alpha$  em que baseamos todas nossas deduções:

$$\begin{aligned} (i) \quad R_\alpha(\vec{u} + \vec{v}) &= R_\alpha(\vec{u}) + R_\alpha(\vec{v}) \quad \forall \vec{u}, \vec{v} \in \mathbf{R}^2; \\ (ii) \quad R_\alpha(t\vec{u}) &= tR_\alpha(\vec{u}) \quad \forall \vec{u} \in \mathbf{R}^2, t \in \mathbf{R}. \end{aligned}$$

*Sugestão: comece provando (ii), que é mais simples; em seguida, considere  $P$  e  $Q$  tais que  $\vec{u} = \overrightarrow{OP}$  e  $\vec{v} = \overrightarrow{OQ}$ , prove a congruência entre os triângulos  $OPQ$  e  $OP'Q'$  e conclua que o ponto médio de  $P'Q'$  é o rodado do ponto médio de  $PQ$ ; use a propriedade (ii) para, daí, concluir que vale a propriedade (i).*



# Capítulo 14

## Homotetias

A **homotetia** de centro  $C$  e razão  $k$  ( $k \geq 0$ ) leva o ponto  $P$  no ponto  $P'$  situado na semi-reta  $CP$  e tal que o comprimento de  $CP'$  é  $k$  vezes o de  $CP$  (faremos ainda, por definição,  $C' = C$ ).

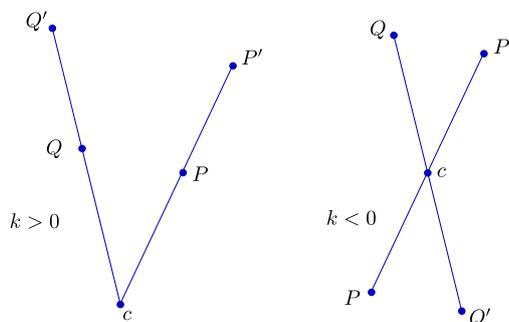


Figura 14.1:

Se  $k \leq 0$ ,  $P'$  será o ponto da reta  $CP$  tal que  $C$  está entre  $P$  e  $P'$ , de forma que o comprimento de  $CP'$  seja  $|k|$  vezes o de  $CP$  (mantemos  $C' = C$ ).

Em termos vetoriais, é claro que, para  $k$  positivo, negativo ou mesmo nulo, temos

$$\overrightarrow{CP'} = k \overrightarrow{CP}.$$

Se o centro  $C$  for a origem do sistema de coordenadas, teremos também, se  $P = (x, y)$  e  $P' = (x', y')$ ,

$$(x', y') = k(x, y).$$

Se o centro  $C$  é um ponto qualquer, de coordenadas  $(a, b)$ , então podemos escrever

$$\overrightarrow{CP'} = (x' - a, y' - b), \quad \overrightarrow{CP} = (x - a, y - b),$$

e, como  $\overrightarrow{CP'} = k \overrightarrow{CP}$ ,

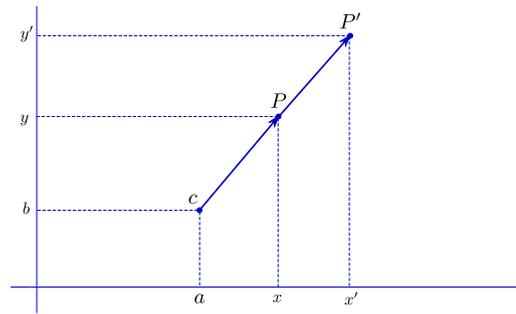


Figura 14.2:

$$(x', y') = (kx + (1 - k)a, ky + (1 - k)b) = k(x, y) + (1 - k)(a, b).$$

Animações similares às que já estudamos para translações e rotações podem ser feitas, por exemplo, mantendo fixo o centro  $C$  e fazendo variar a razão  $k$  com o tempo, isto é, fornecendo uma função  $k(t)$  e considerando, para cada  $t$ , a figura homotética (pela homotetia de centro  $C$  e razão  $k(t)$ ) a uma figura de referência. A função  $k$  deve, é claro, ter o valor 1 no instante inicial e o valor desejado para a razão de homotetia, no instante final.

**Exercício 14.1** *Mostre que a homotetia de centro na origem e razão  $k$  é dada por* 
$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

# Capítulo 15

## Reflexões

Translações e rotações são transformações que preservam distâncias - em consequência, levam cada figura  $F$  em uma figura  $F'$  congruente a  $F$ . Existe ainda uma terceira classe de transformações com essa propriedade, a das **reflexões**.

### 15.1 Reflexão de ponto através de reta passando pela origem

Vejam como expressar em coordenadas a reflexão através de uma reta  $r$  dada. Começemos pelo caso simples em que  $r$  é o eixo dos  $x$ .

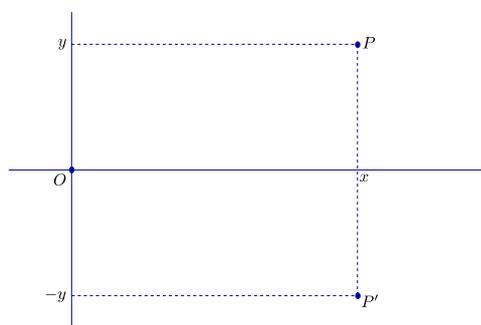


Figura 15.1:

Neste caso, nossa reflexão transforma o ponto  $P = (x, y)$  no ponto  $P' = (x, -y)$ .

No caso em que  $r$  apenas passa pela origem, fazendo um ângulo  $\alpha$  com a horizontal, podemos começar rodando o plano todo de  $-\alpha$ , até que  $r$  se torne horizontal;  $P$  irá parar no ponto  $\bar{P}$ . Em seguida fazemos a reflexão de  $\bar{P}$  através da horizontal e rodamos de volta o ponto  $\bar{P}'$  assim obtido, obtendo finalmente o ponto  $P'$  procurado.

Em termos de coordenadas, se  $P = (x, y)$  teremos  $\bar{P} = (\bar{x}, \bar{y})$ , onde

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} \cos\alpha & \text{sen}\alpha \\ -\text{sen}\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\alpha + y\text{sen}\alpha \\ -x\text{sen}\alpha + y\cos\alpha \end{pmatrix}.$$

$\bar{P}' = (\bar{x}, -\bar{y})$  será então dado por

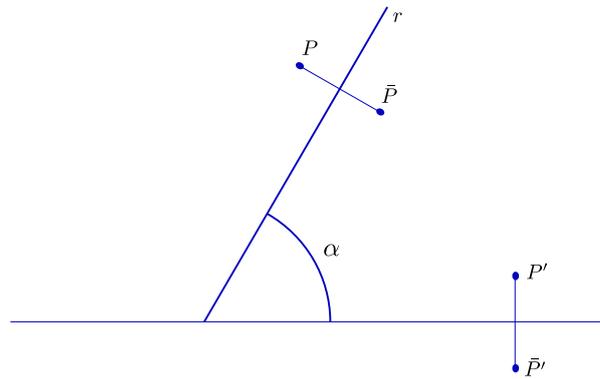


Figura 15.2:

$$\bar{P}' = \begin{pmatrix} \bar{x} \\ -\bar{y} \end{pmatrix} = \begin{pmatrix} x \cos \alpha + y \sin \alpha \\ x \sin \alpha - y \cos \alpha \end{pmatrix}.$$

Podemos agora obter as coordenadas  $(x', y')$  de  $P'$  por

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \bar{x} \\ -\bar{y} \end{pmatrix} = \\ &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x \cos \alpha + y \sin \alpha \\ x \sin \alpha - y \cos \alpha \end{pmatrix} = \\ &= \begin{pmatrix} x(\cos^2 \alpha - \sin^2 \alpha) + y 2 \cos \alpha \sin \alpha \\ x 2 \cos \alpha \sin \alpha - y(\cos^2 \alpha - \sin^2 \alpha) \end{pmatrix}. \end{aligned}$$

Como  $\cos^2 \alpha - \sin^2 \alpha = \cos(2\alpha)$  e  $2 \cos \alpha \sin \alpha = \sin(2\alpha)$ , podemos concluir que

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

## 15.2 Reflexão de vetores

Assim como fizemos no caso das rotações, vamos dar às reflexões um tratamento alternativo, baseado na idéia de **transformação linear**<sup>1</sup>, que torna mais simples a dedução da fórmula.

Consideremos uma reta  $r$  passando pela origem e seja  $S_r$ <sup>2</sup> a transformação que a cada vetor  $\vec{v}$  associa sua imagem refletida através de  $r$ .

<sup>1</sup>Transformações lineares serão objeto de um capítulo à parte, mais à frente

<sup>2</sup>S de *simetria*, para não usarmos o mesmo **R** de *rotação*

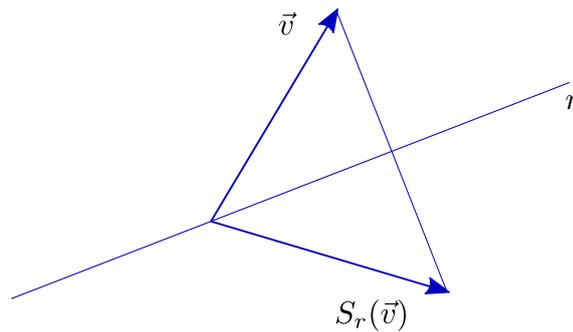


Figura 15.3:

A exemplo das rotações e homotetias,  $S_r$  tem as seguintes propriedades notáveis (sobre as quais o leitor deve pensar um pouco até acreditar, ou demonstrar geometricamente):<sup>3</sup>

$$\begin{aligned} (\beta) S_r(\vec{u} + \vec{v}) &= S_r(\vec{u}) + S_r(\vec{v}) \quad \forall \vec{u}, \vec{v} \in \mathbf{R}^2; \\ (\beta\beta) S_r(t\vec{u}) &= tS_r(\vec{u}) \quad \forall \vec{u} \in \mathbf{R}^2, t \in \mathbf{R}. \end{aligned}$$

Da mesma forma que no caso das rotações, estas propriedades nos permitem obter, para  $\vec{u} = (x, y)$ ,

$$S_r(\vec{u}) = S_r(x\vec{e}_1 + y\vec{e}_2) = xS_r(\vec{e}_1) + yS_r(\vec{e}_2).$$

Ou seja, basta-nos obter as expressões de  $S_r(\vec{e}_1)$  e  $S_r(\vec{e}_2)$ . Sendo, como anteriormente,  $\alpha$  o ângulo de  $r$  com a horizontal, temos:

$$S_r(\vec{e}_1) = (\cos 2\alpha, \sin 2\alpha), \quad S_r(\vec{e}_2) = (\sin 2\alpha, -\cos 2\alpha),$$

o que conduz imediatamente a

$$S_r(x, y) = x(\cos 2\alpha, \sin 2\alpha) + y(\sin 2\alpha, -\cos 2\alpha).$$

Juntando tudo e colocando na forma matricial, obtemos de novo, fazendo  $S_r(x, y) = (x', y')$ ,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

**Exercício 15.1** Verifique que se  $r$  é reta vertical passando pela origem, então a reflexão através de  $r$  é dada por  $S_r(x, y) = (-x, y)$ .

<sup>3</sup> pode, também, não acreditar e apontar um erro

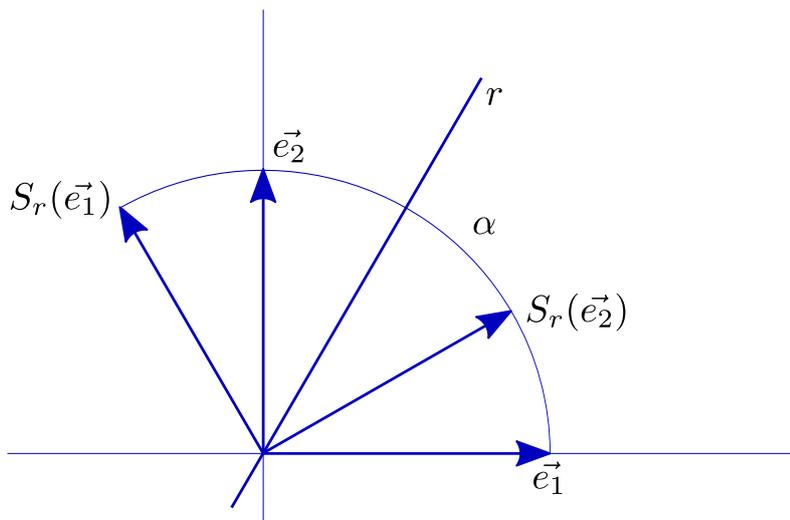


Figura 15.4:

**Exercício 15.2** Mostre que se  $r$  é reta não vertical, então pode ser dada por equação da forma  $y = mx + p$ , onde  $m$  é a tangente do ângulo  $\alpha$  que faz com a horizontal.

**Exercício 15.3** Mostre que o seno e o cosseno de  $2\alpha$  podem ser obtidos a partir da tangente de  $\alpha$ . Isto é, deduza as seguintes fórmulas:

$$\cos 2\alpha = \frac{1 - \operatorname{tg}^2 \alpha}{1 + \operatorname{tg}^2 \alpha}$$

$$\operatorname{sen} 2\alpha = \frac{2 \operatorname{tg} \alpha}{1 + \operatorname{tg}^2 \alpha}$$

*Sugestão:* Use as famosas fórmulas  $\cos 2\alpha = \cos^2 \alpha - \operatorname{sen}^2 \alpha$  e  $\operatorname{sen} 2\alpha = 2 \operatorname{sen} \alpha \cos \alpha$ , multiplique por  $\frac{\cos^2 \alpha}{\cos^2 \alpha}$ , simplifique e depois lembre-se de que  $\cos^2 \alpha = \frac{1}{\sec^2 \alpha} = \frac{1}{1 + \operatorname{tg}^2 \alpha}$ .

**Exercício 15.4** Sirva-se dos resultados dos exercícios anteriores para obter a seguinte fórmula para a reflexão através da reta  $r$  de equação  $y = mx$ :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

**Exercício 15.5** Observe que a translação de  $(0, -p)$  transforma a reta de equação  $y = mx + p$  na reta de equação  $y = mx$ .

**Exercício 15.6** Note que a reflexão através de uma reta qualquer pode ser obtida “trazendo tudo para a origem e depois levando de volta”, a exemplo do que foi feito para rotações. Isto é, podemos começar fazendo uma translação que transforme nossa reta em reta passando pela origem, fazer a reflexão do ponto assim transladado através da nova reta, e depois desfazer a translação.

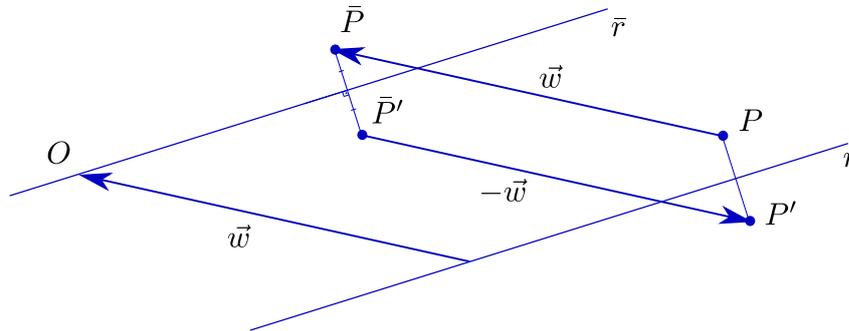


Figura 15.5:

**Exercício 15.7** Sirva-se dos resultados acima para mostrar que se  $(x', y')$  é o refletido de  $(x, y)$  através da reta de equação  $y = mx + p$ , então

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{pmatrix} \begin{pmatrix} x \\ y - p \end{pmatrix} + \begin{pmatrix} 0 \\ p \end{pmatrix}.$$

## 15.3 Animando reflexões

Quando procuramos produzir animações transformando uma figura em outra, obtida por translação, rotação ou homotetia (de razão positiva), pudemos sempre criar, a cada caso, transformações intermediárias (de mesmo tipo) que fossem modificando pouco a pouco a figura inicial até chegar à final. Assim, uma translação pode ser concebida como resultado de uma seqüência de “pequenas” translações, o mesmo ocorrendo com rotações e homotetias.

**Exercício 15.8** Pare e pense nisso.

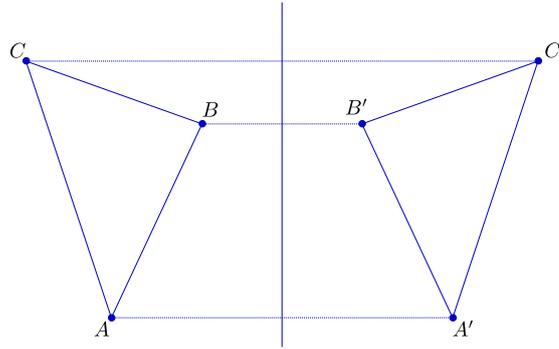


Figura 15.6:

O mesmo não ocorre quando se trata de reflexões. O leitor observará que quando tentamos deslocar uma figura, sem sair do plano, de jeito a transformá-la em sua refletida, não conseguimos fazê-lo guardando sua rigidez. A experiência pode ser feita, sobre a mesa, com qualquer figura plana sem simetrias: não conseguimos “refleti-la” sem tirá-la da mesa. Note que, se a figura for um pedaço de papelão de cores distintas de cada um de seus lados, podemos fazer uma “reflexão” virando-a, o que implica em violar a regra de não sair da mesa; a figura “refletida” terá então cor diferente da original.

Assim, para criarmos as posições intermediárias da figura que desejamos ver refletida, a melhor maneira parece ser movermos cada um de seus pontos sobre o segmento que o une a seu refletido. Vamos tratar disso em um capítulo à parte.

**Exercício 15.9** Note que se  $F$  é uma figura do plano e  $F'$  seu reflexo através da reta  $r$ , se movermos cada ponto  $P$  de  $F$ , a velocidade constante (para cada ponto uma velocidade constante, possivelmente diferente de ponto para ponto) sobre o segmento  $PP'$  que o une a seu reflexo, então no meio do caminho todos os pontos estarão exatamente sobre  $r$ .

**Exercício 15.10** Se  $A = (a, b)$  e  $B = (c, d)$  são pontos do plano, mostre que os pontos do segmento  $AB$  são da forma

$$A + t\vec{AB} = (a, b) + t(c - a, d - b) = (1 - t)(a, b) + t(c, d), \quad t \in [0, 1].$$

# Capítulo 16

## Deformações

As animações de que temos tratado têm quase sempre guardado a rigidez de nossas figuras: estas apenas se deslocam sobre o plano, à exceção do caso das homotetias, em que há também variação de tamanho. É desejável, porém, que possamos criar animações em que a figura original e a final tenham formas distintas.

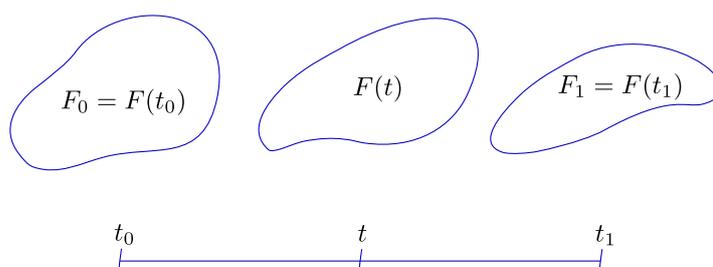


Figura 16.1:

### 16.1 Casos elementares

Chamaremos de **deformação** uma aplicação  $F$  que a cada  $t$  de um intervalo  $[t_0, t_1]$  associa uma figura  $F(t)$ . Diremos que a aplicação  $F$  deforma  $F_0 = F(t_0)$  em  $F_1 = F(t_1)$ .

A idéia básica é extremamente simples: cada ponto  $P_0$  de  $F_0$  deve se transformar em um ponto  $P_1$  de  $F_1$ . A maneira mais natural de conseguir tal efeito é fazer com que nosso ponto caminhe sobre o segmento  $P_0P_1$ , começando em  $P_0$  e terminando em  $P_1$ .

Ora, já aprendemos a fazer isso quando tratamos de translações: se  $P_0 = (x_0, y_0)$  e  $P_1 = (x_1, y_1)$ , fazemos

$$\vec{v} = \overrightarrow{P_0P_1} = (x_1 - x_0, y_1 - y_0)$$

<sup>1</sup>Na realidade, deve-se exigir de  $F$  algum tipo de continuidade, isto é,  $F$  não deve “saltar bruscamente” de uma figura para outra. Para evitar detalhes excessivamente técnicos omitiremos menções explícitas a tal propriedade, que o bom senso deve nos encarregar de observar em cada caso concreto

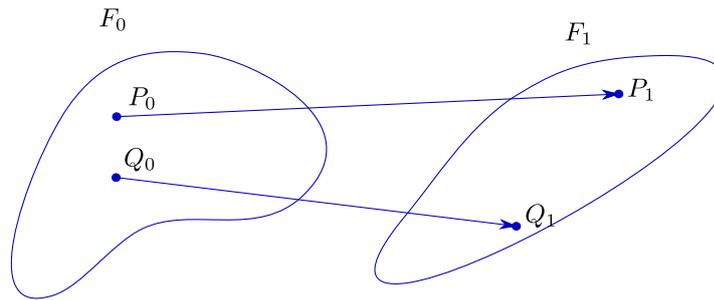


Figura 16.2:

e tomamos, para  $t \in [0, 1]$ ,

$$P(t) = P_0 + t\vec{v} = (x_0, y_0) + t(x_1 - x_0, y_1 - y_0) = ((1 - t)x_0 + tx_1, (1 - t)y_0 + ty_1).$$

**Exercício 16.1** Note que se queremos trabalhar com o intervalo  $[t_0, t_1]$  no lugar de  $[0, 1]$ , basta fazermos

$$P(t) = P_0 + \frac{t - t_0}{t_1 - t_0} \vec{v} = \frac{t_1 - t}{t_1 - t_0} (x_0, y_0) + \frac{t - t_0}{t_1 - t_0} (x_1, y_1).$$

Simple e fácil! A diferença para o caso das translações aparece quando resolvemos aplicar o processo, simultaneamente, a dois ou mais pontos, o que é inevitável em qualquer aplicação séria: se o ponto  $P_0$  se transforma em  $P_1$  ao mesmo tempo em que  $Q_0$  vira  $Q_1$  as translações correspondentes são, em geral, distintas, já que não vamos ter sempre  $P_0\vec{P}_1 = Q_0\vec{Q}_1$ . Cada ponto de nossa figura terá sua própria trajetória, independente das trajetórias dos demais.

Vejam um exemplo simples: deformar um triângulo de vértices  $A, B, C$  em outro, de vértices  $A', B', C'$ . Embora a escolha seja arbitrária (não somos sequer forçados a transformar cada vértice de  $ABC$  em um vértice de  $A'B'C'$ ), é natural levar  $A$  em  $A'$ ,  $B$  em  $B'$  e  $C$  em  $C'$ .

Assim, pelo processo descrito acima, criamos, para cada  $t$  no intervalo desejado, vértices  $A(t), B(t), C(t)$  que nos darão triângulos intermediários  $A(t)B(t)C(t)$ .

**Exercício 16.2** Arbitre coordenadas para  $A, B, C, A', B', C'$  e faça as contas. Se preferir, faça direto no computador.

Outro exemplo simples e ilustrativo consiste em deformar um triângulo  $ABC$  em um quadrilátero  $PQRS$ . Mais uma vez existem infinitas soluções; vejamos uma.

Podemos arbitrar que  $A$  vai em  $P$ ,  $C$  vai em  $S$  e  $B$  vai “se abrir” em  $Q$  e  $R$ . Funciona como se o triângulo  $ABC$  fosse, na verdade, um quadrilátero  $AB_1B_2C$  degenerado,

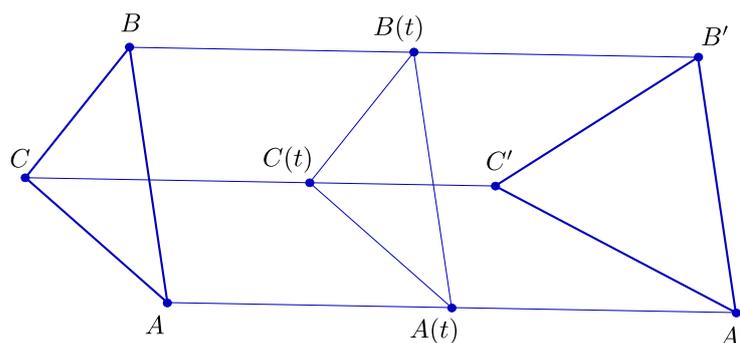


Figura 16.3:

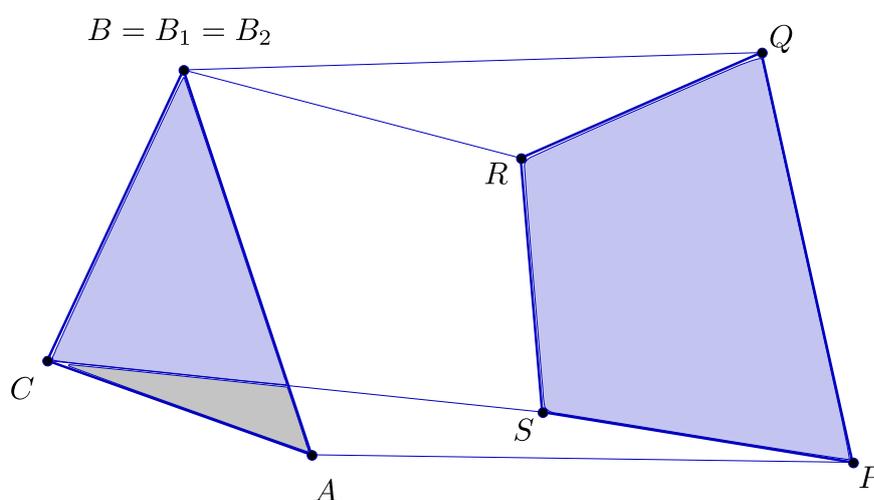


Figura 16.4:

com  $B_1 = B_2 = B$ . Consideramos pois os “quatro” pontos  $A, B_1, B_2, C$  e procedemos normalmente, levando  $A$  em  $P, B_1$  em  $Q, B_2$  em  $R$  e  $C$  em  $S$ .

**Exercício 16.3** *Dê coordenadas numéricas aos pontos e crie a animação correspondente.*

É claro que, em geral, nossos pontos não precisam andar em linha reta, nem a velocidade constante. O leitor pode se divertir criando variadas animações com as ferramentas de que já dispõe.

## 16.2 Deformações em outras dimensões

Vejam uma situação bastante natural: desejamos deformar a imagem dada por uma fotografia,  $F$  em outra,  $F'$ . No computador,  $F$  e  $F'$  serão dois arquivos, cada um com os dados referentes a todos os pontos da tela. Para simplificar, comecemos em preto e branco e digamos que para cada ponto da tela damos dois dados, o primeiro para

o tom de cinza (0 para preto até 1 para branco) e o segundo para o brilho (também expresso por um número).

