

ANÁLISE **em dimensões superiores**

Dietmar A. Salamon
ETH Zürich
12. September 2017

Tradutor do alemão ao português
Joa Weber
UNICAMP

11 de junho de 2020

Prefácio

Nestas notas trata-se de um manuscrito para o curso “*Analysis II*” para alunos do segundo semestre nas áreas da matemática, da física, e da engenharia na ETH Zürich. Serão pressuposto os conteúdos dos cursos “*Analysis I*” e “*Lineare Algebra I*” do primeiro semestre.

O manuscrito foi produzido para o curso Analysis II no D-ITET no primeiro semestre de 2017. Neste curso foram tratado no Capítulo 1 só espaços de Banach de dimensão finita e normas matriz, os resultados de Capítulo 4 foram resumidos sem demonstrações, e Capítulo 7 sobre formas diferenciais e o Teorema de Stokes não foram tratados.

O apêndice contém tanto material que é pressuposto conhecido de outros cursos, como material que pertence ao conteúdo deste curso. Na última categoria pertencem o Apêndice A sobre fecho e interior de subconjuntos de um espaço métrico e o Apêndice B sobre espaços métricos conexos. Com respeito à primeira categoria será pressuposto da *Álgebra Linear* o Apêndice D sobre o determinante e serão pressupostos da *Análise I* partes do Apêndice C sobre espaços métricos compactos assim como o Apêndice E sobre o Teorema do Ponto Fixo de Banach. (O Teorema de Arzelà-Ascoli em Apêndice C nem foi tratado na aula, nem é usado no texto.)

Dia 4 de maio 2017

Dietmar A. Salamon

Prefácio do tradutor

Investigando a literatura para o curso MA/MM-720 “*Análise no \mathbb{R}^n* ” no primeiro semestre de 2019 na UNICAMP encontrei o manuscrito “*Analysis II*” no site do autor. Como é comum o assunto é tratado em duas partes, diferenciação e integração. Gostei muito que o autor trata a parte da diferenciação *livre de coordenadas locais*. Neste parte X é um espaço vetorial de dimensão finita e munido de uma norma qualquer. Na minha opinião \mathbb{R}^n não é um ambiente bom para introduzir a derivada. A razão, talvez surpreendente na primeira vista, é a existência de uma base canônica e de um produto interno canônico. Isso complica o entendimento, não só da derivada. Como assim? Objetos totalmente diferentes como, por exemplo, a derivada e o gradiente de uma função parecem ser iguais. Não poucos alunos pensam que são listas com n entradas e a diferença seria escrever a primeira lista ‘horizontal’ e a segunda ‘vertical’. Neste jeito a significância da derivada é obscurecida. Também é perdido a consciência que o gradiente só existe na presença de um produto interno e que o valor dele é ser um campo vetorial e consequentemente gerar um fluxo, em outras palavras, um sistema dinâmico. Por isso temos substituído o título tradicional, mas bastante constringindo, “*Análise no \mathbb{R}^n* ”, pelo título “*Análise em dimensões superiores*” abrangendo ambas – modernidade e tradição.

O assunto do primeiro parte é análise em espaços vetoriais normados na dimensão finita. Baseado na dimensão finita tais espaços são automaticamente completo, ou seja, toda sequência de Cauchy tem um elemento limite. Um espaço vetorial normado, seja a dimensão finita ou não, com esta propriedade é chamado de *espaço de Banach*. O primeiro parte de análise em dimensões superiores trata de derivada e diferenciação e o segundo parte de integral e integração. O primeiro parte – em contraste extremo ao segundo – abre os braços para dimensão infinita essencialmente sem custo extra.

Para a definição de derivada não precisa-se nenhuma hipótese sobre a dimensão. E a demonstração (longa) do teorema da função inversa é igual

em dimensão finita e infinita com a exceção de uma única aplicação do teorema da inversa limitada, o qual simplesmente citamos. Para não implantar na memória hipóteses desnecessariamente restritivas vamos formular os resultados, em casos sem custo adicional, em termos de espaços de Banach. Consideramos gratuito uma referência, por exemplo, ao teorema da inversa limitada. Isso aplica-se para § 2.1–2.5 e § 3.1–3.2. Começando com § 3.3, sobre subvariedades, a situação muda drasticamente – agora dimensão infinita requer métodos muito diferentes – e por isso vai-se impor a condição de dimensão finita imediatamente.

Analogamente não faz sentido restringir-se a séries de *matrizes quadradas* se todo que importa só é que os operadores lineares e suas composições sejam limitados e o espaço vetorial normado seja completo. Tem-se que entender assim o nível de generalidade em § 1.3–1.6. Em § 1.7 a integral 1-dimensional com valores vetores é disponibilizado para seu uso em Capítulo 2.

No semestre 2019-1 temos tratado na aula todos os assuntos, menos o Capítulo 4. Em Capítulo 5 temos omitido as provas com exceção das Seções 5.7 e 5.9. Temos iniciado o curso revisando os apêndices.

Agradeço muito a grande confiança do autor refletido no fato de permitir a tradução do texto dele da língua alemã ao português, e ainda concordar com a proposta de introduzir a derivada e o teorema da função inversa no ambiente natural – espaços de Banach, seja a dimensão finita ou não.

Em outras palavras, a restrição ao caso de dimensão finita no texto original foi removido pelo tradutor nas hipóteses das Seções 1.7, 2.1–2.5, e 3.1–3.2. Usamos como insumo externo (na dimensão infinita só) uma vez cada um dos seguintes: Os teoremas da inversa limitada e da extensão limitada, e um corolário do teorema de Hahn-Banach. (Localize no índice remissivo.) O tradutor leva toda a responsabilidade.

Destaco que a texto e as provas tratam análise na dimensão finita.

Agradecimentos. Com prazer agradeço aos alunos Luiz Augusto Silvestre Luccas e Matias Zimmermann pela ajuda com a língua portuguesa e erros tipográficos, e à grande maioria da turma do curso MA/MM-720 em 2019-1 na UNICAMP pelo interesse na matéria e uma ambiência muito agradável.

Campinas,
Junho 2020

Joa Weber

Sumário

1	Espaços de Banach e Operadores Lineares	1
1.1	Normas equivalentes	1
1.2	A norma operador	8
1.3	Álgebras de Banach	14
1.4	A série geométrica	21
1.5	A aplicação exponencial	23
1.6	Composição de séries de potências	27
1.7	Integração com valores vetoriais	30
2	Aplicações Diferenciáveis	37
2.1	Conceitos básicos	37
2.2	Diferenciabilidade contínua	49
2.3	A regra da cadeia	53
2.4	O Teorema da Barreira	65
2.5	Derivadas superiores	69
2.6	A série de Taylor	78
2.7	Extremos locais	89
3	Funções Implícitas	101
3.1	O Teorema da Aplicação Inversa	101
3.2	O Teorema da Função Implícita	111
3.3	Subvariedades	118
3.4	Extremas sob restrições	130
4	Campos de Vetores e Fluxos	135
4.1	Existência e unicidade	135
4.2	O fluxo de um campo de vetores	144
4.3	Continuidade do fluxo	147

4.4	Diferenciabilidade do fluxo	151
5	Integrais Múltiplas	159
5.1	A integral de Riemann	159
5.2	Conjuntos Jordan-nulos e continuidade	167
5.3	Somas de Riemann	173
5.4	Integrais iteradas	182
5.5	Conjuntos Jordan-mensuráveis	187
5.6	Conjuntos Jordan-mensuráveis – mais	191
5.7	A fórmula da transformação de variáveis	201
5.8	Integrais impróprias	220
5.9	Divergência	226
6	Integração ao longo de Subvariedades	231
6.1	O volume d-dimensional	231
6.2	Definição da integral	242
6.3	Exemplos	252
6.4	O Teorema de Gauß	262
6.5	Mais exemplos	273
7	Integração de Formas Diferenciais	279
7.1	Formas alternadas	279
7.2	Formas diferenciais	285
7.3	Orientação	294
7.4	Integração	300
7.5	O Teorema de Stokes	305
7.6	Exemplos e aplicações	311
A	Fecho e Interior	319
A.1	Conceitos básicos da topologia	319
A.2	Os conceitos fecho e interior	321
B	Espaços Conexos	325
B.1	A topologia relativa	325
B.2	O conceito conexo	327
B.3	Conjuntos conexo por caminho	330
B.4	Componentes conexos	332
B.5	Exemplos	333

C Espaços Métricos Compactos	335
C.1 O conceito de compacidade	335
C.2 O Teorema de Arzelà–Ascoli	339
D O Determinante	345
D.1 Grupos	345
D.2 Espaços vetoriais	349
D.3 O Determinante	352
D.4 O grupo linear geral	362
E O Teorema de Banach do Ponto Fixo	367
Bibliografia	369
Índice Remissivo	371

Capítulo 1

Espaços de Banach e Operadores Lineares

1.1 Normas equivalentes

Esta primeira seção tem por objetivo recordar algumas definições e resultados conhecidos do curso Análise I. Seja X um espaço vetorial real. Uma **norma** em X é uma aplicação $X \rightarrow \mathbb{R} : x \mapsto \|x\|$ com as seguintes propriedades.

- (i) Para todo $x \in X$ vale $\|x\| \geq 0$ e $\|x\| = 0 \iff x = 0$.
- (ii) Para todo $x \in X$ e todo $\lambda \in \mathbb{R}$ vale $\|\lambda x\| = |\lambda| \|x\|$.
- (iii) Para todos os $x, y \in X$ vale $\|x + y\| \leq \|x\| + \|y\|$.

Um **espaço vetorial normado** é um par $(X, \|\cdot\|)$ composto de um espaço vetorial real X e uma função norma $\|\cdot\|$ em X . Uma classe importante de espaços vetoriais normados são aqueles cuja norma é gerada por um produto interno. Um **produto interno** em X é uma aplicação bilinear simétrica $X \times X \rightarrow \mathbb{R} : (x, y) \mapsto \langle x, y \rangle$, a qual satisfaz a condição $\langle x, x \rangle > 0$ para todos os $x \in X \setminus \{0\}$. Dado tal produto interno, então a fórmula

$$\|x\| := \sqrt{\langle x, x \rangle} \quad \text{para } x \in X \quad (1.1.1)$$

define uma norma e vale a desigualdade de Cauchy–Schwarz $|\langle x, y \rangle| \leq \|x\| \|y\|$ para todos os $x, y \in X$.

Seja $(X, \|\cdot\|)$ um espaço vetorial normado. Ora a fórmula

$$d(x, y) := \|x - y\| \quad \text{para } x, y \in X \quad (1.1.2)$$

define uma função distância $d : X \times X \rightarrow \mathbb{R}$ a qual faz de X um espaço métrico. O espaço vetorial normado $(X, \|\cdot\|)$ é chamado de **espaço de Banach** se o espaço métrico associado (X, d) é **completo**, ou seja, se toda sequência de Cauchy converge. (lembrando que uma sequência $(x_n)_{n \in \mathbb{N}}$ em X é uma **sequência de Cauchy**, se para todo $\varepsilon > 0$ existe um $n_0 \in \mathbb{N}$, tal que para todos os $n, m \in \mathbb{N}$ vale: $n, m \geq n_0 \implies d(x_n, x_m) < \varepsilon$.) Um **espaço Hilbert (real)** é um par $(X, \langle \cdot, \cdot \rangle)$ composto de um espaço vetorial real X e um produto interno $\langle \cdot, \cdot \rangle$ em X , tal que o espaço métrico associado (X, d) é completo com respeito à função distância dada através de (1.1.1) e (1.1.2). Consequentemente todo espaço Hilbert é um espaço de Banach com respeito à função norma (1.1.1).

Ora, seja $(X, \|\cdot\|)$ um espaço vetorial normado. Dado $x, y \in X$ e $r > 0$, chama-se de **intervalo** o segmento $[x - y, x + y] := \{x + ty \mid t \in [-1, 1]\}$ e de **bola (aberta) de raio r com centro x** o conjunto

$$B_r(x) := B_r(x; \|\cdot\|) := B_r(x; X, \|\cdot\|) := \{y \in X \mid \|x - y\| < r\}.$$

Diz-se que subconjunto $U \subset X$ é **aberto** se para todo elemento $x \in U$ existe um $\varepsilon > 0$, tal que a bola $B_\varepsilon(x)$ está contida em U . Denotamos com o símbolo

$$\mathcal{U}(X, \|\cdot\|) := \{U \subset X \mid U \text{ é aberto com respeito } \|\cdot\|\} \quad (1.1.3)$$

a coleção de todos os conjuntos abertos.

Exemplo 1.1.1. Seja \mathbb{R}^n o espaço euclidiano de todos vetores $x = (x_1, \dots, x_n)$ com $x_i \in \mathbb{R}$. Para $1 \leq p < \infty$ a fórmula

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad x \in \mathbb{R}^n \quad (1.1.4)$$

define uma norma no \mathbb{R}^n . A desigualdade triangular para esta norma é chamada de *desigualdade de Minkowski*. No caso de $p > 1$ a prova utiliza a *desigualdade de Hölder* $\langle x, y \rangle = \sum_{i=1}^n x_i y_i \leq \|x\|_p \|y\|_q$ para $x, y \in \mathbb{R}^n$ e $1/p + 1/q = 1$. Para $p = \infty$ a fórmula

$$\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}, \quad x \in \mathbb{R}^n \quad (1.1.5)$$

define uma norma no \mathbb{R}^n .

Exemplo 1.1.2. Uma matriz simétrica $A = A^T \in \mathbb{R}^{n \times n}$ com coeficientes $a_{ij} = a_{ji}$ para $i, j = 1, \dots, n$ é chamada de **positiva definida** se ela satisfaz a condição $x^T A x = \sum_{i,j=1}^n x_i a_{ij} x_j > 0$ para todos os $x \in \mathbb{R}^n \setminus \{0\}$. Toda matriz $(n \times n)$ positiva definida A define um produto interno no \mathbb{R}^n pela fórmula

$$\langle x, y \rangle_A := x^T A y = \sum_{i,j=1}^n x_i a_{ij} y_j.$$

A norma associada de um vetor $x \in \mathbb{R}^n$ é dado pela fórmula $\|x\|_A := \sqrt{x^T A x}$.

Exemplo 1.1.3. Seja (M, d) um espaço métrico. Então o espaço

$$\text{BC}(M) := \{f : M \rightarrow \mathbb{R} \mid f \text{ é contínuo e limitado}\}$$

munido da norma definida através de

$$\|f\|_\infty := \sup_{p \in M} |f(p)| \quad \text{para } f \in \text{BC}(M) \quad (1.1.6)$$

é um espaço de Banach. A completude segue dos fatos que, primeiro, o próprio limite de uma sequência uniformemente convergente de funções contínuas também é contínuo e, segundo, convergência uniforme é equivalente a convergência com respeito à norma do supremo. Caso (M, d) seja compacto (ou seja, toda sequência em M admite uma subsequência convergente), então toda função contínua $f : M \rightarrow \mathbb{R}$ é necessariamente limitada. Neste caso o espaço $\text{BC}(M)$ coincide com o espaço $C(M)$ de todas as funções reais contínuas em M e o espaço $C(M)$ munido da norma do supremo (1.1.6) é um espaço de Banach.

Exemplo 1.1.4. Seja $I = [a, b]$ um intervalo compacto de números reais (com $a < b$). Para $1 \leq p < \infty$ a fórmula

$$\|f\|_p := \left(\int_a^b |f(t)|^p dt \right)^{1/p} \quad \text{para } f \in C(I) \quad (1.1.7)$$

define uma norma no espaço $C(I)$ de todas as funções reais contínuas em I . O espaço $C(I)$ munido da norma (1.1.7) não é completo (Exercício) e daí não é um espaço de Banach. No caso de $p = 2$ o produto interno

$$\langle f, g \rangle := \int_a^b f(t)g(t) dt \quad \text{para } f, g \in C(I) \quad (1.1.8)$$

gera a norma (1.1.7).

Exemplo 1.1.5. Seja $I = [a, b]$ um intervalo compacto de reais (com $a < b$) e seja $\ell \geq 0$ um número inteiro não-negativo. Então o espaço vetorial real

$$C^\ell(I) := \{f : I \rightarrow \mathbb{R} \mid f \text{ é } \ell \text{ vezes continuamente diferenciável}\}$$

munida da norma definida pela fórmula

$$\|f\|_{C^\ell} := \max_{k=0, \dots, \ell} \sup_{a \leq t \leq b} |f^{(k)}(t)|$$

para $f \in C^\ell(I)$ é um espaço de Banach. Para $\ell = 0$ isso é o espaço $C(I)$ das funções contínuas $f : I \rightarrow \mathbb{R}$ com a norma do supremo como no Exemplo 1.1.3.

Definição 1.1.6. Sejam $\|\cdot\|$ e $\|\cdot\|'$ duas funções normas num espaço vetorial real X . Chama-se a função norma $\|\cdot\|$ **equivalente** a $\|\cdot\|'$ (notação $\|\cdot\| \sim \|\cdot\|'$) se existe uma constante $c > 0$, tal que

$$\frac{1}{c} \|x\| \leq \|x\|' \leq c \|x\| \quad \text{para todos os } x \in X. \quad (1.1.9)$$

Comentário 1.1.7. (i) Seja X um espaço vetorial real. Então a “*equivalência de normas*” define uma relação de equivalência no conjunto de todas as normas em X . Ou seja, para todas as normas $\|\cdot\|, \|\cdot\|', \|\cdot\|''$ em X vale

- $\|\cdot\| \sim \|\cdot\|$
- $\|\cdot\| \sim \|\cdot\|' \iff \|\cdot\|' \sim \|\cdot\|$
- $\|\cdot\| \sim \|\cdot\|', \|\cdot\|' \sim \|\cdot\|'' \implies \|\cdot\| \sim \|\cdot\|''$

(ii) Sejam $\|\cdot\|$ e $\|\cdot\|'$ duas normas equivalentes num espaço vetorial real X . Então as duas normas geram os mesmos conjuntos abertos. (Exercício; veja também o Exemplo 1.2.4).

(iii) Sejam $\|\cdot\|$ e $\|\cdot\|'$ duas normas equivalentes num espaço vetorial real X . Como os subconjuntos abertos de X com respeito à duas normas são os mesmos, segue que todas as *afirmações topológicas em X* (portanto afirmações as quais podem ser provadas por meio de conjuntos abertos, por exemplo convergência, continuidade, ser fechado, compacidade), não dependem de qual das duas normas equivalentes utilizamos. Se por exemplo $A \subset X$ é compacto (respectivamente fechado) com respeito da norma $\|\cdot\|$, então A também é compacto (respectivamente fechado) com respeito à norma $\|\cdot\|'$. Além disso, vale que uma sequência $(x_k)_{k \in \mathbb{N}}$ em X é Cauchy com respeito à norma $\|\cdot\|$ (respectivamente converge) se e somente se é Cauchy com respeito à norma $\|\cdot\|'$ (respectivamente converge). Daí $(X, \|\cdot\|)$ é um espaço de Banach se e somente se $(X, \|\cdot\|')$ é um espaço de Banach.

Seja $I = [a, b] \subset \mathbb{R}$ um intervalo compacto com $a < b$. Então o conjunto $U := \{f \in C(I) \mid \sup_I |f| < 1\}$ é aberto com respeito à norma do supremo $\|\cdot\|_\infty$ no Exemplo 1.1.3, mas não com respeito à norma $\|\cdot\|_p$ no Exemplo 1.1.4 para $1 \leq p < \infty$. Consequentemente as normas $\|\cdot\|_\infty$ e $\|\cdot\|_p$ no espaço $C(I)$ não são equivalente. Isto deve-se ao fato que o espaço vetorial das funções contínuas num intervalo compacto não trivial tem dimensão infinita. O próximo teorema afirma que nos espaços vetoriais de dimensão finita duas normas quaisquer são equivalente.

Teorema 1.1.8. *Seja X um espaço vetorial de dimensão finita. Então quaisquer duas normas em X são equivalente.*

Demonstração. A prova tem quatro passos. Seja $n := \dim X \in \mathbb{N}$. Denotamos com $\|\cdot\|_2$ a norma euclidiana no \mathbb{R}^n ; veja (1.1.4) para $p = 2$.

Passo 1. *Seja $\|\cdot\|$ uma norma arbitrária no \mathbb{R}^n . Então existe uma constante $c > 0$, tal que todo $x \in \mathbb{R}^n$ satisfaz a desigualdade $\|x\| \leq c \|x\|_2$.*

Os vetores $e_i := (0, \dots, 0, 1, 0, \dots, 0)$ (com 1 no i -ésimo lugar) para $i = 1, \dots, n$ formam a base canônica do \mathbb{R}^n . Seja $c := (\sum_{i=1}^n \|e_i\|^2)^{1/2} > 0$. Então vale $x = \sum_{i=1}^n x_i e_i$ para todo vetor $x \in \mathbb{R}^n$ e daí

$$\|x\| \leq \sum_{i=1}^n |x_i| \|e_i\| \leq \sqrt{\sum_{i=1}^n |x_i|^2} \sqrt{\sum_{i=1}^n \|e_i\|^2} = c \|x\|_2.$$

Aqui a primeira desigualdade segue da desigualdade triangular para $\|\cdot\|$ e a segunda da desigualdade de Cauchy–Schwarz para a norma euclidiana.

Passo 2. *Seja $\|\cdot\|$ como no passo 1. Então existe uma constante $\delta > 0$, tal que todo $x \in \mathbb{R}^n$ satisfaz a desigualdade $\delta \|x\|_2 \leq \|x\|$.*

A esfera unitária $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ munido da função distância $d(x, y) := \|x - y\|_2$ para $x, y \in \mathbb{S}^{n-1}$ é um espaço métrico compacto através do Teorema de Heine–Borel. Definimos a função $f : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$ pela fórmula $f(x) := \|x\|$ para $x \in \mathbb{S}^{n-1}$. Através de passo 1 vale $|f(x) - f(y)| = \left| \|x\| - \|y\| \right| \leq \|x - y\| \leq c \|x - y\|_2$ para todos os $x, y \in \mathbb{S}^{n-1}$ e daí f é contínuo. Como \mathbb{S}^{n-1} é compacto existe um $x_0 \in \mathbb{S}^{n-1}$, tal que $\inf_{\mathbb{S}^{n-1}} f = f(x_0) =: \delta > 0$. Para $x \in \mathbb{R}^n \setminus \{0\}$ vale então

$$\delta \leq f(\|x\|_2^{-1} x) = \|\|x\|_2^{-1} x\| = \|x\|_2^{-1} \|x\|.$$

Daqui segue a afirmação do Passo 2.

Passo 3. *Quaisquer duas normas no \mathbb{R}^n são equivalente.*

Segundo o Passo 1 e Passo 2, qualquer norma no \mathbb{R}^n é equivalente à norma euclidiana. Daí segue o Passo 3 da parte (i) do comentário 1.1.7.

Passo 4. *Quaisquer duas normas em X são equivalente.*

Escolha uma base f_1, \dots, f_n de X . Então a aplicação

$$\Phi : \mathbb{R}^n \rightarrow X, \quad \Phi\xi := \sum_{i=1}^n \xi_i f_i \quad \text{para } \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$$

é linear e bijetiva, conseqüentemente um isomorfismo de espaços vetoriais. Sejam $\|\cdot\|$ e $\|\cdot\|'$ duas normas em X , então as fórmulas $\|\xi\|_\Phi := \|\Phi\xi\|$ e $\|\xi\|'_\Phi := \|\Phi\xi\|'$ para $\xi \in \mathbb{R}^n$ definem duas normas no \mathbb{R}^n . Tais normas são equivalente segundo o Passo 3, conseqüentemente as duas normas $\|\cdot\|$ e $\|\cdot\|'$ em X também são equivalente. Com isso o Teorema 1.1.8 fica provado. \square

Teorema 1.1.8 tem umas conseqüências importantes.

Corolário 1.1.9. *Todo o espaço vetorial normado de dimensão finita é um espaço de Banach.*

Demonstração. No curso Análise I foi provado que o \mathbb{R}^n munido da norma euclidiana é um espaço de Banach. Daí, segundo o Teorema 1.1.8, o \mathbb{R}^n munido de qualquer norma é um espaço de Banach. Como todo espaço vetorial normado de dimensão finita é isomorfo a \mathbb{R}^n para um $n \in \mathbb{N}_0$, a afirmação segue. \square

Corolário 1.1.10. *Um subconjunto de um espaço vetorial normado de dimensão finita é compacto se e somente se é fechado e limitado.*

Demonstração. Segundo o Teorema de Heine–Borel isto vale para o \mathbb{R}^n com a norma euclidiana. Mas através do Teorema 1.1.8 quaisquer duas normas no \mathbb{R}^n são equivalente, portanto todas as três propriedades (compacidade, ser fechado, ser limitado) de um subconjunto do \mathbb{R}^n são independente da escolha da norma. Portanto a afirmação vale para qualquer norma no \mathbb{R}^n . Como todo espaço vetorial real de dimensão finita é isomorfo a \mathbb{R}^n para um $n \in \mathbb{N}_0$, segue daí a afirmação em geral. \square

Segundo o Corolário 1.1.10 a esfera unitária $S := \{x \in X \mid \|x\| = 1\}$ é compacta em todo espaço vetorial normado de dimensão finita. E o exercício seguinte mostra que esta propriedade caracteriza exatamente os espaços vetoriais normados de dimensão finita.

Exercício 1.1.11. Seja $(X, \|\cdot\|)$ um espaço vetorial real normado, tal que a esfera unitária $S := \{x \in X \mid \|x\| = 1\}$ é compacta. Então X tem dimensão finita. **Dica:** Prova o seguinte.

(i) Um subespaço linear de X de dimensão finita é um subconjunto fechado de X .

(ii) Seja $Y \subset X$ um subconjunto fechado. Então vale para todo $x \in X \setminus Y$ a desigualdade $d(x, Y) := \inf_{y \in Y} \|x - y\| > 0$.

(iii) Seja $Y \subsetneq X$ um subespaço linear fechado de X não igual a X completo. Então existe um vetor $x \in X$, tal que $\|x\| = 1$ e $d(x, Y) \geq \frac{1}{2}$. (Seja $x_0 \in X \setminus Y$. Então existe, segundo (ii), um $y_0 \in Y$, tal que $\|x_0 - y_0\| \leq 2d(x_0, Y)$. Ora, seja $x := \|x_0 - y_0\|^{-1}(x_0 - y_0)$.)

(iv) Se X tem dimensão infinita existe uma sequência $(x_n)_{n \in \mathbb{N}}$ em S , tal que $\|x_n - x_m\| \geq \frac{1}{2}$ para todos os $m, n \in \mathbb{N}$ com $m \neq n$. Então S não é compacto.

Exercício 1.1.12. Seja X um espaço vetorial real de dimensão finita. Chama-se um subconjunto $K \subset X$ **convexo** se para todos os $x, y \in X$ e todo $t \in \mathbb{R}$ a condição

$$x, y \in K, \quad 0 \leq t \leq 1 \quad \implies \quad (1-t)x + ty \in K$$

é satisfeita.

(i) Seja $f : X \rightarrow \mathbb{R}$ uma função. O conjunto aberto

$$U := \{(x, s) \in X \times \mathbb{R} \mid s > f(x)\}$$

é convexo se e somente se a função f é convexa.

(ii) Seja $U \subset X$ um subconjunto aberto, limitado, convexo, tal que para todo $x \in X$ vale o seguinte (simetria de U em respeito à origem):

$$x \in U \quad \iff \quad -x \in U.$$

(Conforme o Teorema 1.1.8 os conceitos “aberto” e “limitado” para subconjuntos de X não dependem da norma usada na definição.) Então a fórmula

$$\|x\| := \inf \{\lambda > 0 \mid \lambda^{-1}x \in U\} \quad \text{para } x \in X$$

define uma função norma em X e U é a bola aberta unitária com respeito a esta norma, em símbolos $U = \{x \in X \mid \|x\| < 1\}$.

Exercício 1.1.13. No espaço $C([0, 1])$ consideramos as normas $\|\cdot\|_p$ do Exemplo 1.1.4 para $1 \leq p \leq \infty$. Para quais p e q existe uma constante $c > 0$, tal que a desigualdade $\|f\|_p \leq c \|f\|_q$ é válida para todo $f \in C([0, 1])$? Mostre que as normas $\|\cdot\|_p$ e $\|\cdot\|_q$ para $p \neq q$ não são equivalente.

1.2 A norma operador

Sejam $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ espaços vetoriais reais normados.

Definição 1.2.1. *Chama-se um operador linear $A : X \rightarrow Y$ limitado se existe uma constante $c \geq 0$, tal que a desigualdade*

$$\|Ax\|_Y \leq c \|x\|_X \quad (1.2.1)$$

vale para todos os $x \in X$. A menor de tais constantes $c \geq 0$ chama-se a **norma** (respectivamente **norma operador**) de A e usa-se a notação

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X}. \quad (1.2.2)$$

(Aqui entende-se o supremo sobre todos os vetores não nulos $x \in X$.) O conjunto dos operadores lineares limitados de X para Y é denotado $\mathcal{L}(X, Y)$.

O termo “operador” é usado na literatura matemática como sinônimo do termo “aplicação”. A origem desta escolha de terminologia surge do fato que muitos, dos mais importantes, exemplos de *operadores diferenciais* ou *operadores integrais* são da física matemática e para o estudo deles foi desenvolvido a teoria das aplicações lineares. Este assunto leva-nos à “*análise funcional*” a qual iremos abordar só marginalmente neste curso.

Corolário 1.2.2. *Seja X de dimensão finita. Então todo operador linear $A : X \rightarrow Y$ é limitado.*

Demonstração. Definimos uma função

$$X \rightarrow \mathbb{R} : x \mapsto \|x\|_A := \|x\|_X + \|Ax\|_Y. \quad (1.2.3)$$

Como A é linear, a função (1.2.3) é uma norma. Como X possui dimensão finita, existe, segundo o Teorema 1.1.8, uma constante $c \geq 1$, tal que para todo $x \in X$ vale a desigualdade $\|x\|_A \leq c \|x\|_X$. Segue imediatamente a desigualdade desejada $\|Ax\|_Y \leq c \|x\|_X$ para todos os $x \in X$. \square

Teorema 1.2.3. *Seja $A : X \rightarrow Y$ um operador linear. Então são equivalente as afirmações seguintes.*

- (i) A é limitado.
- (ii) A é contínuo.
- (iii) A é contínuo no ponto $x = 0$.

Demonstração. Mostramos que (i) implica (ii). Se A é limitado, então A é Lipschitz-contínuo com a constante Lipschitz $\|A\|$, porque vale

$$\begin{aligned} d_Y(Ax_0, Ax_1) &= \|Ax_0 - Ax_1\|_Y = \|A(x_0 - x_1)\|_Y \\ &\leq \|A\| \|x_0 - x_1\|_X = \|A\| d_X(x_0, x_1) \end{aligned}$$

para todos os $x_0, x_1 \in X$; assim A também é contínuo. Que (ii) implica (iii) segue imediato da definição de continuidade.

Mostramos que (iii) implica (i). Seja então A contínuo no ponto $x = 0$. Desta forma segue da definição de continuidade com $\varepsilon = 1$ que existe um $\delta > 0$ (escolha $\delta \leq 1$), tal que todo $x \in X$ com $\|x\|_X \leq \delta$ satisfaz a desigualdade $\|Ax\|_Y \leq 1$. Ora, seja $x \in X \setminus \{0\}$. Então vale

$$\left\| \frac{\delta}{\|x\|_X} x \right\|_X = \delta$$

e por isto

$$1 \geq \left\| \frac{\delta}{\|x\|_X} Ax \right\|_Y = \frac{\delta}{\|x\|_X} \|Ax\|_Y.$$

Segue que $\|Ax\|_Y \leq \delta^{-1} \|x\|_X$ para todo $x \in X$. Então A é limitado. \square

Exemplo 1.2.4. Sejam $\|\cdot\|$ e $\|\cdot\|'$ duas normas num espaço vetorial real X . Consideremos a afirmação

$$\exists c > 0 \forall x \in X : \|x\|' \leq c \|x\|. \quad (1.2.4)$$

Esta condição vale se e somente se o operador linear

$$\text{id} : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|') \quad (1.2.5)$$

é limitado. Segundo o Teorema 1.2.3 isso significa que a aplicação (1.2.5) é contínua. Isto por sua vez significa que a pre-imagem de qualquer conjunto aberto em $(X, \|\cdot\|')$ sob a aplicação (1.2.5) é aberta em $(X, \|\cdot\|)$. Ou seja

$$\mathcal{U}(X, \|\cdot\|') \subset \mathcal{U}(X, \|\cdot\|). \quad (1.2.6)$$

Portanto temos mostrado que as afirmações (1.2.4) e (1.2.6) são equivalentes. Desta forma segue a afirmação

$$\|\cdot\| \sim \|\cdot\|' \iff \mathcal{U}(X, \|\cdot\|) = \mathcal{U}(X, \|\cdot\|')$$

para quaisquer duas normas num espaço vetorial real X (veja também a parte (ii) do Comentário 1.1.7). Segundo o Teorema 1.1.8 isto significa, por sua vez, que a coletividade dos conjuntos abertos é independente da escolha da norma num espaço vetorial real de dimensão finita. Assim todo espaço vetorial real de dimensão finita possui uma *topologia canônica*.

Lema 1.2.5. $\mathcal{L}(X, Y)$ é um espaço vetorial normado.

Demonstração. Temos que mostrar as três afirmações seguintes.

- (a) Para todo $A \in \mathcal{L}(X, Y)$ vale $\|A\| = 0 \implies A = 0$.
- (b) Para todos os $A \in \mathcal{L}(X, Y)$ e $\lambda \in \mathbb{R}$ vale $\lambda A \in \mathcal{L}(X, Y)$ e $\|\lambda A\| = |\lambda| \|A\|$.
- (c) Para $\forall A, B \in \mathcal{L}(X, Y)$ vale $A + B \in \mathcal{L}(X, Y)$ e $\|A + B\| \leq \|A\| + \|B\|$.

As afirmações (a) e (b) seguem imediatamente das definições. Provaremos (c). Sejam $A, B : X \rightarrow Y$ operadores lineares limitados. Então vale para todo $x \in X$ a seguinte desigualdade

$$\begin{aligned} \|(A + B)x\|_Y &= \|Ax + Bx\|_Y \\ &\leq \|Ax\|_Y + \|Bx\|_Y \\ &\leq \|A\| \|x\|_X + \|B\| \|x\|_X \\ &= (\|A\| + \|B\|) \|x\|_X. \end{aligned}$$

Portanto $A + B$ é limitado e para $x \in X \setminus \{0\}$ vale

$$\frac{\|(A + B)x\|_Y}{\|x\|_X} \leq \|A\| + \|B\|.$$

Seguindo o resultado da desigualdade triangular

$$\|A + B\| \leq \|A\| + \|B\|$$

através da definição da norma operador. □

Exemplo 1.2.6. No caso de $X = \mathbb{R}^n$ e $Y = \mathbb{R}^m$, o espaço de todas as aplicações lineares de \mathbb{R}^n para \mathbb{R}^m (Corolário 1.2.2) é $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Podemos identificar ele com o espaço $\mathbb{R}^{m \times n}$ das matrizes reais $(m \times n)$, porque qualquer aplicação linear de \mathbb{R}^n para \mathbb{R}^m é dada pela multiplicação com uma matriz

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

Nesta situação a norma operador da aplicação linear induzida por A com respeito a normas adequadas no \mathbb{R}^n e no \mathbb{R}^m também é chamada de **norma matriz**. Seguem três exemplos.

(i) Com

$$\|x\|_1 := \sum_{j=1}^n |x_j|, \quad \|y\|_1 := \sum_{i=1}^m |y_i|$$

para $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ a norma operador de $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ é o número

$$\|A\|_1 := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|_1}{\|x\|_1} = \max_{j=1, \dots, n} \sum_{i=1}^m |a_{ij}|.$$

(ii) Com

$$\|x\|_\infty := \max_{j=1, \dots, n} |x_j|, \quad \|y\|_\infty := \max_{i=1, \dots, m} |y_i|$$

para $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ a norma operador de $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ é o número

$$\|A\|_\infty := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|_\infty}{\|x\|_\infty} = \max_{i=1, \dots, m} \sum_{j=1}^n |a_{ij}|.$$

(iii) Com

$$\|x\|_2 := \sqrt{\sum_{j=1}^n x_j^2}, \quad \|y\|_2 := \sqrt{\sum_{i=1}^m y_i^2}$$

para $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$ a norma operador de $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ é a raiz do maior autovalor da matriz quadrada $A^T A \in \mathbb{R}^{n \times n}$, ou seja,

$$\|A\|_2 := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|_2}{\|x\|_2} = \max\{\lambda \geq 0 \mid \ker(\lambda^2 \mathbb{1} - A^T A) \neq \{0\}\}.$$

(iv) Para $n > 1$ e $m > 1$ nenhuma das normas operador em (i), (ii), e (iii) é a norma de um produto interno no espaço das matrizes reais ($m \times n$). Para ver isso, consideramos o caso $m = n = 2$ e as matrizes

$$A := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B := \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Então as matrizes A , B , $A + B$, $A - B$ são todas de norma 1 com respeito a qualquer uma das três normas em (i), (ii), e (iii). Portanto vale

$$\|A + B\|^2 + \|A - B\|^2 = 2 \neq 4 = 2\|A\|^2 + 2\|B\|^2$$

mas então a identidade do paralelogramo é violada.

Teorema 1.2.7. *Se Y é um espaço de Banach, então $\mathcal{L}(X, Y)$ também é.*

Demonstração. Temos que mostrar que $\mathcal{L}(X, Y)$ é completo. Seja então $(A_n)_{n \in \mathbb{N}}$ uma sequência de Cauchy em $\mathcal{L}(X, Y)$. Então a desigualdade

$$\|A_n x - A_m x\|_Y \leq \|A_n - A_m\| \|x\|_X$$

mostra que a sequência $(A_n x)_{n \in \mathbb{N}}$ para todo $x \in X$ é uma sequência de Cauchy em Y . Como Y é completo, esta sequência converge e definimos $A : X \rightarrow Y$ através de

$$Ax := \lim_{n \rightarrow \infty} A_n x$$

para $x \in X$. Vamos provar que A é um operador linear limitado e que a sequência A_n converge para A na norma operador.

A é linear. Para $x_0, x_1 \in X$ vale

$$\begin{aligned} A(x_0 + x_1) &= \lim_{n \rightarrow \infty} A_n(x_0 + x_1) \\ &= \lim_{n \rightarrow \infty} A_n x_0 + \lim_{n \rightarrow \infty} A_n x_1 \\ &= Ax_0 + Ax_1. \end{aligned}$$

Analogamente mostra-se que $A(\lambda x) = \lambda Ax$ para todo $\lambda \in \mathbb{R}$ e todo $x \in X$.

A é limitado. Primeiramente segue da desigualdade

$$\left| \|A_n\| - \|A_m\| \right| \leq \|A_n - A_m\|$$

que a sequência $(\|A_n\|)_{n \in \mathbb{N}}$ é uma sequência de Cauchy de números reais e por isso converge. Seja

$$c := \lim_{n \rightarrow \infty} \|A_n\|.$$

Então todo $x \in X$ satisfaz a desigualdade

$$\begin{aligned} \|Ax\|_Y &= \left\| \lim_{n \rightarrow \infty} A_n x \right\|_Y \\ &= \lim_{n \rightarrow \infty} \|A_n x\|_Y \\ &\leq \lim_{n \rightarrow \infty} \|A_n\| \|x\|_X \\ &= c \|x\|_X. \end{aligned}$$

Segue, então, que A é limitado e $\|A\| \leq c$.

A_n **converge para** A . Seja $\varepsilon > 0$. Como A_n é uma sequência de Cauchy, existe um $n_0 \in \mathbb{N}$, tal que para todos os $n, m \in \mathbb{N}$ vale:

$$m > n \geq n_0 \quad \implies \quad \|A_n - A_m\| < \varepsilon.$$

Desta forma segue a desigualdade $\|A_n x - A_m x\|_Y < \varepsilon \|x\|_X$ para todos os $x \in X$ e $m > n \geq n_0$ e daí

$$\|A_n x - Ax\|_Y = \lim_{m \rightarrow \infty} \|A_n x - A_m x\|_Y \leq \varepsilon \|x\|_X$$

para todo $x \in X$ e todo $n \in \mathbb{N}$ com $n \geq n_0$. Dividindo por $\|x\|_X$ com $x \neq 0$ encontraremos

$$\|A_n - A\| = \sup_{x \neq 0} \frac{\|A_n x - Ax\|_Y}{\|x\|_X} \leq \varepsilon$$

para todo $n \in \mathbb{N}$ com $n \geq n_0$. Com isso o Teorema 1.2.7 fica provado. \square

Um caso especial importante é $Y = \mathbb{R}$. O espaço de Banach $X^* := \mathcal{L}(X, \mathbb{R})$ é chamado de **espaço dual** do espaço vetorial normado X . Teorema 1.2.7 mostra que este espaço dual sempre fica completo, mesmo que X não seja.

Exercício 1.2.8. (i) Seja X um espaço vetorial normado de dimensão finita. Então $X^* = \mathcal{L}(X, \mathbb{R})$ é de dimensão finita e $\dim X^* = \dim X$.

(ii) Seja $n \in \mathbb{N}$ e sejam $1 \leq p, q \leq \infty$ com $1/p + 1/q = 1$. Para $y \in \mathbb{R}^n$ definimos a aplicação linear $\Lambda_y : \mathbb{R}^n \rightarrow \mathbb{R}$ através de

$$\Lambda_y(x) := \sum_{i=1}^n x_i y_i = \langle x, y \rangle$$

para $x \in \mathbb{R}^n$. Então a aplicação

$$\mathbb{R}^n \rightarrow (\mathbb{R}^n)^* : y \mapsto \Lambda_y$$

é um isomorfismo de espaços vetoriais e vale

$$\|y\|_q = \sup_{x \neq 0} \frac{|\langle x, y \rangle|}{\|x\|_p} = \sup_{x \neq 0} \frac{|\Lambda_y(x)|}{\|x\|_p}$$

para todos os $y \in \mathbb{R}^n$. (Veja o Exemplo 1.1.1 para $\|x\|_p$ e $\|y\|_q$.)

1.3 Álgebras de Banach

Seja $(X, \|\cdot\|)$ um espaço de Banach. Então o espaço $\mathcal{L}(X) := \mathcal{L}(X, X)$ de todos os operadores lineares limitados de X para X também é um espaço de Banach (veja o Teorema 1.2.7). Este espaço possui uma estrutura de produto $(S, T) \mapsto S \circ T$, dado pela composição. Esta estrutura de produto em geral não é comutativa, nem existe um inverso para todo operador não-nulo $T \in \mathcal{L}(X)$; pois $\mathcal{L}(X)$ não é um corpo; mas este espaço satisfaz todas as outras propriedades de um corpo; um tal espaço chama-se uma álgebra. O conceito de uma álgebra de Banach é útil para cálculos com operadores inversos e matrizes exponenciais.

Definição 1.3.1. *Uma álgebra real é composta de um espaço vetorial real \mathcal{A} , uma aplicação real $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A} : (x, y) \mapsto xy$, (chamada de **produto**) e um elemento $\mathbb{1} \in \mathcal{A}$ (chamado de **1-elemento** ou **o elemento neutro**) as quais satisfazem os seguintes axiomas.*

(i) *O produto é associativo, ou seja, para todos os $x, y, z \in \mathcal{A}$ vale*

$$(xy)z = x(yz).$$

(ii) *O produto é bilinear, ou seja, para todos os $x, y, z \in \mathcal{A}$ e todo $\lambda \in \mathbb{R}$ vale*

$$x(y + z) = xy + xz, \quad (x + y)z = xz + yz, \quad x(\lambda z) = \lambda(xz) = (\lambda x)z.$$

(iii) *Para todo $x \in \mathcal{A}$ vale $\mathbb{1}x = x\mathbb{1} = x$.*

Uma álgebra real normada é composta de uma álgebra real \mathcal{A} e uma norma $\mathcal{A} \rightarrow [0, \infty) : x \mapsto \|x\|$ a qual satisfaz a desigualdade-produto

$$\|xy\| \leq \|x\| \|y\| \tag{1.3.1}$$

para todos os $x, y \in \mathcal{A}$. Uma álgebra de Banach real é uma álgebra real normada na qual toda sequência de Cauchy converge (ou seja, visto como espaço vetorial normado \mathcal{A} é completo, então é um espaço de Banach).

Observe-se que o 1-elemento numa álgebra real \mathcal{A} é determinado unicamente pelo produto. Seja $\mathbb{1}' \in \mathcal{A}$ algum outro elemento que satisfaça a condição $\mathbb{1}'x = x\mathbb{1}' = \mathbb{1}'$ para todos os $x \in \mathcal{A}$, então vale $\mathbb{1}' = \mathbb{1}'\mathbb{1} = \mathbb{1}$. Uma álgebra difere de um corpo pelos seguintes fatos: Primeiro, também possuir uma estrutura de um espaço vetorial (compatível com o produto), segundo, o produto não precisa ser comutativo, e terceiro, nem todo elemento não-nulo de \mathcal{A} precisa ser invertível.

Definição 1.3.2. *Seja \mathcal{A} uma álgebra real. Um elemento $x \in \mathcal{A}$ chama-se **invertível** se existe um elemento $y \in \mathcal{A}$ satisfaça a seguinte condição*

$$xy = yx = \mathbb{1}. \quad (1.3.2)$$

*Caso exista tal elemento, então ele é determinado unicamente por x (pois qualquer outro elemento $y' \in \mathcal{A}$ satisfazendo a condição $y'x = \mathbb{1}$, também satisfaz a identidade $y' = y'\mathbb{1} = y'(xy) = (y'x)y = \mathbb{1}y = y$). Se $x \in \mathcal{A}$ é invertível chama-se o elemento único $y \in \mathcal{A}$, o que satisfaz a condição (1.3.2), de **elemento inverso de x** , denotado $x^{-1} := y$. O grupo dos elementos invertíveis de \mathcal{A} é denotado por*

$$\mathcal{G} := \{x \in \mathcal{A} \mid \text{existe um } y \in \mathcal{A}, \text{ tal que } xy = yx = \mathbb{1}\}.$$

Lema 1.3.3. *Seja \mathcal{A} uma álgebra real com 1-elemento $\mathbb{1}$. Então vale o seguinte.*

- (i) $\mathbb{1} \in \mathcal{G}$ e $\mathbb{1}^{-1} = \mathbb{1}$.
- (ii) Se $x \in \mathcal{G}$ então $x^{-1} \in \mathcal{G}$ e vale $(x^{-1})^{-1} = x$.
- (iii) Se $x, y \in \mathcal{G}$ então $xy \in \mathcal{G}$ e vale $(xy)^{-1} = y^{-1}x^{-1}$.

Demonstração. Estas afirmações seguem imediatamente da definição. □

Exemplo 1.3.4. Os números reais e complexos formam álgebras de Banach com o produto comum e a função absoluto como norma. Um exemplo similar é dado pelos quatérnios. Os **quatérnios** são definidos no espaço vetorial real $\mathbb{H} \cong \mathbb{R}^4$ através de uma estrutura de produto adequada. Escrevemos os elementos de \mathbb{H} como somas formais

$$x = x_0 + \mathbf{i}x_1 + \mathbf{j}x_2 + \mathbf{k}x_3$$

com $x_i \in \mathbb{R}$ e definimos a estrutura de produto tal que valem as equações

$$\mathbf{j}\mathbf{k} = -\mathbf{k}\mathbf{j} = \mathbf{i}, \quad \mathbf{k}\mathbf{i} = -\mathbf{i}\mathbf{k} = \mathbf{j}, \quad \mathbf{i}\mathbf{j} = -\mathbf{j}\mathbf{i} = \mathbf{k}, \quad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$$

e $\mathbf{1}$ é o 1-elemento. A norma em \mathbb{H} é dada por $|x| := \sqrt{\sum_{i=0}^3 x_i^2}$ e vale $|xy| = |x||y|$ para todos os $x, y \in \mathbb{H}$. Uma aplicação bilinear a qual também possui esta propriedade, ainda que não é mais associativo, também existe em dimensão oito (os octónios), mas em nenhuma dimensão exceto 1, 2, 4, 8.

Exemplo 1.3.5. Se X é um espaço de Banach, então $\mathcal{L}(X)$ junto com a norma operador, a composição, e a função identidade é uma álgebra de Banach. A completude segue do Teorema 1.2.7. **Exercício:** A norma operador em $\mathcal{L}(X)$ satisfaz a condição (1.3.1).

Exemplo 1.3.6. O espaço $\mathbb{R}^{n \times n}$ das matrizes quadradas é uma álgebra de Banach sob qualquer norma que satisfaça a condição (1.3.1). Uma tal norma no espaço $\mathbb{R}^{n \times n}$ chama-se **norma matriz**. Por exemplo, pode ser a norma operador associada a uma norma arbitrária no \mathbb{R}^n ; veja o Exemplo 1.2.6. Um outro exemplo seria a norma euclidiana a qual pode ser recuperada através da identificação $\mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}$; neste caso obtém-se

$$\|A\| = \left(\sum_{i,j=1}^n a_{ij}^2 \right)^{1/2}$$

para $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. **Exercício:** Esta norma satisfaz as condições (1.3.1) e $\|Ax\|_2 \leq \|A\| \|x\|_2$ para as matrizes $A \in \mathbb{R}^{n \times n}$ e os vetores $x \in \mathbb{R}^n$.

Exemplo 1.3.7. Seja (X, d) um espaço métrico compacto. Então $C(X)$ munido da norma do supremo é uma álgebra de Banach (veja o Exemplo 1.1.3).

Exemplo 1.3.8. Uma sequência $x = (x_0, x_1, x_2, \dots)$ de números reais é chamada de **absolutamente somável** se a série $\sum_{k=0}^{\infty} x_k$ converge absolutamente. Defina a norma de uma sequência absolutamente somável por

$$\|x\| := \sum_{k=0}^{\infty} |x_k|.$$

Denotamos com ℓ^1 o espaço de todas as sequências reais $(x_k)_{k \in \mathbb{N}_0}$ absolutamente somáveis. O espaço vetorial ℓ^1 munido desta norma é um espaço de Banach. A **convolução** de duas sequências $x, y \in \ell^1$ é definida através de

$$(x * y)_k := \sum_{i=0}^k x_i y_{k-i} \quad \text{para } k \in \mathbb{N}_0.$$

Esta sequência também é absolutamente somável e vale $\|x * y\| \leq \|x\| \|y\|$. (Isto segue do Teorema da reorganização grande em [4].) Consequentemente ℓ^1 é uma álgebra de Banach. **Exercício:** ℓ^1 é completo.

Exercício 1.3.9. Numa álgebra de Banach vale $\|\mathbb{1}\| \geq 1$. Para quais das álgebras de Banach indicadas acima o produto é comutativo? Se \mathcal{A} é uma álgebra de Banach, então \mathcal{A}^2 com a norma $\mathcal{A}^2 \rightarrow \mathbb{R} : (x, y) \mapsto \|x\| + \|y\|$ é um espaço de Banach e a aplicação produto $\mathcal{A}^2 \rightarrow \mathcal{A} : (x, y) \mapsto xy$ é contínua.

Séries em espaços de Banach

Os conceitos de convergência e continuidade foram introduzidos no curso Análise I para espaços métricos arbitrários. Então valem particularmente para funções e seqüências com valores num espaço de Banach ou numa álgebra de Banach. Por causa da estrutura de um espaço vetorial pode-se transferir o conceito da convergência de uma série para séries com valores em espaços de Banach. Se $(X, \|\cdot\|)$ é um espaço de Banach e $(x_k)_{k \in \mathbb{N}_0}$ uma seqüência em X , então chama-se a série $\sum_{k=0}^{\infty} x_k$ de **convergente**, se a seqüência das **somas parciais** $s_n := \sum_{k=0}^n x_k$ converge; neste caso denotamos o limite assim

$$\sum_{k=0}^{\infty} x_k := \lim_{n \rightarrow \infty} \sum_{k=0}^n x_k.$$

Portanto adotaremos a mesma terminologia e o mesmo simbolismo como no caso de séries de números reais ou complexos, ou seja, o termo “ $\sum_{k=0}^{\infty} x_k$ ” tem dois significados: por um lado denota a seqüência das somas parciais $(s_n)_{n \in \mathbb{N}_0}$ e por outro lado denota o limite deles, caso exista. Chamamos a série $\sum_{k=0}^{\infty} x_k$ **absolutamente convergente** se a série $\sum_{k=0}^{\infty} \|x_k\|$ converge. Neste caso também chamamos a seqüência $(x_k)_{k \in \mathbb{N}_0}$ **absolutamente somável**.

Teorema 1.3.10. *Seja $(X, \|\cdot\|)$ um espaço de Banach e seja $(x_k)_{k \in \mathbb{N}_0}$ uma seqüência em X . Se a série $\sum_{k=0}^{\infty} x_k$ converge absolutamente, então ela converge.*

Demonstração. Seja $s_n := \sum_{k=0}^n x_k$ a seqüência das somas parciais e seja $\varepsilon > 0$. Porque a série $\sum_{k=0}^{\infty} x_k$ converge absolutamente, converge a série $\sum_{k=0}^{\infty} \|x_k\|$ em \mathbb{R} . Ou seja, a seqüência das somas parciais $\sigma_n := \sum_{k=0}^n \|x_k\|$ da série das normas converge e consequentemente é uma seqüência de Cauchy. Portanto existe um número natural $n_0 \in \mathbb{N}$, tal que para todos os $m, n \in \mathbb{N}$ vale o seguinte:

$$n > m \geq n_0 \quad \implies \quad \sigma_n - \sigma_m = \|x_{m+1}\| + \cdots + \|x_n\| < \varepsilon.$$

Segue da desigualdade triangular que para todos os $m, n \in \mathbb{N}$ com $n > m \geq n_0$ vale a seguinte desigualdade

$$\|s_n - s_m\| = \|x_{m+1} + \cdots + x_n\| \leq \|x_{m+1}\| + \cdots + \|x_n\| < \varepsilon.$$

Com isso é mostrado que $(s_n)_{n \in \mathbb{N}_0}$ é uma seqüência de Cauchy em X . Como X é um espaço de Banach (então completo), esta seqüência converge, ou seja, esta série $\sum_{k=0}^{\infty} x_k$ converge. Com isso o Teorema 1.3.10 fica provado. \square

Séries de potências em álgebras de Banach

Seja \mathcal{A} uma álgebra de Banach real e

$$P(z) := \sum_{k=0}^{\infty} a_k z^k \quad (1.3.3)$$

uma série de potências com coeficientes reais $a_k \in \mathbb{R}$ e raio de convergência

$$\rho := \rho_P := \frac{1}{\limsup_{n \rightarrow \infty} |a_n|^{1/n}}.$$

Suponhamos que o raio de convergência seja positivo (ou ainda infinito). Agora desejamos substituir a variável z por um elemento da álgebra de Banach \mathcal{A} . Isto faz sentido formalmente, porque numa álgebra de Banach produtos como x^k são definidos. Em particular, definimos $x^0 := \mathbb{1}$ para todo $x \in \mathcal{A}$. Portanto, levando isto em consideração a série resultante em \mathcal{A} será:

$$P_{\mathcal{A}}(x) := \sum_{k=0}^{\infty} a_k x^k. \quad (1.3.4)$$

Ora pode se levantar a questão para quais $x \in \mathcal{A}$ esta série é convergente.

Corolário 1.3.11. *A série (1.3.4) converge para todo $x \in \mathcal{A}$ com $\|x\| < \rho$.*

Demonstração. Seja $x \in \mathcal{A}$ com $\|x\| < \rho$. Segue da condição (1.3.1) na definição de álgebra de Banach que vale a desigualdade

$$\|a_k x^k\| \leq |a_k| \|x\|^k$$

para todo $k \in \mathbb{N}$. Como $\|x\| < \rho$, a série $\sum_{k=0}^{\infty} |a_k| \|x\|^k$ converge. Segundo o teste da comparação, a série $\sum_{k=0}^{\infty} \|a_k x^k\|$ também converge. Mas isso significa exatamente que a série $\sum_{k=0}^{\infty} a_k x^k$ converge absolutamente em \mathcal{A} , e daí segundo o Teorema 1.3.10 converge. \square

Portanto temos mostrado que a formula (1.3.4) define uma aplicação

$$P_{\mathcal{A}} : \mathcal{B}_{\rho} := \{x \in \mathcal{A} \mid \|x\| < \rho\} \rightarrow \mathcal{A} \quad (1.3.5)$$

onde ρ é o raio de convergência de P . O seguinte Lema mostra que esta aplicação é contínua. Desde já notamos que o conceito de diferenciabilidade, e de derivada, transfere-se para funções com valores num espaço de Banach ao pé da letra.

Definição 1.3.12. *Seja $I \subset \mathbb{R}$ um intervalo e $(X, \|\cdot\|)$ um espaço de Banach. Chama-se uma função $f : I \rightarrow X$ **diferenciável no ponto** $t \in I$ com **derivada** $a := f'(t) \in X$ se vale: $\forall \varepsilon > 0 \exists \delta > 0 \forall h \in \mathbb{R}$:*

$$0 < |h| < \delta, \quad t + h \in I \quad \Longrightarrow \quad \frac{\|f(t+h) - f(t) - ha\|}{|h|} < \varepsilon.$$

Equivalentemente isso significa que o quociente de diferença $h^{-1}(f(t+h) - f(t))$ converge em X para a (quando $h \rightarrow 0$). Em outras palavras

$$f'(t) := \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h}$$

caso o limite exista, e neste caso chamamos f diferenciável no ponto t .

Teorema 1.3.13. *Seja \mathcal{A} uma álgebra de Banach real e seja (1.3.3) uma série de potências com coeficientes reais e raio de convergência $\rho > 0$. Então vale o seguinte:*

(i) *Para todo $r < \rho$ a restrição da aplicação (1.3.5) para o subconjunto $\overline{\mathcal{B}}_r := \{x \in \mathcal{A} \mid \|x\| \leq r\}$ é Lipschitz-contínua; vale*

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \leq c_r \|x - y\|, \quad c_r := \sum_{k=1}^{\infty} k |a_k| r^{k-1}$$

para $x, y \in \mathcal{A}$ com $\|x\|, \|y\| \leq r$.

(ii) *Para todo $x \in \mathcal{A}$ a função definida por $t \mapsto P_{\mathcal{A}}(tx)$ no intervalo aberto $(-\rho/\|x\|, \rho/\|x\|)$ é diferenciável e vale*

$$\frac{d}{dt} P_{\mathcal{A}}(tx) = xQ_{\mathcal{A}}(tx), \quad Q(z) := P'(z) := \sum_{k=1}^{\infty} k a_k z^{k-1} \quad (1.3.6)$$

para todos os $t \in \mathbb{R}$ com $|t|\|x\| < \rho$.

Demonstração. Provaremos (i). Vale $x^k - y^k = \sum_{j=1}^k x^{k-j}(x-y)y^{j-1}$ para todo $k \in \mathbb{N}$ e todos os $x, y \in \mathcal{A}$, e daí

$$\|x\| \leq r, \quad \|y\| \leq r \quad \Longrightarrow \quad \|x^k - y^k\| \leq kr^{k-1} \|x - y\|.$$

Daqui resulta

$$\|P_{\mathcal{A}}(x) - P_{\mathcal{A}}(y)\| \leq \sum_{k=0}^{\infty} |a_k| \|x^k - y^k\| \leq \sum_{k=0}^{\infty} k |a_k| r^{k-1} \|x - y\| = c_r \|x - y\|$$

para todos os $x, y \in \mathcal{A}$ com $\|x\| \leq r$ e $\|y\| \leq r$. Com isso (i) fica provado.

Na prova de (ii) atentamos que $Q_{\mathcal{A}}(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$. Daqui vale

$$\begin{aligned}
& \|P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx) - hxQ_{\mathcal{A}}(tx)\| \\
&= \left\| \sum_{k=2}^{\infty} ((t+h)^k - t^k - kht^{k-1}) a_k x^k \right\| \\
&\leq \sum_{k=2}^{\infty} |(t+h)^k - t^k - kht^{k-1}| |a_k| \|x\|^k \\
&\leq \sum_{k=2}^{\infty} \sum_{j=2}^k \binom{k}{j} |h|^j |t|^{k-j} |a_k| \|x\|^k \\
&= \sum_{k=2}^{\infty} \left((|t|+|h|)^k - |t|^k - k|h||t|^{k-1} \right) |a_k| \|x\|^k \\
&= |p(|t|+|h|) - p(|t|) - |h|\dot{p}(|t|)|.
\end{aligned}$$

Aqui p denota a série de potências

$$p(\lambda) := \sum_{k=0}^{\infty} |a_k| \|x\|^k \lambda^k$$

com raio de convergência $\rho/\|x\|$. Seja $t \in \mathbb{R}$ com $|t| < \rho/\|x\|$. Então existem duas constantes $\delta > 0$ e $c > 0$, tal que $|t| + \delta < \rho/\|x\|$ e daí todo $h \in \mathbb{R}$ com $|h| < \delta$ satisfaz a desigualdade

$$|p(|t|+|h|) - p(|t|) - |h|\dot{p}(|t|)| \leq c|h|^2.$$

Daqui segue

$$\|P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx) - hxQ_{\mathcal{A}}(tx)\| \leq c|h|^2$$

para todos os $h \in \mathbb{R}$ com $|h| < \delta$. Com isso temos mostrado que

$$\lim_{h \rightarrow 0} \left\| \frac{P_{\mathcal{A}}((t+h)x) - P_{\mathcal{A}}(tx)}{h} - xQ_{\mathcal{A}}(tx) \right\| = 0$$

e daí segue (ii). Com isso Teorema 1.3.13 fica provado. \square

Consideramos o caso especial $\mathcal{A} := \mathbb{R}^{n \times n}$ com uma norma matriz $\|\cdot\|$. Então segue através de Teorema 1.3.13 que a série matriz $P(A) := \sum_{k=0}^{\infty} a_k A^k$ converge nesta norma para toda matriz $A \in \mathbb{R}^{n \times n}$ com $\|A\| < \rho$. Enquanto, segundo o Teorema 1.1.8, a conclusão (convergência) é independente da escolha da norma, precisamos na hipótese ($\|A\| < \rho$) uma norma matriz.

1.4 A série geométrica

A **série geométrica** tem raio de convergência $\rho = 1$ e representa a função

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \quad \text{para } z \in \mathbb{C} \text{ com } |z| < 1.$$

Segundo o Teorema 1.3.13 recebe-se daí o seguinte resultado.

Teorema 1.4.1. *Seja \mathcal{A} uma álgebra de Banach com $\|\mathbb{1}\| = 1$ e seja $\mathcal{G} \subset \mathcal{A}$ o grupo dos elementos invertíveis. Então vale o seguinte.*

(i) *Seja $x \in \mathcal{A}$ com $\|x\| < 1$. Então $\mathbb{1} - x \in \mathcal{G}$ e*

$$(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k. \quad (1.4.1)$$

(ii) *Seja $x \in \mathcal{G}$ e $y \in \mathcal{A}$, tal que*

$$\|x - y\| \|x^{-1}\| < 1. \quad (1.4.2)$$

Então $y \in \mathcal{G}$ e

$$y^{-1} = x^{-1} \sum_{k=0}^{\infty} (\mathbb{1} - yx^{-1})^k, \quad \|x^{-1} - y^{-1}\| \leq \frac{\|x^{-1}\|^2 \|x - y\|}{1 - \|x^{-1}\| \|x - y\|}. \quad (1.4.3)$$

(iii) *O grupo \mathcal{G} dos elementos invertíveis em \mathcal{A} é um subconjunto aberto de \mathcal{A} e a aplicação $\mathcal{G} \rightarrow \mathcal{G} : x \mapsto x^{-1}$ é contínua.*

Demonstração. Seja $x \in \mathcal{A}$ com $\|x\| < 1$. Então vale

$$\sum_{k=0}^{\infty} \|x^k\| \leq \sum_{k=0}^{\infty} \|x\|^k = \frac{1}{1 - \|x\|} < \infty.$$

Daí converge a série $\sum_{k=0}^{\infty} x^k$, segundo o Teorema 1.3.10. Para $\forall n \in \mathbb{N}$ vale

$$(\mathbb{1} - x) \sum_{k=0}^n x^k = \sum_{k=0}^n x^k (\mathbb{1} - x) = \sum_{k=0}^n x^k - \sum_{k=0}^n x^{k+1} = \mathbb{1} - x^{n+1}$$

e daí segue $(\mathbb{1} - x) \sum_{k=0}^{\infty} x^k = \sum_{k=0}^{\infty} x^k (\mathbb{1} - x) = \mathbb{1}$. Portanto $\mathbb{1} - x \in \mathcal{G}$ e $(\mathbb{1} - x)^{-1} = \sum_{k=0}^{\infty} x^k$ e com isso (i) fica provado.

Ora, sejam $x \in \mathcal{G}$ e $y \in \mathcal{A}$ dados com $\|x - y\| \|x^{-1}\| < 1$. Então vale

$$\|\mathbb{1} - yx^{-1}\| = \|(x - y)x^{-1}\| \leq \|x - y\| \|x^{-1}\| < 1.$$

Segundo parte (i), segue dali $yx^{-1} \in \mathcal{G}$ e $(yx^{-1})^{-1} = \sum_{k=0}^{\infty} (\mathbb{1} - yx^{-1})^k$. Daqui $y \in \mathcal{G}$ e vale

$$y^{-1} = x^{-1}(yx^{-1})^{-1} = x^{-1} \sum_{k=0}^{\infty} (\mathbb{1} - yx^{-1})^k$$

e

$$\begin{aligned} \|x^{-1} - y^{-1}\| &= \left\| x^{-1} \sum_{k=1}^{\infty} (\mathbb{1} - yx^{-1})^k \right\| \\ &\leq \|x^{-1}\| \sum_{k=1}^{\infty} \|\mathbb{1} - yx^{-1}\|^k \\ &= \|x^{-1}\| \frac{\|\mathbb{1} - yx^{-1}\|}{1 - \|\mathbb{1} - yx^{-1}\|} \\ &\leq \frac{\|x^{-1}\|^2 \|x - y\|}{1 - \|x^{-1}\| \|x - y\|}. \end{aligned}$$

Com isto parte (ii) fica provado e parte (iii) segue imediatamente de (ii). Com isso Teorema 1.4.1 fica provado. \square

Exemplo 1.4.2. (i) Consideramos a álgebra de Banach

$$\mathcal{A} = \mathbb{R}^{n \times n}$$

de todas as matrizes reais quadradas $n \times n$. Toda matriz $A \in \mathbb{R}^{n \times n}$ determina uma aplicação linear $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ através de $T_A x := Ax$ para $x \in \mathbb{R}^n$. Segundo a definição, uma matriz $A \in \mathbb{R}^{n \times n}$ é um elemento invertível de \mathcal{A} se e somente se existe uma matriz $B \in \mathbb{R}^{n \times n}$ com a propriedade $AB = BA = \mathbb{1}$. Tal matriz B existe se e somente se a aplicação linear T_A é bijetiva. Neste caso a matriz inversa $A^{-1} = B$ induz a aplicação inversa $T_A^{-1} = T_{A^{-1}}$. Da *Álgebra Linear* sabemos que as matrizes invertíveis são precisamente aquelas com determinante não-nulo (veja o Apêndice D). Por isso o grupo dos elementos invertíveis em $\mathcal{A} = \mathbb{R}^{n \times n}$ é dado por $\text{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\}$. Isto é o **grupo linear geral**. Segundo o Teorema 1.4.1 isto é um subconjunto aberto de $\mathbb{R}^{n \times n}$ e a aplicação $\text{GL}(n, \mathbb{R}) \rightarrow \text{GL}(n, \mathbb{R}) : A \mapsto A^{-1}$ é contínua.

(ii) Seja X um espaço de Banach, e sejam $\mathcal{A} := \mathcal{L}(X)$ e $\text{GL}(X) := \mathcal{G}$. Segundo o Teorema 1.4.1 a aplicação $\text{GL}(X) \rightarrow \text{GL}(X) : A \mapsto A^{-1}$ é contínua.

1.5 A aplicação exponencial

A série de potências $\exp(z) := \sum_{k=0}^{\infty} z^k/k!$ tem raio de convergência $\rho = \infty$ e induz assim uma aplicação exponencial em qualquer uma álgebra de Banach.

Definição 1.5.1. *Seja \mathcal{A} uma álgebra de Banach real. A aplicação exponencial em \mathcal{A} é a aplicação*

$$\exp : \mathcal{A} \rightarrow \mathcal{A}$$

definido através de

$$\exp(x) := e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \text{para } x \in \mathcal{A}. \quad (1.5.1)$$

Segundo o Teorema 1.3.13 a série em (1.5.1) converge absolutamente em todo ponto $x \in \mathcal{A}$.

Teorema 1.5.2. *Seja \mathcal{A} uma álgebra de Banach real. Então a aplicação exponencial $\exp : \mathcal{A} \rightarrow \mathcal{A}$ em (1.5.1) é contínua e tem as seguintes propriedades.*

(i) *Para todo $x \in \mathcal{A}$ e toda sequência $(x_n)_{n \in \mathbb{N}}$ em \mathcal{A} vale*

$$\lim_{n \rightarrow \infty} x_n = x \quad \implies \quad \exp(x) = \lim_{n \rightarrow \infty} \left(\mathbb{1} + \frac{x_n}{n} \right)^n. \quad (1.5.2)$$

(ii) *Vale $\exp(0) = \mathbb{1}$ e, para todos os $x, y \in \mathcal{A}$,*

$$xy = yx \quad \implies \quad \exp(x + y) = \exp(x) \exp(y). \quad (1.5.3)$$

(iii) *Para todo $x \in \mathcal{A}$ vale*

$$\exp(x) \in \mathcal{G}, \quad \exp(x)^{-1} = \exp(-x). \quad (1.5.4)$$

(iv) *Para todo $x \in \mathcal{A}$ e todo $y \in \mathcal{G}$ vale*

$$\exp(yxy^{-1}) = y \exp(x) y^{-1}. \quad (1.5.5)$$

(v) *Para todo $x \in \mathcal{A}$ a função $\mathbb{R} \rightarrow \mathcal{A} : t \mapsto \exp(tx)$ é continuamente diferenciável e a derivada dela é*

$$\frac{d}{dt} \exp(tx) = x \exp(tx). \quad (1.5.6)$$

Demonstração. A continuidade da aplicação exponencial é uma consequência de Teorema 1.3.13 e a afirmação (i) prova-se como no caso de $\mathcal{A} = \mathbb{R}$ (veja [4]). Ora lembramos o argumento da prova de (i). Seja $\varepsilon > 0$ dado e seja $N \in \mathbb{N}$ escolhido, tal que $\|x_n - x\| \leq 1$ para todo número natural $n \geq N$ e

$$\sum_{k=N}^{\infty} \frac{(\|x\| + 1)^k}{k!} < \frac{\varepsilon}{3}. \quad (1.5.7)$$

Então vale $\lim_{n \rightarrow \infty} \binom{n}{k} \frac{1}{n^k} = \frac{1}{k!}$ para todo $k \in \{0, 1, \dots, N-1\}$. Dali existe um número inteiro $n_0 \geq N$, tal que para todos os $n \in \mathbb{N}$ vale o seguinte:

$$n \geq n_0 \quad \implies \quad \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| < \frac{\varepsilon}{3}. \quad (1.5.8)$$

Ora, seja $n \in \mathbb{N}$ com $n \geq n_0$. Então $n \geq N$ e vale

$$\begin{aligned} \left\| \left(\mathbb{1} + \frac{x_n}{n} \right) - e^x \right\| &= \left\| \sum_{k=0}^n \binom{n}{k} \frac{x_n^k}{n^k} - \sum_{k=0}^{\infty} \frac{x^k}{k!} \right\| \\ &\leq \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| + \sum_{k=N}^n \binom{n}{k} \frac{\|x_n\|^k}{n^k} + \sum_{k=N}^{\infty} \frac{\|x\|^k}{k!} \\ &\leq \sum_{k=0}^{N-1} \left\| \binom{n}{k} \frac{x_n^k}{n^k} - \frac{x^k}{k!} \right\| + 2 \sum_{k=N}^{\infty} \frac{(\|x\| + 1)^k}{k!} \\ &< \varepsilon. \end{aligned}$$

Aqui segue o penúltimo passo das desigualdades $\|x_n\| \leq \|x\| + 1$ para $n \geq N$ e $\binom{n}{k} \frac{1}{n^k} \leq \frac{1}{k!}$ para $n \in \mathbb{N}$ com $n \geq k$. O último passo segue de (1.5.7) e (1.5.8). Com isso (i) fica provado.

A afirmação (ii) prova-se igualmente como no caso de $\mathcal{A} = \mathbb{R}$ (veja [4]). Pode-se também deduzir esta afirmação no entanto diretamente de (i), porque se $x, y \in \mathcal{A}$ comutam tal que vale $xy = yx$, então vale

$$\left(\mathbb{1} + \frac{x}{n} \right)^n \left(\mathbb{1} + \frac{y}{n} \right)^n = \left(\mathbb{1} + \frac{x + y + xy/n}{n} \right)^n$$

para todo $n \in \mathbb{N}$. Na passagem ao limite $n \rightarrow \infty$ resulta daí, segundo (i), a equação desejada $\exp(x) \exp(y) = \exp(x + y)$.

Parte (iii) segue imediatamente de parte (ii) com $y = -x$, parte (iv) segue da equação $(yxy^{-1})^k = yx^ky^{-1}$ para $y \in \mathcal{G}$, e parte (v) segue do Teorema 1.3.13. Com isso Teorema 1.5.2 fica provado. \square

Sistemas de equações diferenciais lineares

A aplicação exponencial na álgebra de Banach $\mathcal{A} = \mathbb{R}^{n \times n}$ de todas as matrizes $(n \times n)$ desempenha um papel importante na solução do sistema de n equações diferenciais ordinárias (EDOs) lineares de primeira ordem e em n variáveis. Um tal sistema de equações diferenciais tem a forma

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, & x_1(0) &= x_{01}, \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, & x_2(0) &= x_{02}, \\ &\vdots & & \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n, & x_n(0) &= x_{0n}. \end{aligned} \tag{1.5.9}$$

Aqui os $a_{ij} \in \mathbb{R}$ são coeficientes reais constantes e, para todo i , a variável x_i é uma aplicação continuamente diferenciável $x_i : \mathbb{R} \rightarrow \mathbb{R}$. Pensamos sobre a variável desta aplicação como variável do tempo t e denotamos com $\dot{x}_i = dx_i/dt$ a derivada da função x_i com respeito ao tempo t . Na interpretação física a n -tuplo

$$x(t) := (x_1(t), \dots, x_n(t)) \in \mathbb{R}^n$$

denota o vetor da posição de uma partícula (ou várias partículas simultaneamente) no momento t e a derivada $\dot{x}(t) = (\dot{x}_1(t), \dots, \dot{x}_n(t)) \in \mathbb{R}^n$ denota o vetor velocidade no momento t . Englobando os coeficientes constantes numa matriz $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ podemos escrever o sistema de equações diferenciais (1.5.9) na forma curta

$$\dot{x} = Ax, \quad x(0) = x_0. \tag{1.5.10}$$

Aqui $x_0 := (x_{01}, \dots, x_{0n}) \in \mathbb{R}^n$ denota o **vetor inicial** (ou o vetor posição no momento $t = 0$). Segue da equação (1.5.6) que a única solução da equação diferencial (1.5.10) é dado pela fórmula

$$x(t) := e^{At}x_0 \tag{1.5.11}$$

onde

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} \in \mathbb{R}^{n \times n}$$

denota a matriz exponencial. A convergência desta série segue da convergência da aplicação exponencial para álgebras de Banach arbitrárias \mathcal{A} no caso especial $\mathcal{A} = \mathbb{R}^{n \times n}$ (veja o Teorema 1.5.2).