Se  $N$  é o total de pontos na tela, cada ponto  $P_i$  de  $F$  será caracterizado por duas “coordenadas”, que se organizam no par ordenado  $(x_i, y_i)$ , a primeira indicando o tom de cinza e a segunda o brilho. O mesmo ponto em  $F'$  será dado por  $(x'_i, y'_i)$ . Assim, para deformar  $F$  em  $F'$  basta que se proceda como nos casos mais simples, criando as figuras  $F(t)$ ,  $t \in [0, 1]$ , com os pontos  $P_i(t)$  caracterizados por

$$(x_i(t), y_i(t)) = (1 - t)(x_i, y_i) + t(x_i, y_i).$$

Na prática,  $F$  e  $F'$  são dadas por “vetores” enormes, isto é, com um número de coordenadas igual a  $N$  (o número de pontos da tela) vezes o número de dados necessários para caracterizar cada ponto (que, no caso de imagens a cores, será maior do que dois). Se  $M$  é esse número de coordenadas, podemos dizer que

$$F = (x_1, x_2, x_3, x_4, \dots, x_M),$$

$$F' = (x'_1, x'_2, x'_3, x'_4, \dots, x'_M).$$

Naturalmente escreveremos

$$F(t) = F + t\vec{F}' ,$$

onde

$$\vec{F}' = (x'_1 - x_1, x'_2 - x_2, x'_3 - x_3, x'_4 - x_4, \dots, x'_M - x_M),$$

ou seja,

$$F(t) = ((1 - t)x_1 + tx'_1, (1 - t)x_2 + tx'_2, (1 - t)x_3 + tx'_3, (1 - t)x_4 + tx'_4, \dots, (1 - t)x_M + tx'_M).$$

**Exercício 16.4** Reflita sobre a idéia de que  $F, F'$  e cada  $F(t)$  são pontos de um espaço de  $M$  dimensões.

# Capítulo 17

## Transformações lineares

### 17.1 Definição

Das transformações que estudamos nos capítulos precedentes (translações, rotações, homotetias e reflexões), apenas as translações não são dadas por matrizes. Pudemos também observar que, de novo à exceção das translações, todas gozam de duas propriedades notáveis: se  $T$  designa uma dessas transformações, então

$$\begin{aligned}(i) T(\vec{u} + \vec{v}) &= T\vec{u} + T\vec{v} \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^2; \\(ii) T(t\vec{u}) &= tT\vec{u} \quad \forall \vec{u} \in \mathbb{R}^2, t \in \mathbb{R}.\end{aligned}$$

**Observação:** Para uma transformação com as propriedades (i) e (ii) acima, a notação  $T\vec{u}$  é consagrada e usada de preferência a  $T(\vec{u})$ .

Independente de qualquer outra circunstância, se uma transformação  $T$  goza das propriedades acima, então podemos obter, para  $\vec{u} = (x, y)$ , o valor de  $T\vec{u}$ , a partir de  $T\vec{e}_1$  e  $T\vec{e}_2$ , observando que  $(x, y) = x(1, 0) + y(0, 1) = x\vec{e}_1 + y\vec{e}_2$ :

$$T\vec{u} = T(x, y) = T(x\vec{e}_1 + y\vec{e}_2) = T(x\vec{e}_1) + T(y\vec{e}_2) = xT\vec{e}_1 + yT\vec{e}_2.$$

Assim, se  $T\vec{e}_1 = (a, b)$  e  $T\vec{e}_2 = (c, d)$ , teremos

$$T(x, y) = x(a, b) + y(c, d) = (ax + cy, bx + dy),$$

ou, em notação matricial, fazendo  $T(x, y) = (x', y')$ ,

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = x \begin{pmatrix} a \\ b \end{pmatrix} + y \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ax + cy \\ bx + dy \end{pmatrix} =: \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Uma transformação  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  com as propriedades (i) e (ii) acima é dita uma **transformação linear**

Uma boa forma de visualizar transformações lineares é desenhar à esquerda o sistema de coordenadas canônico (tendo como base os vetores  $\vec{e}_1$  e  $\vec{e}_2$ ) e a direita um outro sistema, tendo como base os vetores  $\vec{v}_1 = T\vec{e}_1$  e  $\vec{v}_2 = T\vec{e}_2$  (supondo que  $T\vec{e}_1$  e  $T\vec{e}_2$  sejam linearmente independentes).

Desenhamos à esquerda o vetor  $\vec{u} = x\vec{e}_1 + y\vec{e}_2$  e à direita sua imagem por  $T$ ,  $T\vec{u} = x\vec{v}_1 + y\vec{v}_2$ . Assim,  $\vec{u}$  e  $T\vec{u}$  têm as mesmas coordenadas, apenas em sistemas diferentes.

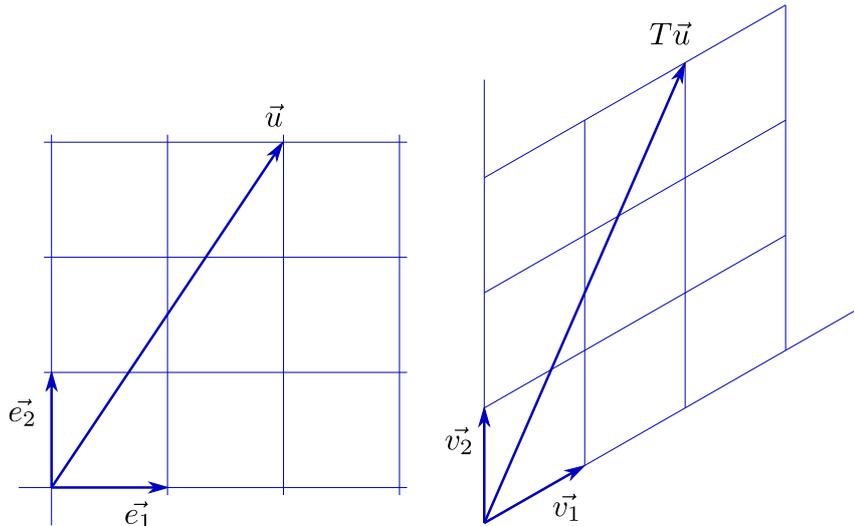


Figura 17.1:

**Exercício 17.1** Seja  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  definida pela matriz  $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ , isto é, se  $T(x,y)=(x',y')$ , então  $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ . Mostre que  $T$  é linear.

**Exercício 17.2** Note que se a transformação linear  $T$  é dada pela matriz

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

então  $a, b, c, d$  são dados por  $(a, b) = T\vec{e}_1$ ,  $(c, d) = T\vec{e}_2$ . Seja  $T'$  a transformação linear dada pela matriz

$$\begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix}.$$

Mostre que a composta  $T'T$  é uma transformação linear e que sua matriz é

$$\begin{pmatrix} a'a + c'b & a'c + c'd \\ b'a + d'b & b'c + d'd \end{pmatrix} =: \begin{pmatrix} a' & c' \\ b' & d' \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix}.^1$$

## 17.2 Transformações preservando distâncias

Os chamados “casos de congruência” nos fornecem a idéia de dar uma olhadinha nas transformações do plano que preservam distâncias (se  $f$  é uma transformação do plano, diz-se que  $f$  **preserva distâncias** se para quaisquer dois pontos  $P$  e  $Q$  a distância entre  $F(P)$  e  $F(Q)$  é igual à de  $P$  a  $Q$ ). Uma transformação que preserva distâncias é também chamada uma **isometria**.

<sup>1</sup>Se o leitor já se deparou com a definição de produto de matrizes e essa lhe pareceu misteriosa, a ideia acima pode servir de motivação

**Teorema:** Se  $f$  é uma isometria do plano, então  $f$  pode ser obtida pela aplicação sucessiva de uma transformação linear, dada por uma reflexão ou uma rotação, e uma translação.

Demonstração: Fixemos no plano um sistema de coordenadas canônico. Se  $O$  é a origem de nosso sistema, seja  $C = f(O)$ .

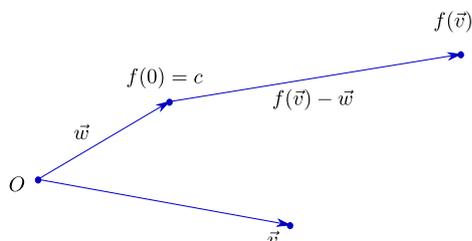


Figura 17.2:

Seja  $\vec{w} = \overrightarrow{OC}$ . Seja  $T_{\vec{w}}$  a translação de  $\vec{w}$ . Se existir um ponto  $A$  do plano tal que  $f(A) = A$ , podemos fixar a origem em  $A$ . Neste caso, temos  $\vec{w} = \vec{0}$  e podemos, claro, dispensar a translação (estamos, como de hábito, identificando pontos e vetores). De qualquer forma, seja  $T$  a transformação do plano dada por

$$T\vec{v} = f(\vec{v}) - \vec{w}.$$

É claro que  $T$  também preserva distâncias. Afiramos que, além disso,  $T$  é uma transformação linear.

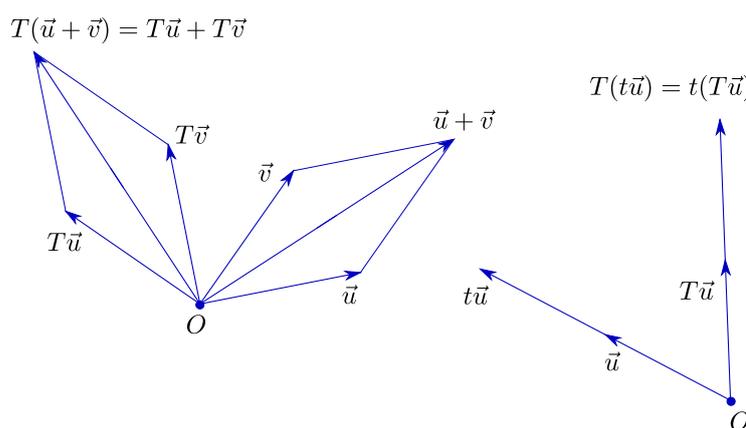


Figura 17.3:

Exercício de Geometria Plana: Olhe as figuras e prove que  $T$  é linear (ou, pelo menos, convença-se de que isso é verdade). Para uma demonstração "vetorial", veja a Proposição ao final do próximo capítulo.

Ora, se  $T$  é linear, vamos determinar  $T$  a partir de  $T\vec{e}_1$  e  $T\vec{e}_2$ . Como  $T$  deve preservar as normas dos vetores (suas distâncias à origem),  $T\vec{e}_1$  será dado por  $T\vec{e}_1 = (\cos\theta, \text{sen}\theta)$  para algum ângulo  $\theta$ .

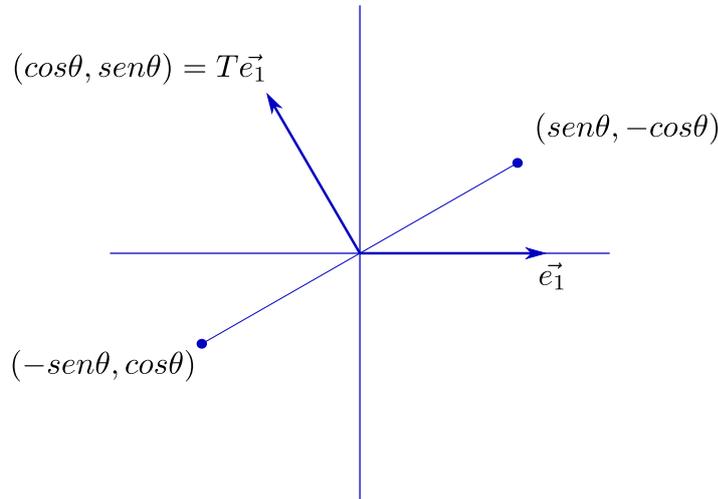


Figura 17.4:

Como  $T\vec{e}_2$  também é unitário e sua distância a  $T\vec{e}_1$  é  $\sqrt{2}$ , só temos duas opções:  $T(\vec{e}_2) = (-\text{sen}\theta, \text{cos}\theta)$  ou  $T\vec{e}_2 = (\text{sen}\theta, -\text{cos}\theta)$ . No primeiro caso  $T$  é uma rotação de  $\theta$ ; no segundo  $T$  é uma reflexão através da reta fazendo ângulo  $\theta/2$  com a horizontal.

Sejam pois  $R$  a rotação de  $\theta$  em torno da origem e  $S$  a reflexão através da reta fazendo ângulo  $\theta/2$  com a horizontal. Acabamos de ver que  $T = R$  ou  $T = S$ . Como  $f(P)$  é obtida aplicando  $T_{\vec{w}}$  a  $T(P)$ , temos que  $f = T_{\vec{w}}R$  ou  $f = T_{\vec{w}}S$ . ■

**Exercício 17.3** *Sejam  $R$  uma rotação de  $\theta$  e  $\vec{w}$  um vetor não nulo. Suponha que  $\theta \neq 0$ . Mostre que existe um ponto  $P$  do plano tal que  $R(P) = P - \vec{w}$ . Conclua que o caso  $f = T_{\vec{w}}R$  (rotação seguida de translação) se desdobra em dois: se o ângulo de rotação é nulo, temos uma translação pura; se, ao contrário, o ângulo de rotação é não nulo, então  $f$  tem um ponto fixo e a translação pode ser dispensada. Conclua que toda isometria do plano é de um dos seguintes tipos: translação, rotação, reflexão ou reflexão seguida de translação.*

# Capítulo 18

## Produto interno

As duas noções métricas que temos manejado são as distâncias e os ângulos. Se para as distâncias pudemos facilmente dar uma versão em termos de coordenadas, o mesmo não foi feito para os ângulos. De fato, se os pontos  $P$  e  $Q$  têm por coordenadas, respectivamente,  $(a, b)$  e  $(c, d)$ , então sua distância é dada por  $\sqrt{(c-a)^2 + (d-b)^2}$

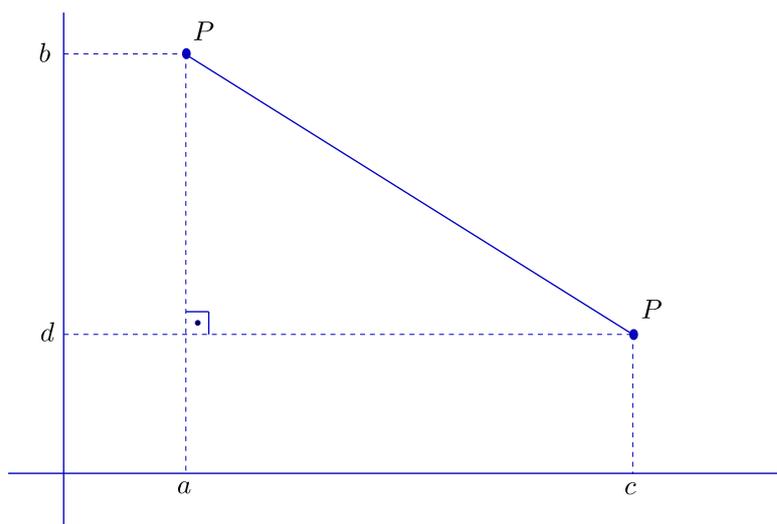


Figura 18.1:

Mas e quanto ao ângulo entre os vetores  $\vec{OP}$  e  $\vec{OQ}$ ? Na realidade, medir ângulos é bem mais difícil do que medir distâncias.<sup>1</sup> Por outro lado, um ângulo é caracterizado por seus seno e cosseno, cujas definições dependem apenas de distâncias. É por aí que podemos atacar a questão.

Sejam  $\vec{u}$  e  $\vec{v}$  dois vetores do plano e seja  $\theta$  o ângulo (o menor dos dois, digamos) entre

---

<sup>1</sup>Dado um arco  $a$  de um círculo de raio  $r$ , o **ângulo** correspondente a  $a$  é definido como a razão entre o comprimento de  $a$  e  $r$ . Note que essa definição deixa claro que ângulo é, como dizem os físicos, adimensional: um ângulo é dado por um número, sem qualquer unidade de medida (podemos dizer que radianos são uma espécie de “faz de conta”)

$\vec{u}$  e  $\vec{v}$ . A **norma** de  $\vec{u}$  e a **norma** de  $\vec{v}$  são dadas por:

$$|\vec{u}| = \sqrt{x_1^2 + y_1^2}, \quad |\vec{v}| = \sqrt{x_2^2 + y_2^2},$$

onde  $\vec{u} = (x_1, y_1)$  e  $\vec{v} = (x_2, y_2)$  (na base canônica).

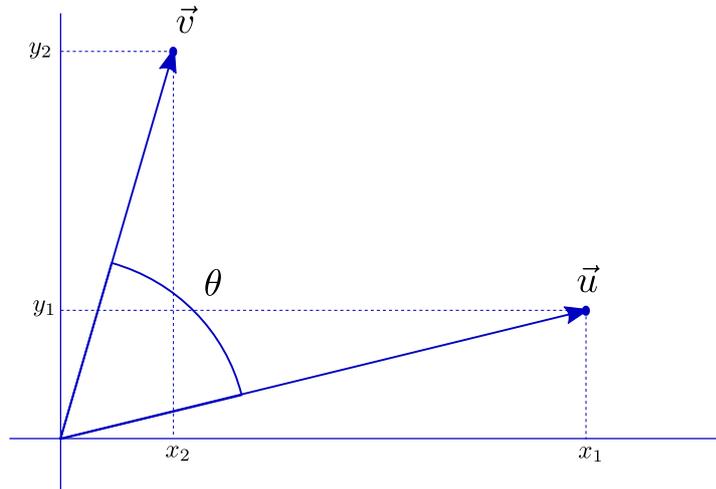


Figura 18.2:

Nosso problema é determinar  $\theta$ . Vamos definir a **projeção escalar** de  $\vec{u}$  na direção de  $\vec{v}$ , que, por preguiça, chamaremos apenas de **projeção** de  $\vec{u}$  na direção de  $\vec{v}$ , por

$$p_{\vec{v}}(\vec{u}) = |\vec{u}| \cos\theta$$

( $\cos\theta$  tem um sinal, que podemos manter para indicar se a projeção cai no sentido de  $\vec{v}$ , caso  $\cos\theta \geq 0$ , ou no sentido contrário). Da mesma forma, a projeção de  $\vec{v}$  na direção de  $\vec{u}$  é dada por

$$p_{\vec{u}}(\vec{v}) = |\vec{v}| \cos\theta.$$

Note que o fato de estarmos trabalhando com projeções positivas e negativas cai bem, pois assim temos:

$$\begin{aligned} (i) p_{\vec{v}}(\vec{u}_1 + \vec{u}_2) &= p_{\vec{v}}(\vec{u}_1) + p_{\vec{v}}(\vec{u}_2) \quad \forall \vec{u}_1, \vec{u}_2 \in \mathbb{R}^2, \\ (ii) p_{\vec{v}}(t\vec{u}) &= t p_{\vec{v}}(\vec{u}) \quad \forall \vec{u} \in \mathbb{R}^2, t \in \mathbb{R}, \end{aligned}$$

as mesmas propriedades valendo para  $p_{\vec{u}}$ .

**Exercício 18.1** *Assegure-se de que entende e concorda com o acima exposto.*

Vamos agora introduzir uma pequena novidade para facilitar a vida mais à frente. Definimos o **produto interno** (também chamado **produto escalar**) de  $\vec{u}$  e  $\vec{v}$  como o número  $\vec{u} \cdot \vec{v}$  (também notado  $\langle \vec{u}, \vec{v} \rangle$ ), dado por

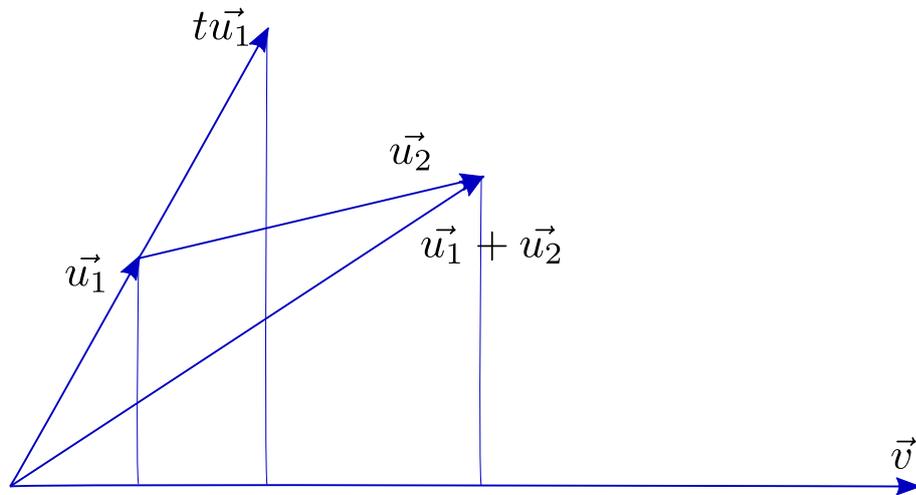


Figura 18.3:

$$\vec{u} \cdot \vec{v} = |\vec{u}| |\vec{v}| \cos \theta.$$

Da definição temos

$$\vec{u} \cdot \vec{v} = |\vec{v}| p_{\vec{v}}(\vec{u}) = |\vec{u}| p_{\vec{u}}(\vec{v}),$$

de onde podemos deduzir que:

$$\begin{aligned} (i) \quad & \langle \vec{u}_1 + \vec{u}_2, \vec{v} \rangle = \langle \vec{u}_1, \vec{v} \rangle + \langle \vec{u}_2, \vec{v} \rangle \quad \forall \vec{u}_1, \vec{u}_2, \vec{v} \in \mathbb{R}^2; \\ (ii) \quad & \langle \vec{u}, \vec{v}_1 + \vec{v}_2 \rangle = \langle \vec{u}, \vec{v}_1 \rangle + \langle \vec{u}, \vec{v}_2 \rangle \quad \forall \vec{u}, \vec{v}_1, \vec{v}_2 \in \mathbb{R}^2; \\ (iii) \quad & \langle t\vec{u}, \vec{v} \rangle = t \langle \vec{v}, \vec{u} \rangle = \langle \vec{u}, t\vec{v} \rangle \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^2, \forall t \in \mathbb{R}. \end{aligned}$$

É claro também que

$$(iv) \quad \langle \vec{u}, \vec{v} \rangle = \langle \vec{v}, \vec{u} \rangle \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^2.$$

**Exercício 18.2** Observe que  $\langle \vec{u}, \vec{u} \rangle = |\vec{u}|^2$  para todo  $\vec{u}$  em  $\mathbb{R}^2$ . Em particular,  $\langle \vec{e}_1, \vec{e}_1 \rangle = \langle \vec{e}_2, \vec{e}_2 \rangle = 1$ . Note ainda que  $\langle \vec{e}_1, \vec{e}_2 \rangle = 0$ .

Das propriedades acima podemos obter a expressão de  $\vec{u} \cdot \vec{v}$  em termos das coordenadas de  $\vec{u}$  e  $\vec{v}$ :

$$\vec{u} \cdot \vec{v} = x_1 x_2 + y_1 y_2.$$

De fato, escrevendo  $\vec{u} = (x_1, y_1) = x_1 \vec{e}_1 + y_1 \vec{e}_2$ ,  $\vec{v} = (x_2, y_2) = x_2 \vec{e}_1 + y_2 \vec{e}_2$ , temos:

$$\begin{aligned}\langle \vec{u}, \vec{v} \rangle &= \langle x_1 \vec{e}_1 + y_1 \vec{e}_2, x_2 \vec{e}_1 + y_2 \vec{e}_2 \rangle = \\ &= x_1 x_2 \langle \vec{e}_1, \vec{e}_1 \rangle + x_1 y_2 \langle \vec{e}_1, \vec{e}_2 \rangle + y_1 x_2 \langle \vec{e}_2, \vec{e}_1 \rangle + y_1 y_2 \langle \vec{e}_2, \vec{e}_2 \rangle. \quad ^2\end{aligned}$$

Como  $\langle \vec{e}_1, \vec{e}_2 \rangle = \langle \vec{e}_2, \vec{e}_1 \rangle = 0$  e  $\langle \vec{e}_1, \vec{e}_1 \rangle = \langle \vec{e}_2, \vec{e}_2 \rangle = 1$ , temos o que afirmamos.

A obtenção da fórmula acima nos fornece meios para tratar algebricamente as questões referentes a ângulos. Se desejamos conhecer o ângulo  $\theta$  entre os vetores não nulos  $\vec{u}=(x_1, y_1)$  e  $\vec{v}=(x_2, y_2)$ , temos a expressão

$$\cos\theta = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{u}, \vec{u} \rangle^{\frac{1}{2}} \langle \vec{v}, \vec{v} \rangle^{\frac{1}{2}}} = \frac{x_1 x_2 + y_1 y_2}{(x_1^2 + y_1^2)^{\frac{1}{2}} (x_2^2 + y_2^2)^{\frac{1}{2}}}.$$

Não custa nada usarmos o produto interno para uma outra demonstração da fórmula

$$\cos(\alpha + \beta) = \cos\alpha \cos\beta - \operatorname{sen}\alpha \operatorname{sen}\beta.$$

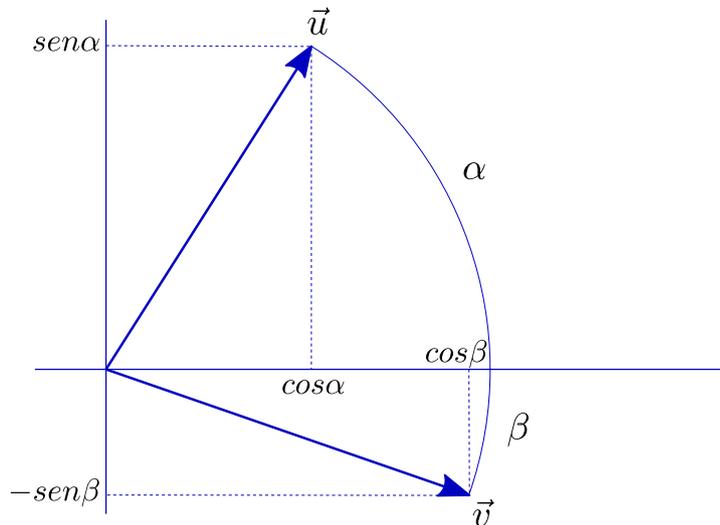


Figura 18.4:

Considerando os vetores  $\vec{u}=(\cos\alpha, \operatorname{sen}\alpha)$  e  $\vec{v}=(\cos\beta, -\operatorname{sen}\beta)$ , temos que o ângulo entre  $\vec{u}$  e  $\vec{v}$  é  $(\alpha+\beta)$  e portanto, da definição de produto interno,

$$\cos(\alpha + \beta) = \vec{u} \cdot \vec{v}$$

(note que  $|\vec{u}| = |\vec{v}| = 1$ ). Por outro lado, a expressão do produto interno em termos de coordenadas nos dá

$$\vec{u} \cdot \vec{v} = \cos\alpha \cos\beta - \operatorname{sen}\alpha \operatorname{sen}\beta,$$

o que prova a fórmula.

<sup>2</sup>Note que nesta passagem usamos as propriedades (i), (ii) e (iii)

**Exercício 18.3** Mostre, sem recorrer a argumentos geométricos, mas usando as propriedades do produto interno, que toda transformação  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  que fixa a origem (isto é:  $f(0,0) = (0,0)$ ) e preserva distâncias (isto é:  $|f(\vec{u}) - f(\vec{v})| = |\vec{u} - \vec{v}| \forall \vec{u}, \vec{v} \in \mathbb{R}^2$ ) é linear. Se não conseguir, veja a proposição abaixo.

**Proposição:** Se  $f$  é uma isometria do plano e, além disso,  $f(\mathbf{0}) = \mathbf{0}$ , então  $f$  é uma transformação linear.

**Demonstração:** Começemos observando que, como  $f(\mathbf{0}) = \mathbf{0}$ ,  $f$  preserva norma (já que  $|f(\vec{u})| = |f(\vec{u}) - f(\mathbf{0})|$ ). Daí vem

$$\begin{aligned} |f(\vec{u})|^2 + |f(\vec{v})|^2 - 2 \langle f(\vec{u}), f(\vec{v}) \rangle &= |f(\vec{u}) - f(\vec{v})|^2 = \\ &= |\vec{u} - \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2 - 2 \langle \vec{u}, \vec{v} \rangle. \end{aligned}$$

Logo, temos  $\langle f(\vec{u}), f(\vec{v}) \rangle = \langle \vec{u}, \vec{v} \rangle$  (entenda isso como triângulos de lados congruentes têm ângulos correspondentes também congruentes, o que implica em dizer que a preservação de distâncias nos dá a preservação de ângulos). Daí vem:

$$\begin{aligned} |f(\vec{u} + t\vec{v}) - (f(\vec{u}) + tf(\vec{v}))|^2 &= \\ &= \langle f(\vec{u} + t\vec{v}) - (f(\vec{u}) + tf(\vec{v})), f(\vec{u} + t\vec{v}) - (f(\vec{u}) + tf(\vec{v})) \rangle = \\ &= |f(\vec{u} + t\vec{v}) - f(\vec{u})|^2 + t^2 |f(\vec{v})|^2 - 2t \langle f(\vec{u} + t\vec{v}) - f(\vec{u}), f(\vec{v}) \rangle = \\ &= |(\vec{u} + t\vec{v}) - \vec{u}|^2 + t^2 |\vec{v}|^2 - 2t \langle (\vec{u} + t\vec{v}) - \vec{u}, \vec{v} \rangle = 0, \end{aligned}$$

o que prova que  $f(\vec{u} + t\vec{v}) = f(\vec{u}) + tf(\vec{v})$ . ■

**Exercício 18.4** Sejam  $\vec{u}$  e  $\vec{v}$  dois vetores distintos. Considere o círculo  $c$  de diâmetro  $|\vec{v} - \vec{u}|$ , passando por  $\vec{u}$  e por  $\vec{v}$ . Note que o centro de  $c$  é  $\frac{1}{2}(\vec{u} + \vec{v})$ . Mostre que  $\vec{w}$  está em  $c$  se, e somente se,

$$\langle \vec{w} - \vec{u}, \vec{w} - \vec{v} \rangle = 0.$$



# Capítulo 19

## A Decomposição em Valores Singulares

### 19.1 O Teorema

Esta seção começa com um vídeo (clique e assista!):

 [Decomposição em Valores Singulares](#)

O vídeo nos mostra a veracidade do seguinte resultado, que é a versão mais simples da Decomposição em Valores Singulares:<sup>1</sup>

**Teorema 19.1** *Se  $T$  é uma transformação linear bijetiva do plano no plano, então existem: dois vetores unitários e ortogonais,  $\vec{v}_1$  e  $\vec{v}_2$ ; outros dois vetores unitários e ortogonais,  $\vec{u}_1$  e  $\vec{u}_2$ ; e dois números estritamente positivos,  $\lambda_1$  e  $\lambda_2$ , tais que*

$$T\vec{v}_1 = \lambda_1\vec{u}_1, \quad T\vec{v}_2 = \lambda_2\vec{u}_2.$$

Se quiser uma outra versão em vídeo, veja

 [Decomposição em Valores Singulares - demonstração](#)

### 19.2 A geometria do Teorema

Observe que, no enunciado do Teorema, fizemos, deliberadamente uma escolha inabitual: poderíamos ter escrito  $T$  é uma transformação linear bijetiva de  $\mathbb{R}^2$  em  $\mathbb{R}^2$ , mas o que escrevemos foi  $T$  é uma transformação linear bijetiva do plano no plano. A razão é a seguinte:  $\mathbb{R}^2$  já vem, de fábrica, com um sistema de coordenadas; o plano, não. O plano vem purinho (mas vamos supor que já veio com uma origem). Se temos  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , já conhecemos a história e a figura:  $T$  é dada pela matriz

---

<sup>1</sup>Este é um resultado central em Álgebra Linear. Voltaremos a ele em um contexto mais geral

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

de modo que o vetor  $\vec{w} = x_1\vec{e}_1 + x_2\vec{e}_2$  é transformado por  $T$ , por meio de

$$T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix},$$

em  $T\vec{w} = x_1\vec{e}_1 + x_2\vec{e}_2$ , com  $\vec{e}_1 = T\vec{e}_1$ ,  $\vec{e}_2 = T\vec{e}_2$ . Note que não podemos esperar, em geral, que os vetores  $\vec{e}_1$  e  $\vec{e}_2$  sejam ortogonais (seria muita sorte!).

FIGURA

O que o Teorema nos diz é que existe um outro sistema de coordenadas, com uma certa base especial,  $\vec{v}_1, \vec{v}_2$  (mantida a origem), tal que  $T\vec{v}_1$  e  $T\vec{v}_2$  são ortogonais (e, para ficar mais certinho, usamos, para as imagens, a base  $\vec{u}_1, \vec{u}_2$ , formada pelos unitários de  $T\vec{v}_1$  e  $T\vec{v}_2$ ).

Assim, se escrevermos  $\vec{w} = y_1\vec{v}_1 + y_2\vec{v}_2$ , teremos

$$T\vec{w} = Ty_1\vec{v}_1 + y_2T\vec{v}_2 = y_1T\vec{v}_1 + y_2T\vec{v}_2 = y_1\lambda_1\vec{u}_1 + y_2\lambda_2\vec{u}_2.$$

Isso significa que, qualquer que seja a transformação linear  $T$ , do plano no plano, se escolhermos adequadamente os sistemas de coordenadas em que vamos representar (ou, deveríamos dizer, de que maneira vamos inclinar a cabeça para olhar) as figuras, de um lado, e suas imagens por  $T$ , do outro lado, seu efeito corresponde a duas simples mudanças de escala.

FIGURA

### 19.3 Agora com matrizes

Suponhamos, que, na base canônica, a matriz de  $T$ ,  $A$ , e os vetores  $\vec{v}_1, \vec{v}_2, \vec{u}_1, \vec{u}_2$  são dados por

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \vec{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix}, \vec{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix}.$$

Como

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} = \lambda_1 \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix}, \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = \lambda_2 \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix},$$

temos (confira)

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}. \quad (*)$$

Como aprendemos lá atrás, o truque de Fourier (página ??) nos ensina que, como a base  $\alpha = \{\vec{v}_1, \vec{v}_2\}$  é ortonormal, então

$$\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Multiplicando, à direita, (\*) por  $\begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}$ , obtemos

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}.$$

Podemos, então, reenunciar nosso teorema, na versão matricial:

**Teorema 19.2 (Decomposição em Valores Singulares)** Se  $A$  é matriz  $2 \times 2$ , então existem matrizes  $U$ ,  $D$  e  $V$ , tais que:

1.  $D$  é **matriz diagonal** (isto é: só tem termos não nulos nas posições  $a_{11}$  e  $a_{22}$ ), com valores estritamente positivos na diagonal;
2.  $U$  e  $V$  são unitárias;
3.  $A = UDV^T$ .

## 19.4 A SVD em termos de mudanças de base

Se você já compreende que as matrizes  $U$  e  $V^T$ , que aparecem no enunciado do Teorema, representam mudanças de base ( $U$  muda de  $\beta = \{\vec{u}_1, \vec{u}_2\}$ , para a canônica;  $V^T$  muda da canônica para  $\alpha = \{\vec{v}_1, \vec{v}_2\}$ ), então talvez seja melhor pular esta seção e a próxima.

Recapitulando: estamos no nosso velho e bom  $\mathbb{R}^2$  e nossa transformação linear  $T$  chegou até nós, como de costume, na forma de u'a matriz  $2 \times 2$ . Assim, se  $\vec{w} = (x_1, x_2) = x_1\vec{e}_1 + x_2\vec{e}_2$ , temos  $T\vec{w} = (y_1, y_2)$ , com

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (1).$$

Conhecemos, também, as bases,  $\alpha = \{\vec{v}_1, \vec{v}_2\}$  e  $\beta = \{\vec{u}_1, \vec{u}_2\}$ , e os números positivos,  $\lambda_1$  e  $\lambda_2$ , referidos na decomposição em valores singulares. Vamos fazer uma coisa meio óbvia: se  $w$  é dado, na base  $\alpha$  (formada por  $\vec{v}_1$  e  $\vec{v}_2$ ) por  $w = s_1\vec{v}_1 + s_2\vec{v}_2$ , então  $Tw$  é dado, na base  $\beta$  (formada por  $\vec{u}_1$  e  $\vec{u}_2$ ), por  $Tw = t_1\vec{u}_1 + t_2\vec{u}_2 = \lambda_1 s_1\vec{u}_1 + \lambda_2 s_2\vec{u}_2$ , de modo que

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad (2).$$

Como vemos, a fórmula (1) engole as coordenadas  $(x_1, x_2)$  de  $\vec{w}$  na base canônica e cospe as de  $T\vec{w}$ ,  $(y_1, y_2)$ , na base canônica.

Já a fórmula (2), engole as coordenadas  $(s_1, s_2)$  de  $\vec{w}$  na base  $\alpha$  e cospe as de  $T\vec{w}$ ,  $(t_1, t_2)$ , na base  $\beta$ .

Deve, é claro, haver um jeito de explicitar a relação entre as duas fórmulas. Vamos procurá-lo. Procuremos tornar efetiva a fórmula (2):

(i) Como o vetor  $\vec{w}$  nos chega por suas coordenadas,  $(x_1, x_2)$ , na base canônica, como obter, a partir destas, suas coordenadas,  $(s_1, s_2)$ , na base  $\alpha$ , para que possamos inseri-las na fórmula (2)?

(ii) Posto que a fórmula (2) nos fornece as coordenadas,  $(t_1, t_2)$ , de  $T\vec{w}$  na base  $\beta$ , como obter, a partir destas, suas coordenadas na base canônica (que são as que nos fornece a fórmula (1))?

Suponhamos conhecidas (o que parece razoável, nas circunstâncias) as coordenadas de  $\vec{v}_1, \vec{v}_2, \vec{u}_1$  e  $\vec{u}_2$  na base canônica:

$$\vec{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix}, \vec{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix}.$$

A questão (ii) é bem simples. Basta escrevermos  $\vec{u}_1$  e  $\vec{u}_2$  na base canônica:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = T\vec{w} = t_1\vec{u}_1 + t_2\vec{u}_2 = t_1 \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix} + t_2 \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.$$

Ou seja, para passar das coordenadas  $(t_1, t_2)$  de *qualquer* vetor na base  $\beta$  para suas coordenadas na base canônica, basta multiplicar o vetor coluna  $\begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$  pela matriz que tem, nas colunas, as coordenadas de  $\vec{u}_1$  e  $\vec{u}_2$ . Observe que esse procedimento independe do fato de serem  $\vec{u}_1$  e  $\vec{u}_2$  unitários e ortogonais.

A questão (ii) também não é difícil, embora exija uma pequena manipulação com o produto escalar (a ortonormalidade das nossas bases, afinal, tem que servir para algo!). Se temos  $\vec{w} = (x_1, x_2)$  (na base canônica) e queremos obter  $s_1$  e  $s_2$  tais que

$$\vec{w} = s_1\vec{v}_1 + s_2\vec{v}_2,$$

poderíamos resolver o sistema linear que surge, se substituirmos  $\vec{v}_1$  e  $\vec{v}_2$  por suas coordenadas na base canônica (que conhecemos). Mas há um truque, descoberto por Fourier<sup>2</sup>: multiplicamos escalarmente por  $\vec{v}_1$  e por  $\vec{v}_2$  os dois lados. Obtemos, multiplicando por  $\vec{v}_1$ :

<sup>2</sup>Fourier, J. B. J., *Théorie analytique de la chaleur*, Didot, Paris, 1822, pode-se dizer, inventou o produto escalar em um contexto absolutamente fantástico para a época: o problema era propagação de calor e os espaços eram de dimensão infinita

$$\langle \vec{w}, \vec{v}_1 \rangle = \langle s_1 \vec{v}_1 + s_2 \vec{v}_2, \vec{v}_1 \rangle = s_1 \langle \vec{v}_1, \vec{v}_1 \rangle + s_2 \langle \vec{v}_2, \vec{v}_1 \rangle.$$

Como  $\vec{v}_1$  e  $\vec{v}_2$  são unitários e ortogonais, obtemos  $s_1 = \langle \vec{w}, \vec{v}_1 \rangle$ . Cálculo análogo conduz a  $s_2 = \langle \vec{w}, \vec{v}_2 \rangle$ . Assim, usando as coordenadas de  $\vec{w}$ ,  $\vec{v}_1$  e  $\vec{v}_2$ , chegamos a

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} v_{11}x_1 + v_{21}x_2 \\ v_{12}x_1 + v_{22}x_2 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Ou seja, para passar das coordenadas  $(x_1, x_2)$  de *qualquer* vetor na base canônica para suas coordenadas na base  $\beta$ , basta multiplicar o vetor coluna  $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$  pela matriz que tem, nas linhas (horizontais), as coordenadas de  $\vec{v}_1$  e  $\vec{v}_2$ . Note que este procedimento só é possível por conta de serem,  $\vec{v}_1$  e  $\vec{v}_2$ , unitários e ortogonais.

Temos, agora, quatro fórmulas. Recapitulando a situação, temos:

1. uma transformação linear  $T$ , dada pela matriz  $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ ;
2. vetores  $\vec{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix}$ ,  $\vec{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix}$ ,  $\vec{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix}$  e  $\vec{u}_2 = \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix}$ , todos unitários, com  $\vec{v}_1$  ortogonal a  $\vec{v}_2$  e  $\vec{u}_1$  ortogonal a  $\vec{u}_2$ ;  $\vec{v}_1$  e  $\vec{v}_2$  formam a base  $\alpha$ ,  $\vec{u}_1$  e  $\vec{u}_2$  formam a base  $\beta$ ;
3. escalares (estritamente positivos)  $\lambda_1$  e  $\lambda_2$  tais que  $T\vec{v}_1 = \lambda_1\vec{u}_1$  e  $T\vec{v}_2 = \lambda_2\vec{u}_2$ ;

Agora, nossas fórmulas

1. Usando a base canônica, fazemos  $\vec{w} = x_1\vec{e}_1 + x_2\vec{e}_2$  e  $T\vec{w} = y_1\vec{e}_1 + y_2\vec{e}_2$ , com:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (1).$$

2. Usando a base  $\alpha$  para  $\vec{w}$  e a base  $\beta$  para  $T\vec{w}$ , fazemos  $\vec{w} = s_1\vec{v}_1 + s_2\vec{v}_2$  e  $T\vec{w} = t_1\vec{u}_1 + t_2\vec{u}_2$ , com:

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad (2).$$

3. Para transformar as coordenadas de  $\vec{w}$  na base canônica nas do mesmo  $\vec{w}$  na base  $\alpha$ , fazemos:

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (3).$$

4. Para transformar as coordenadas de  $T\vec{w}$  na base  $\beta$  nas do mesmo  $t\vec{w}$  na base canônica, fazemos:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \quad (4).$$

Usando, sucessivamente, (4), (2) e (3), vem:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Ora, isto quer dizer que a matriz que corresponde à transformação linear  $T$  (na base canônica) é

$$\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}.$$

Como a fórmula (1) nos diz, exatamente, quem é essa matriz, temos

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}.$$

Vamos, na próxima seção explicitar o trabalho de mudanças de base que acabamos de fazer.

## 19.5 Mudanças de base

Pelo que acabamos de ver, parece útil destacar os procedimentos envolvidos em mudanças de base de e para a base canônica. Se é fato que nossos vetores vêm, de fábrica, em coordenadas na base canônica, também é verdade que, dependendo da situação, pode ser mais conveniente representá-los em outra base. Por outro lado, mesmo se o problema nos exigir o uso de outra base que não a canônica, é razoável que, ao fim das contas, queiramos expressar nossos resultados em coordenadas canônicas. Vamos sempre supor ( parece razoável) que os vetores de qualquer outra base nos são dados a conhecer por suas coordenadas na base canônica.

Começemos pelo mais fácil:

**(1)  $\vec{w}$  está representado na base  $\beta = \{\vec{u}_1, \vec{u}_2\}$  e queremos suas coordenadas na base canônica**

Esta, realmente, é mole. Se  $\vec{w} = t_1\vec{u}_1 + t_2\vec{u}_2$  e  $\vec{u}_1$  e  $\vec{u}_2$  são dados, na base canônica, respectivamente, por

$$\vec{u}_1 = \begin{bmatrix} u_{11} \\ u_{21} \end{bmatrix}, \quad \vec{u}_2 = \begin{bmatrix} u_{12} \\ u_{22} \end{bmatrix},$$

então basta escrever  $\vec{w} = (y_1, y_2)$ , como acabamos de ver,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}.$$

(2)  $\vec{w}$  está representado na base canônica e queremos suas coordenadas na base  $\alpha = \{\vec{v}_1, \vec{v}_2\}$

Pelo que acabamos de ver, se  $\vec{w} = x_1\vec{e}_1 + x_2\vec{e}_2$  e  $\vec{w} = s_1\vec{v}_1 + s_2\vec{v}_2$ , com

$$\vec{v}_1 = \begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix},$$

então

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}.$$

Como temos tudo, menos  $s_1$  e  $s_2$ , trata-se, na verdade, de resolver um sistema linear, ou, de forma mais pernóstica, encontrar a matriz inversa de

$$V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}.$$

Nada de mais, é só resolver um sisteminha  $2 \times 2$ . Mas, não se esqueça, estamos apenas fazendo um aquecimento para enfrentar problemas com um número grande de variáveis, em que sistemas de  $n$  equações e  $n$  incógnitas, resolvidos de forma inocente, custam algo da ordem de  $n^3/3$  operações de ponto flutuante. Se  $n$  é grande, isso pode ser um grande problema. Aqui entra uma das grandes vantagens de trabalhar com bases ortonormais: o truque de Fourier nos permite concluir, diretamente, que

$$V^{-1} = V^T = \begin{bmatrix} v_{11} & v_{21} \\ v_{12} & v_{22} \end{bmatrix}.$$

## 19.6 De novo a Decomposição em Valores Singulares, versão matricial

Voltemos ao

**Teorema 19.3** *Se  $T$  é uma transformação linear bijetiva do plano no plano, então existem: dois vetores unitários e ortogonais,  $\vec{v}_1$  e  $\vec{v}_2$ ; outros dois vetores unitários e ortogonais,  $\vec{u}_1$  e  $\vec{u}_2$ ; e dois números estritamente positivos,  $\lambda_1$  e  $\lambda_2$ , tais que*

$$T\vec{v}_1 = \lambda_1\vec{u}_1, \quad T\vec{v}_2 = \lambda_2\vec{u}_2.$$

Consideremos a matriz de  $T$  (na base canônica):

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

ou seja se  $\vec{w} = (x_1, x_2)$  e  $T\vec{w} = (y_1, y_2)$ , então

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

O que o Teorema nos diz é que, como  $T\vec{v}_1 = \lambda_1\vec{u}_1$  e  $T\vec{v}_2 = \lambda_2\vec{u}_2$ , então, se  $\vec{w} = (x_1, x_2)$  for representado, na base  $\alpha = \{\vec{v}_1, \vec{v}_2\}$ , por  $(s_1, s_2)$ , e  $T\vec{w} = (y_1, y_2)$  for representado, na base  $\beta = \{\vec{u}_1, \vec{u}_2\}$ , por  $(t_1, t_2)$ , então

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}.$$

Ora, o que acabamos de aprender sobre mudanças de base nos diz que as coordenadas  $(y_1, y_2)$  de  $T\vec{w}$  na base canônica são dadas por

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix},$$

de modo que

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}.$$

Também aprendemos que as coordenadas  $s_1, s_2$  de  $\vec{w}$ , na base  $\alpha = \{\vec{v}_1, \vec{v}_2\}$ , são dadas por

$$\begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

de modo que

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

# Capítulo 20

## Otimização

Os amigos Xavier, Yakecan e Zumbi, também conhecidos com  $X$ ,  $Y$  e  $Z$ , serão personagens da história a seguir.

### 20.1 Amendoim torradinho

$X$ ,  $Y$  e  $Z$  vendem amendoim torrado em sociedade. Certo dia,  $X$  vai vender amendoim em Madureira e  $Y$ , no baixo Gávea (Madureira e Gávea são bairros do Rio de Janeiro).  $Z$ , que gerencia a sociedade, ficou, na véspera, pensando quanto amendoim cada um dos dois deveria levar para seu ponto de venda. Chamando de  $x$  a quantidade que caberá a  $X$ , e  $y$  a que caberá a  $Y$ , as limitações são as seguintes:

(i)  $x + y \leq Q$ , sendo  $Q$  a quantidade total de amendoim em estoque (em unidade de massa, convencionemos);

(ii)  $x \leq Q_1$ , sendo  $Q_1$  o máximo de amendoim que  $X$  consegue carregar (em unidade de massa);

(iii)  $y \leq Q_2$ , sendo  $Q_2$  o máximo de amendoim que  $Y$  consegue carregar (em unidade de massa);

(iv)  $p_1x + p_2y \leq P$ , sendo  $p_1$  a quantidade de papel por unidade de massa necessária para embalar amendoim para venda em Madureira e  $p_2$  a necessária para venda na Gávea (por razões de mercado, as embalagens usadas nos dois bairros são diferentes;  $P$  é a quantidade de papel para embalagem em estoque);

(v)  $t_1x + t_2y \leq T$ , sendo  $t_1$  o tempo que Zumbi leva para embalar cada unidade de massa para venda em Madureira e  $t_2$  o tempo que leva para embalar cada unidade de massa para venda na Gávea (Zumbi embala tudo sozinho e só dispõe de um tempo  $T$  para concluir esse serviço).

Pela experiência do grupo, ambos os vendedores venderão, em qualquer circunstância, toda a mercadoria que levarem. O objetivo é determinar  $x$  e  $y$  não negativos, satisfazendo as desigualdades (i), (ii), (iii), (iv) e (v), de tal maneira que o lucro seja máximo. Como o lucro por unidade vendida em Madureira é  $L_1$ , e o lucro por unidade vendida na Gávea é  $L_2$ , Zumbi tem que determinar  $x$  e  $y$  tais que

$$L(x, y) = L_1x + L_2y$$

seja máximo.<sup>1</sup>

**Exercício 20.1** Desenhe a região  $R$  do plano definida pelas condições (i), (ii), (iii), (iv) e (v), além de  $x \geq 0$  e  $y \geq 0$ . Faça hipóteses razoáveis, tipo  $Q_1 < Q$  e  $Q_2 < Q$ . Para dar mais graça, suponha  $p_1Q < P$  e  $p_2Q_2 < P < p_2Q$ . Mostre que a região  $R$ , com essas hipóteses, é um polígono convexo. Mostre que mesmo sem essas hipóteses adicionais  $R$  é um polígono convexo.

**Exercício 20.2** Observe que a função  $L(x, y)$  é constante em cada reta perpendicular ao vetor  $(L_1, L_2)$ . Observe que  $L$  aumenta se caminhamos na direção e no sentido indicados pelo vetor  $(L_1, L_2)$ .

**Exercício 20.3** Dá para notar que o valor máximo de  $L(x, y)$ , para  $(x, y)$  em  $R$ , é atingido em um vértice de  $R$ ?

**Exercício 20.4** Dependendo de  $Q$ ,  $Q_1$ ,  $Q_2$ ,  $L_1$ ,  $L_2$ ,  $p_1$  e  $p_2$ , não seria mais lucrativo mandar Xavier para a Gávea e Yakecan para Madureira? Yakecan talvez não concorde, pois, com seu jeito de índio, faz enorme sucesso entre as meninas do Baixo Gávea,

Zumbi pensou um bocado, desenhou a região  $R$  e viu que era convexa e limitada. Entendeu o seguinte: cada desigualdade do tipo

$$ax + by \leq c,$$

com  $a$ ,  $b$  e  $c$  reais fixos, pode ser reescrita, usando o produto escalar, como

$$\langle (x, y), (a, b) \rangle \leq c,$$

de forma que o conjunto solução,

$$E = \left\{ (x, y) \in \mathbb{R}^2 \mid \langle (x, y), (a, b) \rangle \leq c \right\},$$

corresponde a um semiplano. Como semiplanos são convexos e a interseção de convexos é um convexo, qualquer região do plano definida por desigualdades do tipo  $ax + by \leq c$ , como as que apareceram em seu problema, é convexa. Dependendo das circunstâncias a tal região pode até ser vazia; com sorte, porém, além de não vazia, será limitada. Contou suas descobertas a Xavier e Yakecan. Xavier disse: ora, é um **polígono convexo!**

Zumbi generalizou: depois de amanhã, quando eu também sair para vender, podemos interpretar as quantidades  $x$ ,  $y$  e  $z$  como um ponto do espaço, dado, em coordenadas, por  $(x, y, z)$ . As desigualdades, então, definem semiespaços; o conjunto de desigualdades define um conjunto  $K$ , que é a interseção dos semiespaços correspondentes às desigualdades. Os elementos de  $K$  são os ternos ordenados  $(x, y, z)$  que correspondem às quantidades admissíveis,  $x$ ,  $y$  e  $z$  (isto é, cada terno  $(x, y, z)$  corresponde a uma possível escolha de quantidades de amendoim a serem atribuídas, respectivamente, a  $X$ , a  $Y$  e a  $Z$ ). Como cada semiespaço é um convexo,  $K$  é, também, convexo. E, se for limitado, é um **poliedro**, concluiu Xavier.

<sup>1</sup>A diferença de lucratividade é determinada não só pelas diferenças de custo das embalagens, mas também por diferenças nos preços praticados nos dois bairros, assim como outros fatores, como taxas de extorsão a serem pagas às máfias de cada local, suborno a policiais, etc.)