Exemplo 1.5.3. Para

$$\Lambda := \text{diag}(\lambda_1, \dots, \lambda_n)$$

vale $\Lambda^k = \text{diag}(\lambda_1^k, \dots, \lambda_n^k)$ e assim

$$e^\Lambda = \sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n}).$$

Exemplo 1.5.4. Para $\lambda \in \mathbb{R}$ vale

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \implies e^{At} = e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercício: Use esta fórmula para encontrar as soluções de um sistema adequado de três equações diferenciais de primeira ordem em três variáveis. Generalize a fórmula neste exemplo para quatro e cinco variáveis, e então para matrizes de tamanho arbitrário.

Exemplo 1.5.5. Para

$$H := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

vale $H^{2k} = (-1)^k \mathbb{1}$ e assim

$$\begin{aligned} e^{Ht} &= \left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} \mp \dots \right) \mathbb{1} + \left(t - \frac{t^3}{3!} + \frac{t^5}{5!} \mp \dots \right) H \\ &= \begin{pmatrix} \cos t & -\text{sen } t \\ \text{sen } t & \cos t \end{pmatrix}. \end{aligned}$$

(Veja [4] para a representação como série de potências do seno e do coseno.) Como as matrizes $a\mathbb{1}$ e bH comutam, vale para a matriz

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

a equação

$$e^{tA} = e^{at} \begin{pmatrix} \cos(bt) & -\text{sen}(bt) \\ \text{sen}(bt) & \cos(bt) \end{pmatrix}$$

para todos os $t \in \mathbb{R}$.

1.6 Composição de séries de potências

Sejam $(a_k)_{k \geq 1}$ e $(b_\ell)_{\ell \geq 0}$ duas sequências de números reais, tal que as séries de potências

$$P(z) := \sum_{k=1}^{\infty} a_k z^k, \quad Q(z) := \sum_{\ell=1}^{\infty} b_\ell z^\ell$$

tem raios de convergência positivos, ou seja,

$$\rho_P := \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}} > 0, \quad \rho_Q := \frac{1}{\limsup_{\ell \rightarrow \infty} |b_\ell|^{1/\ell}} > 0.$$

Particularmente vale $P(0) = 0$. Seja $0 < \rho \leq \rho_P$ escolhido, tal que

$$z \in \mathbb{C}, \quad |z| < \rho \quad \implies \quad \sum_{k=1}^{\infty} |a_k| |z|^k < \rho_Q. \quad (1.6.1)$$

Seja $z \in \mathbb{C}$ com módulo $|z| < \rho \leq \rho_P$, então vale $|P(z)| < \rho_Q$ segundo (1.6.1). Assim a composição $Q \circ P : \{z \in \mathbb{C} \mid |z| < \rho\} \rightarrow \mathbb{C}$ é bem definida. O teorema seguinte diz que esta composição mesma pode ser representado por uma série de potências.

Teorema 1.6.1. *Para $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ seja*

$$c_k := \sum_{\ell=0}^k b_\ell \ell! \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{a_i^{m_i}}{m_i!}. \quad (1.6.2)$$

A segunda soma em (1.6.2) é sobre todas as sequências $(m_i)_{i \in \mathbb{N}}$ de número inteiros não-negativos as quais satisfazem as equações $\sum_i i m_i = k$ e $\sum_i m_i = \ell$. Como $m_i \leq k$ para todo i e $m_i = 0$ para $i > k$, esta soma é finita e se anula caso $\ell > k$. Seja

$$R(z) := \sum_{k=0}^{\infty} c_k z^k \quad (1.6.3)$$

a série de potências com coeficientes c_k . Então vale o seguinte.

(i) O raio de convergência $\rho_R := (\limsup_{k \rightarrow \infty} |c_k|^{1/k})^{-1}$ da série de potências R é maior ou igual ρ e $R(z) = Q(P(z))$ para todos os $z \in \mathbb{C}$ com $|z| < \rho$.

(ii) Seja \mathcal{A} uma álgebra de Banach real. Então todo $x \in \mathcal{A}$ com $\|x\| < \rho$ satisfaz a desigualdade $\|P_{\mathcal{A}}(x)\| < \rho_Q$ e a identidade

$$R_{\mathcal{A}}(x) = Q_{\mathcal{A}}(P_{\mathcal{A}}(x)). \quad (1.6.4)$$

Demonstração. Para $k \in \mathbb{N}_0$ e $N \in \mathbb{N}$ sejam os coeficientes reais $c_{k,N}$ dado através de

$$c_{k,N} := \sum_{\ell=0}^N b_{\ell} \ell! \sum_{\substack{m_1+\dots+m_N=\ell \\ \sum_{i=1}^N i m_i=k}} \frac{a_1^{m_1} a_2^{m_2} \dots a_N^{m_N}}{m_1! \dots m_N!}.$$

Então vale $c_{k,N} = c_k$ para $k \leq N$ e $c_{k,N} = 0$ para $k > N^2$ (porque toda N -tuplo $(m_1, \dots, m_N) \in \mathbb{N}_0^N$ com $\sum_i m_i = \ell \leq N$ satisfaz a desigualdade $k = \sum_i i m_i \leq N \sum_i m_i = N\ell \leq N^2$). Particularmente a função

$$R_N(z) := \sum_{k=0}^{\infty} c_{k,N} z^k$$

é um polinômio. Para $N \in \mathbb{N}$ sejam os polinômio P_N e Q_N definido através de

$$P_N(z) := \sum_{k=1}^N a_k z^k, \quad Q_N(z) := \sum_{\ell=0}^N b_{\ell} z^{\ell}$$

para $z \in \mathbb{C}$. Então vale para todos os $z \in \mathbb{C}$ a equação

$$\begin{aligned} R_N(z) &= \sum_{k=0}^{\infty} c_{k,N} z^k \\ &= \sum_{k=0}^{\infty} \sum_{\ell=0}^N b_{\ell} \ell! \sum_{\substack{m_1+\dots+m_N=\ell \\ \sum_{i=1}^N i m_i=k}} \frac{a_1^{m_1} a_2^{m_2} \dots a_N^{m_N}}{m_1! \dots m_N!} z^k \\ &= \sum_{\ell=0}^N b_{\ell} \sum_{m_1+\dots+m_N=\ell} \frac{\ell!}{m_1! \dots m_N!} (a_1 z)^{m_1} (a_2 z^2)^{m_2} \dots (a_N z^N)^{m_N} \\ &= \sum_{\ell=0}^N b_{\ell} (a_1 z + a_2 z^2 + \dots + a_N z^N)^{\ell} \\ &= Q_N(P_N(z)). \end{aligned}$$

Ora, seja $0 < r < \rho$. Então vale $s := \sum_{k=1}^{\infty} |a_k| r^k < \rho_Q$. Além disso, seja $z \in \mathbb{C}$ com $|z| \leq r$. Então $P_N(z)$ converge para $P(z)$ e vale $|P_N(z)| \leq s$

para todo $N \in \mathbb{N}$. Como Q_N converge ao longo da bola $\{z \in \mathbb{C} \mid |z| \leq s\}$ uniformemente para Q , segue então que $Q_N(P_N(z))$ converge para $Q(P(z))$, portanto vale

$$\lim_{N \rightarrow \infty} R_N(z) = Q(P(z)) \quad \text{para todo } z \in \mathbb{C} \text{ com } |z| \leq r. \quad (1.6.5)$$

Substituindo os coeficientes a_k e b_ℓ pelos módulos dos mesmos, teremos

$$|c_k| \leq \bar{c}_k := \sum_{\ell=0}^k |b_\ell| \ell! \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{a_i^{m_i}}{m_i!}$$

$$|c_{k,N}| \leq \bar{c}_{k,N} := \sum_{\ell=0}^N |b_\ell| \ell! \sum_{\substack{m_1 + \dots + m_N = \ell \\ \sum_{i=1}^N i m_i = k}} \frac{|a_1|^{m_1} |a_2|^{m_2} \dots |a_N|^{m_N}}{m_1! \dots m_N!}.$$

Como $\bar{c}_{k,N} \leq \bar{c}_k$ para todos os k, N bem como $\bar{c}_{k,N} = \bar{c}_k$ para todo $k \leq N$ e $\bar{c}_{k,N} = 0$ para todo $k > N^2$, segue dali para qualquer um $N' \in \mathbb{N}$ que

$$N' \leq \sqrt{N} \quad \Longrightarrow \quad \sum_{k=0}^{\infty} \bar{c}_{k,N'} r^k \leq \sum_{k=0}^N \bar{c}_k r^k \leq \sum_{k=0}^{\infty} \bar{c}_{k,N} r^k. \quad (1.6.6)$$

Aqui os dois termos exteriores convergem para

$$\sum_{\ell=0}^{\infty} |b_\ell| \left(\sum_{k=1}^{\infty} |a_k| r^k \right)^\ell = \sum_{\ell=0}^{\infty} |b_\ell| s^\ell$$

quando $N, N' \rightarrow \infty$, através do mesmo argumento como acima.

Consequentemente a diferença $\sum_{k>N} \bar{c}_{k,N} r^k$ dos dois últimos termos em (1.6.6) converge para zero, e para $z \in \mathbb{C}$ com $|z| \leq r$ vale daí

$$\left| R_N(z) - \sum_{k=0}^N c_k z^k \right| = \left| \sum_{k=N+1}^{\infty} c_{k,N} z^k \right| \leq \sum_{k=N+1}^{\infty} \bar{c}_{k,N} r^k \xrightarrow{N \rightarrow \infty} 0. \quad (1.6.7)$$

Segue de (1.6.5) e (1.6.7) que vale a equação

$$R(z) = \sum_{k=0}^{\infty} c_k z^k = \lim_{N \rightarrow \infty} \sum_{k=0}^N c_k z^k = \lim_{N \rightarrow \infty} R_N(z) = Q(P(z))$$

para todo $z \in \mathbb{C}$ com $|z| \leq r$. Como o número real r foi escolhido arbitrariamente no intervalo $0 < r < \rho$, segue daqui que o raio de convergência da série de potências R satisfaz a desigualdade $\rho_R \geq \rho$ e que $R(z) = Q(P(z))$ para todo $z \in \mathbb{C}$ com $|z| < \rho$. Com isso (i) fica provado. Parte (ii) segue do mesmo argumento, no que substitue-se $z \in \mathbb{C}$ por $x \in \mathcal{A}$ e o módulo de um número complexo pela norma em \mathcal{A} . Com isso Teorema 1.6.1 fica provado. \square

Exemplo 1.6.2 (Logaritmo). A série

$$L(z) := \sum_{k=1}^{\infty} \frac{(-1)^{k-1} z^k}{k} = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} \pm \dots$$

tem raio de convergência $\rho_L = 1$ e representa a função $L(z) = \log(1+z)$ para $z \in \mathbb{C}$ com $|z| < 1$. Ou seja, L satisfaz a equação $\exp(L(z)) = 1+z$ para todos os $z \in \mathbb{C}$ com $|z| < 1$. (Veja [2, Seite 44, Übung 1.7.18].) Segundo parte (i) do Teorema 1.6.1, segue daí primeiramente que os coeficientes $c_0 = c_1 = 1$ e $c_k = 0$, para $k \geq 2$, da série de potências $1+z$ são dados pela fórmula (1.6.2) com $a_i = (-1)^{i-1}/i$ e $b_\ell = 1/\ell!$. Ou seja, para todos os $k \in \mathbb{N}_0$ vale

$$\sum_{\ell=0}^k (-1)^{k-\ell} \sum_{\substack{\sum_i m_i = \ell \\ \sum_i i m_i = k}} \prod_{\substack{i \in \mathbb{N} \\ m_i > 0}} \frac{1}{m_i! i^{m_i}} = \begin{cases} 1, & \text{caso } k = 0, 1, \\ 0, & \text{caso } k \geq 2. \end{cases} \quad (1.6.8)$$

Além disso, segue da parte (ii) do Teorem 1.6.1 que, no caso de uma álgebra de Banach \mathcal{A} , todo $x \in \mathcal{A}$ com $\|x\| < 1$ satisfaz a equação

$$\exp\left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1} x^k}{k}\right) = \mathbb{1} + x. \quad (1.6.9)$$

Particularmente vale a equação (1.6.9) para toda matriz $x \in \mathbb{R}^{n \times n}$ com $\|x\| < 1$, onde $\|\cdot\|$ é uma norma matriz qualquer no $\mathbb{R}^{n \times n}$.

1.7 Integração com valores vetoriais

Queremos definir a integral de uma função contínua num intervalo compacto com valores num espaço de Banach X . Para ilustração suponha $\dim X = n < \infty$. Então no caso $X = \mathbb{R}^n$ a ideia é definir a integral simplesmente coordenada por coordenada, ou seja, a integral de uma função contínua

$x = (x_1, \dots, x_n) : [a, b] \rightarrow \mathbb{R}^n$ é o vetor com i -ésima coordenada dada pela integral da função $x_i : [a, b] \rightarrow \mathbb{R}$. Se o espaço vetorial de dimensão n ainda não for munido de uma base, então pode-se, para esta definição, simplesmente escolher uma base a priori. Porém deve-se mostrar que tal definição não depende da escolha da base. Uma outra possibilidade, para a definição da integral de uma função com valores vetoriais num espaço de Banach qualquer, seria construir-lá novamente através da convergência das somas de Riemann. Qualquer que seja (em dimensão finita) o método empregado, deve-se provar que ambos chegam ao mesmo resultado. Isso não é difícil e será provado no seguinte lema, no qual também tratamos o caso geral de um espaço de Banach qualquer (baseado num corolário do Teorema de Hahn-Banach).

Lema 1.7.1. *Seja X um espaço vetorial normado de dimensão n com uma base f_1, \dots, f_n . Sejam $a < b$ números reais, e seja $\xi : [a, b] \rightarrow X$ uma aplicação contínua. Então para todo $x \in X$ são equivalente as seguintes afirmações.*

(i) *Para todo funcional linear contínuo $\Lambda : X \rightarrow \mathbb{R}$ vale*

$$\Lambda(x) = \int_a^b \Lambda(\xi(t)) dt. \quad (1.7.1)$$

(ii) *Sejam $\xi_1, \dots, \xi_n : [a, b] \rightarrow \mathbb{R}$ as funções contínuas definidas através de $\xi(t) =: \sum_{i=1}^n \xi_i(t) f_i$ para $t \in [a, b]$, então vale*

$$x = \sum_{i=1}^n \left(\int_a^b \xi_i(t) dt \right) f_i. \quad (1.7.2)$$

(iii) *O vetor x é o limite*

$$x = \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right). \quad (1.7.3)$$

Demais existe um único vetor $x \in X$ que satisfaz as condições equivalentes. Para um espaço de Banach X tudo mantém-se válido se omitimos base e (ii).

Definição 1.7.2. *Seja X um espaço de Banach. O vetor único $x \in X$ no Lema 1.7.1 é chamado de **integral de ξ sobre $[a, b]$** e denotado*

$$\int_a^b \xi(t) dt := x.$$

Ou seja, a integral é caracterizada pela condição

$$\Lambda \left(\int_a^b \xi(t) dt \right) = \int_a^b \Lambda(\xi(t)) dt \quad (1.7.4)$$

para toda aplicação linear $\Lambda : X \rightarrow \mathbb{R}$.

Demonstração do Lema 1.7.1. Caso $\dim X < \infty$. A idéia é traduzir o caso geral ao caso padrão da integral de Riemann para funções com valores reais. Os tradutores serão funcionais lineares contínuas.

Provaremos (i) \implies (ii). Em dimensão finita um funcional linear é contínuo. Sejam $\Lambda_i : X \rightarrow \mathbb{R}$ os funcionais lineares definidos através de

$$\Lambda_i(f_j) := \delta_{ij}$$

para $i, j = 1, \dots, n$. Então vale $\Lambda_i(\xi(t)) = \Lambda_i(\sum_{j=1}^n \xi_j(t)f_j) = \xi_i(t)$ para todos os t e i e daí, segundo (i),

$$x = \sum_{i=1}^n \Lambda_i(x)f_i = \sum_{i=1}^n \left(\int_a^b \Lambda_i(\xi(t)) dt \right) f_i = \sum_{i=1}^n \left(\int_a^b \xi_i(t) dt \right) f_i.$$

Com isso fica mostrado que o vetor x satisfaz a condição (ii).

Provaremos (ii) \implies (iii). Na situação de (ii) vale

$$x_i := \int_a^b \xi_i(t) dt = \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi_i \left(a + k \frac{b-a}{N} \right)$$

para $i = 1, \dots, n$ segundo [5, Satz 4.1]. Como $\xi(t) = \sum_{i=1}^n \xi_i(t)f_i$ para todo t pela hipótese, e como $x = \sum_{i=1}^n x_i f_i$ segundo (ii), segue daí que

$$\begin{aligned} x &= \sum_{i=1}^n \left(\lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi_i \left(a + k \frac{b-a}{N} \right) \right) f_i \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \sum_{i=1}^n \frac{b-a}{N} \xi_i \left(a + k \frac{b-a}{N} \right) f_i \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \end{aligned}$$

e daí o vetor x satisfaz a condição (iii).

Provaremos (iii) \implies (i). Como todo funcional linear $\Lambda : X \rightarrow \mathbb{R}$ é contínuo segundo o Corolário 1.2.2, segue de (iii) a equação

$$\begin{aligned} \Lambda(x) &= \Lambda \left(\lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \right) \\ &= \lim_{N \rightarrow \infty} \Lambda \left(\sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \right) \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} (\Lambda \circ \xi) \left(a + k \frac{b-a}{N} \right) \\ &= \int_a^b (\Lambda \circ \xi)(t) dt. \end{aligned}$$

Aqui a última igualdade segue de [5, Satz 4.1] para a função contínua $\Lambda \circ \xi : [a, b] \rightarrow \mathbb{R}$ com valores *reais*. Desta forma fica provado que o vetor x satisfaz a condição (i).

A existência e unicidade de um vetor $x \in X$, o qual satisfaz as três condições equivalentes, segue imediatamente da parte (ii). Com isso o Lema 1.7.1 fica provado no caso de $\dim X < \infty$. \square

Demonstração do Lema 1.7.1. Caso $\dim X = \infty$. (iii) \implies (i). Como em dimensão finita.

(i) \implies (iii). Seja $(\mathcal{B}, \|\cdot\|_\infty)$ o espaço Banach das funções limitadas $[a, b] \rightarrow X$ e seja \mathcal{C} o subespaço composto das funções contínuas. Chama-se uma função $f : [a, b] \rightarrow X$ de *tipo escada* se a sua imagem $f([a, b]) = \{v_1, \dots, v_N\} \subset X$ é um conjunto finito. Para todo 'degrau' $f^{-1}(v_k)$ multiplique sua 'altura' $v_k \in X$ com seu comprimento $c_k \in \mathbb{R}$, depois faça a soma sobre todos degraus de f para receber um elemento de X denotado de $I(f)$. O conjunto \mathcal{S} das funções escada $[a, b] \rightarrow X$ é um subespaço de \mathcal{B} e

$$I : \mathcal{S} \rightarrow X, \quad f \mapsto I(f)$$

é uma aplicação linear e contínua (Exercício) a qual, pelo teorema da extensão limitada, estende-se a um operador linear contínuo definido no fecho, ainda denotado $I : \overline{\mathcal{S}} \rightarrow X$. Embora que \mathcal{C} não é contido em \mathcal{S} , é contido no fecho $\overline{\mathcal{S}}$: Como $[a, b]$ é compacto, todo elemento ξ de \mathcal{C} é uniformemente contínuo, então um limite de funções escada. Com efeito $\xi = \lim_{N \rightarrow \infty} \xi_N$, onde ξ_N é a função escada com N degraus do mesmo comprimento $(b-a)/N$ e valor

$v_k := \xi(a + k\frac{b-a}{N})$ ao longo do k -ésimo degrau cujo ponto final seja $a + k\frac{b-a}{N}$.

(Exercício: Inclusões estritas $\mathcal{C} \subset \overline{\mathcal{F}} \subset \mathcal{B}$ de subespaços fechados.)

Agora seja $\xi : [a, b] \rightarrow X$ contínuo e suponha que $x \in X$ satisfaça (i). Existência do limite em (1.7.3) segue da continuidade de I , com efeito

$$\lim_{N \rightarrow \infty} I(\xi_N) = I\left(\lim_{N \rightarrow \infty} \xi_N\right) = I(\xi). \quad (1.7.5)$$

Seja $\Lambda : X \rightarrow \mathbb{R}$ contínuo e linear, então

$$\begin{aligned} \Lambda\left(\lim_{N \rightarrow \infty} I(\xi_N)\right) &= \Lambda\left(\lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} \xi\left(a + k\frac{b-a}{N}\right)\right) \\ &= \lim_{N \rightarrow \infty} \Lambda\left(\sum_{k=1}^N \frac{b-a}{N} \xi\left(a + k\frac{b-a}{N}\right)\right) \\ &= \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{b-a}{N} (\Lambda \circ \xi)\left(a + k\frac{b-a}{N}\right) \\ &= \int_a^b (\Lambda \circ \xi)(t) dt \\ &= \Lambda(x). \end{aligned}$$

Aqui a penúltima igualdade segue de [5, Satz 4.1] para a função contínua $\Lambda \circ \xi : [a, b] \rightarrow \mathbb{R}$ com valores *reais*. A última igualdade usa a hipótese (i). Ora, suponha por absurdo que $x_0 := x - \lim_{N \rightarrow \infty} I(\xi_N)$ não anule-se. Então um corolário conhecido do Teorema de Hahn-Banach associa a $x_0 \neq 0$ um $\Lambda_0 \in X^*$ tal que $\Lambda_0(x_0) = \|x_0\| > 0$. Mas $\Lambda_0(x_0) = 0$ pelas equações seguindo (1.7.5). Com esta contradição a parte (iii) fica provada.

Existência e unicidade de x . Com respeito à existência de um vetor $x \in X$, o qual satisfaz as duas condições equivalentes (i) e (iii), use a parte (iii) como definição. A existência do limite foi mostrado em (1.7.5). Para ver unicidade suponha por absurdo que um $y \in X$, diferente de x , também satisfaça (1.7.1). Neste caso um corolário do Teorema de Hahn-Banach associa ao vetor não-nulo $x - y \in X$ um $\psi \in X^*$, tal que $\psi(x - y) = \|x - y\|$. Então $\psi(x - y) > 0$ como $x - y \neq 0$. Por outro lado, de (1.7.1) para $\Lambda = \psi$ resulta que $\psi(x - y) = \psi(x) - \psi(y) = 0$ anula-se. Contradição.

Com isso o caso $\dim X = \infty$, e assim o Lema 1.7.1, fica provado. \square

No seguinte denotamos de $C([a, b], X)$, para $a < b$, o espaço vetorial das funções contínuas $x : [a, b] \rightarrow X$ onde X é um espaço de Banach.

Lema 1.7.3 (Propriedades da integral). Para $a < b$ vale o seguinte:

(i) A aplicação $C([a, b], X) \rightarrow X : x \mapsto \int_a^b x(t) dt$ é linear.

(ii) Para todo $x \in C([a, b], X)$ e todo $A \in \mathcal{L}(X, Y)$ vale

$$\int_a^b Ax(t) dt = A \int_a^b x(t) dt. \quad (1.7.6)$$

(iii) Para todo $x \in C([a, b], X)$ vale

$$\left\| \int_a^b x(t) dt \right\|_X \leq \int_a^b \|x(t)\|_X dt \quad (1.7.7)$$

(iv) Para todo $x \in C([a, b], X)$ e todo $c \in (a, b)$ vale

$$\int_a^c x(t) dt + \int_c^b x(t) dt = \int_a^b x(t) dt \quad (1.7.8)$$

(v) Se $x : [a, b] \rightarrow X$ é continuamente diferenciável, então vale

$$\int_a^b \dot{x}(t) dt = x(b) - x(a). \quad (1.7.9)$$

(vi) Se $\xi : [a, b] \rightarrow X$ é contínuo, então a função

$$x(t) := \int_a^t \xi(s) ds, \quad a \leq t \leq b \quad (1.7.10)$$

é continuamente diferenciável com $\dot{x}(t) = \xi(t)$ para todos os $t \in [0, 1]$.

Demonstração. Parte (i) segue diretamente da definição da integral.

Para parte (ii) escolhemos um funcional linear $\Lambda : Y \rightarrow \mathbb{R}$. Então segue

$$\Lambda \left(\int_a^b Ax(t) dt \right) = \int_a^b \Lambda(Ax(t)) dt = \Lambda \left(A \int_a^b x(t) dt \right)$$

de (1.7.4) para Λ e para ΛA . Daí segue (ii) pela unicidade no Lema 1.7.1.

Para a demonstração de (iii) usamos a desigualdade

$$\left\| \sum_{k=1}^N \frac{b-a}{N} \xi \left(a + k \frac{b-a}{N} \right) \right\| \leq \sum_{k=1}^N \frac{b-a}{N} \left\| \xi \left(a + k \frac{b-a}{N} \right) \right\|$$

para $N \in \mathbb{N}$. Daí segue (1.7.7) com $N \rightarrow \infty$, segundo parte (iii) do Lema 1.7.1.

As afirmações em (iv), (v), e (vi) seguem diretamente através das afirmações correspondentes para as integrais de funções com valores reais e parte (i) do Lema 1.7.1. Com isso Lema 1.7.3 fica provado. \square

Observações finais

O Teorema sobre a existência e unicidade de soluções de EDOs será provado no Capítulo 4 numa generalidade bem maior (também para equações diferenciais não-lineares). Fórmulas como aquelas dos exemplos na Seção 1.5 nos permitem, em princípio, resolver explicitamente qualquer sistema linear de EDOs de qualquer ordem (transforma-se o sistema primeiramente num sistema da primeira ordem, depois determina-se a decomposição autovalor da matriz A , ou seja, a forma normal de Jordan, e finalmente calcula-se a matriz exponencial da forma normal de Jordan). No caso de equações diferenciais não-lineares, apenas em casos raros é possível encontrar soluções explícitas. Geralmente pode-se no máximo esperar dizer uma coisa sobre o *comportamento qualitativo* das soluções. Isto leva-nos à teoria qualitativa dos **sistemas dinâmicos**, uma área bem particular e empolgante da pesquisa na matemática.

Um assunto bem diferente, que falamos brevemente, são os espaços de Banach e os operadores lineares. Estes conceitos são apresentados no início da **análise funcional**, uma área importante da matemática, tratado em cursos da pós-graduação, a qual desempenha um papel importante e central em muitas partes tanto na matemática (EDPs até topologia e geometria) como na física (por exemplo na mecânica quântica).

Finalmente temos recebido aqui uma curta olhadela nas álgebras de Banach, localizadas na intersecção entre análise, topologia, e álgebra, mas também formando uma área de pesquisa própria em matemática.

Capítulo 2

Aplicações Diferenciáveis

Este capítulo começa na Seção 2.1 com uma introdução ao conceito da diferenciabilidade para funções de uma ou várias variáveis, ou mais geral para aplicações entre espaços de Banach. A folha da Seção 2.3 são as regras básicas para a derivada, como a regra de Leibniz e a regra da cadeia nas dimensões superiores. A Seção 2.4 trata o Teorema da Barreira (um tipo de teorema do valor médio generalizado) e umas das suas consequências. Seguidamente consideramos derivadas superiores (Seção 2.5), discutimos séries de Taylor (Seção 2.6), e investigamos extremos locais (Seção 2.7).

2.1 Conceitos básicos

Ao início destas considerações introdutórias lembramos que uma função $f : \mathbb{R} \rightarrow \mathbb{R}$ num ponto $a \in \mathbb{R}$ é chamado de diferenciável se existe o limite

$$A := \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}. \quad (2.1.1)$$

Em redação matemática concisa isso é a afirmação

$$\exists A \in \mathbb{R} \forall \varepsilon > 0 \exists \delta > 0 \forall h \in \mathbb{R} \left(0 < |h| < \delta \implies \left| \frac{f(a+h) - f(a)}{h} - A \right| < \varepsilon \right). \quad (2.1.2)$$

Em outras palavras, existe um número real A , tal que para todo número real $\varepsilon > 0$ arbitrariamente pequeno, existe um número real $\delta > 0$, tal que todo número real h com $0 < |h| < \delta$ satisfaz a desigualdade $\left| A - \frac{f(a+h) - f(a)}{h} \right| < \varepsilon$.

A questão que queremos tratar nesta seção é como pode-se expandir este conceito de diferenciabilidade do curso Análise I para funções as quais são definidos num subconjunto aberto de um espaço de Banach. Existem, efetivamente, três tais generalizações, as quais vamos discutir em seguida.

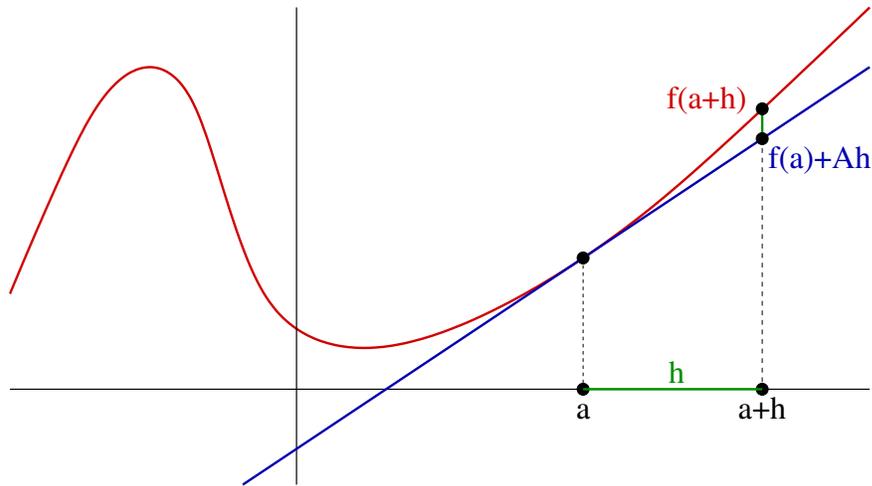


Figura 2.1: A derivada como aproximação linear.

Antes de tudo cumpre observar que não pode-se dividir por vetores, por isso o lado direito em (2.1.1) não generaliza para funções definido em espaços vetoriais. Mas podemos transformar a desigualdade em (2.1.2) numa forma qual faz sentido para elementos de um espaço vetorial normado:

$$\begin{aligned} \left| \frac{f(a+h) - f(a)}{h} - A \right| &= \left| \frac{f(a+h) - f(a) - Ah}{h} \right| \\ &= \frac{|f(a+h) - f(a) - Ah|}{|h|}. \end{aligned} \quad (2.1.3)$$

Ora a condição (2.1.2) para diferenciabilidade diz que este termo torna arbitrariamente pequeno para $|h|$ suficientemente pequeno. Isso tem o significado geométrico que a função $h \mapsto f(a) + Ah$ pode ser interpretada como a tangente ao gráfico de f no ponto a e, além disso, representa uma boa primeira aproximação da função $h \mapsto f(a+h)$ no sentido que a diferença destas duas funções para valores pequenos de h é pequeno em relação a $|h|$ (veja a Figura 2.1).

O conceito de diferenciabilidade

Procurando uma aproximação análoga para funções em conjuntos abertos de um espaço vetorial normado, então o número $A = f'(a)$ no caso de dimensão um deve ser substituído por uma aplicação linear. Para espaços de Banach nós lembramos a notação $\mathcal{L}(X, Y)$ para o conjunto de todas as transformações lineares limitadas de X para Y (Definição 1.2.1).

Definição 2.1.1 (Diferenciabilidade).

Sejam $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ espaços de Banach, seja $U \subset X$ um conjunto aberto, seja $f : U \rightarrow Y$ uma aplicação, e seja $a \in U$. Chama-se a aplicação f **diferenciável no ponto a** , se existe uma aplicação linear limitada $A \in \mathcal{L}(X, Y)$, tal que para todo número real $\varepsilon > 0$, existe um número real $\delta > 0$, tal que para todos os $h \in X$ vale a seguinte afirmação:

$$0 < \|h\|_X < \delta \implies \begin{array}{l} a + h \in U \text{ e} \\ \|f(a+h) - f(a) - Ah\|_Y < \varepsilon \|h\|_X. \end{array} \quad (2.1.4)$$

Escrito como formula lógica isso significa que

$$\forall \varepsilon > 0 \exists \delta > 0 \forall h \in X \left(0 < \|h\|_X < \delta \implies a + h \in U \text{ e } \frac{\|f(a+h) - f(a) - Ah\|_Y}{\|h\|_X} < \varepsilon \right). \quad (2.1.5)$$

A esta definição seja anotado que a definição conhecida do curso Análise I da diferenciabilidade de funções de uma variável real manifesta-se como o caso especial $X = Y = \mathbb{R}$, porque neste caso uma aplicação linear sempre é dado por multiplicação com um número real A .

Segundo, lembramos que uma aplicação linear $A : X \rightarrow Y$ entre espaços vetoriais normados de dimensão finita é, pelo Corolário 1.2.2, automaticamente limitada, o que significa que sua norma operador

$$\|A\| = \sup_{0 \neq x \in X} \frac{\|Ax\|_Y}{\|x\|_X}$$

é finita. Terceiro, mais uma vez seja enfatizado que na definição de diferenciabilidade é exigido a *existência* de uma aplicação linear $A : X \rightarrow Y$ satisfazendo a condição (2.1.5). Isso levanta a questão se existe *só um* tal A . Lema 2.1.3 dá uma resposta positiva a esta questão.

Exercício 2.1.2. Sejam X e Y espaços vetoriais normados de dimensão finita, seja $U \subset X$ um conjunto aberto, seja $f : U \rightarrow Y$ uma aplicação, e seja $a \in U$. Então a afirmação “ f é diferenciável no ponto a ” não depende da escolha das normas em X e Y . **Dica:** Teorema 1.1.8.

Lema 2.1.3. Sejam X, Y, U, a, f como na Definição 2.1.1. Então existe no máximo uma aplicação linear $A \in \mathcal{L}(X, Y)$ a qual satisfaz (2.1.5).

Demonstração. Sejam $A, B : X \rightarrow Y$ duas aplicações lineares as quais satisfaçam (2.1.5), seja $\varepsilon > 0$ dado. Então existe pela hipótese um $\delta > 0$ com

$$B_\delta(a) = \{x \in X \mid \|x - a\|_X < \delta\} \subset U$$

tal que todo vetor $h \in X$ com $0 < \|h\|_X < \delta$ satisfaz as desigualdades

$$\begin{aligned} \|f(a+h) - f(a) - Ah\|_Y &< \varepsilon \|h\|_X, \\ \|f(a+h) - f(a) - Bh\|_Y &< \varepsilon \|h\|_X. \end{aligned}$$

Para todo $h \in X$ com $0 < \|h\|_X < \delta$ resulta daí que

$$\begin{aligned} \|Ah - Bh\|_Y &= \|Ah + f(a) - f(a+h) + f(a+h) - f(a) - Bh\|_Y \\ &\leq \|f(a+h) - f(a) - Ah\|_Y + \|f(a+h) - f(a) - Bh\|_Y \\ &< 2\varepsilon \|h\|_X. \end{aligned}$$

Daí recebemos a desigualdade $\|h\|_X^{-1} \|(A - B)h\|_Y < 2\varepsilon$ para todos os vetores $h \in X$ com $0 < \|h\|_X < \delta$ e daí através de re-escalar ainda para todos os $h \in X \setminus \{0\}$. Consideramos agora o supremo sobre todos os $h \in X \setminus \{0\}$, então recebemos

$$\|A - B\| = \sup_{0 \neq h \in X} \frac{\|(A - B)h\|_Y}{\|h\|_X} \leq 2\varepsilon$$

para cada um $\varepsilon > 0$. Disso segue que $\|A - B\| = 0$ e daí $A = B$. \square

Definição 2.1.4 (Derivada e diferencial). Sejam X, Y espaços de Banach, $U \subset X$ um conjunto aberto e $f : U \rightarrow Y$ uma aplicação, diferenciável no ponto $a \in U$. Chama-se a aplicação linear única $A : X \rightarrow Y$, a qual satisfaz (2.1.5), a **derivada de f no ponto a** , denotada $df(a) := A : X \rightarrow Y$.¹ Se f é diferenciável em todo ponto de U , então chama-se a aplicação

$$df : U \rightarrow \mathcal{L}(X, Y), \quad a \mapsto df(a)$$

do diferencial de f .

¹ Outras símbolos comuns na literatura para a derivada são $Df(a)$, $D_a f$, $d_a f$, $T_a f$, $f'(a)$. Neste texto usamos geralmente a notação $df(a)$.

Lema 2.1.5 (Aplicações diferenciáveis são contínuas).

Sejam X, Y espaços de Banach, seja $U \subset X$ um conjunto aberto e $f : U \rightarrow Y$ uma aplicação, diferenciável no ponto $a \in U$. Então f é contínuo no ponto a .