Nesse caso, o lucro pode ser uma função  $L(x, y, z)$ . Se Xavier vende  $x$  em Madureira, com lucro  $L_1$  por unidade, Yakecan, na Gávea lucra  $L_2$  por unidade e vende  $y$ , e eu, no Centro, vendo  $z$  e consigo  $L_3z$ , nosso lucro total será

$$L(x, y, z) = L_1x + L_2y + L_3z = \langle (x, y, z), (L_1, L_2, L_3) \rangle.$$

E se chutarmos, para começar, quantidades  $x$ ,  $y$  e  $z$ ? sugeriu Xavier. Isso, disse Zumbi! Vamos chamá-las de  $x_0$ ,  $y_0$  e  $z_0$ . Se, partindo do ponto  $(x_0, y_0, z_0)$ , procuramos aumentar o lucro, tem duas coisas claras: andando sobre o plano que passa por  $(x_0, y_0, z_0)$  e é normal a  $(L_1, L_2, L_3)$ , o valor de  $L$  não muda; mas se formos para onde aponta o vetor  $(L_1, L_2, L_3)$ , então  $L$  aumenta!

**Exercício 20.5** *Volte ao caso do plano. Suponha que a região  $R$  é um polígono convexo e chute um ponto de partida,  $(x_0, y_0)$ , no interior de  $R$ . Trace, passando por  $(x_0, y_0)$ , uma reta,  $r$ , perpendicular a  $(L_1, L_2)$ . Agora faça  $r$  ir andando, sempre perpendicular a  $(L_1, L_2)$  (e, portanto, paralelamente à posição inicial), no sentido para o qual  $(L_1, L_2)$  aponta, até o momento em que  $r$  deixa de tocar a região  $R$ . Conclua que o valor máximo de  $L(x, y)$  é obtido em um vértice de  $R$  (note que, mesmo se na posição limite, a reta  $r$  contém um lado de  $R$  e, portanto, todos os pontos desse lado são de máximo, pode-se dizer que o máximo é atingido em um de seus vértices).*

*Faça a mesma coisa em  $\mathbb{R}^3$ : suponha que  $K$  é um poliedro convexo e que queremos maximizar  $L(x, y, z) = L_1x + L_2y + L_3z$  em  $K$ . Chute um ponto de partida,  $(x_0, y_0, z_0)$ , trace por ele um plano,  $\alpha$ , perpendicular a  $(L_1, L_2, L_3)$ ; mova  $\alpha$ , paralelamente à posição inicial, no sentido em que  $L$  aumenta, até a posição limite em que  $\alpha$  deixa de tocar  $K$ . Conclua que o valor máximo de  $L$  é atingido em um vértice de  $K$ .*

No dia seguinte, antes de saírem para o trabalho, Yakecan falou: esta noite tive um sonho. Sonhei que nossa sociedade tinha 623 vendedores. As quantidades de amendoim que cada um devia levar eram 623 números. No começo, estavam, todos os números, em fila, separados por vírgulas, pareciam formar uma enorme cobra. Depois, a cobra começou a se enroscar, se encaracolar...fiquei com medo e saí flutuando, enquanto os números iam se enroscando cada vez mais. Eu fui flutuando, me afastando, e, aos poucos, aquela maçaroca de números, de longe, foi virando apenas um pontinho de um espaço enorme...

Zumbi ficou matutando. Xavier anotou o número, 623, e, sem que os outros vissem, furtivamente, pegou todo o dinheiro que tinham guardado. Chegando a Madureira, jogou tudo no bicho.



# Capítulo 21

## Áreas e determinantes

Até agora somos capazes de medir distâncias e de determinar ângulos através de coordenadas. Vejamos agora como lidar com o cálculo de áreas. Em princípio, se sabemos calcular os comprimentos de dois vetores e o seno (que podemos obter do cosseno) do ângulo entre eles, temos certeza de poder chegar à área de qualquer paralelogramo.

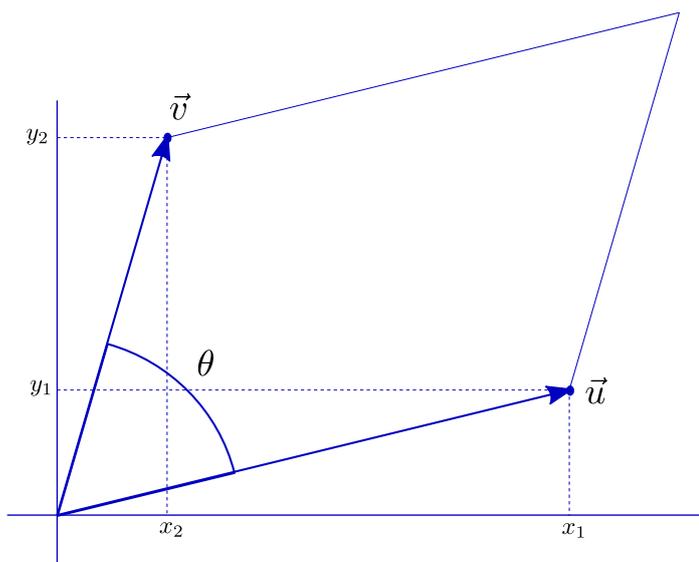


Figura 21.1:

Vamos, porém, partir para uma abordagem direta: tentaremos associar a cada par de vetores,  $\vec{u} = (x_1, y_1)$ ,  $\vec{v} = (x_2, y_2)$ , a área do paralelogramo por eles formado, expressa diretamente em função de  $x_1, y_1, x_2, y_2$ . Veremos, depois de algumas peripécias, que tal área é dada pelo valor absoluto do **determinante**

$$\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1.$$

## 21.1 Orientação

Começemos definindo a **orientação** de um par de vetores. Sejam  $\vec{u}_1, \vec{u}_2$  dois vetores não paralelos e não nulos. Diremos que o par  $\vec{u}_1, \vec{u}_2$  tem orientação positiva se o seno do ângulo  $\theta$  entre  $\vec{u}_1$  e  $\vec{u}_2$ , medido de  $\vec{u}_1$  para  $\vec{u}_2$  no sentido trigonométrico, é positivo (ou, o que é equivalente, se, “para girarmos o ponteiro  $\vec{u}_1$  para o ponteiro  $\vec{u}_2$  pelo menor ângulo, andamos no sentido trigonométrico”).

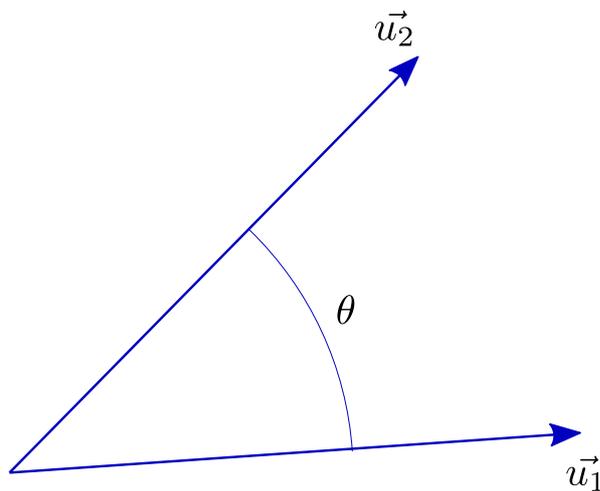


Figura 21.2:

Note que a orientação depende da ordem em que tomamos os vetores, e que se a orientação de  $\vec{u}_1, \vec{u}_2$  é positiva, então a de  $\vec{u}_2, \vec{u}_1$  é negativa. Assim, quando falarmos “a orientação de  $\vec{u}, \vec{v}$ ”, estará sempre implícito que se trata de um par ordenado. Diremos que dois pares de vetores  $\vec{u}_1, \vec{u}_2$  e  $\vec{v}_1, \vec{v}_2$  têm a **mesma orientação** se as respectivas orientações são simultaneamente positivas ou simultaneamente negativas. Assim, por exemplo, o par  $\vec{u}, \vec{v}$  tem orientação positiva se e só se tem a mesma orientação que o par formado pela base canônica,  $\vec{e}_1, \vec{e}_2$ .

**Exercício 21.1** Verifique que se  $\vec{u}, \vec{v}$  tem orientação positiva e  $t$  é um número real não nulo, então  $t\vec{u}, \vec{v}$  e  $\vec{u}, t\vec{v}$  têm orientação positiva se  $t > 0$  e negativa se  $t < 0$ .

**Exercício 21.2** Considere o vetor  $\vec{u} = (x, y)$  identificado com o ponto  $P = (x, y)$ . Considere a reta  $OP$ , coloque-se sobre a origem e olhe para  $P$ . Verifique que o par  $\vec{u}, \vec{v}$  tem orientação positiva se e só se o ponto correspondente a  $\vec{v}$  está à sua esquerda.

**Exercício 21.3** Mostre que  $\vec{u}, \vec{v}$  e  $\vec{u}, \vec{v} + t\vec{u}$  têm a mesma orientação, qualquer que  $t$  seja não nulo.

**Exercício 21.4** Suponha que  $\vec{u}, \vec{v}$  tem orientação positiva. Gire  $\vec{u}$  de um ângulo reto no sentido trigonométrico, obtendo o vetor  $\vec{u}^\perp$ . Mostre que o produto escalar

$$\langle \vec{u}^\perp, \vec{v} \rangle$$

é positivo.

**Exercício 21.5** Sejam  $\vec{u} = (a_{11}, a_{21})$  e  $\vec{v} = (a_{12}, a_{22})$ . Use a observação do exercício anterior para mostrar que  $\vec{u}, \vec{v}$  tem orientação positiva se, e somente se,

$$a_{11}a_{22} - a_{21}a_{12} > 0.$$

## 21.2 Áreas com sinal

Vamos agora definir uma função,  $d$ , que a cada par (ordenado) de vetores  $\vec{u}, \vec{v}$  associa a área do paralelogramo por eles formado.

Fica entendido que se  $\vec{u}$  e  $\vec{v}$  são paralelos (o que inclui a possibilidade de um dos dois ser nulo, ou ambos), então  $d(\vec{u}, \vec{v}) = 0$ . Incluiremos na definição de  $d$ , porém, uma novidade, que a distingue do que comumente chamamos área:  $d(\vec{u}, \vec{v})$  será positiva se o par  $\vec{u}, \vec{v}$  tiver orientação positiva e negativa se a orientação de  $\vec{u}, \vec{v}$  for negativa. É claro que o leitor não é obrigado a aceitar áreas negativas assim à toa, e daremos boas razões algébricas para a ousadia.

A primeira razão algébrica é a seguinte: se  $t$  é positivo, a área do paralelogramo formado por  $t\vec{u}$  e  $\vec{v}$  é  $t$  vezes a do paralelogramo formado por  $\vec{u}$  e  $\vec{v}$ , o que nos leva a conjecturar que

$$d(t\vec{u}, \vec{v}) = td(\vec{u}, \vec{v}).$$

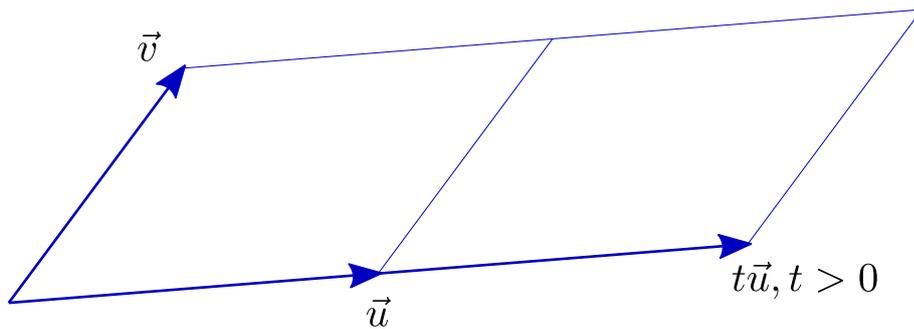


Figura 21.3:

Mas na verdade isso não pode valer para  $t$  negativo, a menos que admitamos valores negativos para  $d$  ou que modifiquemos um pouco a fórmula acima (usando  $|t|$  no lugar de  $t$ ). Podemos ainda notar que o problema que surge diz respeito apenas ao sinal. Ora, se  $d$  troca de sinal quando trocamos a orientação, então a definição que demos está boa, pois  $t$  negativo troca o sinal dos dois lados da igualdade.

O leitor argumentará, talvez, que bastaria escrever  $d(t\vec{u}, \vec{v}) = |t|d(\vec{u}, \vec{v})$ . Poderíamos contra-argumentar dizendo que trabalhar com  $|t|$  é chatíssimo, mas preferimos lançar mão de nossa segunda razão algébrica, que é um verdadeiro canhão.

A figura abaixo nos sugere a seguinte propriedade, pensando em termos de áreas:

$$d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) + d(\vec{u}, \vec{v}_2).$$

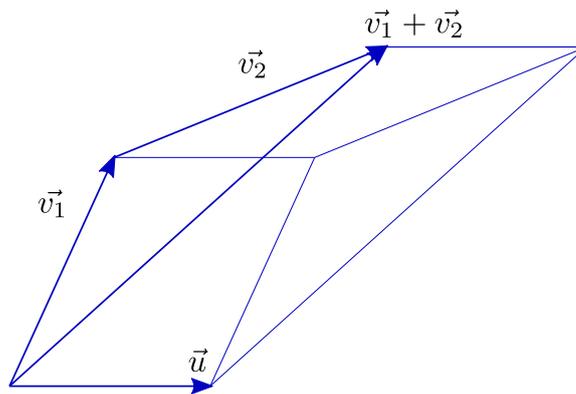


Figura 21.4:

No entanto, a figura seguinte já sugere outra coisa:

$$d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) - d(\vec{u}, \vec{v}_2).$$

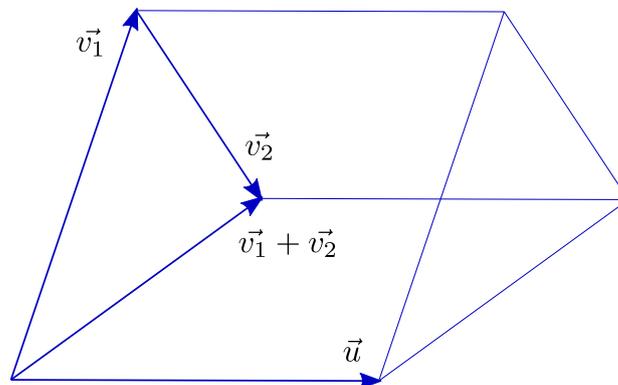


Figura 21.5:

Pois é...Na primeira figura, podemos observar, os pares  $\vec{u}, \vec{v}_1$  e  $\vec{u}, \vec{v}_2$  têm a mesma orientação; já na segunda, as orientações são opostas.

**Exercício 21.6** Pegue papel e lápis e desenhe todos os casos que achar necessários até se convencer de que trabalhando com áreas negativas (isto é, com a definição de  $d$  dada acima) vale a propriedade

$$d(\vec{u}, \vec{v}_1 + \vec{v}_2) = d(\vec{u}, \vec{v}_1) + d(\vec{u}, \vec{v}_2) \quad \forall \vec{u}, \vec{v}_1, \vec{v}_2 \in \mathbb{R}^2.$$

Vamos tratar nossa função  $d$ , agora, lançando mão de certas propriedades notáveis. Vamos ver que tais propriedades caracterizam  $d$  e nos permitem deduzir uma expressão simples para seu cálculo.

$d$  é uma função que a cada par (ordenado)  $\vec{u}, \vec{v}$  de vetores do plano associa um número real,  $d(\vec{u}, \vec{v})$ , com as seguintes propriedades:

$$\begin{aligned} (i) d(\vec{u}, \vec{v}) &= -d(\vec{v}, \vec{u}) & \forall \vec{u}, \vec{v} \in \mathbf{R}^2; \\ (ii) d(t\vec{u}, \vec{v}) &= td(\vec{u}, \vec{v}) & \forall \vec{u}, \vec{v} \in \mathbf{R}^2 \forall t \in \mathbf{R}; \\ (iii) d(\vec{u}, \vec{v}_1 + \vec{v}_2) &= d(\vec{u}, \vec{v}_1) + d(\vec{u}, \vec{v}_2) & \forall \vec{u}, \vec{v}_1, \vec{v}_2 \in \mathbf{R}^2; \\ (iv) d(\vec{e}_1, \vec{e}_2) &= 1. \end{aligned}$$

As propriedades (i), (ii) e (iii) foram discutidas na seção precedente; a propriedade (iv) parece óbvia, mas não teríamos como deduzi-la das demais. Três outras propriedades com as quais contamos podem ser deduzidas de (i), (ii) e (iii):

$$\begin{aligned} (i)' d(\vec{u}, \vec{u}) &= 0 & \forall \vec{u}; \\ (ii)' d(\vec{u}, t\vec{v}) &= td(\vec{u}, \vec{v}) & \forall \vec{u}, \vec{v} \in \mathbf{R}^2 \forall t \in \mathbf{R}; \\ (iii)' d(\vec{u}_1 + \vec{u}_2, \vec{v}) &= d(\vec{u}_1, \vec{v}) + d(\vec{u}_2, \vec{v}) & \forall \vec{u}_1, \vec{u}_2, \vec{v} \in \mathbf{R}^2. \end{aligned}$$

As demonstrações são simples e puramente algébricas:

(i)' segue de  $d(\vec{u}, \vec{u}) = -d(\vec{u}, \vec{u})$  (por (i));

(ii)' se deduz notando que, por (i) e (ii),  $d(\vec{u}, t\vec{v}) = -d(t\vec{v}, \vec{u}) = -td(\vec{v}, \vec{u}) = -t(-d(\vec{u}, \vec{v})) = td(\vec{u}, \vec{v})$ .

**Exercício 21.7** Prove (iii)' usando apenas (i) e (iii).

Vamos agora, sem mais delongas, proceder ao cálculo de  $d(\vec{u}, \vec{v})$ , usando as propriedades acima. Sendo  $\vec{u} = (x_1, y_1) = x_1\vec{e}_1 + y_1\vec{e}_2$ ,  $\vec{v} = (x_2, y_2) = x_2\vec{e}_1 + y_2\vec{e}_2$ , temos:

$$\begin{aligned} d(\vec{u}, \vec{v}) &= d(x_1\vec{e}_1 + y_1\vec{e}_2, x_2\vec{e}_1 + y_2\vec{e}_2) = \\ &= x_1x_2d(\vec{e}_1, \vec{e}_1) + x_1y_2d(\vec{e}_1, \vec{e}_2) + y_1x_2d(\vec{e}_2, \vec{e}_1) + y_1y_2d(\vec{e}_2, \vec{e}_2). \end{aligned}$$

Agora basta substituir

$$d(\vec{e}_1, \vec{e}_1) = 0, d(\vec{e}_2, \vec{e}_2) = 0, d(\vec{e}_1, \vec{e}_2) = 1, d(\vec{e}_2, \vec{e}_1) = -d(\vec{e}_1, \vec{e}_2) = -1$$

para obter

$$d(\vec{u}, \vec{v}) = x_1y_2 - x_2y_1,$$

ou, usando a notação consagrada,

$$d(\vec{u}, \vec{v}) = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} 1.$$

Assim, a área (com sinal) do paralelogramo formado por  $\vec{u} = (x_1, y_1)$  e  $\vec{v} = (x_2, y_2)$  é dada por  $x_1y_2 - x_2y_1$ . Se fizermos questão da área “mesmo”, basta tomarmos o valor absoluto.

### 21.3 O determinante de uma transformação linear

Consideremos a transformação linear  $T$ , dada pela matriz  $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ .

Notemos que

$$T\vec{e}_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix},$$

$$T\vec{e}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}.$$

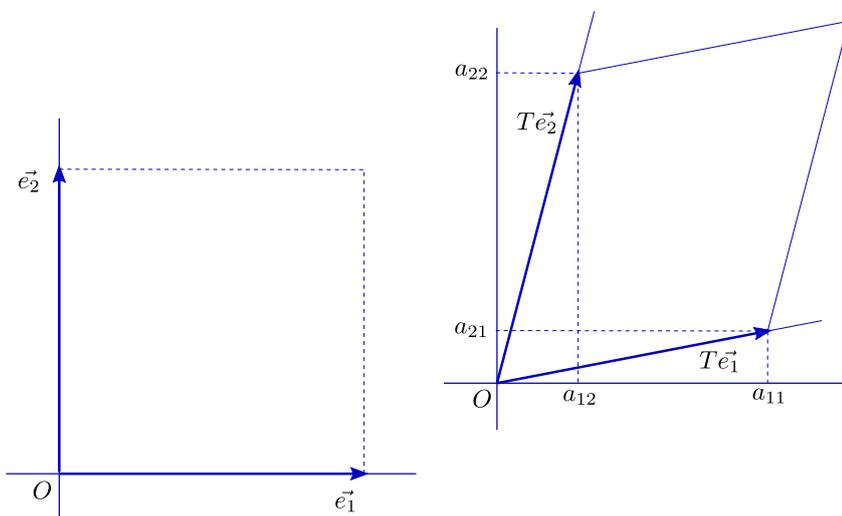


Figura 21.6:

$T$  transforma o quadrado formado por  $\vec{e}_1 = (1, 0)$  e  $\vec{e}_2 = (0, 1)$  no paralelogramo formado por  $T\vec{e}_1 = (a_{11}, a_{21})$  e  $T(\vec{e}_2) = (a_{12}, a_{22})$ . Como a área do quadrado é 1, a relação entre as duas áreas é dada por

---


$$1 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = x_1y_2 - x_2y_1 \text{ é chamado } \mathbf{determinante} \text{ da matriz } \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}$$

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

O número

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

é chamado **determinante** da matriz

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

mas podemos também chamá-lo **determinante da transformação linear**  $T$ . Vejamos seu significado geométrico.

Se  $F$  é uma figura qualquer no plano, consideremos sua imagem por  $T$ ,  $F'$ , e procuremos a relação entre as áreas de  $F'$  e de  $F$ .

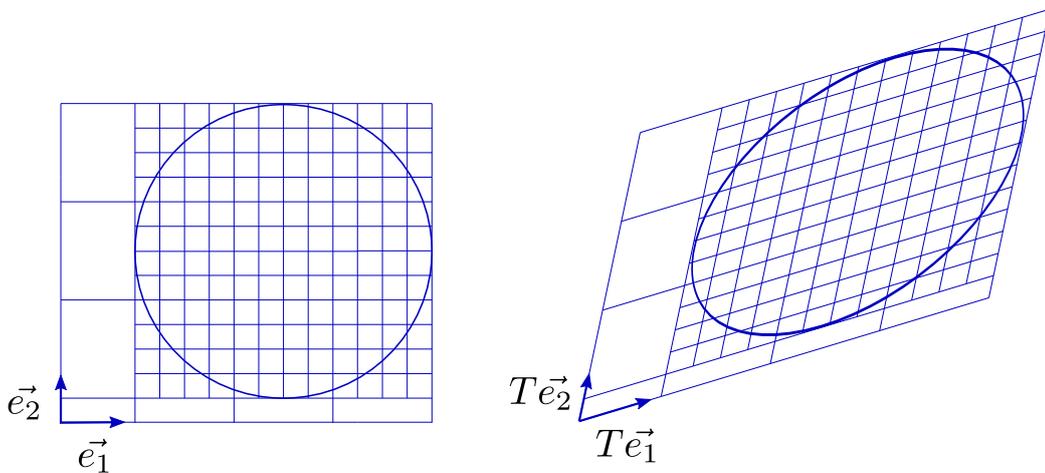


Figura 21.7:

Podemos aproximar a área de  $F$  (por falta ou por excesso) por meio de quadradinhos bem pequeninos, que terão como imagens por  $T$  paralelogramos também pequeninos, que aproximarão  $F'$ . Note que se  $\varepsilon$  é o lado dos quadradinhos, estes serão formados pelos vetores  $\varepsilon\vec{e}_1$  e  $\varepsilon\vec{e}_2$  e terão área  $\varepsilon^2$ . Os correspondentes paralelogramos pequeninos serão formados pelos vetores  $T(\varepsilon\vec{e}_1) = \varepsilon T\vec{e}_1$  e  $T(\varepsilon\vec{e}_2) = \varepsilon T(\vec{e}_2)$ .

**Exercício 21.8** Mostre que a área (com sinal) do paralelogramo formado por  $T(\varepsilon\vec{e}_1) = \varepsilon T\vec{e}_1$  e  $T(\varepsilon\vec{e}_2) = \varepsilon T(\vec{e}_2)$  é  $\varepsilon^2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ . É fácil, visto que  $T\vec{e}_1 = (a_{11}, a_{21})$  e  $T(\vec{e}_2) = (a_{12}, a_{22})$ .

Agora vejamos. Se a aproximação por falta de  $F$  é feita com  $m_\varepsilon$  quadrados e a por excesso com  $n_\varepsilon$  quadrados de área  $\varepsilon^2$ , as correspondentes aproximações de  $F'$  são feitas, por falta e por excesso, respectivamente, com  $m_\varepsilon$  e  $n_\varepsilon$  paralelogramozinhos de área (com sinal)  $\varepsilon^2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ .

Se designarmos por  $|F|$  a área de  $F$  e por  $|F'|$  a área (com sinal) de  $F'$ , teremos que

$$m_\varepsilon\varepsilon^2 \leq |F| \leq n_\varepsilon\varepsilon^2$$

e

$$m_\varepsilon\varepsilon^2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \leq |F'| \leq n_\varepsilon\varepsilon^2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.^2$$

Temos também que, quando  $\varepsilon$  tende a zero, tanto  $m_\varepsilon\varepsilon^2$  como  $n_\varepsilon\varepsilon^2$  tendem a  $|F|$ . Da mesma forma, as correspondentes aproximações para  $|F'|$  tendem a  $|F'|$ . Mas, independente do valor de  $\varepsilon$ , temos sempre

$$\frac{m_\varepsilon\varepsilon^2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{m_\varepsilon\varepsilon^2} = \frac{n_\varepsilon\varepsilon^2 \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}{n_\varepsilon\varepsilon^2} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

A conclusão que se impõe é que

$$\frac{|F'|}{|F|} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Ou seja, se a transformação linear  $T$  é dada pela matriz

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

o número

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

---

<sup>2</sup>Caso  $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \leq 0$ , esta linha deve, é claro, ter seus sinais de desigualdade invertidos

expressa a relação entre a área (com sinal) da imagem por  $T$  de qualquer figura  $F$  e a área de  $F$ . É claro que a relação entre as “áreas mesmo” é dada pelo valor absoluto do determinante, mas não custa nada guardar um eventual sinal negativo para indicar que, neste caso, se o plano de  $F$  é pintado de azul em cima e vermelho embaixo (e portanto vemos  $F$  azul), o de  $F'$  estará pintado de vermelho em cima e azul embaixo (e vemos  $F'$  vermelha).

**Exercício 21.9** Entenda essa última e misteriosa afirmação.

Vejam agora uma interessante consequência das idéias acima. Sejam  $R$  e  $S$  duas transformações lineares de  $\mathbb{R}^2$  em  $\mathbb{R}^2$ , dadas, respectivamente, pelas matrizes

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \text{ e } \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Examinemos a composta de  $R$  e  $S$ , isto é, a transformação  $T$  dada por  $T\vec{v} = S(R\vec{v})$ .