Demonstração. Seja $A := df(a) \in \mathcal{L}(X, Y)$ e seja $\varepsilon > 0$ dado. Escolha $\delta > 0$, tal que vale (2.1.4) e defina

$$\rho := \min \left\{ \delta, \frac{\varepsilon}{2\|A\|}, \frac{1}{2} \right\} > 0.$$

Agora seja $h \in X$ com $\|h\|_X < \rho \leq \delta$. Através de (2.1.4) a soma $a + h \in U$ e satisfaz a desigualdade

$$\begin{aligned} \|f(a+h) - f(a)\|_Y &\leq \|f(a+h) - f(a) - Ah\|_Y + \|Ah\|_Y \\ &\leq \varepsilon \|h\|_X + \|A\| \|h\|_X \\ &< \varepsilon \rho + \|A\| \rho \\ &\leq \varepsilon. \end{aligned}$$

Com isso o Lema 2.1.5 fica provado. \square

Exemplo 2.1.6. Sejam X, Y espaços de Banach e seja $U \subset X$ um conjunto aberto, então toda aplicação constante $f : U \rightarrow Y$ é diferenciável em todo ponto $a \in U$ e a derivada é $df(a) = 0$. O reverso só vale para conjuntos abertos conexos (Apêndice B) e será provado na Seção 2.4.

Exemplo 2.1.7. Sejam dados uma matriz $A \in \mathbb{R}^{m \times n}$ e um vetor $b \in \mathbb{R}^m$. Eles determinam uma aplicação afim $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ através de

$$f(x) := Ax + b \quad \text{para } x \in \mathbb{R}^n.$$

Neste caso vale $f(a+h) - f(a) - Ah = 0$ para todos os $a, h \in \mathbb{R}^n$. Desta forma f é diferenciável em todo ponto $a \in \mathbb{R}^n$ e a derivada é

$$df(a) = A.$$

Observe que aqui identificamos a matriz $A \in \mathbb{R}^{m \times n}$ com a aplicação linear $\mathbb{R}^n \rightarrow \mathbb{R}^m : \xi \mapsto A\xi$ (multiplicação de matriz e vetor). Uma tal identificação é possível sempre quando nossos espaços vetoriais X e Y admitem bases canônicas. Mas vamos também encontrar exemplos nas quais isso não é o caso. O princípio básico é que “a derivada é uma aplicação linear” (a qual pode ser identificado frequentemente, mas não sempre, com uma matriz).

Vale a mesma afirmação para $f(x) := Ax + b$, mesmo argumento, no caso de um operador linear $A \in \mathcal{L}(X, Y)$ entre espaços de Banach e um vetor $b \in Y$.

Exemplo 2.1.8. Seja $S = S^T \in \mathbb{R}^{n \times n}$ uma matriz simétrica e define a aplicação $f : \mathbb{R}^n \rightarrow \mathbb{R}$ através de

$$f(x) := \frac{1}{2}x^T Sx = \frac{1}{2}\langle x, Sx \rangle \quad \text{para } x \in \mathbb{R}^n.$$

Aqui x^T é o vetor linha obtido de x através de transposição e

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i$$

para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ e $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ é o produto interno padrão no \mathbb{R}^n . Para $a, h \in \mathbb{R}^n$ vale

$$\begin{aligned} f(a+h) - f(a) &= \frac{1}{2}(a+h)^T S(a+h) - \frac{1}{2}a^T Sa \\ &= \frac{1}{2}a^T Sh + \frac{1}{2}h^T Sa + \frac{1}{2}h^T Sh \\ &= a^T Sh + \frac{1}{2}h^T Sh \end{aligned}$$

e daí, aplicando a desigualdade de Cauchy–Schwarz, segue que

$$\begin{aligned} |f(a+h) - f(a) - a^T Sh| &= \frac{1}{2} |h^T Sh| \\ &= \frac{1}{2} |\langle h, Sh \rangle| \\ &\leq \frac{1}{2} \|h\| \|Sh\| \\ &\leq \frac{\|S\| \|h\|}{2} \|h\|. \end{aligned}$$

Afinal dado $\varepsilon > 0$, então

$$\delta := \frac{2\varepsilon}{\|S\|} > 0$$

satisfaz a condição (2.1.4) com

$$A := a^T S \in \mathbb{R}^{1 \times n}$$

(visto como aplicação linear de \mathbb{R}^n para \mathbb{R}). Por isso a aplicação f é diferenciável com derivada

$$df(a) = a^T S$$

em todos os pontos $a \in \mathbb{R}^n$.

Derivadas direcionais

Ora nós preocupamos com as duas possibilidades adicionais, já anunciadas no início, de expandir a definição da diferenciabilidade conhecido do curso Análise I para funções de várias variáveis. Trata-se da derivada direcional e das derivadas parciais.

Definição 2.1.9 (Derivada direcional). *Sejam X, Y espaços de Banach, seja $U \subset X$ um conjunto aberto, e sejam dados uma aplicação $f : U \rightarrow Y$ e elementos $a \in U$ e $\xi \in X$. A derivada direcional de f no ponto a na direção ξ é o limite*

$$\partial_{\xi} f(a) := \lim_{t \rightarrow 0} \frac{f(a + t\xi) - f(a)}{t} \in Y \quad (2.1.6)$$

caso exista.

Ora consideramos o caso especial de $X = \mathbb{R}^n$ e $Y = \mathbb{R}^m$ e denotamos a base canônica de \mathbb{R}^n com $e_1, \dots, e_n \in \mathbb{R}^n$, ou seja, o i -ésimo vetor

$$e_i := (0, \dots, 0, 1, 0, \dots, 0), \quad i = 1, \dots, n$$

é uma lista de n números dos quais o i -ésimo é 1 e os outros 0.

Definição 2.1.10 (Derivada parcial). *Seja $U \subset \mathbb{R}^n$ um conjunto aberto e sejam dado uma aplicação $f : U \rightarrow \mathbb{R}^m$ e um elemento $a \in U$. Diz-se que a aplicação f é **parcialmente diferenciável no ponto a** , se as derivadas direcionais*

$$\begin{aligned} \partial_i f(a) &:= \frac{\partial f}{\partial x_i}(a) \\ &:= \partial_{e_i} f(a) \\ &= \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t} \\ &= \left. \frac{d}{dt} \right|_{t=0} f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) \end{aligned} \quad (2.1.7)$$

existem para $i = 1, \dots, n$. Neste caso chama-se os vetores $\partial_i f(a) \in \mathbb{R}^m$ as derivadas parciais de f no ponto a .

Lema 2.1.11. *Sejam X, Y espaços de Banach, seja $U \subset X$ um conjunto aberto, e seja $f : U \rightarrow Y$ diferenciável no ponto $a \in U$. Então existe a derivada direcional $\partial_\xi f(a)$ para todo vetor $\xi \in X$ e é dada através de*

$$\partial_\xi f(a) = df(a)\xi. \quad (2.1.8)$$

Demonstração. Seja $A := df(a) : X \rightarrow Y$ e seja $\xi \in X$. Para $\xi = 0$ decorre a existência da derivada direcional e também a identidade $\partial_\xi f(a) = 0 = A\xi$ direto da definição. Suponhamos então $\xi \neq 0$. Basta mostrar a afirmação

$$\lim_{t \rightarrow 0} \frac{f(a + t\xi) - f(a)}{t} = A\xi. \quad (2.1.9)$$

Então seja dado $\varepsilon > 0$ e seja $\tilde{\varepsilon} := \varepsilon / \|\xi\|_X$. Portanto existe pela definição 2.1.1 um número $\tilde{\delta} > 0$, tal que todo vetor $h \in X$ com $0 < \|h\|_X < \tilde{\delta}$ satisfaz a condição $a + h \in U$ e a desigualdade

$$\|f(a + h) - f(a) - Ah\|_Y < \tilde{\varepsilon} \|h\|_X.$$

Ora, seja $\delta := \tilde{\delta} / \|\xi\|_X$ e seja $t \in \mathbb{R}$ com $0 < |t| < \delta$. Então vale

$$0 < \|t\xi\|_X = |t| \|\xi\|_X < \delta \|\xi\|_X = \tilde{\delta},$$

e daí $a + t\xi \in U$ e

$$\|f(a + t\xi) - f(a) - tA\xi\|_Y < \tilde{\varepsilon} \|t\xi\|_X = |t| \varepsilon$$

e através disso segue que

$$\left\| \frac{f(a + t\xi) - f(a)}{t} - A\xi \right\|_Y < \varepsilon.$$

Com isso ficam provados (2.1.9) e Lema 2.1.11; veja também o Exemplo 2.3.9. \square

Lema 2.1.11 mostra que diferenciabilidade implica a existência das derivadas direcionais. Obviamente a existência das (todas as) derivadas direcionais implica diferenciabilidade parcial. Igualmente sabemos que diferenciabilidade implica continuidade (Lema 2.1.5). Os seguintes exemplos mostram que nenhuma destas implicações pode ser invertido. Também já sabemos do curso Análise I que continuidade ainda não implica diferenciabilidade parcial. (A função contínua $\mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \|x\|$ não é diferenciável no ponto $a = 0$.)

Exemplo 2.1.12. Este exemplo mostra que a existência das derivadas direcionais não implica diferenciabilidade. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função

$$f(x, y) := \frac{x^2 y}{x^2 + y^2} \quad \text{para } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$$

com $f(0, 0) := 0$. Esta função satisfaz a condição

$$f(t\xi, t\eta) = tf(\xi, \eta)$$

para todo $\zeta = (\xi, \eta) \in \mathbb{R}^2$ e todo $t \in \mathbb{R}$. Daí existem todas as derivadas direcionais de f no ponto $a = 0 \in \mathbb{R}^2$ e vale $\partial_\zeta f(0) = f(\zeta)$ para $\zeta \in \mathbb{R}^2$. Como a aplicação $\mathbb{R}^2 \rightarrow \mathbb{R} : \zeta \mapsto \partial_\zeta f(0) = f(\zeta)$ não é linear, pelo Lema 2.1.11 a função f não pode ser diferenciável no ponto $a = 0$.

Exemplo 2.1.13. Este exemplo mostra que a existência das derivadas direcionais ainda não implica continuidade. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função

$$f(x, y) := \frac{x^2 + y^2}{y} \quad \text{se } y \neq 0$$

com $f(x, 0) := 0$. Esta função satisfaz a condição $f(t\zeta) = tf(\zeta)$, para todos os $\zeta \in \mathbb{R}^2$ e $t \in \mathbb{R}$. Consequentemente existem todas as derivadas direcionais de f no ponto $a = 0$. Porém vale $f(\varepsilon, \varepsilon^3) = \varepsilon^{-1} + \varepsilon^3 \geq \varepsilon^{-1}$ para todo $\varepsilon > 0$, e desta forma f é descontínua no ponto $a = 0$. Então, conforme Lema 2.1.5, a função f não é diferenciável no ponto $a = 0$.

Exemplo 2.1.14. Este exemplo mostra que a diferenciabilidade parcial não implica a existência das derivadas direcionais. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função

$$f(x, y) := \frac{xy}{\sqrt{x^2 + y^2}} \quad \text{para } (x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$$

com $f(0, 0) := 0$. Esta função satisfaz a condição

$$f(t\xi, t\eta) = |t|f(\xi, \eta)$$

para todos os $\zeta = (\xi, \eta) \in \mathbb{R}^2$ e $t \in \mathbb{R}$. Por isso existe a derivada direcional de f no ponto $a = 0 \in \mathbb{R}^2$ na direção de $\zeta = (\xi, \eta) \in \mathbb{R}^2$ se e somente se $f(\xi, \eta) = 0$. Mas isso vale para os membros $e_1 = (1, 0)$ e $e_2 = (0, 1)$ da base canônica. Consequentemente f é parcialmente diferenciável no ponto $a = 0$, mas não existem todas as derivadas direcionais de f no ponto $a = 0$.

Exercício: A função f é contínua.

A matriz jacobiana

Agora suponhamos que $U \subset \mathbb{R}^n$ é um conjunto aberto e $f : U \rightarrow \mathbb{R}^m$ uma aplicação, diferenciável no ponto $x_0 \in U$. Então $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma aplicação linear. Como sabemos da Álgebra Linear uma tal aplicação é dada por multiplicação com uma matriz $A \in \mathbb{R}^{m \times n}$. Isso levanta a questão qual matriz representa a derivada $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ da nossa função e como podemos calcular esta matriz. Para responder esta questão precisamos o seguinte Lema para $X = \mathbb{R}^n$ e $Y_1, \dots, Y_m = \mathbb{R}$.

Lema 2.1.15. *Sejam X, Y_1, \dots, Y_m espaços de Banach, seja $U \subset X$ um conjunto aberto, seja*

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : U \rightarrow Y_1 \times \dots \times Y_m =: Y \quad (2.1.10)$$

uma aplicação, e seja $x_0 \in U$ e $\xi \in X$. Então vale o seguinte:

(i) *A derivada direcional $\partial_\xi f(x_0)$ existe se e somente se as derivadas direcionais $\partial_\xi f_j(x_0)$ existem para $j = 1, \dots, m$. Neste caso vale*

$$\partial_\xi f(x_0) = \begin{pmatrix} \partial_\xi f_1(x_0) \\ \vdots \\ \partial_\xi f_m(x_0) \end{pmatrix} \in Y_1 \times \dots \times Y_m. \quad (2.1.11)$$

(ii) *A aplicação $f : U \rightarrow Y$ é diferenciável no ponto x_0 se e somente se as funções $f_1, \dots, f_m : U \rightarrow \mathbb{R}$ são diferenciáveis no ponto x_0 . Neste caso vale*

$$df(x_0) = \begin{pmatrix} df_1(x_0) \\ \vdots \\ df_m(x_0) \end{pmatrix} : X \rightarrow Y_1 \times \dots \times Y_m. \quad (2.1.12)$$

Demonstração. Seja $|\cdot|_j$ a norma em Y_j e aquela em $Y \ni y = (y_1, \dots, y_m)$ é definido assim $\|y\|_Y := \sqrt{\sum_{j=1}^m |y_j|_j^2}$.

Provaremos a parte (i). Para este efeito suponhamos por ora que a derivada direcional de f no ponto x_0 na direção ξ exista e vamos denotar esta derivada direcional assim

$$a := \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} := \partial_\xi f(x_0) \in Y_1 \times \dots \times Y_m.$$

Então vale

$$\left| \frac{f_j(x_0 + t\xi) - f_j(x_0)}{t} - a_j \right| \leq \left\| \frac{f(x_0 + t\xi) - f(x_0)}{t} - a \right\|_Y$$

para todos os $t \in \mathbb{R} \setminus \{0\}$ e todos os $j \in \{1, \dots, m\}$. Como o lado direito converge para zero pela hipótese concluímos que

$$\lim_{t \rightarrow 0} \left| \frac{f_j(x_0 + t\xi) - f_j(x_0)}{t} - a_j \right| = 0 \quad \text{para } j = 1, \dots, m.$$

Consequentemente existe para todo j a derivada direcional de f_j no ponto x_0 e na direção de ξ e vale

$$\partial_\xi f_j(x_0) = \lim_{t \rightarrow 0} \frac{f_j(x_0 + t\xi) - f_j(x_0)}{t} = a_j.$$

Suponhamos, vice versa, que a derivada direcional $a_j := \partial_\xi f_j(x_0)$ existe para todo $j = 1, \dots, m$ e definimos $a := (a_1, \dots, a_m) \in Y$, então vale

$$\left\| \frac{f(x_0 + t\xi) - f(x_0)}{t} - a \right\|_Y = \sqrt{\sum_{i=1}^m \left| \frac{f_i(x_0 + t\xi) - f_i(x_0)}{t} - a_i \right|^2}.$$

O lado direito converge pela hipótese para zero quando $t \rightarrow 0$, e daí segue a existência da derivada direcional $\partial_\xi f(x_0) = a$. Isso prova a parte (i).

Provaremos a parte (ii) com o mesmo argumento. Se f é diferenciável no ponto x_0 com derivada $A := df(x_0) : X \rightarrow Y$, e denotamos de $A_j : X \rightarrow \mathbb{R}$ a composição de A com a projeção de Y para a j -ésima coordenada, então segue a diferenciabilidade de f_j no ponto x_0 com derivada $df_j(x_0) = A_j : X \rightarrow \mathbb{R}$ através da desigualdade

$$\frac{|f_j(x_0 + h) - f_j(x_0) - A_j h|_j}{\|h\|_X} \leq \frac{\|f(x_0 + h) - f(x_0) - Ah\|_Y}{\|h\|_X}.$$

Suponhamos, vice versa, que f_j é diferenciável no ponto x_0 com $A_j := df_j(x_0)$, e definimos a aplicação linear $A : X \rightarrow Y$ como $A\xi := (A_1\xi, \dots, A_m\xi)$ para $\xi \in X$, então segue a diferenciabilidade de f no ponto x_0 através de

$$\frac{\|f(x_0 + h) - f(x_0) - Ah\|_Y^2}{\|h\|_X^2} = \sum_{j=1}^m \left(\frac{|f_j(x_0 + h) - f_j(x_0) - A_j h|_j}{\|h\|_X} \right)^2.$$

Com isso o Lema 2.1.15 fica provado. \square

Lema 2.1.16 (A matriz jacobiana).

Seja $U \subset \mathbb{R}^n$ um conjunto aberto e seja $f : U \rightarrow \mathbb{R}^m$ uma aplicação com as coordenadas f_1, \dots, f_m as quais são diferenciável no ponto $x_0 \in U$. Então a derivada $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é dada por multiplicação com a matriz

$$Df(x_0) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \cdots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} \in \mathbb{R}^{m \times n} \quad (2.1.13)$$

das derivadas parciais.

Demonstração. Seja e_1, \dots, e_n a base canônica do \mathbb{R}^n , então para todo vetor $\xi = \sum_{i=1}^n \xi_i e_i \in \mathbb{R}^n$ temos

$$df(x_0)\xi = \sum_{i=1}^n \xi_i df(x_0)e_i = \sum_{i=1}^n \xi_i \partial_{e_i} f(x_0) = \sum_{i=1}^n \xi_i \frac{\partial f}{\partial x_i}(x_0) = Df(x_0)\xi.$$

Aqui segue a primeira identidade da linearidade de $df(x_0)$, a segunda do Lema 2.1.11, a terceira da definição das derivadas parciais, e a quarta da parte (i) do Lema 2.1.15 assim como da definição da matriz $Df(x_0)$ em (2.1.13). Com isso o Lema 2.1.16 fica provado. \square

Definição 2.1.17. Na situação do Lema 2.1.16 a matriz $Df(x_0) \in \mathbb{R}^{m \times n}$ das derivadas parciais definido pela equação (2.1.13) chama-se a **matriz jacobiana de f no ponto $x_0 \in U$** .

Notamos que o símbolo $Df(x_0)$ da matriz jacobiana foi utilizado com a finalidade de clareza, para a distinção formal entre a derivada como aplicação linear $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e a matriz jacobiana das derivadas parciais, ou seja $Df(x_0) \in \mathbb{R}^{m \times n}$. Como o Lema 2.1.16 fica provado agora, esta distinção torna supérfluo e no seguinte vamos voltar ao nosso símbolo padrão $df(x_0)$ para a derivada, seja como matriz jacobiana ou como aplicação linear. Com efeito, geralmente não será necessário distinguir na notação entre uma aplicação linear de \mathbb{R}^n para \mathbb{R}^m e sua matriz representante.

2.2 Diferenciabilidade contínua

Na Seção 2.1 temos visto que toda função diferenciável é parcialmente diferenciável num subconjunto aberto do \mathbb{R}^n , enquanto o reverso não vale. Mas se supormos que as derivadas parciais existam em todo ponto de um aberto e além disso sejam contínuas, o seguinte teorema mostra que então a nossa função também é diferenciável.

Teorema 2.2.1. *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja dado $x_0 \in U$. Além disso, seja $f : U \rightarrow \mathbb{R}^m$ uma aplicação, parcialmente diferenciável em todo ponto de U , e cujas derivadas parciais $\partial_i f : U \rightarrow \mathbb{R}^m$ para $i = 1, \dots, n$ são contínuas no ponto x_0 . Então f é diferenciável no ponto x_0 .*

Definição 2.2.2. (i) *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto. Uma aplicação $f : U \rightarrow \mathbb{R}^m$ chama-se de **continuamente diferenciável** se é parcialmente diferenciável e suas derivadas parciais $\partial_1 f, \dots, \partial_n f : U \rightarrow \mathbb{R}^m$ são contínuas.*

(ii) *Sejam X, Y espaços de Banach e seja $U \subset X$ um subconjunto aberto. Uma aplicação $f : U \rightarrow Y$ chama-se de **continuamente diferenciável** se é diferenciável e seu diferencial $df : U \rightarrow \mathcal{L}(X, Y)$ é uma aplicação contínua.*

Exercício 2.2.3. No caso de $X = \mathbb{R}^n$ e $Y = \mathbb{R}^m$ as duas definições na Definição 2.2.2 são equivalente uma à outra.

Demonstração do Teorema 2.2.1. Segundo o Lema 2.1.15, é suficiente provar o Teorema para $m = 1$. Ora, seja $f : U \rightarrow \mathbb{R}$ uma função parcialmente diferenciável cujas derivadas são todas contínuas no ponto x_0 . Resta mostrar que f é diferenciável no ponto x_0 . Para isso utilizamos a norma euclidiana

$$\|x\| := \sqrt{x_1^2 + \dots + x_n^2} \quad \text{para } x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

e a base canônica e_1, \dots, e_n do \mathbb{R}^n .

Seja $\varepsilon > 0$ dado. Como as funções $\partial_i f : U \rightarrow \mathbb{R}$ são contínuas no ponto x_0 , existe uma constante $\delta > 0$, tal que para todos os $h \in \mathbb{R}^n$ vale o seguinte:

$$\|h\| < \delta \quad \implies \quad \begin{aligned} &x_0 + h \in U \text{ e, para } i = 1, \dots, n, \\ &|\partial_i f(x_0 + h) - \partial_i f(x_0)| < \varepsilon/n. \end{aligned} \quad (2.2.1)$$

Seja dado um vetor $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ com $0 < \|h\| < \delta$. Então definimos os vetores $x_1, \dots, x_n \in \mathbb{R}^n$ através de

$$x_k := x_0 + h_1 e_1 + \dots + h_k e_k \quad \text{para } k = 1, \dots, n.$$

Note-se que aqui os x_i são todos vetores no \mathbb{R}^n , e *não* as coordenadas de um vetor. Satisfazem as equações

$$x_k = x_{k-1} + h_k e_k \quad \text{para } k = 2, \dots, n, \quad x_n = x_0 + h. \quad (2.2.2)$$

Como $\|x_k - x_0\| \leq \|h\| < \delta$, vale $x_k \in U$ segundo (2.2.1), daí segundo (2.2.2)

$$f(x_0 + h) - f(x_0) = \sum_{k=1}^n (f(x_k) - f(x_{k-1})). \quad (2.2.3)$$

Do mesmo modo vale $\|x_{k-1} + th_k e_k - x_0\| < \delta$ para todos os $t \in [0, 1]$ e todos os $k \in \{1, \dots, n\}$. Segundo (2.2.1), segue dali $x_{k-1} + th_k e_k \in U$ e

$$|\partial_k f(x_{k-1} + th_k e_k) - \partial_k f(x_0)| < \frac{\varepsilon}{n} \quad (2.2.4)$$

para todo $0 \leq t \leq 1$ e todo $k = 1, \dots, n$. Para cada um tal k definimos uma função $\phi_k : [0, 1] \rightarrow \mathbb{R}$ através de

$$\phi_k(t) := f(x_{k-1} + th_k e_k) \quad \text{para } 0 \leq t \leq 1.$$

Dado $k \in \{1, \dots, n\}$, existe segundo o Teorema do Valor Médio um $t_k \in (0, 1)$ com $\phi'_k(t_k) = \phi_k(1) - \phi_k(0)$. Usando $\phi_k(0) = f(x_{k-1})$ e $\phi_k(1) = f(x_k)$ vale

$$h_k \partial_k f(x_{k-1} + t_k h_k e_k) = \phi'_k(t_k) = f(x_k) - f(x_{k-1}). \quad (2.2.5)$$

Usando (2.2.3), (2.2.4), e (2.2.5) encontraremos que

$$\begin{aligned} & \left| f(x_0 + h) - f(x_0) - \sum_{k=1}^n h_k \partial_k f(x_0) \right| \\ &= \left| \sum_{k=1}^n (f(x_{k+1}) - f(x_k) - h_k \partial_k f(x_0)) \right| \\ &= \left| \sum_{k=1}^n h_k (\partial_k f(x_{k-1} + t_k h_k e_k) - \partial_k f(x_0)) \right| \\ &\leq \sum_{k=1}^n |h_k| |\partial_k f(x_{k-1} + t_k h_k e_k) - \partial_k f(x_0)| \\ &< \sum_{k=1}^n \frac{\varepsilon |h_k|}{n} \leq \varepsilon \|h\|. \end{aligned}$$

Com isto segue a afirmação do Teorema 2.2.1 através da definição de diferenciabilidade em Definição 2.1.1. \square

Para muitos exemplos de funções de várias variáveis podemos reduzir a diferenciabilidade, com a ajuda do Teorema 2.2.1, à diferenciabilidade de funções de uma variável, conhecido do curso Análise I.

Exemplo 2.2.4. A função $f : U := (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ definido por

$$f(x, y) := x^y = e^{y \log(x)}$$

é continuamente diferenciável, porque segundo resultados da Análise I é parcialmente diferenciável e as suas derivadas parciais são dadas por

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= e^{y \log(x)} \frac{y}{x} = yx^{y-1}, \\ \frac{\partial f}{\partial y}(x, y) &= e^{y \log(x)} \log(x) = \log(x)x^y \end{aligned}$$

para $x > 0$ e $y \in \mathbb{R}$. Aquelas funções $\frac{\partial f}{\partial x} : U \rightarrow \mathbb{R}$ e $\frac{\partial f}{\partial y} : U \rightarrow \mathbb{R}$ são contínuas, de novo segundo exemplos da Análise I e um resultado da Análise I qual diz que a soma, o produto, e o quociente de duas funções (em espaços métricos arbitrários) também são contínuas. Daqui segue, segundo o Teorema 2.2.1, que nossa função $f : U \rightarrow \mathbb{R}$ é continuamente diferenciável. Segundo o Lema 2.1.5, sua derivada $df(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ (como aplicação linear) é representada pela matriz jacobiana $df(x, y) = (yx^{y-1}, \log(x)x^y) \in \mathbb{R}^{1 \times 2}$ denotada com o mesmo símbolo $df(x, y)$.

Exemplo 2.2.5. Do mesmo modo ve-se que a aplicação $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definida por $f(x) := (x_1 + x_2 + x_3, x_1x_2 + x_3x_1 + x_2x_3, x_1x_2x_3)$ para $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ é continuamente diferenciável. Sua matriz jacobiana no ponto $x \in \mathbb{R}^3$ é

$$df(x) = \begin{pmatrix} 1 & 1 & 1 \\ x_2 + x_3 & x_3 + x_1 & x_1 + x_2 \\ x_2x_3 & x_3x_1 & x_1x_2 \end{pmatrix}.$$

Isso é *um* exemplo só. Segundo o Teorema 2.2.1 e resultados da Análise I, toda aplicação polinomial $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é continuamente diferenciável.

O seguinte teorema refere-se a sequências de funções continuamente diferenciáveis e diz que se convergem uniformemente, junto com as suas derivadas parciais, então o limite também é continuamente diferenciável. Aplicações deste teorema importante encontram-se nas Seções 2.3, 2.5 e 2.6.

Teorema 2.2.6 (Convergência em C^1 ao longo compactos (em C^1_{loc})). *Sejam X, Y espaços de Banach² e seja $U \subset X$ aberto. Seja $f_k : U \rightarrow Y$ uma sequência de aplicações continuamente diferenciáveis, tal que as sequências de aplicações contínuas $(f_k)_{k \in \mathbb{N}}$ e $(df_k)_{k \in \mathbb{N}}$ convergem uniformemente em qualquer compacto $K \subset U$. Considere pontualmente, para $x \in U$, os limites*

$$f(x) := \lim_{k \rightarrow \infty} f_k(x), \quad G(x) := \lim_{k \rightarrow \infty} df_k(x).$$

Então o limite $f : U \rightarrow Y$ é continuamente diferenciável e $df = G$.

Demonstração. Seja $\|\cdot\|_{\mathcal{L}}$ a norma operador em $\mathcal{L}(X, Y)$. Pela hipótese vale

$$\lim_{k \rightarrow \infty} \sup_{x \in K} \|f(x) - f_k(x)\| = 0, \quad \lim_{k \rightarrow \infty} \sup_{x \in K} \|G(x) - df_k(x)\|_{\mathcal{L}} = 0 \quad (2.2.6)$$

ao longo de qualquer subconjunto compacto $K \subset U$. Como o limite de uma sequência uniformemente convergente de aplicações contínuas é contínuo (Análise I), as aplicações $f : U \rightarrow Y$ e $G : U \rightarrow \mathcal{L}(X, Y)$ são contínuas. Resta mostrar que a derivada $df(x)$ existe para todo $x \in U$ e é igual a $G(x)$. Para isso escolhamos $x \in U$ e $\varepsilon > 0$ com $\overline{B_\varepsilon}(x) \subset U$. Seja $e \in X$ um vetor unitário. Pela hipótese a função $[-\varepsilon, \varepsilon] \rightarrow Y : t \mapsto f_k(x + te)$ é continuamente diferenciável com derivada $\frac{d}{dt} f_k(x + te) = df_k(x + te)e$. Daí, segundo (1.7.9), segue

$$f_k(x + te) - f_k(x) = \int_0^t df_k(x + se)e \, ds \quad (2.2.7)$$

para todos os $k \in \mathbb{N}$ e todos os $t \in [-\varepsilon, \varepsilon]$. Além disso, segundo (1.7.7), vale

$$\begin{aligned} \left\| \int_0^t G(x + se)e \, ds - \int_0^t df_k(x + se)e \, ds \right\| &\leq \int_0^t \|G(x + se)e - df_k(x + se)e\| \, ds \\ &\leq \varepsilon \sup_{y \in [x-e, x+e]} \|G(y) - df_k(y)\|_{\mathcal{L}} \cdot \|e\| \end{aligned}$$

para todo $t \in [-\varepsilon, \varepsilon]$ e todo $k \in \mathbb{N}$. Segundo (2.2.6), o lado direito dessa desigualdade converge a zero para $k \rightarrow \infty$, e daí, segundo (2.2.7), segue

$$f(x + te) - f(x) = \int_0^t G(x + se)e \, ds \quad \text{para } -\varepsilon \leq t \leq \varepsilon. \quad (2.2.8)$$

Por isso e segundo (1.7.10), a função $[-\varepsilon, \varepsilon] \rightarrow Y : t \mapsto f(x + te)$ é continuamente diferenciável e vale $df(x)e = \frac{d}{dt}|_{t=0} (f(x + te) - f(x)) = G(x)e$ para todo vetor unitário e . Segue que $df(x) = G(x)$. O Teorema 2.2.6 é provado. \square

² Temos estabelecido integração (Lema 1.7.3) só para funções contínuas no intervalo com valores em dimensão *finita*, mas fica válido no caso de valores num espaço de Banach.

2.3 A regra da cadeia

Ora dedicamos-nos às regras básicas do cálculo diferencial, as quais contêm particularmente a regra de Leibniz e a regra da cadeia. Os seguintes são os resultados principais desta seção.

Teorema 2.3.1 (Regras do cálculo). *Sejam X, Y espaços de Banach, seja $U \subset X$ aberto, seja $x_0 \in U$, e sejam $f, g : U \rightarrow Y$ e $\alpha : U \rightarrow \mathbb{R}$ aplicações as quais são diferenciáveis no mesmo ponto x_0 . Então vale o seguinte.*

(i) *A soma $f + g : U \rightarrow Y$ é diferenciável no ponto x_0 e vale a fórmula*

$$d(f + g)(x_0)\xi = df(x_0)\xi + dg(x_0)\xi \quad (2.3.1)$$

para todos os $\xi \in X$.

(ii) *O produto $\alpha f : U \rightarrow Y$ é diferenciável no ponto x_0 e a sua derivada é dado pela **regra de Leibniz***

$$d(\alpha f)(x_0)\xi = \alpha(x_0)df(x_0)\xi + f(x_0)d\alpha(x_0)\xi \quad (2.3.2)$$

para todos os $\xi \in X$.

(iii) *Se $\alpha(x) \neq 0$ não se anula em nenhum dos pontos $x \in U$, então o quociente $f/\alpha := \alpha^{-1}f : U \rightarrow Y$ é diferenciável no ponto x_0 e vale a fórmula*

$$d\left(\frac{f}{\alpha}\right)(x_0)\xi = \frac{\alpha(x_0)df(x_0)\xi - f(x_0)d\alpha(x_0)\xi}{\alpha(x_0)^2} \quad (2.3.3)$$

para todos os $\xi \in X$.

Demonstração. Veja a página 54. □

Teorema 2.3.2 (Regra da cadeia). *Sejam $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ espaços de Banach, sejam $U \subset X$ e $V \subset Y$ abertos, seja $f : U \rightarrow V$ uma aplicação, diferenciável no ponto $x_0 \in U$, e seja $g : V \rightarrow Z$ uma aplicação, diferenciável no ponto $y_0 := f(x_0) \in V$. Então a composição $g \circ f : U \rightarrow Z$ é diferenciável no ponto x_0 e a sua derivada no ponto x_0 é a composição das derivadas de f e de g , ou seja*

$$d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0) : X \rightarrow Z. \quad (2.3.4)$$

Demonstração. Veja a página 57. □

No caso de $X = \mathbb{R}^n$ e $Y = \mathbb{R}$ a demonstração do Teorema 2.3.1 pode-se gerir em três modos diferentes. Primeiro, possa-se transferir a demonstração do teorema correspondente para funções de várias variáveis. Segundo, possa-se sem perda de generalidade supor $X = \mathbb{R}^n$ e depois, *com* perda de generalidade, supor adicionalmente que f e g são contínuas o que é suficiente para a grande maioria das aplicações. Nesta situação seguem as afirmações do Teorema 2.3.1 imediatamente do Teorema 2.2.1 e dos resultados correspondentes na Análise I sobre a soma, o produto, e o quociente de funções de uma variável real. Particularmente obtêm-se então as fórmulas no Teorema 2.3.1 da consideração das funções $t \mapsto f(x_0 + t\xi)$ e $t \mapsto g(x_0 + t\xi)$. O terceiro método é reduzir Teorema 2.3.1 à regra da cadeia no Teorema 2.3.2, e isto é o método o qual vamos usar aqui.

Teorema 2.3.2 \implies *Teorema 2.3.1* (Caso $Y = \mathbb{R}$ - funções).

Seja $\alpha := g$. Lema 2.1.15 diz que a aplicação $F : U \rightarrow \mathbb{R}^2$ definida assim

$$F(x) := \begin{pmatrix} f(x) \\ g(x) \end{pmatrix} \quad \text{para } x \in U$$

é diferenciável no ponto x_0 com derivada $dF(x_0) : X \rightarrow \mathbb{R}^2$ dada pela fórmula

$$dF(x_0)\xi = \begin{pmatrix} df(x_0)\xi \\ dg(x_0)\xi \end{pmatrix} \quad \text{para } \xi \in X. \quad (2.3.5)$$

Provaremos parte (i). Segundo o Teorema 2.2.1 a função $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida assim

$$h(y, z) := y + z$$

é diferenciável. Segundo (2.1.13) a derivada no ponto $(y, z) \in \mathbb{R}^2$ é

$$dh(y, z)(\eta, \zeta) = \eta + \zeta$$

para $(\eta, \zeta) \in \mathbb{R}^2$. Daí segue, segundo o Teorema 2.3.2, que a função

$$f + g = h \circ F : U \rightarrow \mathbb{R}$$

é diferenciável no ponto x_0 com derivada

$$\begin{aligned} d(f + g)(x_0)\xi &= dh(F(x_0))dF(x_0)\xi \\ &= dh(f(x_0), g(x_0))(df(x_0)\xi, dg(x_0)\xi) \\ &= df(x_0)\xi + dg(x_0)\xi \end{aligned}$$

para $\xi \in X$.

Provaremos parte (ii). Segundo o Teorema 2.2.1 a função $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida assim

$$h(y, z) := yz$$

é diferenciável. Segundo (2.1.13) a derivada no ponto $(y, z) \in \mathbb{R}^2$ é

$$dh(y, z)(\eta, \zeta) = z\eta + y\zeta$$

para $(\eta, \zeta) \in \mathbb{R}^2$. Daí segue, segundo o Teorema 2.3.2, que a função

$$fg = h \circ F : U \rightarrow \mathbb{R}$$

é diferenciável no ponto x_0 com derivada

$$\begin{aligned} d(fg)(x_0)\xi &= dh(F(x_0))dF(x_0)\xi \\ &= dh(f(x_0), g(x_0))(df(x_0)\xi, dg(x_0)\xi) \\ &= g(x_0)df(x_0)\xi + f(x_0)dg(x_0)\xi \end{aligned}$$

para $\xi \in X$.