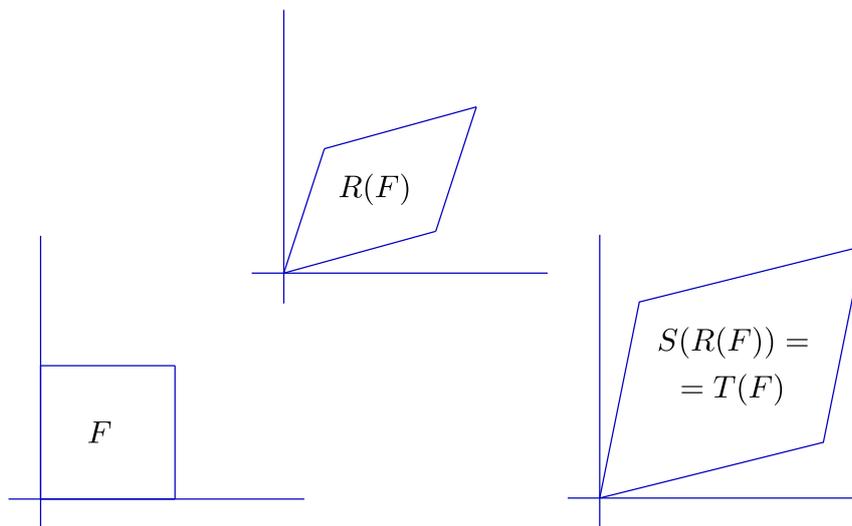


Figura 21.8:

Como vimos no capítulo sobre transformações lineares, a matriz de  $T$  é dada por

$$\begin{aligned} \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} &= \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\ &= \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix}. \end{aligned}$$

<sup>3</sup>

É claro que podemos calcular diretamente o determinante de  $T$  usando a (enorme!) expressão acima. Mas podemos agir de outra forma. Se  $F$  é uma figura do plano, o

<sup>3</sup>Se você pulou o exercício referente a esta propriedade, é hora de atacá-lo

determinante de  $T$  (e de sua matriz) é a relação entre a área (com sinal) de  $T(F)$  e a área de  $F$ . Ora,  $T(F) = S(R(F))$ ; se designarmos por  $a(F)$ ,  $a(R(F))$  e  $a(S(R(F)))$  as áreas (com sinal) de cada uma destas figuras, teremos:

$$\frac{a(R(F))}{a(F)} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \frac{a(S(R(F)))}{a(R(F))} = 4 \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix}.$$

Mas o determinante de  $T$  expressa a relação entre as áreas de  $T(F) = S(R(F))$  e de  $F$ . Portanto,

$$\begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \frac{a(S(R(F)))}{a(F)} = \frac{a(S(R(F)))}{a(R(F))} \frac{a(R(F))}{a(F)}.$$

Isto significa que

$$\begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Acabamos pois de deduzir:

**Teorema:** O determinante da matriz (c) produto das matrizes (a) e (b) é dado pelo produto dos determinantes de (a) e de (b).<sup>5</sup>

**Exercício 21.10** Deixamos, na dedução acima, de considerar o caso em que  $a(R(F))$  é nula. Cuide dele.

**Observação:** Embora nossa apresentação parta de um ponto de vista geométrico, o conceito de determinante surge naturalmente na busca de fórmulas de resolução de sistemas lineares. O exercício a seguir, que só exige contas, é uma primeira pista.

**Exercício 21.11** Considere o sistema linear

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

Mostre que, caso

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0,$$

a solução é dada por

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}; \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

<sup>4</sup>Note que  $R(F)$  é também uma figura do plano!

<sup>5</sup>Este resultado vale para matrizes  $n \times n$ , n qualquer

# Capítulo 22

## O Teorema Espectral

### 22.1 Matrizes simétricas e transformações autoadjuntas

Já andamos esbarrando na necessidade de tomar as **transpostas** de certas matrizes. Voltemos ao assunto. Dada a matriz  $A$ ,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

sua **transposta** é a matriz  $A^T$ ,

$$A^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}.$$

Por motivos óbvios, se  $A^T = A$  (o que equivale a  $a_{12} = a_{21}$ ), a matriz  $A$  é dita **simétrica**.

Uma continha simples vai nos mostrar a seguinte relação:

**Proposição 22.1** *Para toda matriz  $2 \times 2$   $A$  e para quaisquer vetores  $x$  e  $y$  em  $\mathbb{R}^2$ , vale a seguinte relação:*

$$\langle Ax, y \rangle = \langle x, A^T y \rangle.$$

Demonstração: Vamos usar coordenadas:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Agora é uma questão de *jeito de olhar*:

$$\langle Ax, y \rangle = \begin{matrix} (a_{11}x_1 + a_{12}x_2)y_1 \\ + \\ (a_{21}x_1 + a_{22}x_2)y_2 \end{matrix} = x_1 \begin{pmatrix} a_{11}y_1 \\ + \\ a_{21}y_2 \end{pmatrix} + x_2 \begin{pmatrix} a_{12}y_1 \\ + \\ a_{22}y_2 \end{pmatrix} = \langle x, A^T y \rangle.$$

■

Em particular, se  $A$  é simétrica, então, para quaisquer vetores  $x$  e  $y$  em  $\mathbb{R}^2$ ,

$$\langle Ax, y \rangle = \langle x, Ay \rangle.$$

Observe que, se  $A$  é matriz simétrica e  $T$  é a correspondente transformação linear, então, para quaisquer vetores  $x$  e  $y$  em  $\mathbb{R}^2$ ,  $\langle Tx, y \rangle = \langle x, Ty \rangle$ .  $T$  também é dita **simétrica**, mas é mais usual chamá-la de **autoadjunta**. De maneira geral, a transformação linear  $T^*$  associada à transposta de  $A$  é chamada de **adjunta** de  $T$ . Assim, a relação entre  $A$  e  $A^T$  se traduz por

$$\langle Tx, y \rangle = \langle x, T^*y \rangle \quad \forall x \in \mathbb{R}^2, \forall y \in \mathbb{R}^2.$$

A identificação entre a adjunta da transformação linear  $T$  e a transposta de sua matriz, permite deduzir facilmente o seguinte resultado: se  $A$  e  $B$  são matrizes  $2 \times 2$ , então

$$(AB)^T = B^T A^T.$$

O resultado segue, facilmente, do correspondente resultado para transformações lineares: se  $S, T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  são transformações lineares, então

$$(ST)^* = T^* S^*.$$

Esse último fato segue de

$$\langle \vec{u}, (ST)^* \vec{v} \rangle = \langle (ST) \vec{u}, \vec{v} \rangle = \langle T \vec{u}, S^* \vec{v} \rangle = \langle \vec{u}, T^* (S^* \vec{v}) \rangle = \langle \vec{u}, (T^* S^*) \vec{v} \rangle, \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^2.$$

Para concluir, de

$$\langle \vec{u}, (ST)^* \vec{v} \rangle = \langle \vec{u}, (T^* S^*) \vec{v} \rangle, \quad \forall \vec{u}, \vec{v} \in \mathbb{R}^2,$$

que  $(ST)^* = T^* S^*$ , observe que, para qualquer  $\vec{v} \in \mathbb{R}^2$ , teremos

$$\langle \vec{u}, (ST)^* \vec{v} - (T^* S^*) \vec{v} \rangle = 0, \quad \forall \vec{u} \in \mathbb{R}^2.$$

Escolhendo, para cada  $\vec{v}$ ,  $\vec{u} = (ST)^* \vec{v} - (T^* S^*) \vec{v}$ , segue  $|(ST)^* \vec{v} - (T^* S^*) \vec{v}|^2 = 0$ , o que prova que  $(ST)^* \vec{v} - (T^* S^*) \vec{v} = \vec{0} \quad \forall \vec{v} \in \mathbb{R}^2$ .

## 22.2 O Teorema Espectral

Para transformações lineares autoadjuntas, temos uma espécie de Decomposição em Valores Singulares especial, em que não é preciso usar bases diferentes para o domínio e o contradomínio da transformação: o **Teorema espectral**.

**Teorema 22.2** Se  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é transformação linear autoadjunta, então existem dois vetores,  $\vec{e}_1$  e  $\vec{e}_2$ , unitários e ortogonais, e dois números reais,  $\lambda_1$  e  $\lambda_2$ , tais que

$$T\vec{e}_1 = \lambda_1 \vec{e}_1, \quad T\vec{e}_2 = \lambda_2 \vec{e}_2.$$

A demonstração é análoga à da Decomposição em Valores Singulares. Desta vez, vamos fazer por escrito. Uma outra demonstração, em vídeo, pode ser vista em

## Teorema Espectral

Demonstração: Começemos observando que basta-nos encontrar dois vetores,  $\varepsilon_1$  e  $\varepsilon_2$ , unitários e ortogonais, tais que

$$\langle T\varepsilon_1, \varepsilon_2 \rangle = 0.$$

De fato, neste caso, teremos, escrevendo  $T\varepsilon_1 = \lambda_1\varepsilon_1 + \mu\varepsilon_2$ ,

$$0 = \langle T\varepsilon_1, \varepsilon_2 \rangle = \langle \lambda_1\varepsilon_1, \varepsilon_2 \rangle + \langle \mu\varepsilon_2, \varepsilon_2 \rangle = \mu.$$

Como  $T$  é autoadjunta, teremos, também,  $\langle \varepsilon_1, T\varepsilon_2 \rangle = 0$ , de modo que o mesmo raciocínio se aplica a  $\varepsilon_2$ .

Vamos, pois, provar a existência de um tal par de vetores. Para começar, vamos definir, para cada vetor  $w = (x, y)$ , seu ortogonal,  $w^\perp$ , definido por  $w^\perp = (-y, x)$ . É imediato que  $\langle w^\perp, w \rangle = 0$  e que  $w^\perp$  tem norma igual à de  $w$ . Assim, o que queremos é um vetor unitário,  $\varepsilon_1$ , tal que  $\langle w^\perp, T w \rangle = 0$ .

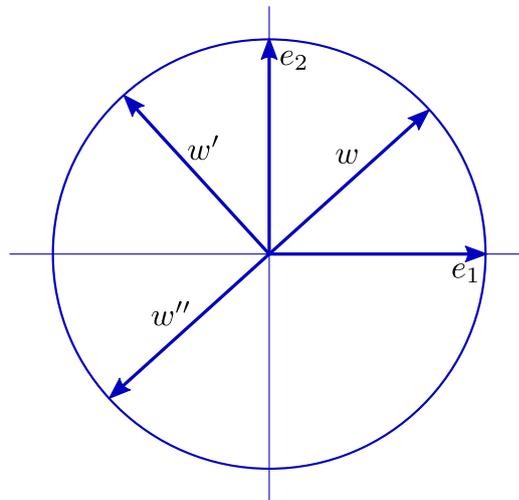


Figura 22.1:  $(w^\perp)^\perp = -w$

Um detalhe interessante é que, para qualquer  $w$  em  $E$ ,  $(w^\perp)^\perp = -w$ , de modo que  $\vec{e}_1^\perp = \vec{e}_2$  e  $\vec{e}_2^\perp = -\vec{e}_1$ . Assim,

$$\langle T\vec{e}_1, \vec{e}_1^\perp \rangle = \langle T\vec{e}_1, \vec{e}_2 \rangle = \langle \vec{e}_1, T\vec{e}_2 \rangle = -\langle T\vec{e}_2, \vec{e}_2^\perp \rangle.$$

Ora, daí segue que a função  $g : E \rightarrow \mathbf{R}$  dada por

$$\langle Tw, w^\perp \rangle$$

muda de sinal quando caminhamos de  $\vec{e}_1$  até  $\vec{e}_2$  sobre o círculo unitário de  $E$  (a rigor, poderíamos ter  $g(\vec{e}_1) = g(\vec{e}_2) = 0$ , mas então,  $\vec{e}_1$  e  $\vec{e}_2$  já resolveriam nosso problema). Ora,  $g$  varia continuamente quando caminhamos de  $\vec{e}_1$  até  $\vec{e}_2$ . Logo, pelo Teorema do Valor Intermediário, existe  $\varepsilon_1$  no círculo unitário tal que  $g(\varepsilon_1) = 0$ , o que demonstra o Teorema.

Para dar um ar mais rigoroso ao argumento acima, considere a aplicação

$$f : [0, \frac{\pi}{2}] \longrightarrow \mathbb{R},$$

dada por  $f(\theta) = g(\cos \theta \vec{e}_1 + \sin \theta \vec{e}_2)$ , e note que:

- (i)  $f$  é contínua;
- (ii)  $f(0) = -f(\frac{\pi}{2})$ ;
- (iii)  $\cos \theta \vec{e}_1 + \sin \theta \vec{e}_2$  está no círculo unitário para todo  $\theta$ .

Logo, existe  $\theta_0$  em  $[0, \frac{\pi}{2}]$  tal que  $f(\theta_0) = 0$ . Nosso  $\varepsilon_1$  é  $\cos \theta_0 \vec{e}_1 + \sin \theta_0 \vec{e}_2$ . ■

Antes de dar a nosso Teorema uma versão matricial, vamos esclarecer um pouco mais a ideia de **matriz de uma transformação linear**.

## 22.3 Matriz de transformação linear em uma base qualquer

Vamos fazer uma observaçãozinha simples, mas importante, na linha do Fourier. Se  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é linear e  $\alpha = \{\vec{e}_1, \vec{e}_2\}$  é uma base para os vetores no plano, podemos, dado um vetor  $\vec{w}$ , escrevê-lo como

$$\vec{w} = s_1 \vec{e}_1 + s_2 \vec{e}_2.$$

Consequentemente,

$$T\vec{w} = s_1 T\vec{e}_1 + s_2 T\vec{e}_2. \quad (1)$$

Imagine agora que sabemos expressar tanto  $T\vec{e}_1$  como  $T\vec{e}_2$  na base  $\alpha$ :

$$T\vec{e}_1 = b_{11}\vec{e}_1 + b_{21}\vec{e}_2; \quad T\vec{e}_2 = b_{12}\vec{e}_1 + b_{22}\vec{e}_2. \quad (2)$$

Isso nos permite, na verdade, obter as coordenadas,  $(t_1, t_2)$ , de  $T\vec{w}$  na base  $\alpha$ : basta substituir (2) em (1). Observe que as contas nos dão, naturalmente, uma expressão matricial:

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} b_{11}s_1 + b_{12}s_2 \\ b_{21}s_1 + b_{22}s_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}.$$

**Definição 22.3** Se  $\alpha = \{\vec{e}_1, \vec{e}_2\}$  é uma base para os vetores no plano e  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é linear, a matriz

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

definida por

$$T\vec{e}_1 = b_{11}\vec{e}_1 + b_{21}\vec{e}_2; \quad T\vec{e}_2 = b_{12}\vec{e}_1 + b_{22}\vec{e}_2,$$

é dita a **matriz de  $T$  na base  $\alpha$** .

Tudo isso é teórico. Vamos nos concentrar no caso em que a base  $\alpha = \{\vec{e}_1, \vec{e}_2\}$  é **ortonormal**, ou seja,  $\vec{e}_1$  e  $\vec{e}_2$  são unitários e ortogonais. Gasta um tempinho e faz o exercício (é só usar o *truque do Fourier*):

**Exercício 22.4** Se  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é linear e  $\alpha = \{\vec{e}_1, \vec{e}_2\}$  é uma base ortonormal, então a matriz de  $T$  na base  $\alpha$  é dada por

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \langle T\vec{e}_1, \vec{e}_1 \rangle & \langle T\vec{e}_2, \vec{e}_1 \rangle \\ \langle T\vec{e}_1, \vec{e}_2 \rangle & \langle T\vec{e}_2, \vec{e}_2 \rangle \end{bmatrix}$$

O resultado do exercício nos garante que, se  $T$  é transformação linear autoadjunta, então sua matriz em qualquer base ortonormal é simétrica.

## 22.4 Teorema Espectral, agora com matrizes

O que o Teorema Espectral nos diz é que, se  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é transformação linear autoadjunta, então existe uma base ortonormal,  $\alpha = \{\vec{e}_1, \vec{e}_2\}$ , tal que a matriz de  $T$  na base  $\alpha$  é **diagonal**matriz!diagonal, ou seja, a matriz de  $T$  na base  $\alpha$  é da forma

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Agora que já somos especialistas em matrizes de mudanças de base, sabemos que a matriz  $Q$  que pega as coordenadas de um vetor na base  $\alpha$  e as transforma nas coordenadas do mesmo vetor na base canônica é ortogonal (já que  $\alpha$  é base ortonormal) e que, portanto, sua inversa (que também muda coordenadas, mas da base canônica para  $\alpha$ ) é sua transposta,  $Q^T$ . Assim, o **Teorema Espectral** pode ser reenunciado em versão matricial.

**Teorema 22.5** Se  $A$  é matriz simétrica  $2 \times 2$ , então existem matrizes  $2 \times 2$ ,  $D$  e  $Q$ , com  $D$  diagonal e  $Q$  ortogonal, tais que

$$A = QDQ^T.$$

## 22.5 O polinômio característico

O Teorema Espectral faz emergir um tipo especial de vetor, em relação a uma dada transformação linear  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ : vetores não nulos  $\vec{v}$  para os quais existe um número  $\lambda$  tal que

$$T\vec{v} = \lambda\vec{v}.$$

Tais vetores são chamados de **vetores próprios**, **autovetores**, ou **vetores característicos** de  $T$ ; os escalares  $\lambda$  correspondentes, recebem o nome de **valores próprios**, **autovalores**, ou **valores característicos** de  $T$ . A denominação mais popular é autovetores (e autovalores). Existe uma caracterização bastante simples dos autovalores, que envolve o determinante., **autovalores**, ou **valores!característicos**

Suponhamos que  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é linear e que  $T\vec{v} = \lambda\vec{v}$ , para um certo  $\vec{v}$  não nulo (observe que, neste caso,  $T(t\vec{v}) = \lambda(t\vec{v})$  para todo  $t$  em  $\mathbb{R}$ ). Temos, então, escrevendo  $\vec{v} = I\vec{v}$ , que  $T\vec{v} = (\lambda I)\vec{v}$ . Na forma matricial,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

ou, juntando um pouco,

$$\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Podemos, então, concluir que  $\lambda$  é autovalor de  $T$  se, e somente se, existe um vetor  $\vec{v}$ , não nulo, tal que  $(T - \lambda I)\vec{v} = 0$ . Agora uma observação crucial:  $(T - \lambda I) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é uma transformação linear que leva um vetor não nulo (nosso autovetor,  $\vec{v}$ ) em  $(0, 0)$ . Não pode, portanto, ser uma bijeção; seu determinante tem que ser nulo. Portanto:

$$\det \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix} = 0.$$

Dito de outra maneira  $\lambda$  tem que ser uma raiz do polinômio (do segundo grau)

$$p(x) = \det \begin{bmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{bmatrix} = x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{21}).$$

**Definição 22.6** O polinômio  $p(x) = \det(T - xI)$ , ou, mais explicitamente,

$$p(x) = \det \begin{bmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{bmatrix} = x^2 - (a_{11} + a_{22})x + (a_{11}a_{22} - a_{12}a_{21}),$$

é chamado de **polinômio característico** da transformação  $T$  (e, também, de sua matriz,  $A$ ).

Pelo que acabamos de ver, está demonstrado o seguinte:

**Proposição 22.7** O número  $\lambda$  é autovalor de  $T$  se, e somente se, é raiz do polinômio característico de  $T$ .

**Observação:** Agora ficou moleza: o polinômio característico é do segundo grau, sabemos achar suas raízes. Logo, temos os autovalores e, conseqüentemente, os autovetores. Tudo bem, divirta-se! Mas o mundo lá fora é cruel: os sistemas e transformações lineares são imensos; os polinômios característicos de verdade são, geralmente, incalculáveis.

## 22.6 O Teorema Espectral e a SVD

Acabamos de ver um método para obter os autovalores de uma transformação linear autoadjunta,  $T$ , e, conseqüentemente, a base ortonormal que diagonaliza  $T$ . Isso também pode nos servir para a Decomposição em valores singulares (SVD, de singular value decomposition). Suponhamos que

$$T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

é uma transformação linear bijetiva. Considere, então, a transformação linear  $T^*T$ . Como  $(T^*T)^* = T^*(T^*)^* = T^*T$ , nossa  $T^*T$  é autoadjunta. Pelo Teorema Espectral, temos uma base ortonormal  $\alpha = \{\vec{u}_1, \vec{u}_2\}$ , de autovetores de  $T^*T$ . Como  $T^*T\vec{u}_1 = \lambda_1\vec{u}_1$ , temos

$$\langle T^*T\vec{u}_1, \vec{u}_2 \rangle = \langle \lambda_1\vec{u}_1, \vec{u}_2 \rangle = \lambda_1 \langle \vec{u}_1, \vec{u}_2 \rangle = 0.$$

Mas, então,

$$\langle T\vec{u}_1, T\vec{u}_2 \rangle = \langle T^*T\vec{u}_1, \vec{u}_2 \rangle = 0,$$

o que mostra que uma base ortonormal de autovetores de  $T^*T$  nos dá uma decomposição em valores singulares de  $T$ .

## 22.7 Exercícios

## Básicos

### 4: Teorema Espectral & Decomposição em Valores Singulares

1. Sejam  $A$  matriz  $2 \times 2$ ,  $x$  e  $y$ , dados por

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

- (a) Olhe para

$$\langle Ax, y \rangle = \begin{pmatrix} (a_{11}x_1 + a_{12}x_2)y_1 \\ + \\ (a_{21}x_1 + a_{22}x_2)y_2 \end{pmatrix} = x_1 \begin{pmatrix} a_{11}y_1 \\ + \\ a_{21}y_2 \end{pmatrix} + x_2 \begin{pmatrix} a_{12}y_1 \\ + \\ a_{22}y_2 \end{pmatrix} = \langle x, A^T y \rangle.$$

(b) Suponha que  $a_{12} = a_{21}$  e que  $\langle Ax, y \rangle = 3$ . Calcule  $\langle x, Ay \rangle$ .

2. Suponha que  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é linear e que

$$T \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} = 3 \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix}, \quad T \begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} = 2 \begin{pmatrix} \sqrt{3}/2 \\ -1/2 \end{pmatrix}.$$

(a) Determine a matriz  $[T]$  de  $T$  na base canônica.

(b) Escreva  $[T]$  como um produto de matrizes,

$$[T] = VDU^T,$$

com  $D$  diagonal positiva e  $U$  &  $V$  ortogonais.

3. Seja  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  linear, dada, na base canônica, pela matriz

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

(a) Calcule os autovalores,  $\lambda_1$  e  $\lambda_2$ , de  $T$ .

(b) Encontre  $u_1$  e  $u_2$ , unitários, tais que  $Tu_1 = \lambda_1 u_1$ ,  $Tu_2 = \lambda_2 u_2$ .

(c) Calcule  $\langle u_1, u_2 \rangle$ .

4. Suponha que  $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  é linear e que

$$T \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix} = -3 \begin{pmatrix} \sqrt{3}/2 \\ 1/2 \end{pmatrix}, \quad T \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix} = 2 \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix}.$$

Calcule o maior valor de  $\langle Tu, u \rangle$  em

$$S^1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1^2 + x_2^2 = 1 \right\}.$$

5. Obtenha matrizes  $2 \times 2$ ,  $U$ ,  $V$  e  $D$ , tais que

$$\begin{pmatrix} 5 & -3 \\ 0 & 4 \end{pmatrix} = UDV^\perp,$$

com  $D$  diagonal positiva e  $U$  &  $V$  ortogonais.

# Capítulo 23

## Números complexos e coordenadas polares

### 23.1 Os complexos

Os números complexos surgem, no século XVI, como um artifício de cálculo, no processo de solução de equações do 3º grau. Expressões do tipo " $a \pm \sqrt{b}$ ", com  $a$  e  $b$  reais,  $b$  negativo, eram aceitas no meio das contas, inicialmente, apenas porque conduziam, eventualmente, a soluções dadas por "números de verdade". Neste sentido, a expressão **número imaginário**, designando raízes quadradas de números negativos, é bastante coerente. É apenas com Euler, já em 1777, que os números complexos passam a ser vistos como entidades da forma  $a + bi$ , com  $a$  e  $b$  reais e  $i$  um "número imaginário", tal que  $i^2 = -1$ . Se  $z$  é o número complexo dado por  $z = a + bi$  (com  $a$  e  $b$  reais)  $a$  é chamado de **parte real** de  $z$  (e notado  $a = \text{Re}(z)$ ) e o número  $b$  de **parte imaginária** de  $z$  (e notado  $\text{Im}(z)$ ). Um número complexo é dito **real** se sua parte imaginária é nula; **imaginário**, se é nula sua parte real. A adição e a multiplicação de números complexos são definidas por:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Pouco depois de Euler, na virada do século XVIII para o XIX, se chega a uma interpretação geométrica dos números complexos e de suas operações. O irlandês William Hamilton, embora não tenha sido o primeiro na geometrização dos complexos, dá, em 1833, uma definição radical. Embora outros já tivessem proposto a interpretação do número complexo  $x + yi$  como um ponto do plano, Hamilton adota o ponto de vista de definir diretamente, em  $\mathbb{R}^2$ , as operações de adição e multiplicação por:

$$(a, b) + (c, d) = (a + c, b + d)$$

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

**Exercício 23.1** *Veja se entendeu. Mostre que, com a notação tradicional,  $0 + 0i$  é neutro para a adição; na notação de Hamilton, o neutro é  $(0,0)$ . Na notação tradicional, convencionou-se que  $a + 0i$  é notado por  $a$  e que  $0 + bi$ , se  $b \neq 0$ , é notado por  $bi$ . Compreenda que se pode passar da definição de Hamilton à*

tradicional, facilmente, observando que  $(a, b) = (a, 0) + (0, b)$  e convencionando que  $(a, 0)$  será notado por  $a$  e  $(0, b)$  por  $bi$ , se  $b \neq 0$ . Temos, naturalmente, que  $(1, 0) = 1$  e  $(0, 1) = i$ . Os números da forma  $(a, 0)$  correspondem aos reais; os da forma  $(0, b)$  são os imaginários.

**Definição:** Dado o número complexo  $z = a + bi$ , seu **conjugado** é o complexo  $\bar{z} = a - bi$ . O **módulo** de  $a + bi$  é

$$|a + bi| = \sqrt{a^2 + b^2} = \sqrt{(a + bi)(a - bi)}.$$

**Exercício 23.2** Sejam  $u = a + bi$  e  $v = c + di$  dois números complexos. Interpretando-os como vetores, mostre que seu produto escalar é  $\text{Re}(u\bar{v}) = \text{Re}(\bar{u}v)$ . Note que  $u\bar{u}$  é sempre real.

**Exercício 23.3** Mostre que o módulo de  $u$  é 1 se, e somente se,  $u\bar{u} = 1$ .

**Exercício 23.4** Mostre que  $u$  é real se, e somente se,  $\bar{u} = u$ . Mostre que  $u$  é imaginário se, e somente se,  $\bar{u} = -u$ .