Provaremos parte (iii). Como $g(x) \neq 0$ para todos os $x \in U$, vale a inclusão

$$F(U) \subset V := \mathbb{R} \times (\mathbb{R} \setminus \{0\}).$$

Segundo o Teorema 2.2.1 a função $h : V \rightarrow \mathbb{R}$ definida assim

$$h(y, z) := \frac{y}{z}$$

é diferenciável. A derivada $dh(y, z) : \mathbb{R}^2 \rightarrow \mathbb{R}$ no ponto $(y, z) \in V$ é

$$dh(y, z)(\eta, \zeta) = \frac{\eta}{z} - \frac{y\zeta}{z^2}$$

para $(\eta, \zeta) \in \mathbb{R}^2$, segundo (2.1.13). Então, segundo o Teorema 2.3.2, a função

$$\frac{f}{g} = h \circ F : U \rightarrow \mathbb{R}$$

é diferenciável no ponto x_0 com derivada

$$\begin{aligned} d\left(\frac{f}{g}\right)(x_0)\xi &= dh(F(x_0))dF(x_0)\xi \\ &= dh(f(x_0), g(x_0))(df(x_0)\xi, dg(x_0)\xi) \\ &= \frac{g(x_0)df(x_0)\xi - f(x_0)dg(x_0)\xi}{g(x_0)^2} \end{aligned}$$

para $\xi \in X$. Com isso o Teorema 2.3.1 fica provado. \square

Demonstração de Teorema 2.3.1 (Caso geral: X, Y). Provaremos parte (i). Sejam g e f diferenciável no mesmo ponto $a := x_0 \in U$. Denotamos $A := dg(a)$ e $B := df(a)$. Dado $\varepsilon > 0$, então existem constantes $\delta_g, \delta_f > 0$ tal que (2.1.4), com $\varepsilon/2$ em vez de ε , vale para g e para f . Neste caso vale (2.1.4) para a aplicação $g + f$, a aplicação linear $A + B$, e as constantes ε e $\delta := \min\{\delta_g, \delta_f\}$. Isso prova que $g + f$ é diferenciável no ponto $x_0 \in U$ com derivada (2.3.1).

Provaremos parte (ii). Sejam $\alpha : U \rightarrow \mathbb{R}$ e $f : U \rightarrow F$ diferenciável no mesmo ponto $a := x_0 \in U$. Denotamos $A := d\alpha(a)$ e $B := df(a)$. Seja $c_a > 0$ uma constante de continuidade de α no ponto a . Dado $\varepsilon > 0$, defina

$$\varepsilon_\alpha := \min \left\{ 4 \|A\|, \frac{\varepsilon}{4} \|f(a)\|_Y \right\}, \quad \varepsilon_f := \min \left\{ \frac{\varepsilon}{8 |\alpha(a)|}, \frac{4 \|A\| \|B\|}{c_a} \right\}.$$

Por convenção $1/0 = \infty$. Então existem constantes $\delta_\alpha, \delta_f > 0$ tal que (2.1.4) vale, respectivamente, para α e ε_α e para f e ε_f . Seja $h \in X$, tal que

$$0 < \|h\|_X < \delta := \min \left\{ \delta_\alpha, \delta_f, \frac{\varepsilon}{4 \|A\| \|B\|} \right\}.$$

Então adicionando três vezes zero encontraremos que

$$\begin{aligned} & \|(\alpha f)(a+h) - (\alpha f)(a) - \alpha(a)Bh - f(a)Ah\|_Y \\ & \leq \|f(a)\|_Y |\alpha(a+h) - \alpha(a) - Ah| + |\alpha(a+h) - \alpha(a) - Ah| \|Bh\|_Y \\ & \quad + |\alpha(a+h)| \|f(a+h) - f(a) - Bh\|_Y + |Ah| \|Bh\|_Y \\ & \leq \left(\varepsilon_\alpha \|f(a)\|_Y + \varepsilon_\alpha \|B\| \|h\|_X \right. \\ & \quad \left. + \varepsilon_f (|\alpha(a)| + c_a \|h\|_X) + \|A\| \|B\| \|h\|_X \right) \|h\|_X \\ & < \varepsilon \|h\|_X. \end{aligned}$$

Os primeiros três dos quatro termos da soma são $\leq \varepsilon/4$ e o último é $< \varepsilon/4$. Isso prova que $g + f$ é diferenciável no ponto $x_0 \in U$ com derivada (2.3.2).

Provaremos parte (iii). Com a ajuda de Exemplo 2.1.6 e da parte (ii) obtemos que $0 = d(1) = d(\alpha \frac{1}{\alpha}) = (d\alpha) \frac{1}{\alpha} + \alpha d(\frac{1}{\alpha})$ e daí $d(\frac{1}{\alpha}) = -\frac{1}{\alpha^2} d\alpha$. Segue desta última fórmula e, primeiramente, de parte (ii) que

$$d \left(f \frac{1}{\alpha} \right) = \frac{1}{\alpha} df + f d \left(\frac{1}{\alpha} \right) = \frac{\alpha df - f d\alpha}{\alpha^2}.$$

Com isso temos provado parte (iii) e Teorema 2.3.1 . □

Demonstração do Teorema 2.3.2. Sejam $A \in \mathcal{L}(X, Y)$ e $B \in \mathcal{L}(Y, Z)$ as aplicações lineares limitadas

$$A := df(x_0), \quad B := dg(y_0).$$

Resta mostrar que a composição

$$g \circ f : U \rightarrow Z$$

é diferenciável no ponto $x_0 \in U$ com derivada

$$d(g \circ f)(x_0) = BA \in \mathcal{L}(X, Z).$$

Para este fim seja $\varepsilon > 0$ dado. Como g é diferenciável no ponto $y_0 = f(x_0) \in V$ com $dg(y_0) = B$, existe uma constante $\rho > 0$, tal que

$$B_\rho(y_0; Y) = \{y \in Y \mid \|y - y_0\|_Y < \rho\} \subset V \quad (2.3.6)$$

e tal que todo elemento $y \in B_\rho(y_0; Y)$ satisfaz a desigualdade

$$\|g(y) - g(y_0) - B(y - y_0)\|_Z \leq \frac{\varepsilon}{2(1 + \|A\|)} \|y - y_0\|_Y. \quad (2.3.7)$$

Como f é diferenciável no ponto $x_0 \in U$ com $df(x_0) = A$, existe uma constante $\delta > 0$, tal que

$$B_\delta(x_0; X) = \{x \in X \mid \|x - x_0\|_X < \delta\} \subset U \quad (2.3.8)$$

e tal que todo elemento $x \in B_\delta(x_0; X)$ satisfaz a desigualdade

$$\|f(x) - f(x_0) - A(x - x_0)\|_Y \leq \frac{\varepsilon}{2(\varepsilon + \|B\|)} \|x - x_0\|_X. \quad (2.3.9)$$

Escolhendo aquele $\delta > 0$ suficientemente pequeno, então vale

$$(1 + \|A\|)\delta < \rho. \quad (2.3.10)$$

Ora, fixe $x \in X$ com $\|x - x_0\|_X < \delta$. Então $x \in U$, segundo (2.3.8), e através de (2.3.9) e (2.3.10) segue com $f(x_0) = y_0$ a desigualdade

$$\begin{aligned} \|f(x) - y_0\|_Y &\leq \|f(x) - f(x_0) - A(x - x_0)\|_Y + \|A(x - x_0)\|_Y \\ &\leq \frac{\varepsilon}{2(\varepsilon + \|B\|)} \|x - x_0\|_X + \|A\| \|x - x_0\|_X \\ &\leq (1 + \|A\|) \|x - x_0\|_X \\ &< (1 + \|A\|)\delta \\ &< \rho. \end{aligned} \quad (2.3.11)$$

Dali segue que $f(x) \in V$, segundo (2.3.6), e que

$$\begin{aligned} & \|g(f(x)) - g(f(x_0)) - BA(x - x_0)\|_Z \\ & \leq \|g(f(x)) - g(y_0) - B(f(x) - y_0)\|_Z + \|B(f(x) - f(x_0) - A(x - x_0))\|_Z \\ & \leq \frac{\varepsilon}{2} \frac{\|f(x) - y_0\|_Y}{1 + \|A\|} + \|B\| \|f(x) - f(x_0) - A(x - x_0)\|_Y \\ & \leq \left(\frac{\varepsilon}{2} + \frac{\varepsilon \|B\|}{2(\varepsilon + \|B\|)} \right) \|x - x_0\|_X \\ & \leq \varepsilon \|x - x_0\|_X. \end{aligned}$$

Aqui o segundo passo segue de (2.3.7) e o terceiro passo de (2.3.9) e (2.3.11). Com isto segue a afirmação do Teorema 2.3.2 da definição de diferenciabilidade em Definição 2.1.1. \square

Corolário 2.3.3. *Sejam $U \subset \mathbb{R}^n$ e $V \subset \mathbb{R}^m$ conjuntos abertos e sejam*

$$f = (f_1, \dots, f_m) : U \rightarrow V, \quad g = (g_1, \dots, g_\ell) : V \rightarrow \mathbb{R}^\ell$$

aplicações continuamente diferenciáveis. Então $g \circ f : U \rightarrow \mathbb{R}^\ell$ é continuamente diferenciável e vale

$$\frac{\partial(g_k \circ f)}{\partial x_i}(x) = \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \frac{\partial f_j}{\partial x_i}(x) \quad (2.3.12)$$

para todo $x \in U$, todo $i = 1, \dots, n$, e todo $k = 1, \dots, \ell$.

Demonstração. Seja $e_i \in \mathbb{R}^n$ o i -ésimo vetor da base canônica. Então vale

$$\begin{aligned} \frac{\partial(g_k \circ f)}{\partial x_i}(x) &= d(g_k \circ f)(x)e_i \\ &= dg_k(f(x))df(x)e_i \\ &= \left(\frac{\partial g_k}{\partial x_1}(f(x)) \cdots \frac{\partial g_k}{\partial x_m}(f(x)) \right) \begin{pmatrix} df_1(x)e_i \\ \vdots \\ df_m(x)e_i \end{pmatrix} \\ &= \sum_{j=1}^m \frac{\partial g_k}{\partial y_j}(f(x)) \frac{\partial f_j}{\partial x_i}(x). \end{aligned}$$

Aqui segue a primeira igualdade do Lema 2.1.11, a segunda do Teorema 2.3.2, a terceira do Lema 2.1.15 e Lema 2.1.16, e a quarta do Lema 2.1.11. Com isto fica provado a equação (2.3.12), e daí segue imediatamente a continuidade das derivadas parciais de $g \circ f$. \square

A equação (2.3.12) diz que a matriz jacobiana de $g \circ f$ no ponto x é o produto da jacobiana de g no ponto $f(x)$ com a jacobiana de f no ponto x . Esta equação também pode ser escrita na abreviatura

$$\frac{\partial z_k}{\partial x_i} = \sum_{j=1}^m \frac{\partial z_k}{\partial y_j} \frac{\partial y_j}{\partial x_i}. \quad (2.3.13)$$

Aqui y_j abrevia $f_j(x)$ e z_k abrevia $g_k(y)$. Esta terminologia é útil em situações quando entende-se as y_j como funções das variáveis x_1, \dots, x_n e as z_k como funções das variáveis y_1, \dots, y_m .

Exemplos

Exemplo 2.3.4 (Gradiente). Seja $U \subset \mathbb{R}^n$ um conjunto aberto e considere uma função continuamente diferenciável $f : U \rightarrow \mathbb{R}$. Então a derivada de f no ponto $x \in U$ é uma aplicação linear $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ a qual pode ser representada através da matriz jacobiana $df(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right) \in \mathbb{R}^{1 \times n}$ (um vetor linha). Frequentemente é útil escrever a derivada com um vetor coluna. Isto corresponde à transição de um funcional linear no \mathbb{R}^n a um vetor do \mathbb{R}^n . No caso da derivada de f chama-se este vetor **o gradiente de f no ponto x** e usa-se o símbolo

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}. \quad (2.3.14)$$

Esta notação é especialmente útil em conexão com o produto interno padrão $\langle \xi, \eta \rangle = \sum_{i=1}^n \xi_i \eta_i$ de dois vetores $\xi, \eta \in \mathbb{R}^n$. Neste caso a derivada direcional de f no ponto $x \in U$ na direção $\xi \in \mathbb{R}^n$ é dada pela fórmula

$$df(x)\xi = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)\xi_i = \langle \nabla f(x), \xi \rangle. \quad (2.3.15)$$

Ora, seja $x : I \rightarrow U$ uma C^1 -curva, então a regra da cadeia resulta na fórmula

$$\frac{d}{dt}f(x(t)) = df(x(t))\dot{x}(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x(t))\dot{x}_i(t) = \langle \nabla f(x(t)), \dot{x}(t) \rangle. \quad (2.3.16)$$

(Aqui o vetor $x = (x_1, \dots, x_n)$ tem *dois significados!* Por um lado x denota as variáveis independentes no conjunto $U \subset \mathbb{R}^n$, por outro lado x denota a aplicação $I \rightarrow U : t \mapsto x(t) = (x_1(t), \dots, x_n(t))$.)

Exemplo 2.3.5. Seja $f : U = (0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ a função

$$f(x, y) := x^y \quad \text{para } x > 0 \text{ e } y \in \mathbb{R}$$

e seja a curva $I = (0, \infty) \rightarrow U : t \mapsto (x(t), y(t))$ dada através de

$$x(t) := y(t) := t \quad \text{para } t > 0.$$

Então vale $f(x(t), y(t)) = t^t$ para $t > 0$ e segundo o Exemplo 2.2.4 vale

$$\frac{\partial f}{\partial x}(x, y) = yx^{y-1}, \quad \frac{\partial f}{\partial y}(x, y) = \log(x)x^y$$

para todos os $x > 0$ e $y \in \mathbb{R}$. Dali a equação (2.3.16) disponibiliza a fórmula

$$\frac{d}{dt}t^t = \frac{\partial f}{\partial x}(x(t), y(t))\dot{x}(t) + \frac{\partial f}{\partial y}(x(t), y(t))\dot{y}(t) = (1 + \log(t))t^t. \quad (2.3.17)$$

Exemplo 2.3.6. Na situação geral do Exemplo 2.3.4 é interessante frequentemente considerar as soluções da equação diferencial

$$\dot{x}(t) = -\nabla f(x(t)). \quad (2.3.18)$$

Para estas soluções a equação (2.3.16) é da forma

$$\frac{d}{dt}f(x(t)) = \langle \nabla f(x(t)), \dot{x}(t) \rangle = -\|\nabla f(x(t))\|^2. \quad (2.3.19)$$

Segundo o teorema do valor médio segue daí que a função definida por $I \rightarrow \mathbb{R} : t \mapsto f(x(t))$ é descendente monótono. **Exercício:** A função $I \rightarrow \mathbb{R} : t \mapsto f(x(t))$ ou é decrescente estritamente monótona no intervalo I completo, ou é constante a solução $I \rightarrow U : t \mapsto x(t)$. Depreende-se disso que a equação diferencial (2.3.18) não possui nenhuma solução periódica não-constante.

Exemplo 2.3.7. Uma outra situação interessante surge no Exemplo 2.3.4 se a função $f : U \rightarrow \mathbb{R}$ é constante ao longo de uma curva $x : I \rightarrow U$, em símbolos $f(x(t)) = c$ para todos os $t \in I$ e uma constante $c \in \mathbb{R}$. Neste caso segue de (2.3.16) a equação

$$0 = \frac{d}{dt}f(x(t)) = \langle \nabla f(x(t)), \dot{x}(t) \rangle.$$

Ou seja, o vetor velocidade $\dot{x}(t)$ fica ortogonal no vetor gradiente de f no ponto $x(t)$.

Exemplo 2.3.8. Um outro uso do Exemplo 2.3.4 resulta das **equações de Newton**

$$\ddot{x}(t) = -\nabla V(x(t)). \quad (2.3.20)$$

Aqui $U \subset \mathbb{R}^n$ é um conjunto aberto e $V : U \rightarrow \mathbb{R}$ uma função potencial suave. O gradiente dela é o *campo de força* e a equação (2.3.20) é o lei de Newton “*Força igual massa vezes aceleração*” (com massa igual 1). A *energia* de uma solução $x : I \rightarrow U$ no tempo t é a soma de energia *cinética* e *potencial* e é definido através de

$$E(t) := \frac{1}{2} \|\dot{x}(t)\|^2 + V(x(t)). \quad (2.3.21)$$

Segundo a regra da cadeia na forma (2.3.16) a derivada temporal da energia é

$$\dot{E}(t) = \langle \dot{x}(t), \ddot{x}(t) \rangle + \langle \nabla V(x(t)), \dot{x}(t) \rangle = \langle \dot{x}(t), \ddot{x}(t) + \nabla V(x(t)) \rangle = 0.$$

Portanto a energia é constante ao longo de toda solução de (2.3.20). No caso especial $n = 3$ com $U = \mathbb{R}^3 \setminus \{0\}$ e

$$V(x) = -\frac{1}{\|x\|}, \quad \nabla V(x) = \frac{x}{\|x\|^3}$$

obter-se o **problema de Kepler**.

Exemplo 2.3.9. Também pode-se deduzir o Lema 2.1.11 ($\partial_\xi f = df \xi$) como caso especial da regra da cadeia. Sejam então X, Y espaços de Banach, seja $U \subset X$ um conjunto aberto, e seja $f : U \rightarrow Y$ uma aplicação a qual é diferenciável no ponto $x \in U$. Além disso, seja $\hat{x} \in X$ dado e seja $I \subset \mathbb{R}$ um intervalo aberto com $0 \in I$, tal que

$$\gamma(t) := x + t\hat{x} \in U$$

para todos os $t \in I$. Segundo o Exemplo 2.1.7 a aplicação $\gamma : I \rightarrow U$ é diferenciável com $\dot{\gamma}(t) = \hat{x}$ para todos os $t \in I$. Daí, segundo o Teorema 2.3.2, a aplicação $f \circ \gamma : I \rightarrow Y$ é diferenciável no ponto $t = 0$ e vale

$$\left. \frac{d}{dt} \right|_{t=0} f(x + t\hat{x}) = \left. \frac{d}{dt} \right|_{t=0} (f \circ \gamma)(t) = df(\gamma(0))\dot{\gamma}(0) = df(x)\hat{x}. \quad (2.3.22)$$

Isto é a equação (2.1.8) no Lema 2.1.11 com $a = x$ e $\xi = \hat{x}$.

Os seguintes exemplos mostram que, ainda em dimensão finita, possa ser absolutamente útil trabalhar com espaços vetoriais gerais, em vez do \mathbb{R}^n . Assim se desocupa-se da necessidade escolher uma base e, as vezes, também é útil usar uma norma diferente da norma euclidiana. Por exemplo as normas no $\mathbb{R}^{m \times n}$ em Exemplo 1.2.6 não são normas de produtos internos.

Exemplo 2.3.10. Seja $\mathcal{A} = \mathcal{L}(X)$, onde X é um espaço de Banach, por exemplo $\mathcal{A} = \mathbb{R}^{n \times n}$ munido de uma das normas no Exemplo 1.2.6, e seja a aplicação $g : \mathcal{A} := \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ dado através de composição, ou seja

$$g(A, B) = AB \quad \text{para } A, B \in \mathcal{A} = \mathcal{L}(X).$$

Caso $\mathcal{A} = \mathbb{R}^{n \times n}$: O g é diferenciável, segundo o Exemplo 2.1.8, e Lema 2.1.15. Sua derivada $dg(A, B) \in \mathcal{L}(\mathcal{A} \times \mathcal{A}, \mathcal{A})$ no ponto $(A, B) \in \mathcal{A} \times \mathcal{A}$ pode ser determinada com ajuda do Exemplo 2.3.9. Para $(\widehat{A}, \widehat{B}) \in \mathcal{A} \times \mathcal{A}$ vale

$$\begin{aligned} dg(A, B)(\widehat{A}, \widehat{B}) &= \left. \frac{d}{dt} \right|_{t=0} f(A + t\widehat{A}, B + t\widehat{B}) \\ &= \left. \frac{d}{dt} \right|_{t=0} (A + t\widehat{A})(B + t\widehat{B}) \\ &= \left. \frac{d}{dt} \right|_{t=0} (AB + t(\widehat{A}B + A\widehat{B}) + t^2\widehat{A}\widehat{B}) \\ &= \widehat{A}B + A\widehat{B}. \end{aligned}$$

Caso geral: O g é diferenciável e a derivada é da mesma forma (Exercício).

Exemplo 2.3.11. Para \mathcal{A} do Exemplo 2.3.10 a aplicação $\phi_k : \mathcal{A} \rightarrow \mathcal{A}$ com

$$\phi_k(A) := A^k \quad \text{para } A \in \mathcal{A} = \mathcal{L}(X)$$

é continuamente diferenciável e a sua derivada $d\phi_k(A) \in \mathcal{L}(\mathcal{A}, \mathcal{A})$ é

$$d\phi_k(A)\widehat{A} = \sum_{j=0}^{k-1} A^j \widehat{A} A^{k-1-j} \quad \text{para } A, \widehat{A} \in \mathcal{A}. \quad (2.3.23)$$

Demonstração: Para $k = 1$ a função $\phi_1 = \text{id} : \mathcal{A} \rightarrow \mathcal{A}$ satisfaz a afirmação, segundo o Exemplo 2.1.7. Se ϕ_k satisfaça a afirmação para um $k \in \mathbb{N}$, se $\psi_k : \mathcal{A} \rightarrow \mathcal{A} \times \mathcal{A}$ seja definido através de $\psi_k(A) := (A, \phi_k(A))$ para $A \in \mathcal{A}$, e se $g : \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ seja a aplicação do Exemplo 2.3.10, então, segundo o Teorema 2.3.2, a função $\phi_{k+1} = g \circ \psi_k$ é continuamente diferenciável com derivada $d\phi_{k+1}(A)\widehat{A} = dg(\psi_k(A))d\psi_k(A)\widehat{A}$ para $A, \widehat{A} \in \mathcal{A}$. Daí também segue a fórmula (2.3.23) para $k + 1$ (Exercício).

Séries de potências de operadores lineares (por exemplo matrizes)

Seja a_0, a_1, a_2, \dots uma sequência de números reais, tal que

$$\rho := \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}} > 0. \quad (2.3.24)$$

Seja X um espaço de Banach e $\mathcal{A} := \mathcal{L}(X)$; veja o Exemplo 1.3.6 para $\mathcal{A} = \mathbb{R}^{n \times n}$. Ora, seja $U := \{A \in \mathcal{A} \mid \|A\| < \rho\}$, e seja $f : U \rightarrow \mathcal{A}$ a aplicação

$$f(A) := \sum_{k=0}^{\infty} a_k A^k \quad (2.3.25)$$

para $A \in U$. Que esta série converge absolutamente para todo $A \in U$ foi provado no Corolário 1.3.11, e que a aplicação $f : U \rightarrow X$ definida através de (2.3.25) é contínua foi provado no Teorema 1.3.13.

Teorema 2.3.12. *A aplicação $f : U \rightarrow \mathcal{A}$ em (2.3.25) é continuamente diferenciável e sua derivada $df(A) \in \mathcal{L}(\mathcal{A}, \mathcal{A})$ no ponto $A \in U$ é*

$$df(A)\hat{A} := \sum_{k=1}^{\infty} a_k \sum_{j=0}^{k-1} A^j \hat{A} A^{k-1-j} \quad \text{para } \hat{A} \in \mathcal{A}. \quad (2.3.26)$$

Demonstração. Para $\ell \in \mathbb{N}$ seja $f_\ell : U \rightarrow \mathcal{A}$ a aplicação definida por

$$f_\ell(A) := \sum_{k=0}^{\ell} a_k \phi_k(A) = \sum_{k=0}^{\ell} a_k A^k \quad \text{para } A \in U.$$

Segundo Exemplo 2.3.11 e Teorema 2.3.1 f_ℓ é continuamente diferenciável e

$$df_\ell(A) = \sum_{k=1}^{\ell} a_k \Phi_k(A), \quad \Phi_k(A)\hat{A} := \sum_{j=0}^{k-1} A^j \hat{A} A^{k-1-j}.$$

Para $0 < r < \rho$ seja $B_r := \{A \in \mathcal{A} \mid \|A\| \leq r\}$. Então vale

$$\|\Phi_k(A)\hat{A}\| \leq \sum_{j=0}^{k-1} \|A^j \hat{A} A^{k-1-j}\| \leq k \|A\|^{k-1} \|\hat{A}\| \leq kr^{k-1} \|\hat{A}\|$$

para $A \in B_r, \hat{A} \in \mathcal{A}$, e daqui $\sum_{k=1}^{\infty} \|a_k \Phi_k(A)\|_{\mathcal{L}(\mathcal{A}, \mathcal{A})} \leq \sum_{k=1}^{\infty} k |a_k| r^{k-1} < \infty$ para $A \in B_r$ segundo (2.3.24). Com isto segue (da prova) do Teorema 1.3.10 que a sequência de aplicações $df_\ell : U \rightarrow \mathcal{L}(\mathcal{A}, \mathcal{A})$ converge uniformemente ao longo de todo compacto $K \subset U$. Como $f_\ell : U \rightarrow \mathcal{A}$ também converge uniformemente para f ao longo de todo compacto $K \subset U$, segundo o Teorema 2.2.6 o limite $f : U \rightarrow \mathcal{A}$ é continuamente diferenciável com derivada (2.3.26). \square

Exemplo 2.3.13. Segundo o Teorema 2.3.12 a aplicação exponencial

$$\exp : \mathcal{A} \rightarrow \mathcal{A}, \quad \exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!}, \quad (2.3.27)$$

onde $\mathcal{A} = \mathcal{L}(X)$, por exemplo $\mathcal{A} = \mathbb{R}^{n \times n}$, é continuamente diferenciável e a derivada é dada através da fórmula

$$d \exp(A) \widehat{A} = \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{j=0}^{k-1} A^j \widehat{A} A^{k-1-j} \quad (2.3.28)$$

para $A, \widehat{A} \in \mathcal{A}$. Particularmente vale

$$d \exp(0) \widehat{A} = \widehat{A}$$

para todos os $\widehat{A} \in \mathcal{A}$. Vamos ver em Corolário 4.4.4 que a aplicação exponencial é suave. As derivadas superiores serão introduzidas na Seção 2.5.

Exemplo 2.3.14. Seja de novo $\mathcal{A} = \mathcal{L}(X)$, onde X é um espaço de Banach, e seja $\text{GL}(X) := \mathcal{G}$ a subálgebra dos elementos invertíveis (veja o Exemplo 1.4.2). Então a aplicação

$$\text{Inv} : \text{GL}(X) \rightarrow \text{GL}(X), \quad \text{Inv}(A) := A^{-1} \quad (2.3.29)$$

é continuamente diferenciável com derivada

$$d \text{Inv}(A) \widehat{A} = -A^{-1} \widehat{A} A^{-1} \quad (2.3.30)$$

para todos os $A \in \text{GL}(n, \mathbb{R})$ e $\widehat{A} \in \mathcal{A}$. **Demonstração:** Para $A \in \text{GL}(X)$ e $\widehat{A} \in \mathcal{A}$ com $\|\widehat{A}\| \|A^{-1}\| < 1$ vale, segundo a parte (ii) do Teorema 1.4.1, que

$$A + \widehat{A} \in \text{GL}(X), \quad (A + \widehat{A})^{-1} = A^{-1} \sum_{k=0}^{\infty} (-\widehat{A} A^{-1})^k.$$

Ou seja, $\text{Inv}(A + \widehat{A}) = A^{-1} f(\widehat{A} A^{-1}) = L \circ f \circ R(\widehat{A})$ para $\|\widehat{A}\| < \|A^{-1}\|^{-1}$, onde as aplicações lineares $L, R : \mathcal{A} \rightarrow \mathcal{A}$ são definidas através de $L(Z) := A^{-1} Z$ e $R(\widehat{A}) := \widehat{A} A^{-1}$ e a aplicação $f : \{Y \in \mathcal{A} \mid \|Y\| < 1\} \rightarrow \mathcal{A}$ através de $f(Y) := \sum_{k=0}^{\infty} (-Y)^k$. Segundo o Teorema 2.3.12 a aplicação f é continuamente diferenciável com $df(0) \widehat{Y} = -\widehat{Y}$ para $\widehat{Y} \in \mathcal{A}$. Com isso a afirmação segue da regra da cadeia no Teorema 2.3.2 para a aplicação $L \circ f \circ R$.

2.4 O Teorema da Barreira

Suponhamos nesta seção que X e Y são espaços de Banach.³ Para os operadores lineares $A \in \mathcal{L}(X, Y)$ utilizamos sempre a norma operador

$$\|A\| := \|A\|_{\mathcal{L}(X, Y)} := \sup_{0 \neq x \in X} \frac{\|Ax\|_Y}{\|x\|_X}.$$

(Veja a Seção 1.2.) Pode-se ver o teorema da barreira como uma extensão do teorema do valor médio para aplicações diferenciáveis entre espaços vetoriais normados de dimensão finita.

Teorema 2.4.1 (Teorema da Barreira). *Seja $U \subset X$ aberto e $f : U \rightarrow Y$ uma aplicação continuamente diferenciável.*

(i) *Seja $K \subset U$ um subconjunto compacto convexo, então vale*

$$\|f(x_0) - f(x_1)\|_Y \leq c_K \|x_0 - x_1\|_X \quad (2.4.1)$$

para todos os $x_0, x_1 \in K$ e com a constante $c_K := \sup_{x \in K} \|df(x)\|$.

(ii) *Seja $K \subset U$ um subconjunto compacto. Então existe um $c > 0$, tal que*

$$\|f(x_0) - f(x_1)\|_Y \leq c \|x_0 - x_1\|_X \quad (2.4.2)$$

para todos os $x_0, x_1 \in K$.

Demonstração. Veja página 66. □

A demonstração do Teorema 2.4.1 precisa a definição da integral de uma função contínua num intervalo compacto com valores num espaço de Banach X . Suponha $\dim X = n < \infty$. Então no caso de $X = \mathbb{R}^n$ a ideia é definir a integral simplesmente coordenada por coordenada, ou seja, a integral de uma função contínua $x = (x_1, \dots, x_n) : [a, b] \rightarrow \mathbb{R}^n$ é o vetor cuja i -ésima coordenada é dada pela integral da função $x_i : [a, b] \rightarrow \mathbb{R}$. Se nosso espaço vetorial de dimensão n ainda não seja munido de uma base, então pode-se para esta definição simplesmente escolher uma base primeiramente. Mas então deve-se mostrar que tal definição não depende da escolha da base. Uma outra possibilidade para a definição da integral de uma função com valores num espaço de Banach qualquer seria construir-la de novo através da convergência das somas de Riemann. Quer-se usar (em dimensão finita) ambos métodos, então deve-se mostrar que ambos lidam ao mesmo resultado. Isso não é difícil e já foi provado em Lema 1.7.1.

³ Temos estabelecido integração (Lema 1.7.3) só para funções contínuas no intervalo com valores em dimensão *finita*, mas fica válido no caso de valores num espaço de Banach.

Demonstração do Teorema 2.4.1. Provaremos parte (i). Seja ora $K \subset U$ um subconjunto compacto convexo e seja $c_K := \sup_{x \in K} \|df(x)\|$. Então $c_K < \infty$ como K é compacto. Sejam $x_0, x_1 \in K$. Então $x_0 + t(x_1 - x_0) \in K$, como K é convexo. Definimos agora a função $y : [0, 1] \rightarrow Y$ através de

$$y(t) := f(x_0 + t(x_1 - x_0)) \quad \text{para } 0 \leq t \leq 1.$$

Esta função é continuamente diferenciável segundo o Teorema 2.3.2 com derivada

$$\dot{y}(t) = df(x_0 + t(x_1 - x_0))(x_1 - x_0).$$

Segue daí

$$\|\dot{y}(t)\|_Y \leq \|df(x_0 + t(x_1 - x_0))\| \|x_1 - x_0\|_X \leq c_K \|x_1 - x_0\|_X. \quad (2.4.3)$$

Além disso segue dali, segundo a parte (v) do Lema 1.7.3, a equação

$$f(x_1) - f(x_0) = y(1) - y(0) = \int_0^1 \dot{y}(t) dt.$$

Dali segue em contrapartida, segundo a parte (iii) do Lema 1.7.3, a desigualdade

$$\|f(x_1) - f(x_0)\|_Y \leq \int_0^1 \|\dot{y}(t)\|_Y dt \leq c_K \|x_1 - x_0\|_X.$$

Aqui o último passo segue de (2.4.3) e com isso a parte (i) fica provada.

Provaremos parte (ii) com um argumento indireto. Seja então $K \subset U$ um subconjunto compacto tal que não vale a afirmação da parte (ii).

Então existem duas seqüências $(a_k)_{k \in \mathbb{N}}$ e $(b_k)_{k \in \mathbb{N}}$ com

$$\|f(a_k) - f(b_k)\|_Y > k \|a_k - b_k\|_X \quad \text{para todos os } k \in \mathbb{N}. \quad (2.4.4)$$

Como K é compacto podemos supor, através de escolher uma subsequência adequada, que os limites

$$a := \lim_{k \rightarrow \infty} a_k, \quad b := \lim_{k \rightarrow \infty} b_k \quad (2.4.5)$$

existam. Estes limites são necessariamente em K e – como veremos – coincidem.

Para isso consideramos a diferença $a_k - b_k$. Segundo (2.4.4) vale

$$\begin{aligned} \|a_k - b_k\|_X &< \frac{1}{k} \|f(a_k) - f(b_k)\|_Y \\ &\leq \frac{1}{k} (\|f(a_k)\|_Y + \|f(b_k)\|_Y) \\ &\leq \frac{2}{k} \sup_{x \in K} \|f(x)\|_Y. \end{aligned}$$

Como K é compacto, temos que

$$\sup_{x \in K} \|f(x)\| < \infty$$

e daí

$$\|a - b\|_X = \lim_{k \rightarrow \infty} \|a_k - b_k\|_X = 0.$$

Portanto vale

$$a = b \in K. \quad (2.4.6)$$

Ora, no caso de $\dim X < \infty$ escolhemos um $\varepsilon > 0$, tal que

$$B := \{x \in X \mid \|x - a\|_X \leq \varepsilon\} \subset U.$$

No caso de um espaço *Banach* X qualquer, seja B o fecho do envelope convexo de K . Então em todo caso B é um subconjunto compacto de U . Dali

$$c_B := \sup_{x \in B} \|df(x)\| < \infty$$

e, segundo a parte (i), vale

$$\|f(x) - f(x')\|_Y \leq c_B \|x - x'\|_X \quad \text{para todos } x, x' \in B. \quad (2.4.7)$$

Segundo (2.4.5) e (2.4.6) existe uma constante $k_0 \in \mathbb{N}$ tal que $a_k, b_k \in B$ para todos os $k \in \mathbb{N}$ com $k \geq k_0$. Segue daí, segundo (2.4.4) e (2.4.7), a desigualdade

$$k \|a_k - b_k\|_X < \|f(a_k) - f(b_k)\|_Y \leq c_B \|a_k - b_k\|_X$$

para todo número inteiro $k > k_0$. Para $k \geq \max\{k_0, c_B\}$ isto é uma contradição. Com isso o Teorema 2.4.1 fica provado. \square

Com a ajuda do Teorema da Barreira estamos agora na posição de provar o reverso da observação que funções constantes têm derivada nula. Precisamos para isso o conceito de um espaço conexo em Apêndice B.

Teorema 2.4.2. *Seja $U \subset X$ um subconjunto aberto, conexo, não-vazio e seja $f : U \rightarrow Y$ uma aplicação diferenciável cuja derivada $df(x) = 0$ anula-se em todo ponto $x \in U$. Então f é constante.*

Demonstração. Se U é convexo, então a afirmação segue diretamente da desigualdade (2.4.1) no Teorema 2.4.1 com $c_K = 0$. No caso geral nosso argumento é o seguinte.

Seja $x_0 \in U$ e $y_0 := f(x_0)$. Resta mostrar que $f(x) = y_0$ para todos os $x \in U$. Para isso definimos o conjunto

$$A := \{x \in U \mid f(x) = y_0\}.$$

Este conjunto obviamente não é vazio, porque x_0 é um elemento de A . Além disso f é contínua segundo o Lema 2.1.5, e dali $A = f^{-1}(y_0)$ é um subconjunto U -fechado de U . Em outras palavras, A é fechado com respeito à topologia relativa de U (veja o Lema B.1.2).

Provaremos agora que A é um subconjunto aberto de U . Ora, seja $x \in A$. Como U é aberto existe uma constante $\varepsilon > 0$ tal que

$$B_\varepsilon(x) = \{x' \in X \mid \|x - x'\| < \varepsilon\} \subset U.$$

Se $x' \in B_\varepsilon(x)$, então $x + t(x' - x) \in B_\varepsilon(x) \subset U$ para todo $t \in [0, 1]$. Segue daí, segundo a parte (v) do Lema 1.7.3, a equação

$$\begin{aligned} f(x') - y_0 &= f(x') - f(x) \\ &= \int_0^1 \frac{d}{dt} f(x + t(x' - x)) dt \\ &= \int_0^1 df(x + t(x' - x))(x' - x) dt \\ &= 0. \end{aligned}$$

Aqui o primeiro passo segue da hipótese $x \in A$, o segundo passo da diferenciabilidade da composição $[0, 1] \rightarrow Y : t \mapsto f(x + t(x' - x))$ no Teorema 2.3.2 assim como da parte (v) do Lema 1.7.3, o terceiro passo da fórmula para a derivada no Teorema 2.3.2, e o último passo do fato que a derivada de f se anula pela hipótese. Assim temos mostrado $B_\varepsilon(x) \subset A$. Então A é um subconjunto não-vazio de U , tanto aberto como fechado com respeito à topologia relativa de U . Como U é conexo, vale $A = U$, o que prova Teorema 2.4.2. \square

2.5 Derivadas superiores

Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja $f : U \rightarrow \mathbb{R}^m$ uma aplicação, parcialmente diferenciável em todo U .

Então as derivadas parciais

$$\partial_i f = \frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}^m, \quad i = 1, \dots, n,$$

mesmas são aplicações de U para \mathbb{R}^m e podemos considerar o caso que eles são de novo parcialmente diferenciável. Para isso lembramos a definição da derivada parcial com respeito à i -ésima variável x_i . Num ponto $x = (x_1, \dots, x_n)$ de U ela é dada através de

$$\begin{aligned} \partial_i f(x) &= \frac{\partial f}{\partial x_i}(x) \\ &= \lim_{t \rightarrow 0} \frac{f(x + te_i) - f(x)}{t} \\ &= \left. \frac{d}{dt} \right|_{t=0} f(x + te_i) \\ &= \left. \frac{d}{dt} \right|_{t=0} f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) \end{aligned}$$

onde $e_i \in \mathbb{R}^n$ é o i -ésimo vetor unitário padrão. Ou seja, recebe-se a i -ésima derivada parcial no ponto x no modo que fixa-se as coordenadas x_j com $j \neq i$ e diferencia-se no ponto x_i a função de uma variável real surgindo daí.

Definição 2.5.1. *Uma aplicação $f : U \rightarrow \mathbb{R}^m$ definida num subconjunto aberto $U \subset \mathbb{R}^n$ chama-se de **duas vezes parcialmente diferenciável** se ela é parcialmente diferenciável e as suas derivadas parciais $\partial_j f : U \rightarrow \mathbb{R}^m$ também são parcialmente diferenciável para $j = 1, \dots, n$. Se uma tal aplicação $f : U \rightarrow \mathbb{R}^m$ é duas vezes parcialmente diferenciável, então para $i, j = 1, \dots, n$ as aplicações $\partial_i \partial_j f : U \rightarrow \mathbb{R}^m$, definidas através de*

$$\begin{aligned} \partial_i \partial_j f(x) &:= \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \\ &:= \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x) \\ &:= \left. \frac{d}{dt} \right|_{t=0} \frac{\partial f}{\partial x_j}(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) \end{aligned} \tag{2.5.1}$$

para $x \in U$, são chamadas das **segundas derivadas parciais** de f .

Desta definição surge naturalmente a questão de qual efeito para as derivadas parciais tem uma permutação da ordem delas. Consideramos um exemplo.

Exemplo 2.5.2. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função definida através de $f(0, 0) = 0$ e

$$f(x, y) := \frac{x^3y - xy^3}{x^2 + y^2}$$

para $x^2 + y^2 \neq 0$. Esta função é parcialmente diferenciável. As suas derivadas parciais anulam-se no ponto $x = y = 0$, e são dadas através de

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= \frac{(3x^2y - y^3)(x^2 + y^2) - 2x(x^3y - xy^3)}{(x^2 + y^2)^2} \\ &= \frac{3x^4y + 3x^2y^3 - x^2y^3 - y^5 - 2x^4y + 2x^2y^3}{(x^2 + y^2)^2} \\ &= \frac{x^4y + 4x^2y^3 - y^5}{(x^2 + y^2)^2} \\ \frac{\partial f}{\partial y}(x, y) &= \frac{x^5 - 4x^3y^2 - xy^4}{(x^2 + y^2)^2}. \end{aligned}$$

para $x^2 + y^2 \neq 0$. A segunda fórmula segue da primeira através de trocar x e y . As derivadas parciais $\partial f/\partial x$ e $\partial f/\partial y$ são contínuas (também no ponto $x = y = 0$), e daí f é continuamente diferenciável. Particularmente vale que

$$\frac{\partial f}{\partial x}(x, 0) = 0, \quad \frac{\partial f}{\partial x}(0, y) = -y, \quad \frac{\partial f}{\partial y}(x, 0) = x, \quad \frac{\partial f}{\partial y}(0, y) = 0.$$

Segue dali que f é (também no ponto $x = y = 0$) duas vezes parcialmente diferenciável com

$$\frac{\partial^2 f}{\partial x \partial x}(0, 0) = 0, \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1, \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1, \quad \frac{\partial^2 f}{\partial y \partial y}(0, 0) = 0.$$

As segundas derivadas parciais mistas não coincidem portanto no ponto origem $x = y = 0$. Além disso calcular as segundas derivadas parciais em $\mathbb{R}^2 \setminus \{(0, 0)\}$ mostra que elas não são contínuas no ponto $x = y = 0$ (Exercício).

O seguinte Teorema de Hermann Amandus Schwarz⁴ mostra que a assimetria das segundas derivadas parciais mistas no exemplo 2.5.2 é um caso especial, não pode surgir se as segundas derivadas parciais não só existem em todo lado, mas também são contínuas.

Teorema 2.5.3 (Schwarz). *Seja $U \subset \mathbb{R}^n$ um aberto e seja $f : U \rightarrow \mathbb{R}^m$ uma aplicação parcialmente diferenciável, seja dado um ponto $x_0 \in U$ e índices $i, j \in \{1, \dots, n\}$ com $i \neq j$. Suponhamos que $\partial f / \partial x_j : U \rightarrow \mathbb{R}^m$ é parcialmente diferenciável com respeito à variável x_i em todo o ponto de U , e que a aplicação*

$$\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} : U \rightarrow \mathbb{R}^m$$

é contínua no ponto x_0 . Então a aplicação $\partial f / \partial x_i : U \rightarrow \mathbb{R}^m$ é parcialmente diferenciável com respeito à variável x_j no ponto x_0 e vale

$$\frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}(x_0) = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x_0).$$

Demonstração. Segundo o Lema 2.1.15 é suficiente considerar o caso $m = 1$. Além disso as afirmações do teorema só afetam a restrição da aplicação f no subconjunto de U no qual todas as variáveis, menos x_i e x_j , são fixadas. Por isso, também é suficiente considerar o caso $n = 2$. Então suponhamos do início que $U \subset \mathbb{R}^2$ é um conjunto aberto e usamos a notação (x, y) para as variáveis em U . Mesmo suponhamos que $f : U \rightarrow \mathbb{R}$ é uma função parcialmente diferenciável, que existe a segunda derivada parcial

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial y}(x + h, y) - \frac{\partial f}{\partial y}(x, y) \right) \quad (2.5.2)$$

para todos os $(x, y) \in U$, e que a função $\frac{\partial^2 f}{\partial x \partial y} : U \rightarrow \mathbb{R}$ é contínua no ponto $(x_0, y_0) \in U$. O valor neste ponto denotamos de

$$a_0 := \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0). \quad (2.5.3)$$

Resta mostrar a equação

$$\lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial x}(x_0, y_0 + h) - \frac{\partial f}{\partial x}(x_0, y_0) \right) = a_0. \quad (2.5.4)$$

⁴ Hermann Amandus Schwarz foi um aluno de Ernst Eduard Kummer e Karl Weierstraß. Ele foi Professor na ETH Zürich desde 1869 até 1875 e no ano 1914 foi atribuído-lê a grandeza de doutor honores causa da ETH Zürich.

Seja $\varepsilon > 0$ dado. Como a função $\partial^2 f / \partial x \partial y : U \rightarrow \mathbb{R}$ é contínua no ponto (x_0, y_0) e la tem valor a_0 , existe uma constante $\delta > 0$, tal que para todos os $(x, y) \in \mathbb{R}^n$ vale a seguinte afirmação:

$$\begin{aligned} |x - x_0| < \delta, \\ |y - y_0| < \delta \end{aligned} \implies (x, y) \in U \text{ e } \left| \frac{\partial^2 f}{\partial x \partial y}(x, y) - a_0 \right| < \varepsilon. \quad (2.5.5)$$

Com esta escolha de δ vamos provar em três passos que todo $y \in \mathbb{R}$, tal que $0 < |y - y_0| < \delta$ satisfaz a desigualdade

$$\left| \frac{\partial f}{\partial x}(x_0, y) - \frac{\partial f}{\partial x}(x_0, y_0) - a_0(y - y_0) \right| \leq \varepsilon |y - y_0|. \quad (2.5.6)$$

Como $\varepsilon > 0$ foi qualquer, assim fica provada a equação desejada (2.5.4).

Passo 1. Para todos os $x, y \in \mathbb{R}$ com $0 < |x - x_0| < \delta$ e $|y - y_0| < \delta$ vale

$$\left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y) - a_0(x - x_0) \right| < \varepsilon |x - x_0|. \quad (2.5.7)$$

Fixamos dois números reais x, y com

$$0 < |x - x_0| < \delta, \quad |y - y_0| < \delta.$$

Suponhamos para já que $x_0 < x < x_0 + \delta$. Então, segundo (2.5.5), $(\xi, y) \in U$ para todo $\xi \in [x_0, x]$. Pela hipótese é diferenciável em todo o ponto a função

$$[x_0, x] \rightarrow \mathbb{R} : \xi \mapsto \frac{\partial f}{\partial y}(\xi, y).$$

Sua derivada no ponto $\xi \in [x_0, x]$ é o número $\partial^2 f / \partial x \partial y(\xi, y)$. Segundo o Teorema do Valor Médio existe então um número $x_0 < \xi < x$ com

$$\frac{\partial^2 f}{\partial x \partial y}(\xi, y) = \frac{\frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y)}{x - x_0}.$$

Ora pegamos a diferença com a_0 e pegamos o módulo. Com $|\xi - x_0| < \delta$ e $|y - y_0| < \delta$ segue então, segundo (2.5.5), a desigualdade desejada (2.5.7). No caso de $x < x_0$ argumenta-se igualmente com a mesma função no intervalo $[x, x_0]$. Com isso o Passo 1 fica provado.

Passo 2. Para todos os $x, y \in \mathbb{R}$ com $0 < |x - x_0| < \delta$ e $0 < |y - y_0| < \delta$ vale

$$\left| \frac{f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0)}{(x - x_0)(y - y_0)} - a_0 \right| < \varepsilon. \quad (2.5.8)$$

Fixamos dois números reais x, y com

$$0 < |x - x_0| < \delta, \quad 0 < |y - y_0| < \delta.$$

Suponhamos para já que $y_0 < y < y_0 + \delta$. Então, segundo (2.5.5), $(x, \eta) \in U$ para todo $\eta \in [y_0, y]$ e definimos $g : [y_0, y] \rightarrow \mathbb{R}$ através de

$$g(\eta) := f(x, \eta) - f(x_0, \eta) \quad \text{para } y_0 \leq \eta \leq y.$$

Esta função é diferenciável pela hipótese. Segundo o Teorema do Valor Médio existe então um número $y_0 < \eta < y$ com

$$g'(\eta) = \frac{g(y) - g(y_0)}{y - y_0}.$$

Segundo a definição de g pode-se escrever esta equação na forma

$$\frac{\partial f}{\partial y}(x, \eta) - \frac{\partial f}{\partial y}(x_0, \eta) = \frac{f(x, y) - f(x_0, y) - f(x, y_0) + f(x_0, y_0)}{y - y_0}.$$

Ora dividimos por $x - x_0$, pegamos a diferença com a_0 , e pegamos o módulo. Com $0 < |x - x_0| < \delta$ e $|\eta - y_0| < \delta$ segue então, segundo (2.5.7) no Passo 1, a desigualdade desejada (2.5.8). No caso de $y < y_0$ argumenta-se igualmente usando a mesma função g no intervalo $[y, y_0]$. Com isso o Passo 2 fica provado.

Passo 3. Todo $y \in \mathbb{R}$ com $0 < |y - y_0| < \delta$ satisfaz (2.5.6).

Fixamos um número real y com $0 < |y - y_0| < \delta$ e escrevemos a desigualdade (2.5.8) no Passo 2 na forma

$$\left| \frac{f(x, y) - f(x_0, y)}{x - x_0} - \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} - a_0(y - y_0) \right| < \varepsilon |y - y_0|$$

para todos os $x \in \mathbb{R}$ com $0 < |x - x_0| < \delta$. Com a transição ao limite $x \rightarrow x_0$ recebe-se daí a desigualdade desejada

$$\left| \frac{\partial f}{\partial x}(x_0, y) - \frac{\partial f}{\partial x}(x_0, y_0) - a_0(y - y_0) \right| \leq \varepsilon |y - y_0|.$$

Com isso o Passo 3 e o Teorema 2.5.3 ficam provados. \square

Definição 2.5.4. *Sejam $\ell, m, n \in \mathbb{N}$ e seja $U \subset \mathbb{R}^n$ um subconjunto aberto.*

(i) *Chama-se uma aplicação $f : U \rightarrow \mathbb{R}^m$ de **duas vezes continuamente diferenciável** se é continuamente diferenciável e as suas derivada parciais*

$$\partial_i f : U \rightarrow \mathbb{R}^m, \quad i = 1, \dots, n,$$

são continuamente diferenciáveis.

(ii) *Chama-se uma aplicação $f : U \rightarrow \mathbb{R}^m$ de **ℓ -vezes continuamente diferenciável**, se é continuamente diferenciável e as suas derivada parciais*

$$\partial_i f : U \rightarrow \mathbb{R}^m, \quad i = 1, \dots, n,$$

são $(\ell - 1)$ -vezes continuamente diferenciável.

(iii) *O conjunto das aplicações de U para \mathbb{R}^m as quais são ℓ -vezes continuamente diferenciáveis denotamos de*

$$C^\ell(U, \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m \mid f \text{ é } \ell\text{-vezes continuamente diferenciável}\}.$$

*Os elementos de $C^\ell(U, \mathbb{R}^m)$ também são chamados de **aplicação C^ℓ** .*

(vi) *Sejam X, Y espaços de Banach e $V \subset X$ um subconjunto aberto. Chama-se uma aplicação $f : V \rightarrow Y$ de **ℓ -vezes continuamente diferenciável**, ou uma **aplicação C^ℓ** , se f é diferenciável em todo ponto $x \in V$ e o seu diferencial $df : V \rightarrow \mathcal{L}(X, Y)$ é $(\ell - 1)$ -vezes diferenciável. (Veja a Definição 2.2.2 para o caso $\ell = 1$.) Definimos*

$$d^2 f := d(df) : V \rightarrow \mathcal{L}(X, \mathcal{L}(X, Y))$$

*e no caso de $\ell \geq 2$ definimos indutivamente $d^k f := d(d^{k-1} f)$. Chamamos f de **suave** ou **da classe C^∞** , se f é ℓ -vezes continuamente diferenciável para todo $\ell \in \mathbb{N}$. Denota-se de $C^\ell(V, Y)$ o conjunto dos C^ℓ aplicações $f : V \rightarrow Y$.*

Definição 2.5.5 (Derivadas parciais superiores). *Sejam $\ell, m, n \in \mathbb{N}$, seja $U \subset \mathbb{R}^n$ um subconjunto aberto, e seja $f : U \rightarrow \mathbb{R}^m$ uma aplicação C^ℓ . Para $i_1, \dots, i_\ell \in \{1, \dots, n\}$ a **derivada parcial***

$$\frac{\partial^\ell f}{\partial x_{i_1} \cdots \partial x_{i_\ell}} : U \rightarrow \mathbb{R}^m$$

da ordem ℓ é definido indutivamente através de

$$\frac{\partial^\ell f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_\ell}} := \frac{\partial}{\partial x_{i_1}} \frac{\partial^{\ell-1} f}{\partial x_{i_2} \cdots \partial x_{i_\ell}}.$$

Comentário 2.5.6 (Aplicações da classe C^ℓ). Segundo a Definição 2.5.4 e a Definição 2.5.5, uma aplicação $f : U \rightarrow \mathbb{R}^m$ é ℓ -vezes continuamente diferenciável se e somente se *todas suas derivadas parciais até a ordem ℓ existem e são contínuas*. Isto significa que para toda sequência de índices $i_1, \dots, i_\ell \in \{1, \dots, n\}$ (de comprimento ℓ), primeiro a derivada parcial $\partial f / \partial x_{i_\ell}$ existe em todo lugar e é contínua, segundo ela é de novo parcialmente diferenciável com respeito $x_{i_{\ell-1}}$ e a segunda derivada parcial $\partial^2 f / \partial x_{i_{\ell-1}} \partial x_{i_\ell}$ resultando daí é contínua, terceiro então ela é parcialmente diferenciável com respeito $x_{i_{\ell-2}}$ e a terceira derivada $\partial^3 f / \partial x_{i_{\ell-2}} \partial x_{i_{\ell-1}} \partial x_{i_\ell}$ resultando daí é contínua, ... patati-patatá ... até $\partial^\ell f / \partial x_{i_1} \cdots \partial x_{i_\ell}$ da ordem ℓ a qual também é contínua.

Comentário 2.5.7 (Multi-índices). Como, segundo o Teorema 2.5.3, a colocação das derivadas parciais de uma aplicação C^ℓ não afeta o resultado, faz sentido introduzir uma notação simplificada a qual atende repetições entre os índices i_1, \dots, i_ℓ . Como só depende de quantas vezes aparece um índice $i \in \{1, \dots, n\}$ numa tal sequência de índices, associamos com qualquer tal índice uma *frequência* α_i . Esta frequência é um elemento do conjunto dos números naturais estendidos

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}.$$

Assim recebemos um **multi-índice**

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}_0^n.$$

A **ordem** de um tal multi-índice é o número

$$|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n \in \mathbb{N}_0.$$

Para $f \in C^{|\alpha|}(U, \mathbb{R}^m)$ a derivada parcial $\partial^\alpha f : U \rightarrow \mathbb{R}^m$ é definida através da fórmula

$$\begin{aligned} \partial^\alpha f &:= \frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}} \\ &:= \frac{\partial^{|\alpha|} f}{\partial x_1 \cdots \partial x_1 \partial x_2 \cdots \partial x_2 \cdots \partial x_n \cdots \partial x_n}. \end{aligned} \tag{2.5.9}$$

Na última expressão aparece o termo ∂x_i exatamente α_i -vezes, onde α_i também pode-se anular. Por definição $\partial^{(0, \dots, 0)} f := f$.

Definição 2.5.8 (Aplicações suaves). *Sejam $m, n \in \mathbb{N}$ e seja $U \subset \mathbb{R}^m$ um conjunto aberto. Uma aplicação $f : U \rightarrow \mathbb{R}^m$ é chamada de **suave** se é continuamente diferenciável quaisquer vezes, ou seja, se ela pertence ao conjunto $C^\ell(U, \mathbb{R}^m)$, e assim é ℓ -vezes continuamente diferenciável, para todo $\ell \in \mathbb{N}$. Uma aplicação suave $f : U \rightarrow \mathbb{R}^m$ também é chamada de **aplicação C^∞** . Denotamos de*

$$C^\infty(U, \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m \mid f \text{ é suave}\} = \bigcap_{\ell \in \mathbb{N}} C^\ell(U, \mathbb{R}^m)$$

o conjunto das aplicações suaves de U para \mathbb{R}^m .

O seguinte teorema resume propriedades importantes de aplicações C^ℓ .

Teorema 2.5.9. *Seja $\ell \in \mathbb{N} \cup \{\infty\}$, sejam X, Y, Z espaços de Banach, e sejam $U \subset X$ e $V \subset Y$ abertos. Então vale o seguinte.*

- (i) *Uma aplicação $f = (f_1, \dots, f_m) : U \rightarrow Y^m$ (com $m \in \mathbb{N}$) é C^ℓ se e somente se $f_i : U \rightarrow Y$ é uma função C^ℓ para todo $i \in \{1, \dots, m\}$.*
- (ii) *Se $f, g \in C^\ell(U, Y)$, então $f + g \in C^\ell(U, Y)$.*
- (iii) *Se $f \in C^\ell(U, Y)$ e $\alpha \in C^\ell(U, \mathbb{R})$, então $\alpha f \in C^\ell(U, Y)$.*
- (iv) *Se $f \in C^\ell(U, Y)$ e $\alpha \in C^\ell(U, \mathbb{R} \setminus \{0\})$, então $f/\alpha := \alpha^{-1}f \in C^\ell(U, Y)$.*
- (v) *Se $f : U \rightarrow V$ e $g : V \rightarrow Z$ são duas aplicações C^ℓ , então a composição $g \circ f : U \rightarrow Z$ também é uma aplicação C^ℓ .*
- (vi) *Se $U \subset X$ é aberto e $f_\nu : U \rightarrow Y$ é uma sequência de funções C^ℓ , tal que a sequência de funções $(d^k f_\nu)_{k \in \mathbb{N}}$ converge uniformemente para toda ordem $k \in \{0, \dots, \ell\}$ em todo compacto $K \subset U$, então $f := \lim_{k \rightarrow \infty} f_\nu : U \rightarrow \mathbb{R}$ é uma função C^ℓ e vale $d^k f = \lim_{k \rightarrow \infty} d^k f_\nu$ para todo $k \in \{0, \dots, \ell\}$.*

Demonstração. Provaremos parte (i) com indução. Para $\ell = 1$ a afirmação segue da parte (ii) do Lema 2.1.15 a qual garante que $df = (df_1, \dots, df_m)$ junto com o fato que df é contínua se e somente se cada uma componente é. Se $\ell \geq 2$ e vale a parte (i) para aplicações $C^{\ell-1}$, então concluímos daí que $df : U \rightarrow Y^m$ é uma aplicação $C^{\ell-1}$ se e somente se $df_1, \dots, df_m : U \rightarrow Y$ são aplicações $C^{\ell-1}$. Com isso a parte (i) segue através de indução sobre ℓ .

Parte (ii) segue da parte (i) do Teorema 2.3.1 através de indução sobre ℓ usando que $d(f + g) = df + dg$. Igualmente funciona a prova da parte (iii) usando a parte (ii), e do Teorema 2.3.1 a regra de Leibniz $d(\alpha f) = \alpha df + f d\alpha$.

Igualmente prova-se (iv) através de indução sobre ℓ usando as partes (ii) e (iii) já provadas e a fórmula $d(f/\alpha) = (\alpha df - f d\alpha)/\alpha^2$ em Teorema 2.3.1.

O argumento de indução para a prova de parte (v) usa a regra da cadeia $d(g \circ f) = dg(f) \circ df$ do Teorema 2.3.2.

Parte (vi) segue através de indução do Teorema 2.2.6. \square

Exercício 2.5.10. Quais fórmulas para produtos e composições resultam das regras de Leibniz e da cadeia para as derivadas parciais superiores?

Exemplo 2.5.11. Todo multi-índice $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ determina uma função suave $f_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}$, definida através de

$$f_\alpha(x) := x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Para as derivadas parciais de f_α com respeito a x_1 recebemos as fórmulas usuais

$$\begin{aligned} \frac{\partial f_\alpha}{\partial x_1}(x) &= \alpha_1 x_1^{\alpha_1-1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \\ \frac{\partial^2 f_\alpha}{\partial x_1^2}(x) &= \alpha_1(\alpha_1 - 1) x_1^{\alpha_1-2} x_2^{\alpha_2} \cdots x_n^{\alpha_n}, \\ &\vdots \\ \frac{\partial^{\beta_1} f_\alpha}{\partial x_1^{\beta_1}}(x) &= \alpha_1(\alpha_1 - 1) \cdots (\alpha_1 - \beta_1 + 1) x_1^{\alpha_1 - \beta_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \\ &= \frac{\alpha_1!}{(\alpha_1 - \beta_1)!} x_1^{\alpha_1 - \beta_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n} \end{aligned}$$

para $0 \leq \beta_1 \leq \alpha_1$. Para as derivadas parciais com respeito às outras variáveis recebe-se obviamente as mesmas fórmulas. Finalmente resulta daí com

$$\alpha! := \alpha_1! \cdot \alpha_2! \cdots \alpha_n!$$

a equação

$$\partial^\beta f_\alpha(x) = \frac{\alpha!}{(\beta - \alpha)!} x^{\alpha - \beta} \quad (2.5.10)$$

para todo multi-índice $\beta = (\beta_1, \dots, \beta_n) \in \mathbb{N}_0^n$ com $\beta_i \leq \alpha_i$ para $i = 1, \dots, n$. Se do outro lado $\beta \in \mathbb{N}_0^n$ é um multi-índice com $\beta_i > \alpha_i$ para um i , então $\partial^\beta f_\alpha \equiv 0$. Particularmente segue daí a fórmula

$$\partial^\beta f_\alpha(0) = \begin{cases} \alpha!, & \text{caso } \beta = \alpha, \\ 0, & \text{caso } \beta \neq \alpha, \end{cases} \quad (2.5.11)$$

para as derivadas parciais de f_α no ponto $x = 0$.

2.6 A série de Taylor

Para multi-índices $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ e vetores $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ temos visto em Comentário 2.5.7 e em Exemplo 2.5.11 como é útil introduzir as abreviações

$$\begin{aligned} |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n \\ \alpha! &:= \alpha_1! \cdot \alpha_2! \cdot \dots \cdot \alpha_n! \\ x^\alpha &:= x_1^{\alpha_1} x_2^{\alpha_2} \cdot \dots \cdot x_n^{\alpha_n}. \end{aligned} \quad (2.6.1)$$

Aqui usamos no caso de $\alpha_i = 0$ as convenções usuais $0! := 1$ e $s^0 := 1$ para qualquer número real s . Ora, seja dada uma aplicação

$$\mathbb{N}_0^n \rightarrow \mathbb{R}^m : \alpha \mapsto a_\alpha$$

a qual associa com todo multi-índice $\alpha \in \mathbb{N}_0^n$ um vetor $a_\alpha \in \mathbb{R}^m$, tal que só um número finito dos a_α são não-nulos. Neste caso a fórmula

$$f(x) := \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha x^\alpha, \quad (2.6.2)$$

para $x \in \mathbb{R}^n$, define uma aplicação suave $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, porque trata-se de uma soma finita e cada um somando é uma aplicação suave. Toda aplicação desta forma chama-se de **polinômio**. A equação (2.5.10) no Exemplo 2.5.11 mostra que as derivadas parciais do polinômio (2.6.2) são dadas através de

$$\partial^\beta f(x) = \sum_{\alpha_i \geq \beta_i} \frac{\alpha!}{(\alpha - \beta)!} a_\alpha x^{\alpha - \beta} \quad (2.6.3)$$

para $x \in \mathbb{R}^n$ e $\beta \in \mathbb{N}_0^n$, onde a soma corre sobre todos os multi-índices $\alpha \in \mathbb{N}_0^n$ com $\alpha_i \geq \beta_i$ para todo $i = 1, \dots, n$. Se valorizamos a equação (2.6.3) na origem $x = 0$ encontraremos o coeficiente

$$a_\beta = \frac{\partial^\beta f(0)}{\beta!} \quad (2.6.4)$$

para todo $\beta \in \mathbb{N}_0^n$. Assim pode-se recuperar os coeficientes do polinômio f das derivadas parciais no ponto $x = 0$. Como a aplicação $x \mapsto f(x_0 + x)$ também é um polinômio, recebemos a equação

$$f(x) = \sum_{\alpha \in \mathbb{N}_0^n} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (2.6.5)$$

para todo polinômio f e todos os $x, x_0 \in \mathbb{R}^n$. Isso lida à seguinte definição.

Definição 2.6.1 (Polinômio de Taylor). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $x_0 \in U$, e seja $f \in C^\ell(U, \mathbb{R}^m)$ com $\ell \in \mathbb{N}_0$. O polinômio de Taylor da ordem ℓ de f no ponto x_0 é o polinômio $T_{x_0}^\ell f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ definido para $x \in \mathbb{R}^n$ através de*

$$(T_{x_0}^\ell f)(x) := \sum_{|\alpha| \leq \ell} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha. \quad (2.6.6)$$

Entende-se a soma em (2.6.6) a todos os multi-índices $\alpha \in \mathbb{N}_0^n$ com $|\alpha| \leq \ell$.

Se f é um polinômio, então a equação (2.6.5) mostra que f coincide com todos seus polinômios de Taylor da ordem correspondente. Geralmente isto não é o caso. Porém o polinômio de Taylor da ordem ℓ de uma função f no ponto x_0 é – num sentido ainda a ser descrito – a *melhor aproximação* de f perto de x_0 . Para isto é útil encontrar uma fórmula adequada para a diferença entre uma função e o seu polinômio de Taylor.

Teorema 2.6.2 (O termo restante no desenvolvimento de Taylor).

Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja $f \in C^{\ell+1}(U, \mathbb{R}^m)$ com $\ell \in \mathbb{N}_0$. Se $x, \xi \in \mathbb{R}^n$ com $x + t\xi \in U$ para $0 \leq t \leq 1$, então vale

$$f(x+\xi) - \sum_{|\alpha| \leq \ell} \frac{\partial^\alpha f(x)}{\alpha!} \xi^\alpha = \int_0^1 (\ell+1)(1-t)^\ell \sum_{|\alpha|=\ell+1} \frac{\partial^\alpha f(x+t\xi)}{\alpha!} \xi^\alpha dt. \quad (2.6.7)$$

Demonstração. Veja a página 83. □

Corolário 2.6.3. *Seja $f \in C^{\ell+1}(U, \mathbb{R}^m)$ como no Teorema 2.6.2 e $x_0 \in U$. Escolha um número $r > 0$ com $\overline{B}_r(x_0) \subset U$, e defina*

$$c_{\ell+1} := \sup_{x \in \overline{B}_r(x_0)} \sum_{|\alpha|=\ell+1} \frac{\|\partial^\alpha f(x)\|}{\alpha!}. \quad (2.6.8)$$

Então todo elemento $x \in \overline{B}_r(x_0)$ satisfaz a desigualdade

$$\|f(x) - (T_{x_0}^\ell f)(x)\| \leq c_{\ell+1} \|x - x_0\|^{\ell+1}. \quad (2.6.9)$$

Demonstração. Segundo o Teorema 2.6.2 e Lema 1.7.3 vale

$$\begin{aligned} \|f(x) - (T_{x_0}^\ell f)(x)\| &\leq \int_0^1 (\ell+1)(1-t)^\ell \sum_{|\alpha|=\ell+1} \frac{\|\partial^\alpha f(x_0+t\xi)\|}{\alpha!} |\xi^\alpha| dt \\ &\leq \int_0^1 (\ell+1)(1-t)^\ell dt c_{\ell+1} \|\xi\|^{\ell+1} = c_{\ell+1} \|\xi\|^{\ell+1}. \end{aligned}$$

Aqui temos usado a desigualdade $|\xi^\alpha| \leq \|\xi\|^{|\alpha|}$. □

Para a estimativa (2.6.9) em Corolário 2.6.3 é de significância que a aplicação f é $(\ell + 1)$ -vezes continuamente diferenciável. Se f só seja ℓ -vezes continuamente diferenciável, então recebe-se uma estimativa mais fraca para a propriedade de aproximação do polinômio de Taylor. Mas esta estimativa ainda é suficiente para caracterizar o polinômio de Taylor unicamente (Comentário 2.6.7).

Corolário 2.6.4. *Sejam $f \in C^\ell(U, \mathbb{R}^m)$ e $x_0 \in U$ como na Definição 2.6.1. Então vale*

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - (T_{x_0}^\ell f)(x)\|}{\|x - x_0\|^\ell} = 0. \quad (2.6.10)$$

Demonstração. Para $\xi \in \mathbb{R}^n$ com

$$x_0 + t\xi \in U \quad \text{para } 0 \leq t \leq 1,$$

e $x := x_0 + \xi$ segue do Teorema 2.6.2 a equação

$$\begin{aligned} f(x) - (T_{x_0}^\ell f)(x) &= f(x) - (T_{x_0}^{\ell-1} f)(x) - \sum_{|\alpha|=\ell} \frac{\partial^\alpha f(x_0)}{\alpha!} \xi^\alpha \\ &= \int_0^1 \ell(1-t)^{\ell-1} \sum_{|\alpha|=\ell} \frac{\partial^\alpha f(x_0 + t\xi) - \partial^\alpha f(x_0)}{\alpha!} \xi^\alpha dt. \end{aligned}$$

Ora, seja dado $\varepsilon > 0$. Como todas as derivadas parciais de f da ordem ℓ são contínuas, existe um $\delta > 0$ com $B_\delta(x_0) \subset U$ e

$$x \in B_\delta(x_0) \quad \implies \quad \sum_{|\alpha|=\ell} \frac{\|\partial^\alpha f(x) - \partial^\alpha f(x_0)\|}{\alpha!} < \varepsilon.$$

Para $x \in B_\delta(x_0)$ e $\xi := x - x_0$ recebemos daí usando Lema 1.7.3 a desigualdade

$$\begin{aligned} \|f(x) - (T_{x_0}^\ell f)(x)\| &\leq \int_0^1 \ell(1-t)^{\ell-1} \sum_{|\alpha|=\ell} \frac{\|\partial^\alpha f(x_0 + t\xi) - \partial^\alpha f(x_0)\|}{\alpha!} |\xi^\alpha| dt \\ &\leq \int_0^1 \ell(1-t)^{\ell-1} dt \varepsilon \|\xi\|^\ell \\ &= \varepsilon \|x - x_0\|^\ell. \end{aligned}$$

Com isso o Corolário 2.6.3 fica provado. \square

Comentário 2.6.5. No caso de $\ell = 1$ temos

$$(T_{x_0}^1 f)(x_0 + \xi) = f(x_0) + \sum_{i=1}^n \partial_i f(x_0) \xi_i = f(x_0) + df(x_0)\xi.$$

Dali pode-se escrever a equação (2.6.10) na forma

$$\lim_{\xi \rightarrow 0} \frac{\|f(x_0 + \xi) - f(x_0) - df(x_0)\xi\|}{\|\xi\|} = 0.$$

Isso é exatamente a definição da diferenciabilidade.

Comentário 2.6.6. Seja $p : \mathbb{R}^n \rightarrow \mathbb{R}^m$ um polinômio da forma

$$p(x) = \sum_{|\alpha| \leq \ell} a_\alpha x^\alpha$$

para $x \in \mathbb{R}^n$. Se p satisfaça a condição

$$\lim_{x \rightarrow 0} \frac{\|p(x)\|}{\|x\|^\ell} = 0, \quad (2.6.11)$$

então todos os coeficientes do polinômio são nulos. **Demonstração:** No caso de $\ell = 0$ o polinômio é constante, e segundo (2.6.11) com $\ell = 0$ aquela constante só pode ser nulo. Ora, seja $\ell \geq 1$ e seja satisfeita a afirmação para $\ell - 1$. Defina

$$q(x) := \sum_{|\alpha| \leq \ell-1} a_\alpha x^\alpha \quad \text{para } x \in \mathbb{R}^n.$$

Então vale $\|q(x) - p(x)\| \leq c \|x\|^\ell$ para todos os $x \in \mathbb{R}^n$ e uma constante adequada $c > 0$. Segue dali $\lim_{x \rightarrow 0} \|x\|^{1-\ell} \|q(x)\| = 0$ e daí $q \equiv 0$, segundo a hipótese da indução. Ou seja, p é um polinômio homogêneo do grau ℓ , e por isso p satisfaz a condição $p(tx) = t^\ell p(x)$ para todos os $x \in \mathbb{R}^n$ e $t \in \mathbb{R}$. Dali segue de (2.6.11) para todo $x \in \mathbb{R}^n \setminus \{0\}$ a equação

$$0 = \lim_{t \rightarrow 0} \frac{\|p(tx)\|}{\|tx\|^\ell} = \lim_{t \rightarrow 0} \frac{|t|^\ell \|p(x)\|}{|t|^\ell \|x\|^\ell} = \frac{\|p(x)\|}{\|x\|^\ell}.$$

Portanto $p \equiv 0$ e daí $a_\alpha = 0$ para todos os α , segundo (2.6.4).

Comentário 2.6.7. Segue do Comentário 2.6.6 que o polinômio de Taylor $T_{x_0}^\ell f$ de uma aplicação $f : U \rightarrow \mathbb{R}^m$ da classe C^ℓ é unicamente determinado através da condição (2.6.10) em Corolário 2.6.4. (Para $\ell = 1$ veja também o Comentário 2.6.5 e o Lema 2.1.3.)

A prova do Teorema 2.6.2 base-se nos seguintes dois lemas. O primeiro é um caso especial da regra da cadeia para derivadas superiores. O segundo é um resultado da Análise I e fornece a forma integral do termo restante para o desenvolvimento de Taylor numa variável real.

Lema 2.6.8. *Sejam $f \in C^\ell(U, \mathbb{R}^m)$ e $x_0 \in U$ como na Definição 2.6.1. Então vale*

$$\left. \frac{d^k}{dt^k} \right|_{t=0} f(x_0 + t\xi) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x_0) \xi^\alpha \quad (2.6.12)$$

para todos os $\xi \in \mathbb{R}^n$ e $k = 0, 1, \dots, \ell$.