A interpretação dos complexos como elementos de  $\mathbb{R}^2$  nos permite tomá-los, conforme nossa conveniência, ora como pontos, ora como vetores do plano (note que a adição de complexos corresponde à de vetores e que podemos multiplicá-los, como os vetores, por números reais:  $t(x + yi) = (t, 0)(x, y) = (tx, ty) = tx + tyi$ ). Mas a grande novidade, o que distingue os complexos de meros vetores, claro, é a possibilidade de multiplicá-los. O exercício a seguir é incontornável.

**Exercício 23.5** Sejam  $(a, b)$  e  $(x, y)$  números complexos. Observe que o produto  $(a, b)(x, y) = (ax - by, ay + bx)$  corresponde à multiplicação do vetor  $(x, y)$  pela matriz

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

ou seja,

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix}.$$

Feito o exercício, resta observar que

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \sqrt{a^2 + b^2} \begin{pmatrix} \frac{a}{\sqrt{a^2 + b^2}} & \frac{-b}{\sqrt{a^2 + b^2}} \\ \frac{b}{\sqrt{a^2 + b^2}} & \frac{a}{\sqrt{a^2 + b^2}} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Ora, a matriz

$$\begin{pmatrix} \frac{a}{\sqrt{a^2 + b^2}} & \frac{-b}{\sqrt{a^2 + b^2}} \\ \frac{b}{\sqrt{a^2 + b^2}} & \frac{a}{\sqrt{a^2 + b^2}} \end{pmatrix}$$

corresponde à rotação do ângulo  $\theta$  formado pelo vetor  $(a, b)$  com o eixo horizontal (no sentido trigonométrico, do eixo para o vetor); o número  $r = \sqrt{a^2 + b^2}$  é exatamente

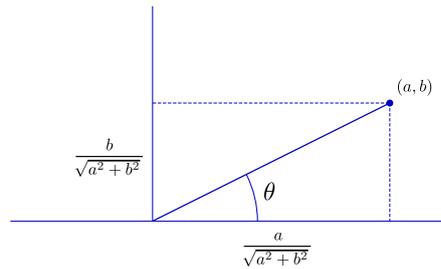


Figura 23.1:

a norma do vetor  $(a, b)$  (neste caso, dizemos também que  $r$  é o **módulo** do número complexo  $a + bi$ ). Assim, multiplicar  $x + yi$  por  $a + bi$  corresponde a rodar  $(x, y)$  de  $\theta$  e multiplicar o resultado por  $r$ .

**Definição:** Dado o par ordenado  $(a, b)$  de números reais, com  $(a, b) \neq (0, 0)$ , o par  $(r, \theta)$ , com

$$\begin{aligned} a &= r \cos \theta \\ b &= r \sin \theta, \end{aligned}$$

é dito uma representação de  $(a, b)$  em **coordenadas polares**. Costuma-se dizer, embora  $\theta$  esteja definido apenas a menos de um múltiplo inteiro de  $2\pi$ , que  $r$  e  $\theta$  são "as coordenadas polares" de  $(a, b)$ .

**Exercício 23.6** Suponha que os números complexos  $z_1$  e  $z_2$  sejam dados por  $z_1 = (r_1 \cos \theta_1, r_1 \sin \theta_1)$  e  $z_2 = (r_2 \cos \theta_2, r_2 \sin \theta_2)$ . Mostre, efetuando diretamente a multiplicação e usando as relações

$$\cos(\theta_1 + \theta_2) = \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2,$$

$$\sin(\theta_1 + \theta_2) = \sin \theta_1 \cos \theta_2 + \cos \theta_1 \sin \theta_2,$$

que  $z_1 z_2 = (r_1 r_2 \cos(\theta_1 + \theta_2), r_1 r_2 \sin(\theta_1 + \theta_2))$ . Ou seja: o produto de dois números complexos é o número complexo obtido multiplicando os módulos e somando os ângulos.

Uma outra maneira de entender as coordenadas polares é dizer que todo número complexo  $z$ , com  $z \neq 0$ , pode ser escrito como

$$z = |z|u, \quad u = \frac{z}{|z|}.$$

**Exercício 23.7** Sejam  $z$  um número complexo não nulo e  $n$  um número natural (também não nulo). Mostre, escrevendo  $z$  em coordenadas polares, que existem exatamente  $n$  números complexos,  $w_1, \dots, w_n$  tais que  $w_j^n = z$ . Mostre que esses números, ditos **raízes enésimas** de  $z$ , estão sobre os vértices de um polígono regular de centro em 0.

**Exercício 23.8** Seja  $u$  um número complexo tal que  $|u| = 1$ . Escrevendo  $u = (\cos \theta, \sin \theta)$ , determine os  $\theta$  para os quais  $\{u^n, n \in \mathbb{Z}\}$  é finito. Mostre que, para os demais valores de  $\theta$ , o conjunto  $\{u^n, n \in \mathbb{Z}\}$  é denso no círculo unitário (isto é, para todo  $z$  com  $|z| = 1$  e para todo  $\varepsilon > 0$  existe  $n \in \mathbb{Z}$  tal que  $|u^n - z| < \varepsilon$ ).

**Exercício 23.9** Observe que, se  $n$  é um inteiro positivo e  $c$  é o círculo definido por  $|z| = r$ , então a imagem de  $c$  pela aplicação  $z \mapsto z^n$  é o círculo de raio  $r^n$  percorrido  $n$  vezes.

**Exercício 23.10** Considere o polinômio  $p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$ , com  $n > 0$ . Mostre que, para todo  $\varepsilon > 0$ , existe  $r_0$  tal que, se  $r > r_0$ , então

$$|z| = r \Rightarrow \frac{1}{|z^n|} |p(z) - z^n| < \varepsilon.$$

Note que isto significa que, se quisermos representar na tela de um computador, simultaneamente, as imagens do círculo  $|z| = r$  por  $z \mapsto z^n$  e por  $z \mapsto p(z)$ , teremos, para  $r$  suficientemente grande, imagens coincidentes.

## 23.2 Inversões

Sejam  $O$  um ponto do plano e  $R$  um número real positivo. Consideremos a transformação  $P \mapsto P'$ , definida, para  $P' \neq O$ , por:  $P'$  é o ponto da semirreta  $\overrightarrow{OP}$  tal que o produto das distâncias  $OP'$  e  $OP$  é  $R^2$ .  $P \mapsto P'$  é dita a **inversão** em relação ao círculo  $c$  de centro  $O$  e raio  $R$ .

**Exercício 23.11** Mostre que a inversão em relação a  $c$  é uma bijeção (do plano menos  $O$  em si mesmo) que deixa invariantes os pontos de  $c$  e traz para dentro de  $c$  os pontos que estão fora, ao mesmo tempo em que leva para fora os que estão dentro (podemos, também, dizer que leva  $O$  para o infinito e traz o infinito para  $O$ ).

**Exercício 23.12** Seja  $r$  uma reta passando por  $O$ . Mostre que a inversão transforma  $r \setminus \{O\}$  em  $r \setminus \{O\}$ .

**Exercício 23.13** Suponha que nosso plano tem um sistema de coordenadas canônico, com origem em  $O$ . Identificando cada ponto  $P$  com seu vetor posição  $\overrightarrow{OP}$ , mostre que a inversão de  $P$  em relação ao círculo de centro  $O$  e raio  $R$  é  $P'$  tal que

$$\overrightarrow{OP'} = \frac{R^2}{|\overrightarrow{OP}|^2} \overrightarrow{OP}.$$

Conclua que a inversão em relação ao círculo de centro  $O$  e raio  $R$  é obtida pela inversão de  $P$  em relação ao círculo de centro  $O$  e raio 1 seguida por uma homotetia de razão  $R^2$ .

As inversões têm três propriedades notáveis não tão evidentes como as enunciadas nos exercícios acima.

1. Inversões transformam retas que não passam por  $O$  em círculos que passam por  $O$  (e, conseqüentemente, círculos que passam por  $O$  em retas que não passam por  $O$ ); em ambos os casos, os círculos são tomados com o ponto  $O$  excluído.

2. Inversões transformam círculos que não passam por  $O$  em círculos (que não passam por  $O$ ).
3. Inversões preservam ângulos.

O leitor está convidado a dar demonstrações puramente geométricas das três propriedades. Seguiremos em outra direção. Seja  $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ . Nosso propósito, aqui, é destacar a relação entre as inversões e a transformação

$$\begin{aligned} \mathbb{C}^* &\longrightarrow \mathbb{C}^*. \\ z &\longmapsto \frac{1}{z}. \end{aligned}$$

**Exercício 23.14** Considere, em  $\mathbb{R}^2$ , a inversão em relação ao círculo de raio 1 e centro na origem. Mostre que a imagem de  $z = (x, y)$  é

$$\frac{1}{x^2 + y^2}(x, y) = \frac{1}{\bar{z}}.$$

Conclua que  $z \mapsto \frac{1}{z}$  é a inversão em relação ao círculo  $|z| = 1$ , seguida da reflexão em relação ao eixo real.

Dado que as propriedades que pretendemos demonstrar, claramente, são invariantes por homotetias e que a inversão em relação ao círculo de centro  $O$  e raio  $R$  é obtida pela inversão de  $P$  em relação ao círculo de centro  $O$  e raio 1 seguida por uma homotetia de razão  $R^2$ , podemos demonstrá-las para o círculo  $|z| = 1$ . Como também não são alteradas por reflexões (em relação a retas passando por  $O$ ), concluímos que basta prová-las para a aplicação  $z \mapsto \frac{1}{z}$ . Desta forma, podemos, nas demonstrações, usar livremente propriedades algébricas dos números complexos.

**Propriedade 1:** Inversões transformam retas que não passam por  $O$  em círculos que passam por  $O$  (e, conseqüentemente, círculos que passam por  $O$  em retas que não passam por  $O$ ); em ambos os casos, exclui-se dos círculos o ponto  $O$ .

Demonstração: Vamos fazer a demonstração para a transformação  $z \mapsto \frac{1}{z}$ . Seja  $r$  uma reta do plano complexo que não passa por 0. Podemos representar os pontos de  $r$  na forma  $u + tv$ ,  $t$  real, sendo  $u$  e  $v$  números complexos fixos, com  $\operatorname{Re}(u\bar{v}) = 0$  (isto corresponde a tomar o vetor  $v$  com a direção de  $r$  e  $u$  normal a  $r$ ; note que a parte real de  $u\bar{v}$  é o produto escalar dos vetores  $u$  e  $v$ ). Afirmamos que a imagem de  $r$  é o círculo de centro  $(2u)^{-1}$  e raio  $(2|u|)^{-1}$ . De fato, para todo  $t$  em  $\mathbb{R}$ :

$$\left| \frac{1}{u + tv} - \frac{1}{2u} \right| = \left| \frac{u - tv}{2u(u + tv)} \right| = \frac{1}{2|u|} \left| \frac{u - tv}{u + tv} \right| = \frac{1}{2|u|}$$

( $u + tv$  e  $u - tv$  têm o mesmo módulo, já que  $u$  e  $v$  são perpendiculares). ■

**Propriedade 2:** Inversões transformam círculos que não passam por  $O$  em círculos (que não passam por  $O$ ).

Demonstração: De novo, vamos fazer a demonstração para a transformação  $z \mapsto \frac{1}{z}$ . Dado um círculo  $c$  que não passa pela origem, podemos tomar um vetor unitário  $u$  (que trataremos, também, como número complexo) tal que a reta  $tu, t \in \mathbb{R}$ , passe pelo centro de  $c$ . A reta em

questão corta  $c$  em dois pontos,  $au$  e  $bu$  (com  $a$  e  $b$  reais). Como o segmento ligando os dois deve ser um diâmetro, temos que  $z$  pertence a  $c$  se, e somente se,  $z - au$  e  $z - bu$  são ortogonais, ou seja (lembre-se de que o produto escalar de dois vetores  $v$  e  $w$ , tomados como números complexos, é  $Re(v\bar{w})$ ):

$$Re((z - au)(\bar{z} - b\bar{u})) = 0.$$

Se o queremos provar é, de fato, verdade, a imagem de  $c$  deve ter um diâmetro ligando  $\frac{1}{au}$  a  $\frac{1}{bu}$ . Basta, então, provar que  $z$  está em  $c$  se, e só se,  $z^{-1}$  está no círculo com esse diâmetro, ou seja:

$$\frac{1}{z} - \frac{1}{au} \perp \frac{1}{z} - \frac{1}{bu}.$$

Calculando o produto escalar entre os dois, temos a parte real de

$$\frac{(au - z)(b\bar{u} - \bar{z})}{auzb\bar{u}\bar{z}}.$$

Como o numerador é  $(z - au)(\bar{z} - b\bar{u})$ , cuja parte real é nula, e o denominador é real, temos o que queríamos. ■

### Propriedade 3: Inversões preservam ângulos.

Demonstração: Esta é uma propriedade que decorre de um resultado bem geral da teoria de funções de variável complexa. Para tranquilizar o leitor, não vamos roubar. Façamos uma demonstração geométrica. A observação crucial é a seguinte: se  $r$  é uma reta que não passa por  $O$ , o círculo que resulta de sua inversão (geométrica, sem usar números complexos) é tangente, em  $O$ , a uma reta paralela a  $r$  (é, basicamente, uma questão de simetria). Assim, se duas retas,  $r$  e  $s$ , se cortam em  $P$  (e nenhuma das duas passa por  $O$ ), os círculos correspondentes se cortam em  $P'$  e em  $O$  (e, é claro, se cortam em  $P'$  e em  $O$  segundo ângulos iguais). Mas as tangentes em  $O$  são paralelas a  $r$  e a  $s$ , o que mostra que o ângulo em  $O$  entre os dois círculos é igual ao ângulo entre  $r$  e  $s$  (em  $P$ ). Se uma das retas,  $r$ , digamos, passa por  $O$ , então podemos supor que a outra,  $s$ , não passa por  $O$ . Assim,  $r$  se transforma em  $r$  e  $s$  se transforma em um círculo,  $c$ , passando por  $O$ , de forma que  $r$  e  $c$  se cortam em  $P'$  e em  $O$ , segundo ângulos iguais. Como a tangente a  $c$  em  $O$  é paralela a  $s$ , o resultado segue. ■

**Exercício 23.15** Sejam  $a, b, c$  e  $d$  números complexos tais que  $ad - bc \neq 0$ . Suponha, também, para ter graça, que  $c \neq 0$ . Mostre que a **transformação de Möbius**  $\tau$  dada por

$$\tau : z \mapsto \frac{az + b}{cz + d}$$

leva retas que não passam por  $-\frac{d}{c}$  em círculos que passam por  $\frac{a}{c}$ , leva círculos que não passam por  $-\frac{d}{c}$  em círculos que não passam por  $\frac{a}{c}$ , retas que passam por  $-\frac{d}{c}$  em retas que passam por  $\frac{a}{c}$  e círculos que passam por  $-\frac{d}{c}$  em retas que não passam por  $\frac{a}{c}$ . Mostre, também, que  $\tau$  preserva ângulos. Sugestão: faça

$$\frac{az + b}{cz + d} = \frac{a}{c} + \frac{bc - ad}{c} \frac{1}{cz + d}$$

e conclua que nossa transformação é composta de homotetias, rotações, translações, reflexões e inversões.

# Capítulo 24

## Equações Polinomiais

### 24.1 Equações polinomiais no Renascimento

O processo de resolução de equações polinomiais do 3º grau é um interessante capítulo da história da Matemática. Embora as fórmulas envolvidas não tenham, aos olhos de hoje, nada de sensacional, o feito, em pleno Renascimento italiano, representa um significativo triunfo: num momento em que ênfase era dada ao retorno aos conhecimentos da antiguidade, os novos sábios europeus podiam enfim apresentar algo que escapara a seus antecessores. De fato, não se tem registro de que alguém, antes de **Scipione del Ferro** (6.II.1465-5.XI.1526), professor da Universidade de Bolonha a partir de 1496, tivesse sido capaz de resolver equações polinomiais gerais de grau superior a 2.

Scipione del Ferro quase levou para o túmulo o processo de que sempre guardou segredo, mas, em seu leito de morte, revelou-o a um discípulo não muito brilhante, Antonio Maria Fior. Tempos depois, a notícia de que existia uma fórmula para resolver equações do 3º grau instiga **Niccolo Tartaglia** (1500-1557), que, por sua vez, obtém de forma independente a solução. A rivalidade Fior-Tartaglia culmina em um *duelo* matemático (1535), em que cada um dos dois propõe ao outro 30 problemas. A vitória de Tartaglia é acachapante:  $30 \times 0$ .

A fama de Tartaglia corre a Itália...Em 1539, **Girolamo Cardano** (1501-1576), em companhia de seu assistente **Ludovico Ferrari** (1522-1560), convence-o a ensinar-lhes o método. Tartaglia, porém, exige que os dois jurem, sobre relíquias sagradas, jamais revelar o segredo. Pouco tempo depois (1540), Ferrari obtém a fórmula da resolução das equações do 4º grau. Ocorre, porém, que o método de Ferrari passa pela redução a uma equação do 3º grau; impedidos pelo juramento sagrado, Cardano e Ferrari não teriam, pois, como divulgar a novidade. Mas uma nova informação vai permitir que o façam, sem terem que pagar o preço do fogo eterno.

Cardano e Ferrari ficam sabendo que Scipione del Ferro deixara, em poder de seu genro Annibale della Nave, anotações que poderiam conter o método de resolução das equações do 3º grau. Bons de conversa, visitam della Nave, que concorda em mostrar-lhes os manuscritos do sogro; esses, de fato, contêm o que buscavam. Considerando que, a partir daí, obtivera a fórmula de del Ferro diretamente da fonte e sem nada jurar,

Cardano se sente livre para publicar, em seu livro **Ars Magna**, tanto a fórmula de del Ferro-Tartaglia como a de Ferrari (que, durante muito tempo, foram conhecidas como as **fórmulas de Cardano**).<sup>1</sup>

## 24.2 A fórmula de del Ferro-Tartaglia

Apresentamos a seguir, sem a preocupação de reproduzir o *Ars Magna*, um método de resolução das equações do 3º grau.<sup>2</sup> Começemos com a equação

$$x^3 + ax^2 + bx + c = 0.$$

Uma primeira substituição,

$$x = y - \frac{a}{3},$$

elimina o termo de 2º grau, conduzindo a uma equação da forma

$$y^3 + py + q = 0.$$

**Exercício 24.1** *Faça a substituição e as contas.*

Uma segunda substituição, menos evidente,<sup>3</sup>

$$y = w - \frac{p}{3w},$$

conduz a

$$(w^3)^2 + qw^3 - \frac{p^3}{27} = 0.$$

**Exercício 24.2** *Faça a substituição e as contas.*

Resolvendo essa última, encontramos dois valores para  $w^3$ , que nos dão, cada um, três valores (complexos) para  $w$ ; esses seis valores de  $w$  devem agora ser substituídos em

$$y = w - \frac{p}{3w},$$

<sup>1</sup>Pode parecer exagerado o zelo de Cardano e Ferrari em manter o juramento feito a Tartaglia, que sequer tinha fama de bom moço. Ainda mais se levarmos em conta a desenvoltura com que juras de amor eterno, feitas diante de um padre na própria casa de Deus, são deixadas para trás nos dias de hoje. Mas são outros tempos...no século XVI, quem quebrasse um tal juramento tinha a certeza de ir para o inferno

<sup>2</sup>Sinalizemos que nossos heróis se beneficiaram dos métodos algébricos desenvolvidos pelos árabes (aí entendidos todos os povos sob domínio árabe durante a Idade Média), que chegaram à Europa a partir do século XII. Também há registro de manuscritos florentinos que, um século antes de del Ferro, já apresentavam a redução da equação do 3º grau a uma outra, equivalente, sem termo de 2º grau

<sup>3</sup>conhecida como substituição de **Vieta**

para, de posse dos correspondentes  $y$ , obtermos os valores de  $x$  por meio de

$$x = y - \frac{a}{3}.$$

**Exercício 24.3** *Mostre que os seis valores de  $w$  se agrupam em três pares, cada par produzindo um único valor para  $y$ .*

## 24.3 A fórmula de Ferrari

Partindo da equação

$$x^4 + ax^3 + bx^2 + cx + d = 0,$$

chegamos, fazendo a substituição

$$x = z - \frac{a}{4},$$

a uma equação da forma

$$z^4 + pz^2 + qz + r = 0.$$

Nossa equação é equivalente, para qualquer  $u$ , a

$$z^4 + z^2u + \frac{u^2}{4} - z^2u - \frac{u^2}{4} + pz^2 + qz + r = 0,$$

ou seja,

$$\left(z^2 + \frac{u}{2}\right)^2 - \left[(u-p)z^2 - qz + \left(\frac{u^2}{4} - r\right)\right] = 0.$$

Ora, o polinômio do segundo grau em  $z$

$$(u-p)z^2 - qz + \left(\frac{u^2}{4} - r\right)$$

terá duas raízes iguais se

$$q^2 - 4(u-p)\left(\frac{u^2}{4} - r\right) = 0.$$

Como esta última é uma equação do terceiro grau em  $u$ , podemos, usando o método que acabamos de aprender, achar-lhe uma raiz  $u$  (se os coeficientes da equação original forem todos reais, podemos, inclusive, escolher  $u$  real). Assim, nossa equação se torna, obtido um tal  $u$  (que só depende de  $p, q$  e  $r$ ),

$$\left(z^2 + \frac{u}{2}\right)^2 - \left(\sqrt{u-p}z - \frac{q}{2\sqrt{u-p}}\right)^2 = 0,$$

cujas raízes são as das duas equações do segundo grau em  $z$ :

$$z^2 + \frac{u}{2} + \sqrt{u-p}z - \frac{q}{2\sqrt{u-p}} = 0$$

e

$$z^2 + \frac{u}{2} - \sqrt{u-p}z + \frac{q}{2\sqrt{u-p}} = 0.$$

Assim, fatoramos nosso polinômio,  $z^4 + pz^2 + qz + r = 0$ , como produto de dois polinômios do segundo grau, o que nos permite: primeiro, concluir que tem quatro raízes; segundo, calculá-las.

## 24.4 Problemas inerentes à solução de del Ferro-Tartaglia

Se a dedução da fórmula não ocupa mais do que uma página, as dificuldades associadas a sua aplicação, ao tempo de del Ferro e Tartaglia, são imensas. Àquela época não existiam sequer os números negativos; muito menos os complexos. Para dar uma primeira noção dos problemas a que a aplicação do método acima poderia conduzir, no século XVI, partamos de uma equação do 3º grau com três raízes reais, distintas e positivas, algo assim como  $(x-1)(x-2)(x-3) = 0$ . Como nossa primeira substituição é uma simples translação, a cada valor de  $x$  corresponde um único  $y$ , e vice-versa. Desta forma, a nova equação,

$$y^3 + py + q = 0,$$

terá, igualmente, três raízes reais distintas. Notemos, porém, que, sendo o coeficiente do termo do segundo grau nulo e igual a menos soma das raízes, teremos, agora, necessariamente, pelo menos uma raiz negativa.

**Exercício 24.4** *Mostre que o coeficiente do termo de segundo grau do polinômio  $(x-a)(x-b)(x-c)$  é  $-(a+b+c)$ .*

Mas esse é apenas um probleminha, perto do que ainda vem...

Examinemos a função  $f(y) = y^3 + py + q$  e vejamos sob que condições suas raízes são reais e distintas. Ora, como

$$\lim_{y \rightarrow -\infty} f(y) = -\infty \text{ e } \lim_{y \rightarrow +\infty} f(y) = +\infty,$$

precisaremos que o gráfico de  $f$ , vindo de  $-\infty$ , *suba* até um valor máximo local positivo e, em seguida, *desça* até um valor mínimo local negativo, antes de voltar a crescer até  $+\infty$ . Calculando a derivada, temos

$$f'(y) = 3y^2 + p,$$

com zeros em

$$\pm \sqrt{\frac{-p}{3}}.$$

Precisamos, pois, que  $p$  seja negativo<sup>4</sup>. Para que o máximo local seja positivo, é preciso que

$$f\left(-\sqrt{\frac{-p}{3}}\right) > 0.$$

Para que o mínimo local seja negativo, por outro lado, a condição é

$$f\left(\sqrt{\frac{-p}{3}}\right) < 0.$$

A primeira desigualdade nos conduz a

$$\frac{2p}{3} \sqrt{\frac{-p}{3}} < q;$$

a segunda, a

$$q > -\frac{2p}{3} \sqrt{\frac{-p}{3}}.$$

Isso equivale a

$$4p^3 + 27q^2 < 0.$$

Assim, se a equação original,

$$x^3 + ax^2 + bx + c = 0,$$

tiver três raízes reais distintas (que podemos até supor positivas), teremos, após a primeira substituição, uma equação da forma

$$y^3 + py + q = 0,$$

com

$$4p^3 + 27q^2 < 0.$$

A segunda substituição nos conduz à equação

$$(w^3)^2 + qw^3 - \frac{p^3}{27} = 0.$$

---

<sup>4</sup>mais um probleminha, mas basta passar o termo  $py$  para o outro lado da equação

Para podermos resolvê-la sem passar pelos números complexos (que não integravam o universo de del Ferro e Tartaglia), seria preciso que

$$4p^3 + 27q^2 \geq 0.$$

Ora, pelo que acabamos de ver, isso jamais ocorrerá, se a equação original tem três raízes reais distintas.

## 24.5 A invenção dos números complexos

Embora o *Ars Magna* já mencione a possibilidade de levar adiante os cálculos, mesmo que a solução da bicúbica conduza à raiz quadrada de um número negativo, quem vai efetivamente escrever o último capítulo dessa história italiana é **Raffaele Bombelli**. Bombelli (1526-1572) publica em 1572 sua **Algebra**, livro em que, pela primeira vez, aparecem números negativos, com a regra dos sinais, e as operações com **números complexos** (adição, subtração e multiplicação), um embrião de notação algébrica sucinta e a utilização dos complexos na solução de equações do 3<sup>o</sup> grau.

Com a possibilidade de operar com números **negativos** e com *números* correspondentes a raízes quadradas de números negativos (que Descartes, em 1637, chamará de **imaginários**), as dificuldades que apresentamos na seção anterior podem, finalmente, ser ultrapassadas.

Assim, em um primeiro momento, a razão para operar com *números* envolvendo raízes quadradas de números negativos era apenas possibilitar a conclusão dos cálculos que conduziram, nos casos *bons*, a soluções reais para as equações do 3<sup>o</sup> grau: as raízes quadradas de números negativos, no fim das contas, se cancelavam, de forma que os *rastros* dos números imaginários utilizados eram apagados. Logo, porém, as propriedades dos complexos foram desenvolvidas e se começa a perceber que faz sentido, pelo menos de um ponto de vista puramente algébrico, falar em raízes complexas para polinômios que, outrora, levavam a equações dadas como sem solução. Em 1629, **Albert Girard** já apresenta o primeiro enunciado do **Teorema Fundamental da Álgebra**, consolidando a ideia de que todo polinômio de grau  $n$  tem direito a suas  $n$  raízes: se não há  $n$  reais, as faltantes devem ser buscadas entre os complexos.