Demonstração. Para $k = 0$ ambos os lados da equação são $f(x_0)$ e para $k = 1$ isto é a regra da cadeia

$$\left. \frac{d}{dt} \right|_{t=0} f(x_0 + t\xi) = \sum_{i=1}^n \partial_i f(x_0) \xi_i = df(x_0) \xi.$$

Seja ora satisfeita a equação (2.6.12) para um $k \in \{1, \dots, \ell - 1\}$. Então segue

$$\begin{aligned} \left. \frac{d^{k+1}}{dt^{k+1}} \right|_{t=0} f(x_0 + t\xi) &= \left. \frac{d}{dt} \right|_{t=0} \left. \frac{d^k}{dt^k} f(x_0 + t\xi) \right|_{t=0} \\ &= \left. \frac{d}{dt} \right|_{t=0} \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x_0 + t\xi) \xi^\alpha \\ &= \sum_{|\alpha|=k} \frac{k!}{\alpha!} \sum_{i=1}^n \partial^{\alpha+e_i} f(x_0) \xi^{\alpha+e_i} \\ &= \sum_{|\beta|=k+1} \left(\sum_{\substack{i=1 \\ \beta_i \geq 1}}^n \frac{k!}{(\beta - e_i)!} \right) \partial^\beta f(x_0) \xi^\beta \\ &= \sum_{|\beta|=k+1} \frac{(k+1)!}{\beta!} \partial^\beta f(x_0) \xi^\beta. \end{aligned}$$

Aqui a última igualdade segue da fórmula

$$\frac{|\beta|}{\beta_1! \cdots \beta_n!} = \sum_{\substack{i=1 \\ \beta_i \geq 1}}^n \frac{1}{\beta_1! \cdots \beta_{i-1}! (\beta_i - 1)! \beta_{i+1}! \cdots \beta_n!}.$$

Com isso o Lema 2.6.8 fica provado. \square

Lema 2.6.9. *Seja $u : [0, 1] \rightarrow \mathbb{R}^m$ uma função da classe $C^{\ell+1}$. Então vale*

$$u(1) - \sum_{k=0}^{\ell} \frac{u^{(k)}(0)}{k!} = \int_0^1 \frac{(1-t)^\ell}{\ell!} u^{(\ell+1)}(t) dt. \quad (2.6.13)$$

Demonstração. Para $\ell = 0$ a equação (2.6.13) é o Teorema Fundamental do Cálculo. Se $\ell \geq 1$, e vale a fórmula para $\ell - 1$, então recebemos com

$$\phi(t) := -\frac{(1-t)^\ell}{\ell!}, \quad \psi(t) := u^{(\ell)}(t) - u^{(\ell)}(0)$$

a cadeia de equações

$$\begin{aligned} u(1) - \sum_{k=0}^{\ell} \frac{u^{(k)}(0)}{k!} &= \int_0^1 \frac{(1-t)^{\ell-1}}{(\ell-1)!} u^{(\ell)}(t) dt - \frac{u^{(\ell)}(0)}{\ell!} \\ &= \int_0^1 \frac{(1-t)^{\ell-1}}{(\ell-1)!} (u^{(\ell)}(t) - u^{(\ell)}(0)) dt \\ &= \int_0^1 \phi'(t) \psi(t) dt \\ &= - \int_0^1 \phi(t) \psi'(t) dt \\ &= \int_0^1 \frac{(1-t)^\ell}{\ell!} u^{(\ell+1)}(t) dt. \end{aligned}$$

Com isso o Lema 2.6.9 fica provado. \square

Demonstração Teorema 2.6.2. Seja $u(t) := f(x + t\xi)$ para $0 \leq t \leq 1$. Então

$$u^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x + t\xi) \xi^\alpha$$

segundo o Lema 2.6.8. Segue daí (usando Lema 2.6.9)

$$\begin{aligned} f(x + \xi) - \sum_{|\alpha| \leq \ell} \frac{\partial^\alpha f(x)}{\alpha!} \xi^\alpha &= u(1) - \sum_{k=0}^{\ell} \frac{1}{k!} u^{(k)}(0) \\ &= \int_0^1 \frac{(1-t)^\ell}{\ell!} u^{(\ell+1)}(t) dt \\ &= \int_0^1 (\ell+1)(1-t)^\ell \sum_{|\alpha|=\ell+1} \frac{\partial^\alpha f(x + t\xi)}{\alpha!} \xi^\alpha dt. \end{aligned}$$

Com isso o Teorema 2.6.2 fica provado. \square

Definição 2.6.10 (Série de Taylor).

Sejam $m, n \in \mathbb{N}$, seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $f \in C^\infty(U, \mathbb{R}^m)$, e seja $x_0 \in U$. A **série de Taylor de f no ponto x_0** é a *série formal*

$$(T_{x_0}^\infty f)(x) := \sum_{\alpha \in \mathbb{N}_0^n} \frac{\partial^\alpha f(x_0)}{\alpha!} (x - x_0)^\alpha \quad (2.6.14)$$

para $x \in \mathbb{R}^n$.

Nesta definição tem-se que atentar duas coisas. Primeiro, trata-se no lado direito da equação (2.6.14) de uma expressão puramente formal, e com isso nada é declarado sobre a convergência. De fato pode-se construir já para $n = 1$ funções suaves $f : \mathbb{R} \rightarrow \mathbb{R}$ cujas séries de Taylor tem raio de convergência nulo no ponto $x_0 = 0$, e assim divergem para todo $x \neq 0$ (veja a parte (ii) do Exemplo 2.6.11). Segundo, a expressão no lado direito de (2.6.14) na verdade não é uma “série”, porque o conjunto de índices \mathbb{N}_0^n não é o conjunto dos números naturais, mas uma *família (possivelmente não-somável)* de vetores no \mathbb{R}^m como em [4]. Entretanto o termo “série de Taylor” tem-se entronizado na literatura. Efetivamente recebe-se uma série, se engloba-se primeiramente todos os somandos da ordem k numa soma finita, e então soma-se sobre $k \in \mathbb{N}_0$. O conceito da convergência relevante é sem embargo a somabilidade absoluta em (2.6.14).

A isto agregam-se ora duas questões, a saber primeiro, se pode-se formular condições nas quais a série de Taylor numa vizinhança aberta do ponto x_0 é somável absolutamente e uniformemente, e segundo, caso haja convergência, se a assim definida função $T_{x_0}^\infty f$ coincide numa vizinhança do ponto x_0 com f mesmo. Aqui também existem, já no caso de $n = 1$, contraexemplos importantes (veja a parte (i) do Exemplo 2.6.11). Um critério para uma resposta positiva para ambas questões fornece a estimativa no Corolário 2.6.3 (veja o Corolário 2.6.13).

Exemplo 2.6.11. (i) A função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida através de $f(x) := e^{-1/x^2}$ para $x \neq 0$ e $f(0) := 0$ é suave e tem a série de Taylor $(T_0^\infty f)(x) = 0$.

(ii) A função $f : \mathbb{R} \rightarrow \mathbb{R}$ definida através de $f(x) := \int_0^\infty (1 + x^2 t)^{-1} e^{-t} dt$ é suave e tem a série de Taylor $(T_0^\infty f)(x) = \sum_{k=0}^\infty (-1)^k k! x^k$.

No seguinte teorema sejam $n, m \in \mathbb{N}$ e usamos as notações

$$B_r := \{x \in \mathbb{R}^n \mid \|x\| < r\}, \quad \bar{B}_r := \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$$

para $r > 0$.

Teorema 2.6.12 (Aplicações analíticas).

Seja dada uma aplicação $\mathbb{N}_0^n \rightarrow \mathbb{R}^m : \alpha \mapsto a_\alpha$ com

$$\rho := \frac{1}{\limsup_{k \rightarrow \infty} c_k^{1/k}} > 0, \quad c_k := \sum_{|\alpha|=k} \|a_\alpha\|. \quad (2.6.15)$$

Então vale o seguinte.

(i) A série

$$f(x) := \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha x^\alpha \quad (2.6.16)$$

converge absolutamente para todos os $x \in B_\rho$ e a convergência é uniforme em todo subconjunto compacto de B_ρ .

(ii) A aplicação $f : B_\rho \rightarrow \mathbb{R}^m$ em (i) é suave com as derivadas parciais

$$\partial^\beta f(x) := \sum_{\alpha_i \geq \beta_i} \frac{\alpha!}{(\alpha - \beta)!} a_\alpha x^{\alpha - \beta} \quad (2.6.17)$$

para $x \in B_\rho$ e $\beta \in \mathbb{N}_0^n$. A série (2.6.17) também converge absolutamente para todo $\beta \in \mathbb{N}_0^n$ e todo $x \in B_\rho$, e para todo β fixado a convergência é uniforme em todo subconjunto compacto de B_ρ .

(iii) A série de Taylor de f no ponto $x_0 = 0$ converge absolutamente e uniformemente em todo subconjunto compacto de B_ρ e coincide com a função f , ou seja $(T_0^\infty f)(x) = f(x)$ para todos os $x \in B_\rho$.

Demonstração. A demonstração tem quatro passos.

Passo 1. Seja $p \in \mathbb{N}_0$ e $0 < r < \rho$. Então vale

$$\sum_{k=p}^{\infty} k^{np} c_k r^{k-p} < \infty. \quad (2.6.18)$$

Como $0 < r < \rho$, vale $r \limsup_{k \rightarrow \infty} c_k^{1/k} < 1$. Como a sequência $k^{np/k} r^{-p/k}$ converge a 1 para $k \rightarrow \infty$, segue daí $\limsup_{k \rightarrow \infty} k^{np/k} c_k^{1/k} r^{1-p/k} < 1$. Ora escolhamos um número real λ com

$$\limsup_{k \rightarrow \infty} k^{np/k} c_k^{1/k} r^{1-p/k} < \lambda < 1.$$

Então existe $k_0 \in \mathbb{N}$, tal que todo $k \in \mathbb{N}$ com $k \geq k_0$ satisfaz a desigualdade

$$k^{np/k} c_k^{1/k} r^{1-p/k} \leq \lambda.$$

Isto da $k^{np} c_k r^{k-p} \leq \lambda^k$ para todos os $k \in \mathbb{N}$ com $k \geq k_0$ e daí segue Passo 1 de teste da comparação para séries de números reais.

Passo 2. *Provaremos parte (i).*

Para $k \in \mathbb{N}_0$ seja $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ o polinômio homogêneo

$$f_k(x) := \sum_{|\alpha|=k} a_\alpha x^\alpha \quad (2.6.19)$$

da ordem k . Seja $r > 0$. Então vale

$$\|f_k(x)\| \leq \sum_{|\alpha|=k} \|a_\alpha\| |x^\alpha| \leq \sum_{|\alpha|=k} \|a_\alpha\| \|x\|^k \leq c_k r^k$$

para todo $x \in \overline{B}_r$ e todo $k \in \mathbb{N}_0$. Se $0 < r < \rho$, então vale $\sum_{k=0}^{\infty} c_k r^k < \infty$ segundo o Passo 1, e daí segue, segundo o Teorema 1.3.10, a convergência absoluta e uniforme da série $f = \sum_{k=1}^{\infty} f_k$ em \overline{B}_r . Com isso o Passo 2 fica provado.

Passo 3. *Provaremos parte (ii).*

Para $k \in \mathbb{N}_0$ seja $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ o polinômio homogêneo em (2.6.19). Seja $\beta \in \mathbb{N}_0^n$ e $p := |\beta|$. Então $\partial^\beta f_k \equiv 0$ para $k = 0, 1, \dots, p-1$, e para $k \geq p$ vale

$$\partial^\beta f_k(x) := \sum_{\substack{|\alpha|=k \\ \alpha_i \geq \beta_i}} \frac{\alpha!}{(\beta - \alpha)!} a_\alpha x^{\alpha - \beta} \quad (2.6.20)$$

para todo $x \in \mathbb{R}^n$, segundo (2.6.3). Como $\frac{\alpha!}{(\beta - \alpha)!} \leq k^{np}$ para todos os $\alpha \in \mathbb{N}_0^n$ com $|\alpha| = k$ e $\alpha_i \geq \beta_i$ para $i = 1, \dots, n$, segue daí a desigualdade

$$\sup_{x \in \overline{B}_r} \|\partial^\beta f_k(x)\| \leq k^{np} \sup_{x \in \overline{B}_r} \sum_{\substack{|\alpha|=k \\ \alpha_i \geq \beta_i}} \|a_\alpha\| \|x\|^{|\alpha - \beta|} \leq k^{np} c_k r^{k-p}$$

para todos os $k \in \mathbb{N}_0$ com $k \geq p$. Dali segue, segundo o Passo 1 e o Teorema 1.3.10, a convergência absoluta e uniforme da série $g_\beta := \sum_{k=p}^{\infty} \partial^\beta f_k$ em \overline{B}_r para todos os $0 < r < \rho$. Segundo o Teorema 2.2.6 segue daí, através de indução, que $f : B_\rho \rightarrow \mathbb{R}^m$ é uma aplicação suave com as derivadas parciais $\partial^\beta f = g_\beta$. Com isso o Passo 3 fica provado.

Passo 4. *Provaremos parte (iii).*

A equação (2.6.17) na parte (ii) nós dá a fórmula $a_\alpha = \partial^\alpha f(0)/\alpha!$ para todos os $\alpha \in \mathbb{N}_0^n$. Dali segue Passo 4 e com isso o Teorema 2.6.12 fica provado. \square

Corolário 2.6.13 (Convergência da série de Taylor).

Sejam $m, n \in \mathbb{N}$, seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $f \in C^\infty(U, \mathbb{R}^m)$, e sejam $x_0 \in U$ e $r > 0$ com $\overline{B}_r(x_0) \subset U$ e

$$r < \rho := \frac{1}{\limsup_{k \rightarrow \infty} c_k^{1/k}}, \quad c_k := \sup_{x \in \overline{B}_r(x_0)} \sum_{|\alpha|=k} \frac{\|\partial^\alpha f(x)\|}{\alpha!}. \quad (2.6.21)$$

Então a série de Taylor (2.6.14) converge, junto com todas suas derivadas parciais, absolutamente e uniformemente em $\overline{B}_r(x_0)$ e vale $f(x) = (T_{x_0}^\infty f)(x)$ para todos os $x \in \overline{B}_r(x_0)$.

Demonstração. Segundo o Corolário 2.6.3 vale a desigualdade

$$\sup_{x \in \overline{B}_r(x_0)} \|f(x) - (T_{x_0}^k f)(x)\| \leq c_{k+1} r^{k+1}$$

para todos os $k \in \mathbb{N}$. Como $\sum_{k=0}^{\infty} c_k r^k < \infty$ segundo (2.6.21) e Passo 1 na demonstração do Teorema 2.6.12, temos $\lim_{k \rightarrow \infty} c_{k+1} r^{k+1} = 0$, e segue daí

$$f(x) = \lim_{k \rightarrow \infty} (T_{x_0}^k f)(x) = (T_{x_0}^\infty f)(x)$$

para todos os $x \in \overline{B}_r(x_0)$ e a convergência é uniforme. A convergência absoluta e uniforme dos polinômios de Taylor $T_{x_0}^k f$ e de todas suas derivadas parciais na bola $\overline{B}_r(x_0)$ segue do Teorema 2.6.12 com $a_\alpha := \partial^\alpha f(x_0)/\alpha!$. Com isso o Corolário 2.6.13 fica provado. \square

Definição 2.6.14 (Aplicações analíticas).

Sejam $m, n \in \mathbb{N}$ e seja $U \subset \mathbb{R}^n$ um subconjunto aberto. Uma aplicação suave $f \in C^\infty(U, \mathbb{R}^m)$ chama-se de **analítica (real)** se para todo $x_0 \in U$ existe um número real $r > 0$ com $\overline{B}_r(x_0) \subset U$ o qual satisfaz (2.6.21).

Segundo o Corolário 2.6.13 as séries de Taylor de uma aplicação analítica f , junto com todas as derivadas parciais, convergem absolutamente e uniformemente em bolas suficientemente pequenas para f , respectivamente as derivadas parciais de f . Vice-versa toda função como no Teorema 2.6.12 é analítica real. O estudo das funções analíticas reais é uma subárea importante da matemática o qual entranha profundamente em questões da geometria, álgebra, e análise. Este assunto nem pode, nem deve, ser mais profundado aqui. Em vez disso consideramos neste ponto só dois exemplos elementares.

Exemplo 2.6.15. No caso de $m = 1$ chama-se o número ρ in (2.6.15) o **raio de convergência** da série f em (2.6.16). No contraste ao caso $n = 1$ possa acontecer plenamente que a série de Taylor f em (2.6.16) converge numa área muito maior como só B_ρ . Como exemplo consideramos a função

$$f(x) := \frac{1}{1 - x_1 - x_2 - \cdots - x_n} \quad (2.6.22)$$

no conjunto aberto

$$U := \{x \in \mathbb{R}^n \mid |x_1 + x_2 + \cdots + x_n| < 1\}.$$

Esta função pode ser representada como a série geométrica

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} (x_1 + x_2 + \cdots + x_n)^k \\ &= \sum_{k=0}^{\infty} \sum_{|\alpha|=k} \frac{k!}{\alpha!} x^\alpha \\ &= \sum_{\alpha \in \mathbb{N}_0^n} \frac{|\alpha|!}{\alpha!} x^\alpha. \end{aligned} \quad (2.6.23)$$

Aqui a série na primeira fórmula no lado direito converge absolutamente para todos os x na região não limitada U , enquanto o último termo representa uma família somável só para $x \in \mathbb{R}^n$ com $\|x\| < 1$. Este último termo também é a série de Taylor de f no ponto $x_0 = 0$ e o raio da convergência é $\rho = 1$ (Exercício).

Exemplo 2.6.16. Seja $U := \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ e seja $f : U \rightarrow \mathbb{R}$ a função

$$f(x, y) := x^y = e^{y \log(x)}. \quad (2.6.24)$$

Os polinômios de Taylor da ordem 2 e 3 de f nos pontos $x_0 = (1, 0)$ e $x_0 = (1, 1)$ são dados pelas fórmulas

$$\begin{aligned} (T_{(1,0)}^2 f)(x, y) &= 1 + (x - 1)y \\ (T_{(1,0)}^3 f)(x, y) &= 1 + (x - 1)y - \frac{1}{2}(x - 1)^2 y \\ (T_{(1,1)}^2 f)(x, y) &= 1 + (x - 1) + (x - 1)(y - 1) \\ (T_{(1,1)}^3 f)(x, y) &= 1 + (x - 1) + (x - 1)(y - 1) + \frac{1}{2}(x - 1)^2 (y - 1). \end{aligned}$$

(Demonstração: Exercício.)

2.7 Extremos locais

Nesta seção denotamos de $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$ sempre a norma euclidiana de um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, e para $x \in \mathbb{R}^n$ e $r > 0$ denotamos de $B_r(x) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$ a bola aberta no \mathbb{R}^n com respeito à norma euclidiana.

Definição 2.7.1 (Extremo local). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja $f : U \rightarrow \mathbb{R}$ uma função contínua. Um elemento $x_0 \in U$ chama-se de*

- **mínimo local de f** , se existe um $\delta > 0$ com

$$x \in B_\delta(x_0) \quad \implies \quad x \in U \text{ e } f(x) \geq f(x_0)$$

- **mínimo local estrito de f** , se existe um $\delta > 0$ com

$$x \in B_\delta(x_0) \setminus \{x_0\} \quad \implies \quad x \in U \text{ e } f(x) > f(x_0)$$

- **máximo local de f** , se existe um $\delta > 0$ com

$$x \in B_\delta(x_0) \quad \implies \quad x \in U \text{ e } f(x) \leq f(x_0)$$

- **máximo local estrito de f** , se existe um $\delta > 0$ com

$$x \in B_\delta(x_0) \setminus \{x_0\} \quad \implies \quad x \in U \text{ e } f(x) < f(x_0).$$

Um elemento $x_0 \in U$ chama-se de **extremo local de f** , se x_0 é um mínimo ou máximo local de f .

Teorema 2.7.2. *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $f : U \rightarrow \mathbb{R}$ uma função contínua, e seja $x_0 \in U$ um extremo local de f . Se f é diferenciável no ponto x_0 , então vale*

$$df(x_0) = 0.$$

Demonstração. Seja x_0 mínimo local de f , seja $\delta > 0$ como na Definição 2.7.1, seja $\xi \in \mathbb{R}^n \setminus \{0\}$, e defina a função $\phi : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ através de $\varepsilon := \delta / \|\xi\|$ e $\phi(t) := f(x_0 + t\xi)$ para $-\varepsilon < t < \varepsilon$. Então ϕ é contínuo, possui um mínimo local no ponto $t = 0$, e é diferenciável no ponto $t = 0$ segundo o Teorema 2.3.2. Então $\phi'(0) = 0$, segundo um teorema da Análise I.

Eis aqui mais uma vez o argumento da Análise I. Para $t \geq 0$ recebe-se

$$\frac{\phi(-t) - \phi(0)}{-t} \leq 0 \leq \frac{\phi(t) - \phi(0)}{t}$$

e segue daí

$$\phi'(0) = \lim_{t \searrow 0} \frac{\phi(-t) - \phi(0)}{-t} \leq 0 \leq \lim_{t \searrow 0} \frac{\phi(t) - \phi(0)}{t} = \phi'(0).$$

Por isso, segundo a regra da cadeia, vale $df(x)\xi = \phi'(0) = 0$. No caso de um máximo local argumenta-se igualmente ou rebe-se o resultado no que substitue-se f por $-f$. Com isso o Teorema 2.7.2 fica provado. \square

Definição 2.7.3 (Ponto crítico). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja $f : U \rightarrow \mathbb{R}$ uma função continuamente diferenciável. Chama-se um elemento $x_0 \in U$ de **ponto crítico de f** se $df(x_0) = 0$.*

Com esta terminologia Teorema 2.7.2 diz que todo extremo local de uma função continuamente diferenciável é um ponto crítico. Como já sabemos do curso Análise I, não todo ponto crítico é um extremo local. Por exemplo, o ponto $x_0 = 0$ é um ponto crítico da função $f(x) = x^3$ de uma variável real, mas $x_0 = 0$ não é um extremo local. Aqui são uns exemplos em dimensão dois.

Exemplo 2.7.4. A função contínua $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida através da fórmula $f(x, y) := \sqrt{x^2 + y^2}$ para $(x, y) \in \mathbb{R}^2$ possui na origem $(x_0, y_0) = (0, 0)$ um mínimo local, mas exatamente neste ponto não é diferenciável. Conseqüentemente não pode-se aplicar Teorema 2.7.2 a este exemplo.

Exemplo 2.7.5. A função

$$f(x, y) := x^2 + y^2$$

possui um mínimo local na origem e isto também é o único ponto no qual a derivada de f se anula. A função $f(x, y) := -x^2 - y^2$ possui um máximo local na origem e igualmente isto é o único ponto crítico de f . A função

$$f(x, y) = xy$$

também possui seu único ponto crítico na origem, o qual porém não é um extremo local neste exemplo (veja a Figura 2.2). Isto é um exemplo de um assim chamado *ponto de sela*.

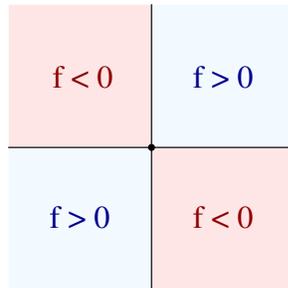


Figura 2.2: O conjunto de nível crítico de um ponto de sela.

Exemplo 2.7.6. A função

$$f(x, y) = xy(x + y)$$

também possui seu único ponto crítico na origem e isto não é um extremo local. A função $f_\lambda(x, y) = (xy - \lambda)(x + y)$ possui para $\lambda > 0$ exatamente dois pontos críticos nos pontos $\pm(\sqrt{\lambda}, \sqrt{\lambda})$ e também para $\lambda < 0$ exatamente dois pontos críticos nos pontos $\pm(\sqrt{|\lambda|}, -\sqrt{|\lambda|})$. Para $\lambda \neq 0$ isto são pontos de sela. Para $\lambda = 0$ trata-se de um assim chamado ponto crítico *degenerado* (veja a Figura 2.3).

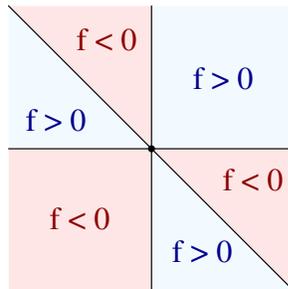


Figura 2.3: Um ponto crítico degenerado.

Exemplo 2.7.7. A função

$$f_\lambda(x, y) := x^3 - 3\lambda x + y^2$$

possui para $\lambda < 0$ nenhum ponto crítico, para $\lambda = 0$ exatamente um ponto crítico no ponto $(0, 0)$, e para $\lambda > 0$ exatamente dois pontos críticos nos pontos $(-\sqrt{\lambda}, 0)$ (ponto de sela) e $(\sqrt{\lambda}, 0)$ (mínimo local).

Exemplo 2.7.8. A função $f(x, y) := \sin(2\pi x) + \sin(2\pi y)$ no \mathbb{R}^2 possui um número infinito de pontos críticos. Considerando f no entanto como função no toro $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$, no que identifica-se dois elementos (x, y) e (x', y') no \mathbb{R}^2 um com o outro se a sua diferença $(x' - x, y' - y) \in \mathbb{Z}^2$ é um vetor com coeficientes nos números inteiros, então recebe-se uma função com exatamente quatro pontos críticos (um mínimo local, um máximo local, e dois pontos de sela).

Para um entendimento melhor destes e de muitos outros exemplos é útil considerar a *matriz hessiana* num ponto crítico.

Definição 2.7.9 (Matriz hessiana). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja $f : U \rightarrow \mathbb{R}$ uma função da classe C^2 . Para $x \in U$ chama-se a matriz*

$$d^2f(x) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (2.7.1)$$

das segundas derivadas parciais a **matriz hessiana de f no ponto x** . Segundo o Teorema 2.5.3 a matriz hessiana é simétrica.

Exemplo 2.7.10. A matriz hessiana da função $f(x, y) = x^2 + y^2$ na origem (um mínimo local) é a matriz positiva definida

$$d^2f(x) := \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Para $f(x, y) = -x^2 - y^2$ recebe-se na origem (um máximo local) uma matriz hessiana negativa definida, e para $f(x, y) = xy$ recebe-se na origem (um ponto de sela) a matriz hessiana indefinida

$$d^2f(x) := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Isto não é um acidente, como vamos ver no seguinte teorema. A matriz hessiana da função $f(x, y) = xy(x + y)$ na origem (num ponto crítico degenerado) é a matriz nula.

Teorema 2.7.11 (Critérios para extremos locais). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $f : U \rightarrow \mathbb{R}$ uma função duas vezes continuamente diferenciável, e seja $x_0 \in U$ um ponto crítico de f . Então vale o seguinte.*

(i) *Se x_0 é um mínimo local, então a matriz hessiana $d^2f(x_0)$ é positiva semi-definida, ou seja, para todos os $\xi \in \mathbb{R}^n$ vale $\xi^T d^2f(x_0)\xi \geq 0$.*

(ii) *Se a matriz hessiana $d^2f(x_0)$ é positiva definida, ou seja, para todos os $\xi \in \mathbb{R}^n$ com $\xi \neq 0$ vale $\xi^T d^2f(x_0)\xi > 0$, então x_0 é um mínimo local estrito.*

(iii) *Se x_0 é um máximo local, então a matriz hessiana $d^2f(x_0)$ é negativa semi-definida, ou seja, para todos os $\xi \in \mathbb{R}^n$ vale $\xi^T d^2f(x_0)\xi \leq 0$.*

(iv) *Se a matriz hessiana $d^2f(x_0)$ é negativa definida, ou seja, para todos os $\xi \in \mathbb{R}^n$ com $\xi \neq 0$ vale $\xi^T d^2f(x_0)\xi < 0$, então x_0 é um máximo local estrito.*

Demonstração. Provaremos a parte (i). Seja dado $\xi := (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$. Então existe um número $\varepsilon > 0$, tal que $x_0 + t\xi \in U$ para todo $t \in \mathbb{R}$ com $|t| < \varepsilon$. Definimos a função $\phi : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ através de

$$\phi(t) := f(x_0 + t\xi) \quad \text{para } -\varepsilon < t < \varepsilon.$$

Segundo o Teorema 2.3.2 a função é duas vezes continuamente diferenciável com

$$\begin{aligned} \phi'(t) &= \sum_{j=1}^n \partial_j f(x_0 + t\xi) \xi_j, \\ \phi''(t) &= \sum_{j=1}^n \sum_{i=1}^n \partial_i \partial_j f(x_0 + t\xi) \xi_i \xi_j = \xi^T d^2f(x_0 + t\xi) \xi. \end{aligned} \tag{2.7.2}$$

Como ϕ possui no ponto $t = 0$ um mínimo local, vale $\phi''(0) \geq 0$, segundo um teorema da Análise I. Eis aqui mais uma vez o argumento. Seja $0 < \delta < \varepsilon$ escolhido tal que $\phi(0) \leq \phi(t)$ para $-\delta \leq t \leq \delta$. Então, segundo o Teorema do Valor Médio, existe para todo $k \in \mathbb{N}$ um número $0 < t_k \leq \delta/k$ com

$$\phi'(t_k) = \frac{\phi(\delta/k) - \phi(0)}{\delta/k} \geq 0.$$

(Pode-se escolher o maior tal t_k no intervalo $0 < t \leq \delta/k$, caso quer-se evitar o axioma da escolha, versão contável.) Como $\phi'(0) = 0$, segue daí usando (2.7.2) a desigualdade

$$\xi^T d^2f(x_0)\xi = \phi''(0) = \lim_{t \searrow 0} \frac{\phi'(t)}{t} = \lim_{k \rightarrow \infty} \frac{\phi'(t_k)}{t_k} \geq 0.$$

Com isso a parte (i) fica provada. Observe-se que este argumento não utiliza a continuidade das segundas derivadas parciais de f .

Provaremos a parte (ii) em dois passos.

Passo 1. *Seja $A = A^T \in \mathbb{R}^{n \times n}$ uma matriz simétrica positiva definida, ou seja, para todo $\xi \in \mathbb{R}^n \setminus \{0\}$ vale $\xi^T A \xi > 0$. Então existe uma constante $\varepsilon > 0$, tal que todo vetor $\xi \in \mathbb{R}^n$ satisfaz a desigualdade $\xi^T A \xi \geq \varepsilon \|\xi\|^2$.*

A esfera unitária $\mathbb{S}^{n-1} := \{\xi \in \mathbb{R}^n \mid \|\xi\| = 1\}$ é um subconjunto compacto do \mathbb{R}^n segundo o Teorema de Heine-Borel. Como a função $g : \mathbb{S}^{n-1} \rightarrow (0, \infty)$ definida através de $g(\xi) := \xi^T A \xi$ é contínua, existe, segundo um teorema da Análise I, um elemento $\xi_0 \in \mathbb{R}^n$ com $g(\xi_0) \leq g(\xi)$ para todos os $\xi \in \mathbb{S}^{n-1}$. Segue daí

$$\|\xi\|^{-2} \xi^T A \xi = (\|\xi\|^{-1} \xi)^T A (\|\xi\|^{-1} \xi) \geq \xi_0^T A \xi_0 =: \varepsilon > 0$$

para todos os $\xi \in \mathbb{R}^n \setminus \{0\}$ e com isso o Passo 1 fica provado.

Passo 2. *Provaremos a parte (ii).*

Como a matriz hessiana de f no ponto x_0 é positiva definida, existe, segundo o Passo 1, uma constante $\varepsilon > 0$ com

$$\xi^T d^2f(x_0)\xi \geq \varepsilon \|\xi\|^2 \quad \text{para todos os } \xi \in \mathbb{R}^2. \quad (2.7.3)$$

Segundo o Corolário 2.6.4 existe ora um número $\delta > 0$ com $B_\delta(x_0) \subset U$, tal que todo vetor $\xi \in \mathbb{R}^n$ com $0 < \|\xi\| < \delta$ satisfaz a desigualdade

$$\frac{|f(x_0 + \xi) - (T_{x_0}^2 f)(x_0 + \xi)|}{\|\xi\|^2} < \frac{\varepsilon}{4}. \quad (2.7.4)$$

Com $(T_{x_0}^2 f)(x_0 + \xi) = f(x_0) + df(x_0)\xi + \frac{1}{2}\xi^T d^2f(x_0)\xi$ e $df(x_0) = 0$ segue daí que para todo vetor $\xi \in \mathbb{R}^n$ com $0 < \|\xi\| < \delta$ vale a desigualdade

$$\begin{aligned} f(x_0 + \xi) - f(x_0) &= \frac{1}{2}\xi^T d^2f(x_0)\xi + f(x_0 + \xi) - (T_{x_0}^2 f)(x_0 + \xi) \\ &\geq \frac{1}{2}\xi^T d^2f(x_0)\xi - |f(x_0 + \xi) - (T_{x_0}^2 f)(x_0 + \xi)| \\ &\geq \frac{\varepsilon}{4} \|\xi\|^2. \end{aligned}$$

No último passo temos utilizado as desigualdades (2.7.3) e (2.7.4). Com isso a parte (ii) fica provada.

As afirmações em (iii) e (iv) seguem diretamente de (i) e (ii) no que substitue-se f por $-f$. Com isso o Teorema 2.7.11 fica provado. \square

Exemplo 2.7.12. A função dado por $f(x, y) := 3x^4 + y^2$ no \mathbb{R}^2 possui no ponto $(x_0, y_0) = (0, 0)$ um mínimo local estrito, no entanto a sua matriz hessiana não é positiva definida neste ponto.

Exemplo 2.7.13. A função

$$f(x, y) := 3x^4 + y^2 - 4x^2y = (y - x^2)(y - 3x^2)$$

possui no ponto $(x_0, y_0) = (0, 0)$ um ponto crítico com a seguinte propriedade. Restrita-se a função a uma reta qualquer passando a origem, então possui a função restrita um mínimo local estrito na origem. Mais precisamente, definimos para um vetor $(x, y) \in \mathbb{R}^2 \setminus \{0, 0\}$ arbitrário a função $\phi_{x,y} : \mathbb{R} \rightarrow \mathbb{R}$ através de

$$\phi_{x,y}(t) := f(tx, ty) = 3x^4t^4 - 4x^2yt^3 + t^2y^2$$

para $t \in \mathbb{R}$, então $t_0 = 0$ é um mínimo local estrito de $\phi_{x,y}$ (tanto no caso de $y = 0$ como na caso $y \neq 0$). Embora a origem não é um mínimo local de f , porque vale $f(\varepsilon, 2\varepsilon^2) < 0 < f(\varepsilon, 0)$ para todo $\varepsilon > 0$.

Exercício 2.7.14. Seja $f : U \rightarrow \mathbb{R}$ uma função duas vezes continuamente diferenciável num subconjunto aberto $U \subset \mathbb{R}^n$ e seja $x_0 \in U$ um ponto crítico de f . Se existam dois vetores $\xi, \eta \in \mathbb{R}^n$ com $\xi^T d^2f(x_0)\xi < 0 < \eta^T d^2f(x_0)\eta$, então x_0 não é um extremo local.

Exercício 2.7.15. Seja $f : U \rightarrow \mathbb{R}$ continuamente diferenciável e seja o gradiente $\nabla f : U \rightarrow \mathbb{R}^n$ diferenciável no ponto $x_0 \in U$. Então vale

$$\lim_{\xi \rightarrow 0} \frac{|f(x_0 + \xi) - f(x_0) - df(x_0)\xi - \frac{1}{2}\xi^T d^2f(x_0)\xi|}{\|\xi\|^2} = 0. \quad (2.7.5)$$

Segue daí que todas as afirmações do Teorema 2.7.11 ficam válidas nestas hipóteses mais fracas. (Confira (2.7.5) com (2.6.10).)

Exercício 2.7.16. Seja $A = A^T \in \mathbb{R}^{n \times n}$ uma matriz positiva semi-definida.

(i) Para todos os $\xi \in \mathbb{R}^n$ vale $\langle \xi, A\xi \rangle = 0 \iff A\xi = 0$.

(ii) A matriz A é positiva definida se e somente se $\det(A) \neq 0$.

Dica: Se $\langle \xi, A\xi \rangle = 0$, então a função $\mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \langle x, Ax \rangle$ possui um mínimo local no ponto $x = \xi$. Para a parte (ii) veja o Teorema D.3.7.

Exercício 2.7.17. Seja $U \subset \mathbb{R}^n$ um conjunto aberto e seja $f : U \rightarrow \mathbb{R}$ uma função da classe C^2 . Um ponto crítico $x_0 \in U$ de f chama-se de **não-degenerado**, se a matriz hessiana $d^2f(x_0)$ possui um determinante não-nulo. Seja $x_0 \in U$ um mínimo local de f e um ponto crítico não-degenerado de f . Então x_0 é um mínimo local estrito de f .

Funções convexas

Primeiramente lembramos que chama-se um subconjunto $U \subset \mathbb{R}^n$ de **convexo**, se U também contava, junto com quaisquer dois pontos $x_0, x_1 \in U$, o segmento $[x_0, x_1]$ entre eles, ou seja, o conjunto de todas as *combinações convexas* $(1-t)x_0 + tx_1$ para $0 \leq t \leq 1$. Usamos no seguinte o conceito da Lipschitz-continuidade local (Definição 4.1.1) o que será discutido com mais detalhe só no capítulo 4 (veja o Lema 4.1.2).

Definição 2.7.18 (Funções convexas). *Seja $U \subset \mathbb{R}^n$ um subconjunto convexo. Uma função $f : U \rightarrow \mathbb{R}$ chama-se de **convexo**, se ela satisfaça para todos os $x_0, x_1 \in U$ e todos os $0 \leq t \leq 1$ a desigualdade*

$$f((1-t)x_0 + tx_1) \leq (1-t)f(x_0) + tf(x_1).$$

*Chama-se f de **estritamente convexo**, se f satisfaça a desigualdade*

$$f((1-t)x_0 + tx_1) < (1-t)f(x_0) + tf(x_1)$$

para todos os $x_0, x_1 \in U$ com $x_0 \neq x_1$ e todos os $0 < t < 1$.

Teorema 2.7.19. *Seja $U \subset \mathbb{R}^n$ um subconjunto convexo e seja $f : U \rightarrow \mathbb{R}$ uma função convexa. Então vale o seguinte.*

- (i) *f é localmente Lipschitz-contínuo.*
- (ii) *Se f é diferenciável no ponto $x_0 \in U$ com $df(x_0) = 0$, então vale*

$$f(x_0) \leq f(x) \quad \text{para todos os } x \in U.$$

- (iii) *O conjunto $\{x \in U \mid f(x) = \inf_U f\}$ é convexo. (Pode ser vazio também.)*

Demonstração. Na prova da parte (i) utilizamos a norma $\|x\|_\infty := \max_i |x_i|$ assim como a norma euclidiana $\|x\| := \sqrt{\sum_i x_i^2}$ para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, e denotamos a base padrão do \mathbb{R}^n com e_1, \dots, e_n . Ora, seja $x_0 \in U$ e escolha $r > 0$ com

$$Q := \{x \in \mathbb{R}^n \mid \|x - x_0\|_\infty \leq 3r\} \subset U.$$

Provaremos em três passos que a restrição de f à bola fechada

$$B := \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\} \subset Q \subset U$$

é Lipschitz-contínuo.