A concepção que hoje temos dos complexos, porém, não vem pronta com Bombelli. A primeira tentativa conhecida de lhes dar uma interpretação geométrica é feita por **Wallis**, apenas em 1673, mas a identificação entre  $\mathbb{C}$  e os pontos de um plano vai esperar muito mais: o símbolo  $i$  e a notação  $a + bi$  foram introduzidos por **Euler** em 1777; em 1799 **Wessel** publica um trabalho em que, finalmente, os complexos são vistos como pontos do plano. A divulgação, porém, é lenta; outros, entre os quais **Gauss** e **Hamilton**, trabalham, de forma independente, com as mesmas ideias, de forma que, apenas nos anos 1830, a identificação entre  $\mathbb{C}$  e  $\mathbb{R}^2$  se torna, de fato, corrente. A expressão **números complexos** é usada pela primeira vez por Gauss, em 1831.

## Capítulo 25

# O Teorema Fundamental da Álgebra

Já no começo do século XVII se conjecturava que qualquer polinômio deveria ter raízes, se não reais, pelo menos, complexas. Esse resultado é o que conhecemos hoje como **Teorema Fundamental da Álgebra**. O caminho até uma demonstração foi longo. A primeira tentativa séria foi feita por **D'Alembert** (1746) - na França, o teorema é conhecido como teorema de D'alembert. Mas só em 1799 Gauss prova o teorema, em sua tese de doutorado; o próprio Gauss apresenta, mais tarde, outras três demonstrações. A demonstração que vamos esboçar, a seguir, está apoiada em ideias topológicas que, esperamos, parecem razoáveis. Antes, o enunciado.

**Teorema Fundamental da Álgebra:** Todo polinômio de coeficientes complexos e grau maior ou igual a um tem raiz em  $C$ .

A própria ideia de buscar uma demonstração do Teorema Fundamental da Álgebra implica na escolha de um caminho diferente do seguido por Del Ferro, Tartaglia, Cardano e Ferrari: renuncia-se à obtenção de uma fórmula geral que resolva todas as equações polinomiais; o que se procura é apenas assegurar que as soluções existem.

Para esboçar a demonstração, fixaremos o grau do polinômio (um inteiro  $m \geq 1$ ) e os coeficientes ( $m + 1$  números complexos,  $a_m, \dots, a_0$ ). Para simplificar as coisas, podemos, sem perda de generalidade, supor  $a_m = 1$  e  $a_0 \neq 0$ . Nosso polinômio, então, será dado por

$$p(z) = z^m + a_{m-1}z^{m-1} + \dots + a_1z + a_0.$$

**Exercício 25.1** *Observe que, se os coeficientes são reais e o grau é ímpar, temos, considerando apenas valores reais de  $z$*

$$\lim_{z \rightarrow -\infty} p(z) = -\infty, \quad \lim_{z \rightarrow \infty} p(z) = \infty,$$

*de modo que, como  $p(z)$  não pode ir, continuamente, de  $-\infty$  a  $+\infty$ , é certo que  $p$  tem ao menos uma raiz real.*

**1ª ideia:**  $p$  pode ser visto como uma função do plano no plano (podemos, conforme a conveniência do momento, pensar o plano como  $\mathbb{R}^2$  ou como  $C$ ).

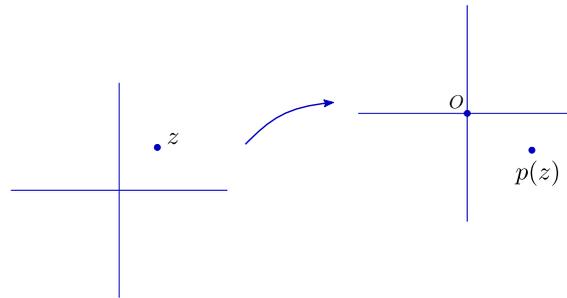


Figura 25.1: 1ª ideia:  $p$  como função de  $\mathbb{R}^2$  em  $\mathbb{R}^2$

Achar uma raiz para  $p$  é, claro, encontrar  $z$  tal que  $p(z) = 0$ . Isso pode ser feito *varrendo* todo o domínio: tomamos todos os pontos  $z$  do plano e checamos se  $p(z) = 0$ . É claro que isso não parece muito razoável...

**2ª ideia:** A imagem por  $p$  de uma curva contínua fechada é uma curva contínua fechada.

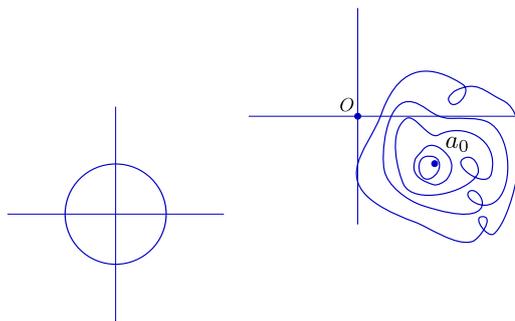


Figura 25.2: 2ª ideia: imagem por  $p$  de curva fechada

**3ª ideia:** O plano pode ser *varrido* por meio de círculos concêntricos, de centro na origem e raio crescente; a imagem de cada círculo  $\gamma_r$  será, então uma curva fechada,  $c_r$ , que se move continuamente sobre o plano, em função da variação do raio.

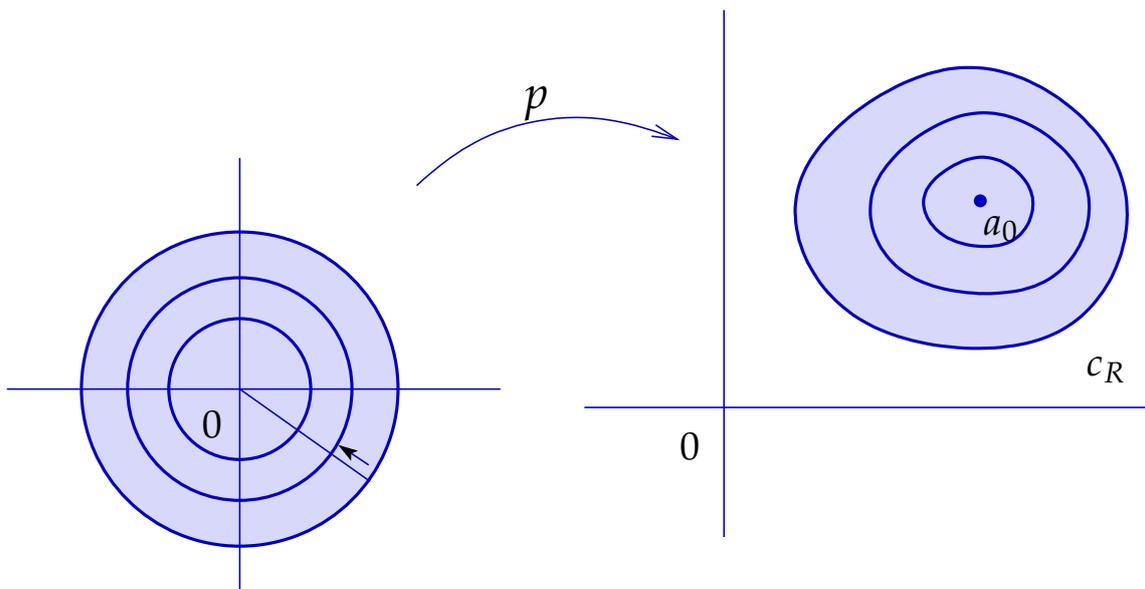


Figura 25.3: 3ª ideia: imagens por  $p$  dos círculos de centro na origem

Demonstrar a existência de uma raiz equivale, então, a demonstrar a existência de um círculo cuja imagem por  $p$  passe pela origem

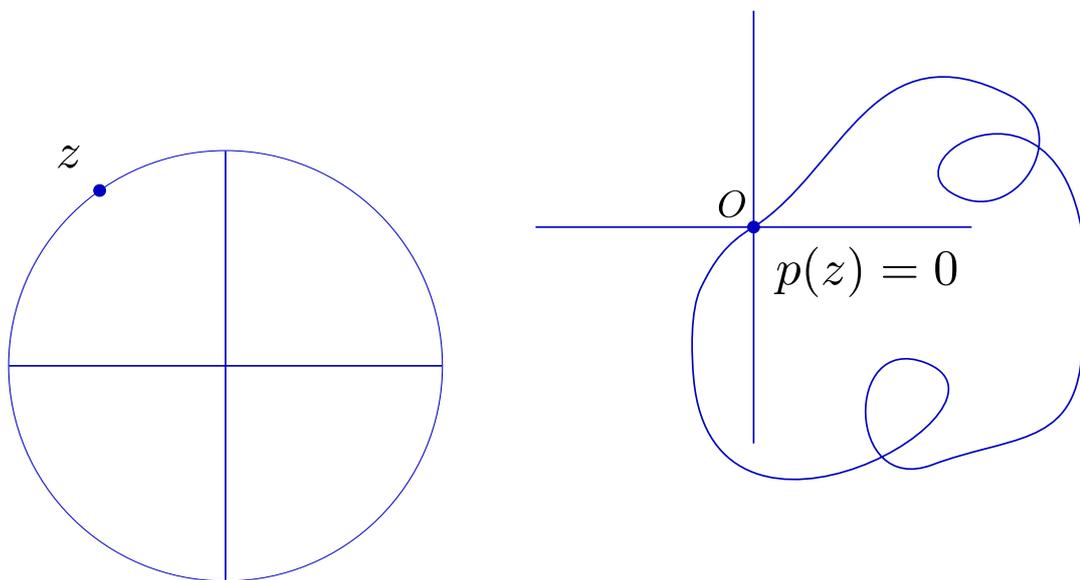


Figura 25.4: raiz de  $p$

A questão é: será que, sendo  $c_r$  a imagem por  $p$  do círculo de raio  $r$  e centro na origem, haverá sempre um raio  $r$  tal que  $c_r$  passa por 0?

Podemos, neste ponto, conceber o seguinte procedimento *experimental*: construímos um programa que, para cada polinômio  $p$  a coeficientes complexos, produza uma animação que mostre a variação da imagem por  $p$  dos círculos de centro na origem e raio  $r$ , acompanhando o crescimento de  $r$ . Vejamos o que nos dão os experimentos.

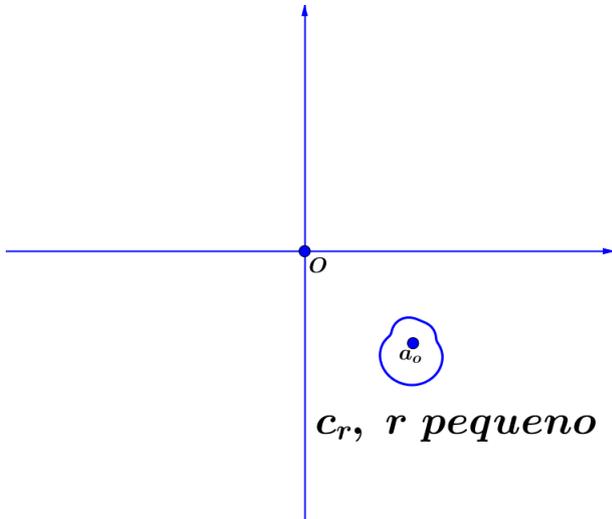


Figura 25.5: Imagem de círculo de raio  $r = 0,35$  pelo polinômio  $p(z) = z^5 + z^4 + (1 + i)z^3 + 3z^2 + iz + 3 - 4i$

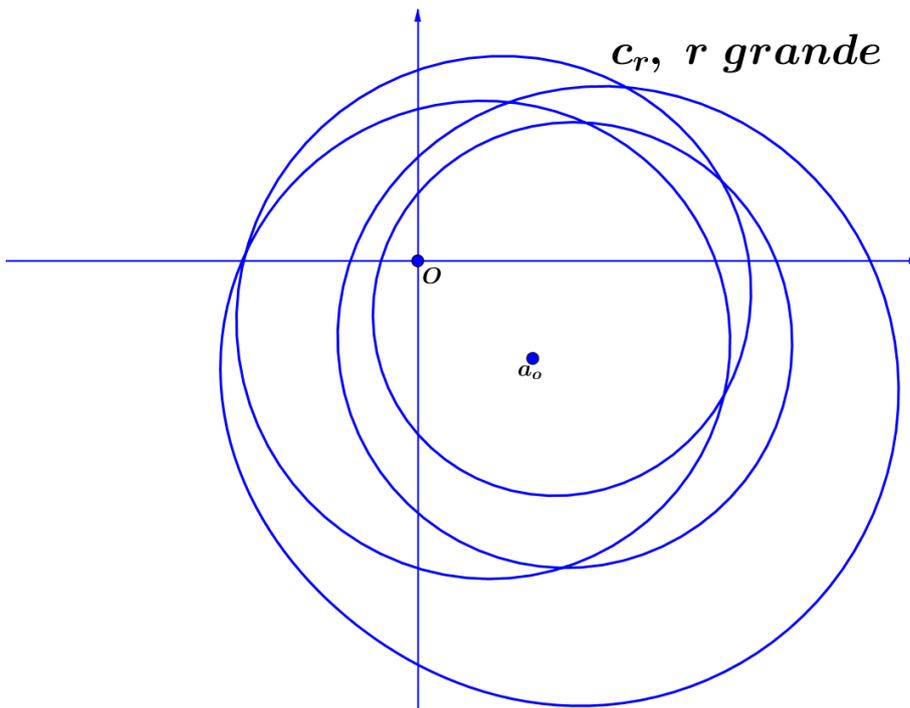


Figura 25.6: Imagem de círculo de raio  $r = 1,04$  pelo polinômio  $p(z) = z^5 + z^4 + (1 + i)z^3 + 3z^2 + iz + 3 - 4i$

O que os experimentos parecem mostrar é: quando  $r$  é pequeno,  $c_r$  fica próximo ao ponto  $a_0$ , que é o termo independente de  $z$ ; já quando  $r$  é grande,  $c_r$  envolve a origem  $O$ .

Ora, não é possível que uma curva contínua e fechada vá, de uma posição em que não envolve  $O$  para outra em que envolve  $O$ , sem passar por  $O$ . Essa variação do **número de voltas** de  $c_r$  em torno de  $O$  parece ser a chave para uma demonstração de nosso Teorema.



# Capítulo 26

## O Número de Voltas

### 26.1 O número de Voltas

Se  $c : [a, b] \rightarrow \mathbb{R}^2$  é uma **curva fechada** (isto é, se  $c(a) = c(b)$ ) que não passa pelo ponto  $P$ , notaremos por  $n(c, P)$  o **número de voltas** que  $c$  dá em torno de  $P$ . Tal número (também chamado de **índice** de  $c$  em relação a  $P$  e notado  $\gamma(c, P)$ ) é um inteiro, positivo de  $c$  gira no sentido trigonométrico, negativo se gira no horário.

Note que o número de voltas depende da parametrização.

**Exercício 26.1** Considere, para cada  $n$  inteiro (positivo, negativo ou nulo), a curva  $c_n : [0, 2\pi] \rightarrow \mathbb{R}^2$  dada por  $c_n(t) = (\cos(nt), \sin(nt))$ . Note que o **traço** de  $c_n$  (isto é, o subconjunto de  $\mathbb{R}^2$  formado pelos pontos  $(\cos(nt), \sin(nt))$ , com  $t$  em  $[0, 2\pi]$ ) é o círculo de equação  $x^2 + y^2 = 1$  (para qualquer  $n$ , desde que  $n \neq 0$ ). Qual é o número de voltas de  $c_n$  em torno da origem? E de  $P = (\sqrt{2}/2, -\sqrt{2}/2)$ ? E de  $P = (1, 2)$ ? Considere, também, o caso  $n = 0$ .

Vamos fazer farto uso de um **princípio fundamental**: se a curva  $c$  e o ponto  $P$  se movem (continuamente no plano sem se cruzarem (isto é, de forma que  $P$  nunca esteja sobre o traço de  $c$ ), então  $n(c, P)$ , o número de voltas de  $c$  em torno de  $P$ , não varia. Para enunciá-lo de forma um pouco mais precisa (mesmo sem dar uma definição precisa de número de voltas), vamos a algumas definições.

### 26.2 Curvas e Homotopias

**Definição 26.2** Uma **curva fechada** é uma função contínua  $c : [a, b] \rightarrow X$ , tal que  $c(b) = c(a)$ . Salvo menção em contrário, suporemos  $X = \mathbb{R}^2$ . Quando nos aprofundarmos um pouco mais no assunto, exigiremos um pouco mais (tipicamente, que  $c$  tenha derivada contínua).

**Definição 26.3** Dadas duas curvas (fechadas ou não),  $c_0$  e  $c_1$ , contínuas,  $c_1, c_0 : [a, b] \rightarrow \mathbb{R}^2$ , a **distância** entre  $c_1$  e  $c_0$  é o número real

$$d(c_1, c_0) = \max \{|c_1(t) - c_0(t)|, t \in [a, b]\}.$$

Ou seja: a distância entre  $c_1$  e  $c_0$  é a maior das distâncias entre os correspondentes pontos  $c_1(t)$  e  $c_0(t)$ , com  $t$  no intervalo (comum) de definição das curvas.

Para dizer que uma curva se movimenta da posição  $c_0$  até a posição  $c_1$  (cada posição, neste caso, é uma curva) criamos uma função que a cada valor de um parâmetro  $s$ , variando de 0 a 1, associa uma curva fechada  $c_s$ . Embora possa parecer um pouco assustador, o fato é que  $c_s$  se move no conjunto  $\mathcal{C}$  das curvas fechadas no plano. Deixaremos implícito que, nesse caso, todas as curvas estão definidas no mesmo intervalo,  $[a, b]$ .

**Definição 26.4** Dadas as curvas fechadas  $c_0, c_1 : [a, b] \rightarrow X$ , uma **homotopia** entre  $c_0$  e  $c_1$  em  $X$  é uma função contínua que a cada  $s$  em um certo intervalo  $[\alpha, \beta]$  associa uma curva fechada  $c_s : [a, b] \rightarrow X$ , com  $c_\alpha = c_0$  e  $c_\beta = c_1$ . A continuidade da aplicação  $s \mapsto c_s$  no valor  $s_0$  do parâmetro, deve ser entendida, caso precisemos de uma definição rigorosa, por:

$$\forall \varepsilon > 0 \exists \delta > 0 \mid |s - s_0| < \delta \implies d(c_s, c_{s_0}) < \varepsilon.$$

**Observação:** Tipicamente, uma homotopia é dada por uma função

$$H : [\alpha, \beta] \times [a, b] \rightarrow X,$$

contínua e tal que:

- (i)  $H(s, a) = H(s, b)$ , para todo  $s$  em  $[\alpha, \beta]$ ;
- (ii)  $H(\alpha, t) = c_0(t)$  e  $H(\beta, t) = c_1(t)$  para todo  $t$  em  $[a, b]$ .

Neste caso, fazemos, claro,  $c_s(t) = H(s, t)$ , para todo  $s$  em  $[\alpha, \beta]$  e para todo  $t$  em  $[a, b]$ .

**Exercício 26.5** Mostre que podemos sempre supor que  $s$  varia no intervalo  $[0, 1]$ , fazendo  $s' = (s - \alpha) / (\beta - \alpha)$ .

## 26.3 O princípio fundamental

**Princípio Fundamental do Número de Voltas:** Sejam  $P$  um ponto de  $\mathbb{R}^2$ ,  $c_0, c_1 : [a, b] \rightarrow \mathbb{R}^2 \setminus \{P\}$  curvas fechadas. Se existe homotopia em  $\mathbb{R}^2 \setminus \{P\}$  entre  $c_0$  e  $c_1$ , então  $n(c_1, P) = n(c_0, P)$ .

Podemos dar ainda uma versão forte:

**Princípio Fundamental do Número de Voltas, versão forte:** Sejam  $c_0, c_1 : [a, b] \rightarrow \mathbb{R}^2$  curvas fechadas. Se  $P : [0, 1] \rightarrow \mathbb{R}^2$  é contínua e existe homotopia em  $s \mapsto c_s$  entre  $c_0$  e  $c_1$ , com  $s \in [0, 1]$  e tal que  $P(s)$  não está no traço de  $c_s$ , qualquer que seja  $s$  em  $[0, 1]$ , então  $n(c_1, P(1)) = n(c_0, P(0))$ .

**Exercício 26.6** Mostre que o princípio forte pode ser reduzido ao anterior (em que o ponto  $P$  permanece fixo): a ideia é que, se  $c$  é uma curva fechada e  $P$  é um ponto, ambos no plano, então  $n(c, P) = n(\bar{c}, O)$ , sendo  $O$  a origem e  $\bar{c}(t) = c(t) - O$  (como não definimos o número de voltas, vamos tomar esse fato como óbvio).

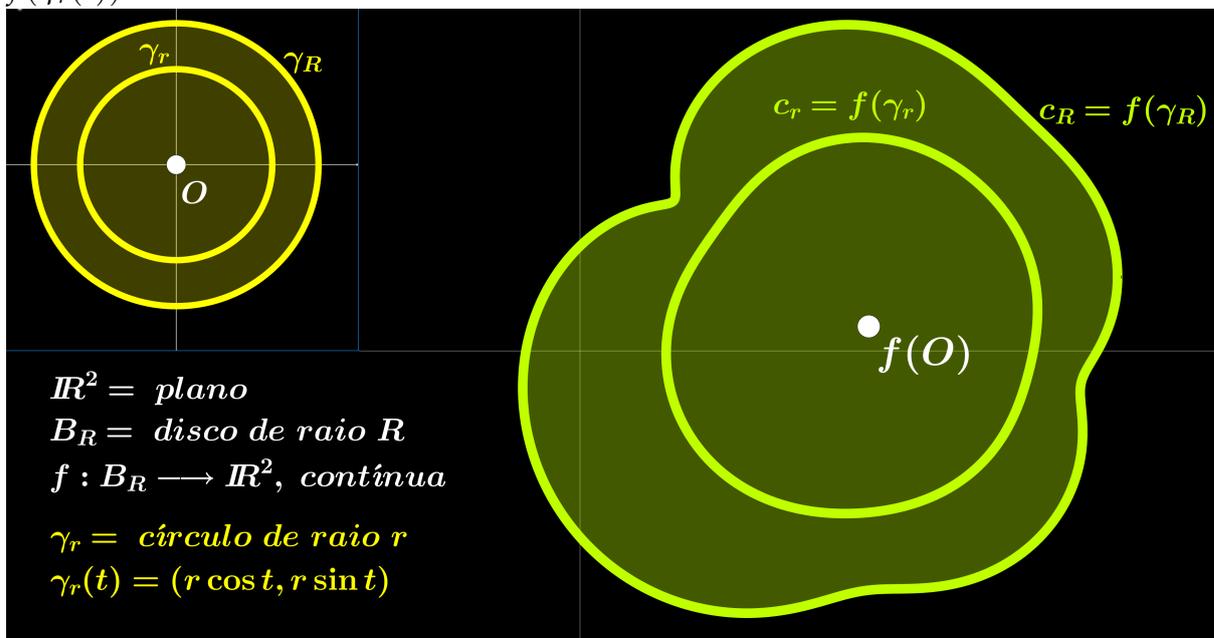
**Exercício 26.7** Compreenda que o princípio forte significa, exatamente, que: se a curva  $c$  e o ponto  $P$  se movem (continuamente no plano sem se cruzarem (isto é, de forma que  $P$  nunca esteja sobre o traço de  $c$ ), então  $n(c, P)$ , o número de voltas de  $c$  em torno de  $P$ , não varia.

## 26.4 Lema de Brouwer

Chamaremos de **Lema de Brouwer** uma consequência do Princípio Fundamental baseada na seguinte ideia: se  $B_R(O)$  é o disco de centro  $O$  e raio  $R$  e  $f : B_R(O) \rightarrow \mathbb{R}^2$  é uma função contínua, então podemos, a partir de  $f$ , criar uma homotopia. Fazemos o seguinte:

(i) para cada  $r$  em  $[0, R]$ , criamos o círculo parametrizado  $\gamma_r : [0, 2\pi] \rightarrow \mathbb{R}^2$ , dado por  $\gamma_r(t) = (r \cos t, r \sin t)$ ;

(ii) usando  $f$ , associamos a cada  $\gamma_r$  a curva  $c_r : [0, 2\pi] \rightarrow \mathbb{R}^2$ , dada por  $c_r(t) = f(\gamma_r(t))$ .



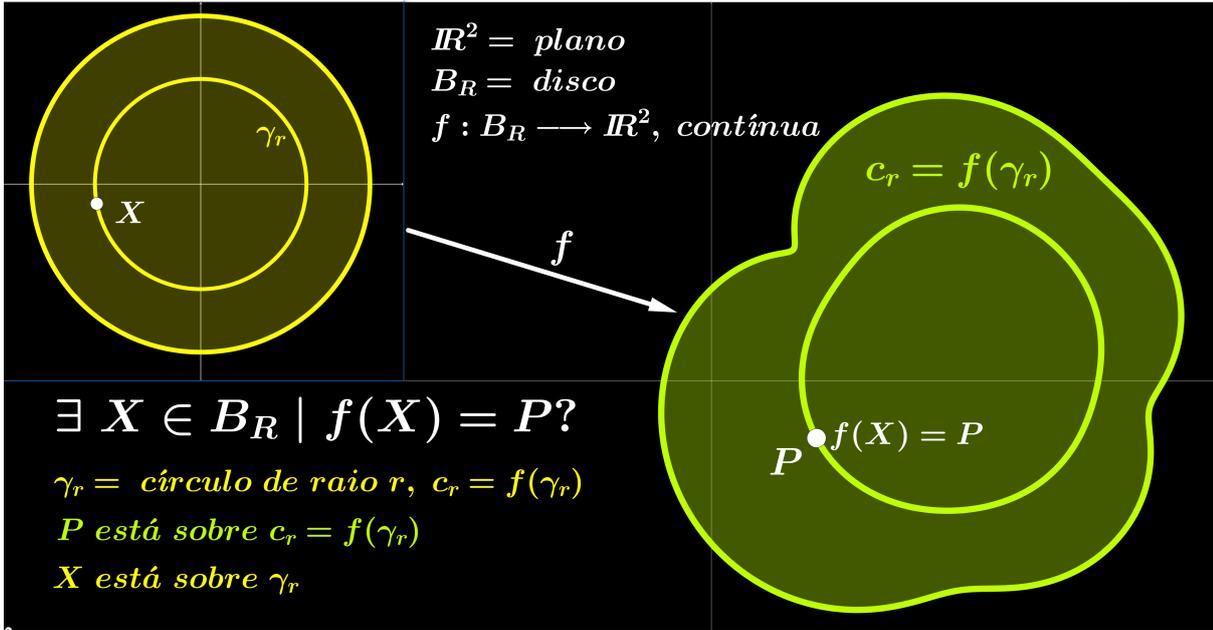
**Exerc cio 26.8** Note que  $r \mapsto c_r$ , com  $r$  variando de 0 a  $R$ ,   uma homotopia entre a curva  $c_0$  que fica parada no ponto  $f(O)$  e a curva  $c_R$  que   a imagem por  $f$  do c rculo de centro  $O$  e raio  $R$ , parametrizada por  $t \mapsto f(R \cos t, R \sin t)$ .

**Lema 26.9** (de Brouwer) Sejam  $B_R(O)$  o disco de centro  $O$  e raio  $R$  e  $f : B_R(O) \rightarrow \mathbb{R}^2$  uma fun o cont nua. Seja  $c_R : [0, 2\pi] \rightarrow \mathbb{R}^2$  dada por  $c_R(t) = f(R \cos t, R \sin t)$ ,  $t \in [0, 2\pi]$ . Se  $P$  n o est  no tra o de  $c_R$  e  $n(c_R, P) \neq 0$ , ent o existe um ponto  $X$  em  $B_R(O)$  tal que  $f(X) = P$ .

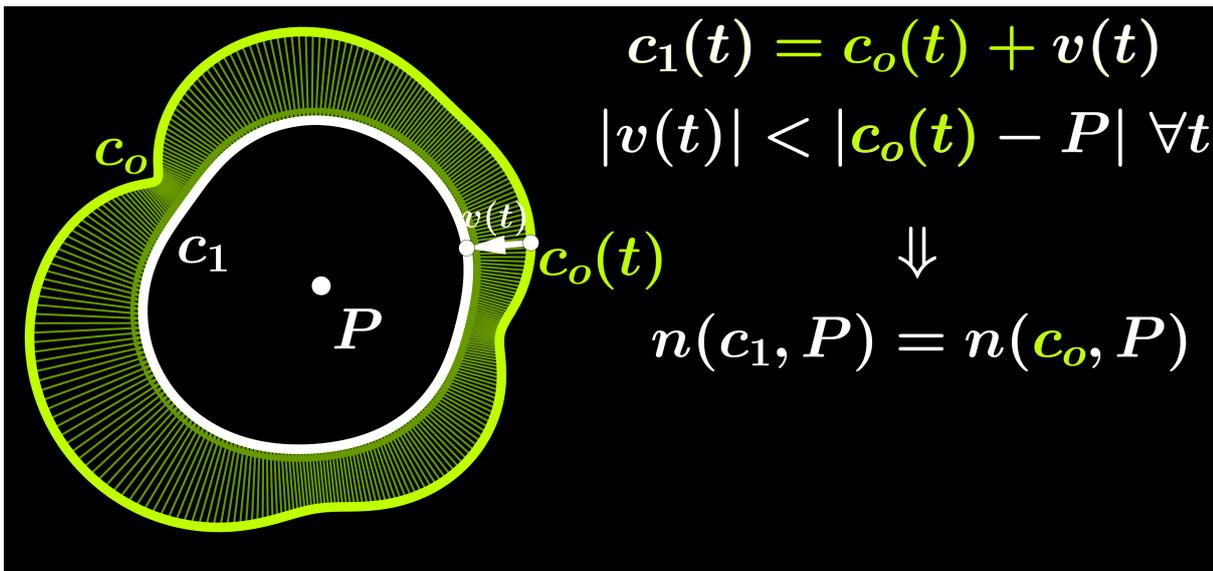
**Demonstra o:** Como acabamos de ver, a aplica o  $r \mapsto c_r$ , com  $c_r(t) = f(r \cos t, r \sin t)$ ,  $t \in [0, 2\pi]$ ,   uma homotopia entre a curva  $c_0$  e  $c_R$ . Se n o existisse  $X$  em  $B_R(O)$  tal que  $f(X) = P$ , ter amos, pelo Princ pio Fundamental,  $n(c_R, P) = n(c_0, P)$ . Mas, como  $c_0$  se reduz a um ponto, temos, trivialmente,  $n(c_0, P) = 0$ , enquanto que, por hip tese,  $n(c_R, P) \neq 0$ .

de

Brouwer.PNG



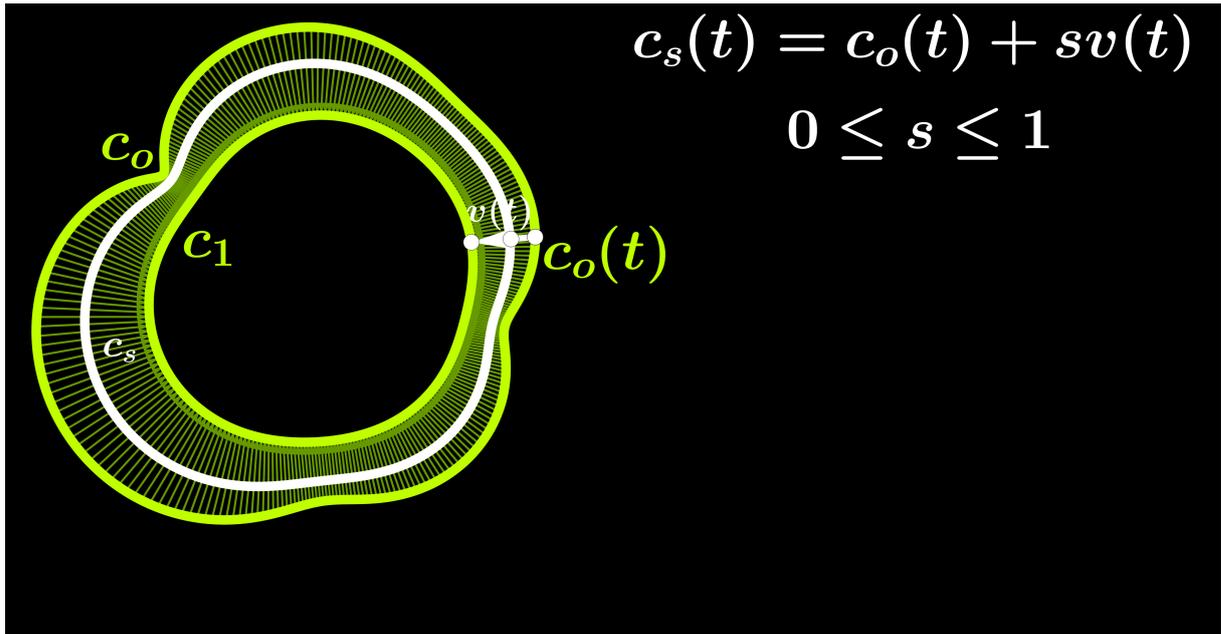
### 26.5 Lema de Rouch\u00e9



Veremos agora uma segunda consequ\u00eancia do Princ\u00edpio Fundamental do N\u00famero de Voltas: se temos uma curva  $c_0$  e um ponto  $P$  que n\u00e3o est\u00e1 sobre  $c_0$  e somamos a cada ponto de  $c_0$ ,  $c_0(t)$ , um vetor  $v(t)$  tal que  $|v(t)| < |c_0(t) - P|$ , obtendo uma nova curva,  $c_1$ , ent\u00e3o  $n(c_1, P) = n(c_0, P)$ .

Para a demonstra\u00e7\u00e3o, basta notar que as curvas  $c_s$ , dadas por  $c_s(t) = c_0(t) + sv(t)$ , com  $s$  em  $[0, 1]$ , n\u00e3o passam por  $P$ , j\u00e1 que, para  $0 \leq s \leq 1$ ,

$$|c_s(t) - P| = |c_0(t) + sv(t) - P| \geq |c_0(t) - P| - s|v(t)| \geq |c_0(t) - P| - |v(t)| > 0 \quad \forall t.$$



Assim, temos uma homotopia entre  $c_0$  e  $c_1$  sem passar por  $P$ . O princípio fundamental, então, nos garante que  $n(c_1, P) = n(c_0, P)$ .

O nome Lema de Brouwer deve ser tomado como uma homenagem e uma alusão ao fato de que é a ideia básica para demonstrar resultados associados ao nome de Brouwer. Desconheço indícios que associem Brouwer a esse resultado. Acredito que Brouwer, inclusive, ficaria desgostoso com a demonstração que apresento para "seu" Lema. O resultado que aqui chamamos de Lema de Rouché é conhecido, no contexto da Teoria das funções de variável complexa, como Teorema de Rouché.



## Capítulo 27

# Demonstração do Teorema Fundamental da Álgebra

Vamos, agora, mais fortes, retomar a demonstração do nosso Teorema

**Teorema 27.1 (Teorema Fundamental da Álgebra):** *Todo polinômio não constante a coeficientes complexos tem raiz em  $\mathbb{C}$ .*<sup>1</sup>

**Demonstração:** Nossa demonstração vai utilizar o número de voltas e aplicar os dois lemas do capítulo anterior. Começemos, pois, fixando nosso polinômio,

$$p(z) = a_m z^m + a_{m-1} z^{m-1} + \cdots + a_1 z + a_0,$$

com  $a_0, \dots, a_m$  números complexos,  $m > 0$ ,  $a_m \neq 0$ . Vamos, também, supor  $a_0 \neq 0$ , pois, caso contrário, 0 seria raiz de  $p$  e não haveria mais nada a provar. Vamos, agora, considerar, para cada real  $r$ , com  $r \geq 0$ , o círculo  $\gamma_r$  de raio  $r$  e centro em 0, dado por

$$\gamma_r(t) = r(\cos t + i \sin t, t \in [0, 2\pi]).$$

Para cada  $r$ , a imagem de  $\gamma_r$  por  $p$  é a curva  $c_r$ , dada por

$$c_r(t) = p(\gamma_r(t)), t \in [0, 2\pi].$$

Estamos, pois, no contexto do **Lema de Brouwer**. Fixamos  $R$  "grande" e olhamos para a função contínua  $p : B_R(O) \rightarrow \mathbb{R}^2 = \mathbb{C}$  (sendo  $O$  o zero dos números complexos, que é identificado a  $(0, 0)$  em  $\mathbb{R}^2$ ), dada pela restrição ao disco  $B_R(O)$  do polinômio  $p$ . O Lema de Brouwer nos garante, então, que, se  $n(c_R, 0) \neq 0$ , então existe  $x$  em  $B_R(O)$  tal que  $p(x) = 0$ .

Nossa demonstração estará concluída, portanto, se mostrarmos que, para  $R$  suficientemente grande,  $n(c_R, 0) \neq 0$ . Vamos, na verdade, usando o **Lema de Rouché**, mostrar que, para  $R$  grande, teremos  $n(c_R, 0) = m$  ( $m$  é o grau do nosso querido polinômio  $p$ ).

Para demonstrar essa última afirmação, vamos, para simplificar, supor  $a_m = 1$  (isto não muda nada, é como dividir por  $a_m$  a equação  $p(z) = 0$ ). Vamos, ainda, reescrever nosso polinômio  $p$  como

---

<sup>1</sup>Ou, como disse o Profeta Malaquias: *o polinômio complexo, se tem  $x$ , tem raiz*

$$p(z) = z^m + q(z), \quad q(z) = a_{m-1}z_{m-1} + \dots + a_0.$$

É importante notar que, se  $|z|$  é suficientemente grande, então  $|z^m| > |q(z)|$ .

**Exercício 27.2** Mostre que, se  $|z| > 1$  e  $|z| > |a_{m-1}| + |a_{m-2}| + \dots + |a_0|$ , então  $|z^m| > |q(z)|$ .

A consequência disto é que, se  $R$  é grande o suficiente para que  $|z| = R$  implique em  $|z^m| > |q(z)|$ , podemos, escrevendo

$$c_R(t) = \gamma_R(t)^m + q(\gamma_R(t)),$$

cair no contexto do Lema de Rouché: rebatizamos de  $c_0$  a curva dada por  $\gamma_R(t)^m$ , que nada mais é que  $c_0(t) = z^m$ , com  $z = R(\cos t + i \sin t)$ ,  $t \in [0, 2\pi]$  e fazemos  $v(t) = q(\gamma_R(t))$ ,  $t \in [0, 2\pi]$ . Note que, como  $|\gamma_R(t)| = R$ , qualquer que seja  $t$  em  $[0, 2\pi]$ , temos

$$|v(t) < |c_0(t)| \quad \forall t \in [0, 2\pi].$$

Ora, bem sabemos que nossa  $c_0$  dá  $m$  voltas em torno de 0. O **Lema de Rouché** nos dá, então, que  $n(c_R, 0) = m$ , como queríamos demonstrar.

**Escólio:** Para uma abordagem mais focada na tela do computador, suponha que  $r$  seja muito grande e que queiramos ver, inteira, a curva  $c_r$ . Ora, para  $|z| = r$ ,  $r$  grande, o maior termo em  $p(z) = z^m + \dots + a_0$  é  $z^m$ ; se  $r$  for muito grande, mesmo, a diferença entre  $p(z)$  e  $z^m$  pode ser minúscula, face a  $z^m$ . Esse é o significado de

$$\lim_{|z| \rightarrow \infty} \frac{|p(z) - z^m|}{|z^m|} = 0.$$

**Exercício 27.3** Entenda isso. Note que  $|p(z) - z^m| \leq |a_{m-1}||z|^{m-1} + \dots + |a_0|$  e que, portanto,

$$\frac{|p(z) - z^m|}{|z^m|} \leq \frac{|a_{m-1}||z|^{m-1} + \dots + |a_0|}{|z|^m} = \frac{|a_{m-1}|}{|z|} + \dots + \frac{|a_0|}{|z|^m} \xrightarrow{|z| \rightarrow \infty} 0.$$

**Exercício 27.4** Entenda que, se escolhermos uma escala adequada para que  $c_r$  apareça no monitor, poderemos até ter, se  $r$  for bem grande,  $p(z)$  e  $z^m$  ocupando o mesmo pixel.

**Exercício 27.5** Lembre-se de que a imagem, pela função  $z \rightarrow z^m$ , do círculo de centro na origem e raio  $r$ , percorrido uma vez, é o círculo de centro na origem e raio  $r^m$ , percorrido  $m$  vezes.

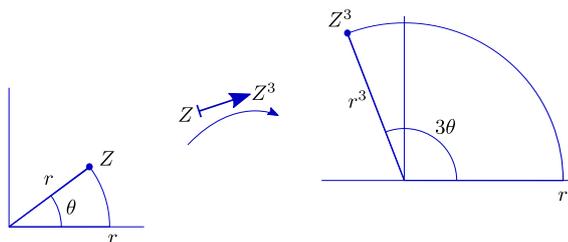


Figura 27.1: imagem do círculo de centro em 0 e raio  $r$  por  $z \mapsto z^3$

**Exercício 27.6** Pense que, quando percorremos uma vez o círculo  $\gamma_r$ , de raio  $r$ , o círculo de raio  $r^m$ , percorrido  $m$  vezes, corresponde à órbita de um planeta em torno de um sol situado na origem; a imagem de  $\gamma_r$  por  $p$  corresponde, então à órbita de um satélite. A distância entre o satélite e o planeta, dada por

$$|a_{m-1}z^{m-1} + \dots + a_0|,$$

é extremamente pequena, se comparada a  $r^m$ , de modo que o planeta é, também, forçado a dar  $m$  voltas em torno da origem.

Juntando tudo, concluímos que, quando  $r$  vai de 0 a infinito,  $n(c_r, 0)$  passa de 0 a  $m$ , o que só pode ocorrer se, para algum  $r$ ,  $\gamma_r$  passar pela origem. Logo, existe  $z$  em  $\mathbb{C}$  tal que  $p(z) = 0$ .

**Software:** Esta seção vem acompanhada do software TFA. Com ele você pode digitar os coeficientes de seu polinômio predileto e visualizar as ideias apresentadas acima. O link a seguir, quando funciona, abre o software TFA. [TFA](#)

Se o link não funciona, você pode baixar o TFA.jar diretamente do link [TFA.jar](#)



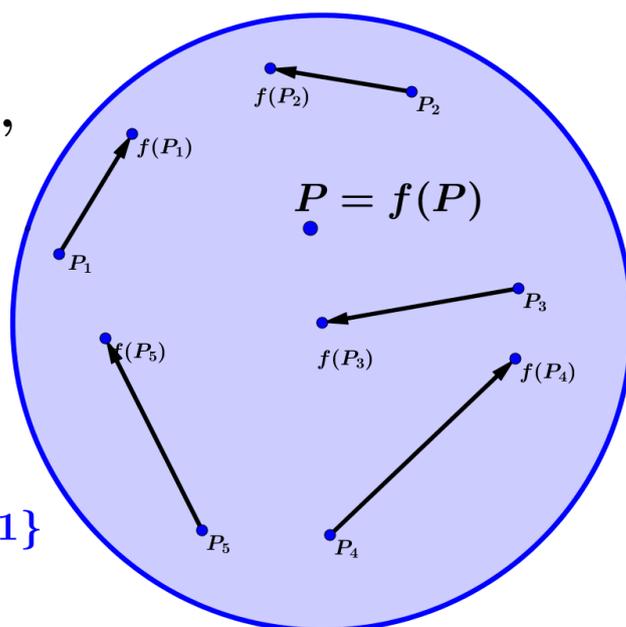
# Capítulo 28

## O Teorema de Brouwer

Este curto capítulo tem como propósito aproveitar as ideias que desenvolvemos anteriormente para demonstrar um famoso teorema.

*Teorema de Brouwer :*  
*Se  $f : B \rightarrow B$  é contínua,*  
*então existe  $P$  em  $B$*   
*tal que  $f(P) = P$ .*

$$B = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$$



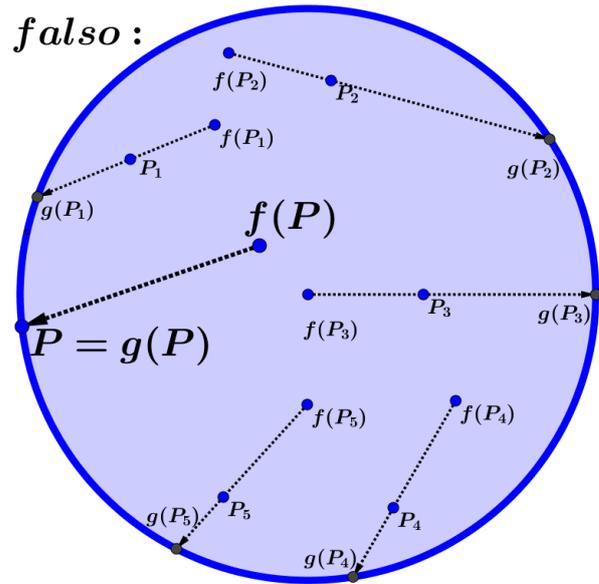
O que aconteceria se o Teorema de Brouwer fosse falso?

Se o Teorema de Brouwer fosse falso :

existiria  $f : B \rightarrow B$  contínua,  
com  $f(P) \neq P \forall P \in B$ . Seja,  
então,  $g : B \rightarrow S$ , dada por

$$g(X) = S \cap \text{semirreta } \overset{\text{-----}}{f(X)X} .$$

$g : B \rightarrow S$  seria  
contínua e tal que  
 $g(P) = P \forall P \in S$ .



$$S = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

Ocorre, porém, que uma tal  $g$  não pode existir

**Teorema 28.1** (de irretratabilidade): Não existe  $g : B \rightarrow S$ , contínua e tal que  $g(P) = P$  para todo  $P$  em  $S$

**Demonstração:** Uma tal  $g$  seria uma função contínua de  $B$  no plano tal que a imagem por  $g$  do círculo  $\gamma(t) = (\cos t, \sin t) \ t \in [0, 2\pi]$  dá uma volta em torno de  $(0, 0)$ . O Lema de Brouwer nos garante que, neste caso, existiria  $X$  em  $B$  tal que  $g(X) = (0, 0)$ , o que contraria a hipótese, já que  $g$  leva  $B$  em  $S$ . ■

**Definição 28.2** Dois subconjuntos  $A$  e  $B$  do plano são ditos **homeomorfos** se existe uma bijeção contínua

$$h : A \longrightarrow B$$

com inversa,  $h^{-1} : B \rightarrow A$ , contínua.  $h$ , caso exista é dita um **homeomorfismo** entre  $A$  e  $B$ .

**Corolário 28.3** Se  $K$  é homeomorfo a  $B$  e  $\varphi : K \rightarrow K$  é contínua, então existe  $Q$  em  $K$  tal que  $\varphi(Q) = Q$ .

**Demonstração:** Considere  $h : B \rightarrow K$  homeomorfismo e defina  $f : B \rightarrow B$  por  $f(X) = h^{-1}(\varphi(h(X)))$ . Como é composta de funções contínuas,  $f$  é contínua. Segue do Teorema de Brouwer que existe  $P$  em  $B$  tal que  $f(P) = P$ . Basta, então, tomar  $Q = h(P)$ . ■

**Exercício 28.4** Encontre subconjuntos do plano homeomorfos a  $B$ .

# Capítulo 29

## Iterações

### 29.1 De Ferrari a Galois

Encontrados métodos para a resolução, por radicais, das equações do terceiro e do quarto graus, a bola da vez passou a ser a equação do quinto grau. Muita água rolou por baixo da ponte, passaram-se quase 300 anos. Os italianos ainda deram contribuições importantes: em 1770, Lagrange (Giuseppe Lodovico Lagrangia, 1736-1813) chama a atenção para o fato de que os processos de solução das equações de grau inferior ao quinto envolvem um truque, com permutações das raízes, que não funciona nas do quinto grau; em 1799, Paolo Ruffini (1765-1822) afirma a impossibilidade.

**Teorema de Ruffini:** *Não é possível, em geral, resolver por radicais as equações do quinto grau.*

A prova de Ruffini não é convincente; considera-se que o primeiro a, de fato, provar a impossibilidade foi o norueguês Niels Henrik Abel (1802-1829), em 1824.

Mas o trabalho mais espetacular é o do francês Évariste Galois (1811-1832). Morto em um duelo aos 20 anos, Galois deixa um testamento matemático que só vai ser reconhecido mais de dez anos depois. Resumidamente, o trabalho de Galois mostra que se pode associar, a cada polinômio, um grupo de permutações (de suas raízes), chamado **grupo de Galois** do polinômio em questão. O polinômio terá soluções por radicais se, e somente se, seu grupo de Galois for **solúvel** (sem entrar em detalhes: se  $G$  é o grupo de Galois do polinômio  $p$ ,  $p$  terá solução por radicais se e só se houver uma cadeia  $G = G_0 \triangleright G_1 \triangleright \dots \triangleright G_n = I$ , sendo  $I$  o grupo trivial, e cada  $G_j$  subgrupo normal de  $G_{j-1}$ , com  $G_{j-1}/G_j$  cíclico). O Teorema de Ruffini decorre, então, do fato que existem polinômios, a partir do quinto grau, cujos grupos de Galois não são solúveis.

### 29.2 Soluções aproximadas

As fórmulas de resolução por radicais para as equações do segundo, do terceiro e do quarto graus podem ser tranquilizadoras, mas, se considerarmos que um número é dado por sua representação decimal (ou em qualquer outra base), temos que

reconhecer que o que nos fornecem, mesmo no caso de equações tão simples como  $x^2 - 2 = 0$ , são indicações de como podemos obter sequências de aproximações das raízes. Afinal, números irracionais, como  $\sqrt{2}$ , serão sempre representados, qualquer que seja a base adotada para o sistema de numeração, por sequências infinitas de algarismos.

Tudo bem,  $\sqrt{2}$  é irracional, mas podemos, dada a unidade, construí-lo com régua e compasso, como hipotenuza de um triângulo retângulo de catetos iguais a 1. Mas há outros números obtidos por radicais, como  $\sqrt[3]{2}$ , que não têm a mesma sorte (a impossibilidade de se construir  $\sqrt[3]{2}$  com régua e compasso, embora elaborada, se demonstra com ferramentas algébricas relativamente simples). De qualquer forma, é razoável considerarmos que resolver uma equação  $p(x) = 0$ , na verdade, consiste em:

(i) fornecer uma fórmula que permita calcular, para cada natural  $n$ , um número racional  $a_n$ , de forma que a sequência dos  $a_n$  satisfaça

$$\lim_{n \rightarrow \infty} a_n = a,$$

com  $p(a) = 0$ ;

(ii) fornecer uma segunda fórmula que permita, para cada real positivo  $\varepsilon$ , determinar  $n$  tal que  $|a_n - a| < \varepsilon$  (e, sendo mais exigentes, tal que, também,  $|p(a_n)| < \varepsilon$ ).

## 29.3 A iteração de Newton

### videoaula O método de Newton

Nesta direção, **Isaac Newton** (1642-1727), apresenta um método iterativo para o cálculo aproximado das raízes de polinômios de qualquer grau (a convergência do método, porém, depende de uma boa escolha do *chute* inicial). A ideia de Newton é extremamente simples. Considere o polinômio

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$$

e chute uma raiz aproximada inicial  $x_0$ . A raiz verdadeira é  $x_0 + h$ ; substituindo-a na equação, temos, desenvolvendo cada potência de  $(x_0 + h)$ ,

$$\begin{aligned} 0 = p(x_0 + h) &= p(x_0) + n a_n x_0^{n-1} h + (n-1) a_{n-1} x_0^{n-2} h + \\ &+ \dots + 2 a_2 h + a_1 h + O(h^2) = p(x_0) + p'(x_0) h + O(h^2), \end{aligned}$$

chamando  $n a_n x_0^{n-1} + (n-1) a_{n-1} x_0^{n-2} + \dots + 2 a_2 + a_1$  de  $p'(x_0)$  e representando por  $O(h^2)$  os termos que envolvem potências de  $h$  superiores à primeira. A ideia é que, se

$x_0$  é uma boa aproximação, então  $h$  é pequeno, e  $O(h^2)$  é desprezível. Podemos, então, tomar  $h_1$  como solução de

$$p(x_0) + p'(x_0)h = 0,$$

o que nos dá um novo chute,

$$x_1 = x_0 + h_1 = x_0 - \frac{p(x_0)}{p'(x_0)}.$$

Reiterando repetidas vezes, com a fórmula

$$x_{n+1} = x_n - \frac{p(x_n)}{p'(x_n)},$$

podemos esperar que

$$\lim_{n \rightarrow \infty} x_n = x,$$

com  $p(x) = 0$ .

Newton não tinha computadores...nós temos. Ainda hoje, o método de Newton é a base para boa parte dos processos de cálculo aproximado de raízes das mais variadas equações (afinal, trocando em miúdos, o método utiliza a aproximação linear de  $p(x+h)$  por  $p(x) + p'(x)h$ , ou seja, o processo de diferenciação).

**Exercício 29.1** Faça um desenho mostrando que o método consiste em aproximar  $p(x)$  por  $p(x_0) + p'(x_0)(x - x_0)$ .

**Exercício 29.2** Sejam  $a > 1$ ,  $n > 1$  e  $p(x) = x^n - a$ . Mostre que, para qualquer  $x_0 \geq a$ , o método converge.

## 29.4 A Ferradura de Smale

 [videoaula Ferradura de Smale](#)

 [video Ferradura de Smale](#)

Vamos, nesta seção, apresentar uma situação, importante na Teoria dos Sistemas Dinâmicos, conhecida como a **Ferradura de Smale**. Embora as situações em que aparecem possam de nível matemático bem acima do deste curso, é possível, despidendo do contexto, apresentá-la em um nível razoavelmente elementar.