Passo 1. Vale $M := \sup_{x \in Q} f(x) < \infty$.

Seja $\mathcal{E} \subset \mathbb{R}^n$ o conjunto finito de índices

$$\mathcal{E} := \{\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \mathbb{R}^n \mid \varepsilon_i \in \{-1, +1\} \text{ para } i = 1, \dots, n\}$$

e para $\varepsilon \in \mathcal{E}$ seja

$$v_\varepsilon := x_0 + 3r \sum_{i=1}^n \varepsilon_i e_i \in Q.$$

Isto são os *pontos extremais* do cubo Q e o máximo de f sobre estes pontos é o número $M := \max\{f(v_\varepsilon) \mid \varepsilon \in \mathcal{E}\}$. Ora, seja $x \in Q$. Então existe uma aplicação $\lambda : \mathcal{E} \rightarrow [0, 1]$ com $x = \sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) v_\varepsilon$ e $\sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) = 1$. A existência de uma tal aplicação pode ser mostrada através de indução sobre a dimensão n . Como f é convexo, vale

$$f(x) \leq \sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) f(v_\varepsilon) \leq \sum_{\varepsilon \in \mathcal{E}} \lambda(\varepsilon) M = M$$

e com isso o Passo 1 fica provado.

Passo 2. Para todos os $x \in Q$ vale $f(x) \geq m := 2f(x_0) - M$.

Se $x \in Q$, então também $2x_0 - x \in Q$. Como $x_0 = \frac{1}{2}x + \frac{1}{2}(2x_0 - x)$, segue da convexidade de f a desigualdade $2f(x_0) \leq f(x) + f(2x_0 - x) \leq f(x) + M$ e daí $f(x) \geq 2f(x_0) - M = m$. Com isso o Passo 2 fica provado.

Passo 3. Para todos os $x, y \in B$ vale $|f(x) - f(y)| \leq \frac{M-m}{2r} \|x - y\|$.

Sejam $x, y \in B$ com $x \neq y$. Então vale $0 < \|y - x\| \leq 2r$. Seja

$$t := \frac{\|y - x\|}{2r}, \quad \xi := \frac{y - x}{t}.$$

Então vale $\|\xi\| = 2r$, então $\|x + \xi - x_0\| \leq 3r$ e segue daí $x + \xi \in Q$. Além disso vale $y = x + t\xi = (1 - t)x + t(x + \xi)$ e $0 \leq t \leq 1$ e dali

$$\begin{aligned} f(y) &\leq (1 - t)f(x) + tf(x + \xi) \\ &= f(x) + \frac{f(x + \xi) - f(x)}{2r} \|y - x\| \\ &\leq f(x) + \frac{M - m}{2r} \|y - x\|. \end{aligned}$$

A desigualdade $f(x) \leq f(y) + \frac{M-m}{2r} \|y - x\|$ segue através de trocar os papéis de x e y . Com isso o Passo 3 e a parte (i) são provados.

Provaremos parte (ii). Seja f diferenciável no ponto x_0 . Então vale

$$f(x) \geq f(x_0) + df(x_0)(x - x_0) \quad (2.7.6)$$

para todos os $x \in U$. Para a demonstração desta desigualdade fixamos um elemento $x \in U$ e definimos $\xi := x_0 - x$. Então vale $x_0 + t\xi \in U$ para todo $t > 0$ suficientemente pequeno. Com $x_0 - \xi = x$ e

$$x_0 = \frac{t}{1+t}(x_0 - \xi) + \frac{1}{1+t}(x_0 + t\xi)$$

resulta daí a desigualdade

$$f(x_0) \leq \frac{t}{1+t}f(x) + \frac{1}{1+t}f(x_0 + t\xi).$$

Dali segue

$$f(x_0) + tf(x_0) \leq tf(x) + f(x_0 + t\xi)$$

e assim

$$f(x_0) - f(x) \leq \frac{f(x_0 + t\xi) - f(x_0)}{t}.$$

Com a transição ao limite $t \rightarrow 0$ segue a desigualdade afirmada (2.7.6). Se $df(x_0) = 0$, então $f(x_0) \leq f(x)$ para todos os $x \in U$ segundo (2.7.6). Com isso a parte (ii) fica provada. Parte (iii) segue diretamente da definição de convexidade e com isso o Teorema 2.7.19 fica provado. \square

Exemplo 2.7.20. A função $f : [0, 1] \rightarrow \mathbb{R}$ definida através de $f(x) := 0$ para $0 < x < 1$ e $f(0) := f(1) := 1$ é convexa, porém descontínua ao longo do bordo. Portanto na parte (i) do Teorema 2.7.19 não pode-se abrir mão da hipótese que U é aberta.

A parte (ii) do Teorema 2.7.19 diz que todo ponto crítico de uma função convexa continuamente diferenciável $f : U \rightarrow \mathbb{R}$ num subconjunto aberto convexo $U \subset \mathbb{R}^n$ é um mínimo absoluto de f . Vice-versa o Teorema 2.7.2 diz que todo mínimo absoluto de f é um ponto crítico. Consequentemente segue da parte (iii) do Teorema 2.7.19 que o conjunto dos pontos críticos de qualquer função $f : U \rightarrow \mathbb{R}$ continuamente diferenciável e convexa é um subconjunto convexo de U . O seguinte teorema caracteriza as funções convexas duas vezes continuamente diferenciáveis por meio das propriedades da sua matriz hessiana.

Teorema 2.7.21. *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função duas vezes continuamente diferenciável. Então vale o seguinte.*

(i) *f é convexo se e somente se a matriz hessiana de f é positiva semi-definida em todo ponto, ou seja, se a desigualdade $\xi^T d^2f(x)\xi \geq 0$ é satisfeita para todos os $x \in U$ e todos os $\xi \in \mathbb{R}^n$.*

(ii) *Se a matriz hessiana de f é positiva definida em todo ponto, ou seja, se a desigualdade $\xi^T d^2f(x)\xi > 0$ é satisfeita para todos os $x \in U$ e todos os $\xi \in \mathbb{R}^n \setminus \{0\}$, então f é estritamente convexo.*

Demonstração. Se f é convexo e $x_0 \in U$, neste caso também é convexo a função $g : U \rightarrow \mathbb{R}$ definida através de

$$g(x) := f(x) - f(x_0) - df(x_0)(x - x_0).$$

Esta função g possui um ponto crítico no ponto x_0 . Segundo a parte (ii) do Teorema 2.7.19 o ponto x_0 é daí um mínimo local de g , e segundo a parte (i) do Teorema 2.7.11 a matriz hessiana $d^2g(x_0) = d^2f(x_0)$ é daí positiva semi-definida.

Vice-versa suponhamos que a matriz hessiana $d^2f(x)$ é positiva semi-definida para todo $x \in U$. Sejam $x_0, x_1 \in U$ e seja $\phi : [0, 1] \rightarrow \mathbb{R}$ a função

$$\phi(t) := f((1-t)x_0 + tx_1), \quad 0 \leq t \leq 1.$$

Segundo o Teorema 2.3.2 esta função é duas vezes continuamente diferenciável e possui, segundo (2.7.2), a segunda derivada

$$\phi''(t) = (x_1 - x_0)^T d^2f((1-t)x_0 + tx_1)(x_1 - x_0) \geq 0.$$

Então, segundo o Teorema do Valor Médio, a primeira derivada $\phi' : [0, 1] \rightarrow \mathbb{R}$ é crescente e monótona e segue dali para todo $t \in [0, 1]$ a desigualdade

$$\phi(t) - \phi(0) = t \int_0^1 \phi'(ts) ds \leq t \int_0^1 \phi'(s) ds = t(\phi(1) - \phi(0)). \quad (2.7.7)$$

Consequentemente $\phi(t) \leq (1-t)\phi(0) + t\phi(1)$ e com isso a parte (i) fica provada.

Se para todo $x \in U$ a matriz hessiana $d^2f(x)$ é positiva definida e $x_0 \neq x_1$, então no argumento acima vale $\phi''(t) > 0$, e dali ϕ' é crescente e estritamente monótona, e daí a desigualdade em (2.7.7) é estrita para $0 < t < 1$. Com isso também a parte (ii) do Teorema 2.7.21 fica provada. \square

Exemplo 2.7.22. (i) A função $f(x) := 1/x$ para $x > 0$ é estritamente convexa, mas não possui nenhum ponto crítico. Ela não é Lipschitz-contínua globalmente.

(ii) A função $f(x) := x^4$ é estritamente convexa, mas a sua “matriz hessiana” (neste caso simplesmente o número real $f''(x)$) anula-se no ponto $x = 0$.

(iii) A função $f(x, y) := x^2$ é convexa, mas não estritamente convexa, e o conjunto dos seus pontos críticos é o eixo vertical $C := \{(x, y) \in \mathbb{R}^2 \mid x = 0\}$.

(iv) A função $f(x, y) := |x|$ é convexa, mas não é estritamente convexa, e é Lipschitz-contínua globalmente. Ela possui o mesmo conjunto C de mínimas globais como em (iii), mas não é diferenciável (exatamente nos pontos de C).

Exercício 2.7.23. Seja $U \subset \mathbb{R}^n$ um conjunto aberto convexo e $f : U \rightarrow \mathbb{R}$ uma função convexa duas vezes continuamente diferenciável. Se f possua um ponto crítico não-degenerado $x_0 \in U$, então vale

$$f(x_0) < f(x) \quad \text{para todos os } x \in U \setminus \{x_0\}$$

e f não possui outros pontos críticos. **Dica:** Exercício 2.7.17 e Teorema 2.7.19.

Capítulo 3

Funções Implícitas

O capítulo debruça-se sobre o estudo das soluções de uma equação da forma

$$f(x) = y,$$

onde $f : U \rightarrow Y$ é uma aplicação continuamente diferenciável entre espaços de Banach X e Y e onde $U \subset X$ é um subconjunto aberto. A Seção 3.1 trata do caso no que ambos espaços vetoriais tem a mesma dimensão e fornece um critério para a existência local e a unicidade da solução $x \in U$ assim como sua dependência diferenciável de y . Isto é o *Teorema da Aplicação Inversa*, o Teorema mais importante deste capítulo. Suponha para ilustração que X, Y sejam de dimensão finita. Então a Seção 3.2 trata do caso $\dim(X) > \dim(Y)$ e fornece condições nas quais o conjunto das soluções pode ser representado localmente como o gráfico de uma aplicação diferenciável. Isto é o *Teorema da Função Implícita* o qual pode ser reduzido facilmente ao Teorema da Aplicação Inversa. Este Teorema lida ao conceito de uma *subvariedade do \mathbb{R}^n* o qual será tratado na Seção 3.3. A Seção 3.4 ocupa-se com a questão como encontrar os extremos de uma função diferenciável numa tal subvariedade e isso lida à noção de *multiplicadores de Lagrange*.

3.1 O Teorema da Aplicação Inversa

Primeiramente lembramos que no ambiente dos espaços euclidianos temos introduzido a noção de uma aplicação C^ℓ via as derivadas parciais, veja a Definição 2.5.4. Como há vários exemplos de aplicações suaves definidas em subconjuntos abertos de espaços de Banach, ainda no caso de dimensão finita,

as quais não são munidos de bases canônicas, é útil falar de aplicações C^ℓ sem escolher anteriormente uma base. Repetimos da Definição 2.5.4 como pode-se formular isto indutivamente. Sejam X, Y espaços de Banach, seja $U \subset X$ um subconjunto aberto, e seja $\ell \in \mathbb{N}$. Chamamos uma aplicação $f : U \rightarrow Y$ de **ℓ -vezes continuamente diferenciável**, ou de **aplicação C^ℓ** , se é diferenciável em todo ponto $x \in U$ e a sua derivada $df : U \rightarrow \mathcal{L}(X, Y)$ é $(\ell - 1)$ -vezes continuamente diferenciável. (Veja a Definição 2.2.2 para o caso $\ell = 1$.) Definimos

$$d^2f := d(df) : U \rightarrow \mathcal{L}(X, \mathcal{L}(X, Y))$$

e no caso de $\ell \geq 2$ definimos indutivamente $d^k f := d(d^{k-1} f)$. Chamamos f de **suave** ou **da classe C^∞** , se f é ℓ -vezes continuamente diferenciável para todo $\ell \in \mathbb{N}$. Denota-se de $C^\ell(U, Y)$ o conjunto dos C^ℓ aplicações $f : U \rightarrow Y$.

Definição 3.1.1 (Difeomorfismo). *Sejam X e Y espaços de Banach e sejam $U \subset X$ e $V \subset Y$ subconjuntos abertos. Seja $\ell \in \mathbb{N} \cup \{\infty\}$. Uma aplicação $f : U \rightarrow V$ é chamado de **difeomorfismo da classe C^ℓ** ou de **C^ℓ -difeomorfismo** se f é bijetiva e tanto $f : U \rightarrow V$ como a aplicação inversa $f^{-1} : V \rightarrow U$ é uma aplicação C^ℓ . Um difeomorfismo da classe C^∞ também é chamado de **difeomorfismo suave**. Chama-se dois conjuntos U e V de **difeomorfo** se existe um difeomorfismo suave $f : U \rightarrow V$.*

Lema 3.1.2. *Sejam $U \subset X$ e $V \subset Y$ subconjuntos abertos de espaços de Banach, e seja $f : U \rightarrow V$ um difeomorfismo da classe C^1 . Então são iguais as dimensões $\dim(X) = \dim(Y)$, para todo $x \in U$ a derivada $df(x) : X \rightarrow Y$ é uma aplicação (operador) linear limitada invertível, e para todo $y \in V$ vale*

$$df^{-1}(y) = df(f^{-1}(y))^{-1}. \quad (3.1.1)$$

Demonstração. Seja $g := f^{-1} : V \rightarrow U$ a aplicação inversa de f , seja $x \in U$, e sejam $y \in V$, $A \in \mathcal{L}(X, Y)$, e $B \in \mathcal{L}(Y, X)$ dados através de

$$y := f(x), \quad A := df(x), \quad B := dg(y).$$

Como $g \circ f = \text{id}_U$ e $f \circ g = \text{id}_V$, encontraremos da regra da cadeia, Teorema 2.3.2, as equações

$$\begin{aligned} B \circ A &= dg(f(x)) \circ df(x) = d(g \circ f)(x) = \mathbb{1}_X \\ A \circ B &= df(g(y)) \circ dg(y) = d(f \circ g)(y) = \mathbb{1}_Y. \end{aligned}$$

Segue daí que A é invertível e $A^{-1} = B$. Particularmente X e Y tem a mesma dimensão (pode ser infinita) e com isso Lema 3.1.2 fica provado. \square

Exemplo 3.1.3. A função exponencial $\exp : \mathbb{R} \rightarrow (0, \infty)$ é um difeomorfismo suave cuja aplicação inversa, denotada de $\log := \exp^{-1} : (0, \infty) \rightarrow \mathbb{R}$, chama-se de *função logaritmo (natural)*. Que ambas funções são suave foi provado no curso Análise I. As representações delas como séries de potências mostram que são analíticas, como pode-se deduzir de Corolário 2.6.13.

Exemplo 3.1.4. A função $\mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^3$ é suave e bijetiva, porém anula-se a derivada no ponto $x = 0$, e sua aplicação inversa $\mathbb{R} \rightarrow \mathbb{R} : y \mapsto \text{sign}(y)|y|^{1/3}$ não é diferenciável no ponto $y = 0$.

Exemplo 3.1.5. Seja $I \subset \mathbb{R}$ um intervalo aberto e $f : I \rightarrow \mathbb{R}$ uma função continuamente diferenciável com $f'(x) > 0$ para todo $x \in I$. Então $J := f(I)$ é um intervalo segundo o Teorema do Valor Intermediário, e aberto porque f não pode possuir nenhum extremo local. Além disso, a função $f : I \rightarrow J$ é monótona e estritamente crescente segundo o Teorema do Valor Médio, e com isso é bijetivo. A aplicação inversa $f^{-1} : J \rightarrow I$ é diferenciável segundo a versão em dimensão 1 do Teorema da Aplicação Inversa no curso Análise I.

Exemplo 3.1.6. Seja $\|\cdot\|$ a norma euclidiana no \mathbb{R}^n . A bola unitária $U := \{x \in \mathbb{R}^n \mid \|x\| < 1\}$ é difeomorfa ao \mathbb{R}^n . Um difeomorfismo suave $f : U \rightarrow \mathbb{R}^n$ e a sua aplicação inversa são dadas através de

$$f(x) = \frac{x}{\sqrt{1 - \|x\|^2}}, \quad f^{-1}(y) = \frac{y}{\sqrt{1 + \|y\|^2}}$$

para $x \in U$ e $y \in \mathbb{R}^n$.

Exemplo 3.1.7. O semi-plano acima $\mathbb{H} := \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$ é difeomorfo ao disco unitário $\mathbb{D} := \{\zeta \in \mathbb{C} \mid |\zeta| < 1\}$. Um difeomorfismo $f : \mathbb{H} \rightarrow \mathbb{D}$ suave e sua aplicação inversa são dados através de

$$f(z) = \frac{z - \mathbf{i}}{z + \mathbf{i}}, \quad f^{-1}(\zeta) = \mathbf{i} \frac{1 + \zeta}{1 - \zeta}$$

para $z \in \mathbb{H}$ e $\zeta \in \mathbb{D}$.

Exemplo 3.1.8. Sejam $\mathcal{A} := \mathcal{L}(X)$ e $\text{GL}(X) := \mathcal{G}$ (Exemplo 2.3.14). Então a aplicação

$$\text{Inv} : \text{GL}(X) \rightarrow \text{GL}(X), \quad \text{Inv}(A) := A^{-1}$$

é contínua, segundo o Teorema 1.4.1, e continuamente diferenciável, segundo o Exemplo 2.3.14. Como a aplicação Inv é bijetiva e coincide com sua aplicação inversa, assim a inversa Inv é um difeomorfismo da classe C^1 . Que Inv ainda é um difeomorfismo suave será provado no seguinte lema.

Lema 3.1.9. *A aplicação $\text{Inv} : \text{GL}(X) \rightarrow \text{GL}(X)$ é suave.*

Demonstração. Segundo o Exemplo 2.3.14 a aplicação Inv é continuamente diferenciável e a derivada $d\text{Inv}(A) \in \mathcal{L}(\mathcal{A}, \mathcal{A})$ onde $\mathcal{A} = \mathcal{L}(X)$ é dada através de

$$d\text{Inv}(A)\widehat{A} = -A^{-1}\widehat{A}A^{-1} = -\text{Inv}(A)\widehat{A}\text{Inv}(A) \quad (3.1.2)$$

no ponto $A \in \text{GL}(X) = \mathcal{G}$ e para $\widehat{A} \in \mathcal{A}$. Neste momento é útil introduzir a seguinte notação. Quaisquer duas matrizes $B_0, B_1 \in \mathcal{A}$ determinam uma aplicação linear $\Lambda(B_0, B_1) : \mathcal{A} \rightarrow \mathcal{A}$ através da fórmula

$$(\Lambda(B_0, B_1))(\widehat{A}) := -B_0\widehat{A}B_1$$

para $\widehat{A} \in \mathcal{A}$. Recebemos assim uma aplicação

$$\Lambda : \mathcal{A}^2 = \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{L}(\mathcal{A}, \mathcal{A})$$

cujos diferencial $d\Lambda : \mathcal{A}^2 \rightarrow \mathcal{L}(\mathcal{A}^2, \mathcal{L}(\mathcal{A}, \mathcal{A}))$ é dado através da fórmula

$$(B_0, B_1) \mapsto \left[(\widehat{B}_0, \widehat{B}_1) \mapsto -\widehat{B}_0\widehat{A}B_1 - B_0\widehat{A}\widehat{B}_1 \right]$$

(Exercício). Note-se que o diferencial $d\Lambda$ é uma aplicação linear. Segundo o Exemplo 2.1.7 uma aplicação linear é diferenciável e possui uma derivada constante (neste caso com valores em $\mathcal{L}(\mathcal{A}^2, \mathcal{L}(\mathcal{A}^2, \mathcal{L}(\mathcal{A}, \mathcal{A})))$), e toda aplicação constante é suave segundo o Exemplo 2.1.6. Então Λ é suave. Ora pode-se escrever a equação (3.1.2) na forma

$$d\text{Inv}(A) = \Lambda(\text{Inv}(A), \text{Inv}(A)). \quad (3.1.3)$$

Ou seja, a derivada da aplicação inversão $\text{Inv} : \mathcal{G} \rightarrow \mathcal{G}$ pode ser escrito como composição

$$d\text{Inv} = \Lambda \circ \iota \circ \text{Inv} : \mathcal{G} \rightarrow \mathcal{L}(\mathcal{A}, \mathcal{A}) \quad (3.1.4)$$

onde a aplicação

$$\iota : \mathcal{A} \rightarrow \mathcal{A} \times \mathcal{A}$$

é definida através de $\iota(B) := (B, B)$ para $B \in \mathcal{A}$. Como ι é uma aplicação linear, também é suave. Ora, se Inv é uma aplicação C^ℓ , então segue da equação (3.1.4) e do Teorema 2.5.9 que $d\text{Inv}$ também é uma aplicação C^ℓ e assim Inv é uma aplicação $C^{\ell+1}$. Como já sabemos que Inv é uma aplicação C^1 , temos mostrado através de indução que Inv é suave. Com isso Lema 3.1.9 fica provado. \square

Lema 3.1.10. *Sejam $U \subset X$ e $V \subset Y$ subconjuntos abertos de espaços de Banach e seja $f : U \rightarrow V$ um C^1 -difeomorfismo. Se f é uma aplicação da classe C^ℓ para um $\ell \in \mathbb{N} \cup \{\infty\}$, então a aplicação inversa $f^{-1} : V \rightarrow U$ também é C^ℓ , e assim f é um C^ℓ -difeomorfismo.*

Demonstração. Segundo o Lema 3.1.2 podemos identificar X e Y linearmente, então suavemente, através de $df(x) \in \mathcal{L}(X, Y)$ para qualquer $x \in U$ com inversa $df^{-1}(f(x)) \in \mathcal{L}(Y, X)$. Seja então $X = Y$. A fórmula

$$df^{-1}(y) = df(f^{-1}(y))^{-1} = \text{Inv}(df(f^{-1}(y))) \quad \text{para } y \in V$$

mostra que podemos escrever a derivada da aplicação inversa $f^{-1} : V \rightarrow U$ como composição

$$V \xrightarrow{f^{-1}} U \xrightarrow{df} \text{GL}(X) \xrightarrow{\text{Inv}} \text{GL}(X).$$

Ora, argumentamos com indução sobre ℓ . Para $\ell = 1$ não tem nada para provar. Suponhamos então que $\ell \geq 2$, que a afirmação do Lema vale para $\ell - 1$, e que $f : U \rightarrow V$ é uma aplicação C^ℓ bijetiva cuja aplicação inversa é continuamente diferenciável. Então $f^{-1} : V \rightarrow U$ é $C^{\ell-1}$, segundo a hipótese da indução, e $df : U \rightarrow \text{GL}(X)$ é $C^{\ell-1}$ pela hipótese. Como, segundo o Lema 3.1.9, a aplicação $\text{Inv} : \text{GL}(X) \rightarrow \text{GL}(X)$ é suave, a parte (v) do Teorema 2.5.9 implica que a composição $df^{-1} = \text{Inv} \circ df \circ f^{-1} : V \rightarrow \text{GL}(X)$ é $C^{\ell-1}$ e daí f^{-1} é uma aplicação C^ℓ . Com isso Lema 3.1.10 fica provado. \square

Com estas preparações somos dispostos para o teorema principal deste capítulo. Como a prova não depende de se a dimensão é finita ou infinita (modulo um único uso de um resultado avançado – o teorema da inversa limitada), vamos formular o seguinte teorema para espaços de Banach gerais. Se quiser, escolha $X = \mathbb{R}^n$ e $Y = \mathbb{R}^m$.

Teorema 3.1.11 (Teorema da Aplicação Inversa). *Sejam X e Y espaços de Banach, seja $\Omega \subset X$ um subconjunto aberto, seja $f : \Omega \rightarrow Y$ uma aplicação C^ℓ com $\ell \in \mathbb{N} \cup \{\infty\}$, e seja $x_0 \in \Omega$. Suponhamos que a derivada de f no ponto x_0 , ou seja $df(x_0) \in \mathcal{L}(X, Y)$, é uma aplicação (linear) bijetiva. Então existe um subconjunto aberto $U \subset \Omega$ com $x_0 \in U$, tal que $V := f(U)$ é um subconjunto aberto de Y e tal que a restrição*

$$f|_U : U \rightarrow V$$

é um difeomorfismo da classe C^ℓ .

Demonstração. Veja a página 109. \square

A demonstração do Teorema 3.1.11 resta no seguinte lema. Neste lema usamos a norma operador

$$\|A\| := \sup_{0 \neq \xi \in X} \frac{\|A\xi\|}{\|\xi\|}$$

para os endomorfismos (aplicações lineares introspetivas) $A : X \rightarrow X$ de um espaço vetorial normado $(X, \|\cdot\|)$. Além disso, lembramos que a bola aberta com centro $x_0 \in X$ e raio $r > 0$ é denotada com o símbolo $B_r(x_0) := \{x \in X \mid \|x - x_0\| < r\}$.

Lema 3.1.12. *Seja $(X, \|\cdot\|)$ um espaço de Banach e sejam dado $x_0 \in X$ e duas constantes $r > 0$ e $0 < \gamma < 1$. Além disso, seja $\psi : B_r(x_0) \rightarrow X$ uma aplicação continuamente diferenciável com*

$$\|\mathbb{1} - d\psi(x)\| \leq \gamma \quad \text{para todos os } x \in B_r(x_0). \quad (3.1.5)$$

Então a aplicação $\psi : B_r(x_0) \rightarrow X$ é injetiva, seu imagem $\psi(B_r(x_0))$ é aberto, valem as inclusões

$$B_{r(1-\gamma)}(\psi(x_0)) \subset \psi(B_r(x_0)) \subset B_{r(1+\gamma)}(\psi(x_0)) \quad (3.1.6)$$

e a inversa $\psi^{-1} : \psi(B_r(x_0)) \rightarrow B_r(x_0)$ é continuamente diferenciável.

Demonstração. Segundo a hipótese e Teorema 2.3.1 a aplicação

$$\phi : B_r(x_0) \rightarrow X, \quad \phi(x) := x - \psi(x),$$

é continuamente diferenciável com derivada $d\phi(x) = \mathbb{1} - d\psi(x)$ para todo $x \in B_r(x_0)$. Segundo (3.1.5) a derivada de ϕ satisfaz dali a desigualdade $\|d\phi(x)\| \leq \gamma$ para todo $x \in B_r(x_0)$, e segundo o Teorema 2.4.1, segue daí

$$\|\phi(x) - \phi(x')\| \leq \gamma \|x - x'\| \quad (3.1.7)$$

para todos os $x, x' \in B_r(x_0)$. Em contrapartida seguem daí as desigualdades

$$\begin{aligned} \|\psi(x) - \psi(x')\| &= \|x - x' - \phi(x) + \phi(x')\| \\ &\leq \|x - x'\| + \|\phi(x) - \phi(x')\| \\ &\leq (1 + \gamma) \|x - x'\|, \end{aligned} \quad (3.1.8)$$

$$\begin{aligned} \|\psi(x) - \psi(x')\| &= \|x - x' - \phi(x) + \phi(x')\| \\ &\geq \|x - x'\| - \|\phi(x) - \phi(x')\| \\ &\geq (1 - \gamma) \|x - x'\| \end{aligned} \quad (3.1.9)$$

para todos os $x, x' \in B_r(x_0)$. Segundo (3.1.9) a aplicação $\psi : B_r(x_0) \rightarrow X$ é injetiva. Vamos provar as outras afirmações em três passos.

Passo 1. *Provaremos (3.1.6).*

A inclusão $\psi(B_r(x_0)) \subset B_{r(1+\gamma)}(\psi(x_0))$ segue diretamente de (3.1.8). Resta mostrar a inclusão $B_{r(1-\gamma)}(\psi(x_0)) \subset \psi(B_r(x_0))$. Dado $y \in B_{r(1-\gamma)}(\psi(x_0))$, então $\|y - \psi(x_0)\| < r(1 - \gamma)$ e daí

$$0 \leq s := \frac{\|y - \psi(x_0)\|}{1 - \gamma} < r. \quad (3.1.10)$$

Com isto definimos o conjunto não-vazio

$$K := \{x \in X \mid \|x - x_0\| \leq s\} \subset B_r(x_0)$$

que é um espaço métrico completo sob a função distância $d(x, x') := \|x - x'\|$ para $x, x' \in K$. Ora, seja a aplicação $f_y : K \rightarrow X$ definida através de

$$f_y(x) := y + \phi(x) \quad \text{para } x \in K.$$

Então todo $x \in K$ satisfaz a desigualdade

$$\begin{aligned} \|f_y(x) - x_0\| &= \|y + \phi(x) - x_0\| \\ &= \|\phi(x) - \phi(x_0) + y - \psi(x_0)\| \\ &\leq \|\phi(x) - \phi(x_0)\| + \|y - \psi(x_0)\| \\ &\leq \gamma \|x - x_0\| + (1 - \gamma)s \\ &\leq \gamma s + (1 - \gamma)s \\ &= s. \end{aligned}$$

Aqui o segundo passo segue da equação $x_0 = \phi(x_0) + \psi(x_0)$, o terceiro da desigualdade triangular, o quarto de (3.1.7) e (3.1.10), e o quinto da definição do conjunto K . Com isto temos mostrado que $f_y(K) \subset K$. Além disso, vale

$$\|f_y(x) - f_y(x')\| = \|\phi(x) - \phi(x')\| \leq \gamma \|x - x'\|$$

para todos os $x, x' \in K$, segundo (3.1.7). Portanto $f_y : K \rightarrow K$ é uma contração num espaço métrico completo não-vazio e possui daí, segundo o Teorema do Ponto Fixo de Banach (Teorema E.0.7), um ponto fixo único

$$x = f_y(x) = y + \phi(x) \in K \subset B_r(x_0).$$

Este ponto fixo satisfaz a equação $\psi(x) = x - \phi(x) = y$ e com isso Passo 1 fica provado.

Passo 2. O conjunto imagem $\psi(B_r(x_0)) = \{\psi(x) \mid x \in B_r(x_0)\}$ é aberto.

Seja dado $y_1 \in \psi(B_r(x_0))$, escolha $x_1 \in B_r(x_0)$, tal que $\psi(x_1) = y_1$. Então para $\varepsilon := r - \|x_1 - x_0\| > 0$ vale $B_\varepsilon(x_1) \subset B_r(x_0)$ segundo a desigualdade triangular. Segundo o Passo 1 segue daí $B_{\varepsilon(1-\gamma)}(y_1) \subset \psi(B_\varepsilon(x_1)) \subset \psi(B_r(x_0))$ e com isso Passo 2 fica provado.

Passo 3. A aplicação inversa ψ^{-1} é continuamente diferenciável.

A aplicação $\psi : B_r(x_0) \rightarrow \psi(B_r(x_0))$ é bijetiva segundo (3.1.9), e a mesma desigualdade mostra que ψ^{-1} é Lipschitz-contínua. Vamos mostrar que ψ^{-1} é diferenciável. Neste caso, segundo o Lema 3.1.2, sua derivada é a composição

$$\psi(B_r(x_0)) \xrightarrow{\psi^{-1}} B_r(x_0) \xrightarrow{d\psi} \text{GL}(X) \xrightarrow{\text{Inv}} \text{GL}(X)$$

e daí é contínua. (Aqui $\text{GL}(X) \subset \mathcal{L}(X, X)$ denota o grupo das aplicações lineares limitadas bijetivas $A : X \rightarrow X$.)

Diferenciabilidade de ψ^{-1} num ponto $y_1 \in \psi(B_r(x_0))$: Seja $x_1 \in B_r(x_0)$ com $\psi(x_1) = y_1$, e seja $\Psi := d\psi(x_1)$. Segundo (3.1.5) vale $\|\mathbb{1} - \Psi\| \leq \gamma < 1$, e daí Ψ é invertível segundo o Teorema 1.4.1 com $\Psi^{-1} = \sum_{k=0}^{\infty} (\mathbb{1} - \Psi)^k$. Dali

$$\|\Psi^{-1}\| \leq \sum_{k=0}^{\infty} \|(\mathbb{1} - \Psi)^k\| \leq \sum_{k=0}^{\infty} \|\mathbb{1} - \Psi\|^k \leq \sum_{k=0}^{\infty} \gamma^k = \frac{1}{1 - \gamma}. \quad (3.1.11)$$

Ora, seja dado $\varepsilon > 0$. Segundo a definição da derivada existe um $\rho > 0$ com $\rho \leq r - \|x_0 - x_1\|$, tal que todo $x \in B_\rho(x_1) \subset B_r(x_0)$ satisfaz a desigualdade

$$\|\psi(x) - \psi(x_1) - \Psi(x - x_1)\| \leq \varepsilon(1 - \gamma)^2 \|x - x_1\|. \quad (3.1.12)$$

Com $\delta := \rho(1 - \gamma)$ vale $B_\delta(y_1) \subset \psi(B_\rho(x_1))$ segundo o Passo 1.

Ora, seja dado $y \in X$ com $\|y - y_1\| < \delta$. Então existe um $x \in B_\rho(x_1)$ com $\psi(x) = y$. Usando (3.1.12), (3.1.11), e (3.1.9) (exatamente neste ordem) encontraremos dali a cadeia de desigualdades

$$\begin{aligned} \|\psi^{-1}(y) - \psi^{-1}(y_1) - \Psi^{-1}(y - y_1)\| &= \|\Psi^{-1}(y - y_1 - \Psi(x - x_1))\| \\ &\leq \|\Psi^{-1}\| \|\psi(x) - \psi(x_1) - \Psi(x - x_1)\| \\ &\leq \|\Psi^{-1}\| \varepsilon(1 - \gamma)^2 \|x - x_1\| \\ &\leq \varepsilon(1 - \gamma) \|x - x_1\| \\ &\leq \varepsilon \|y - y_1\|. \end{aligned}$$

Segue daí que ψ^{-1} é diferenciável no ponto y_1 e $d\psi^{-1}(y_1) = \Psi^{-1}$. Com isso Passo 3 e Lema 3.1.12 são provados. \square

Demonstração do Teorema 3.1.11. Pela hipótese a aplicação linear

$$A := df(x_0) : X \rightarrow Y$$

é bijetiva e limitada. Em dimensão finita A^{-1} já é limitada e em dimensão infinita isto é garantido pelo Teorema da Inversa Limitada (o qual é baseado no Teorema de Baire). Ora, definimos a aplicação $\psi : \Omega \rightarrow X$ através de

$$\psi(x) := A^{-1}f(x) \quad \text{para } x \in \Omega.$$

Então ψ é continuamente diferenciável segundo o Teorema 2.3.2 e vale

$$d\psi(x) - \mathbb{1} = A^{-1}df(x) - \mathbb{1} = A^{-1}(df(x) - df(x_0))$$

para todos os $x \in \Omega$. Ora, seja

$$\varepsilon := \frac{1}{2\|A^{-1}\|} > 0, \quad \|A^{-1}\| = \sup_{0 \neq \eta \in Y} \frac{\|A^{-1}\eta\|_X}{\|\eta\|_Y}.$$

Como a aplicação $df : \Omega \rightarrow \mathcal{L}(X, Y)$ é contínua pela hipótese, existe uma constante $\delta > 0$ com $B_\delta(x_0) \subset \Omega$ tal que todo $x \in B_\delta(x_0)$ satisfaz a desigualdade $\|df(x) - df(x_0)\| < \varepsilon$. Daí segue para todo $x \in B_\delta(x_0)$ a desigualdade

$$\begin{aligned} \|d\psi(x) - \mathbb{1}\| &= \|A^{-1}(df(x) - df(x_0))\| \\ &\leq \|A^{-1}\| \|df(x) - df(x_0)\| \\ &< \|A^{-1}\| \varepsilon \\ &= \frac{1}{2}. \end{aligned}$$

Dali, segundo o Lema 3.1.12 com $\gamma = 1/2$, o subconjunto $\psi(B_\delta(x_0)) \subset X$ é aberto e $\psi : B_\delta(x_0) \rightarrow \psi(B_\delta(x_0))$ é um C^1 -difeomorfismo. Ora, definimos

$$U := B_r(x_0) \subset \Omega, \quad V := A(\psi(B_\delta(x_0))) \subset Y.$$

Então $V = \{A\psi(x) \mid x \in U\} = f(U)$ é um subconjunto aberto de Y , porque a aplicação linear $A^{-1} : Y \rightarrow X$ é contínua e V a pre-imagem do subconjunto aberto $\psi(B_\delta(x_0)) \subset X$ sob A^{-1} . Ora, a aplicação

$$f|_U = A\psi|_U : U \rightarrow V$$

é bijetiva com a aplicação inversa

$$(f|_U)^{-1}(y) = (\psi|_U)^{-1}(A^{-1}y) \quad \text{para } y \in V.$$

Então $(f|_U)^{-1} : V \rightarrow U$ é continuamente diferenciável, segundo o Lema 3.1.12 e o Teorema 2.3.2. Como f é pela hipótese uma aplicação C^ℓ , segue do Lema 3.1.10 que $(f|_U)^{-1} : V \rightarrow U$ também é uma aplicação C^ℓ . \square

Corolário 3.1.13 (Imagem aberta). *Sejam X, Y espaços de Banach, seja $U \subset X$ um conjunto aberto, e seja $f : U \rightarrow Y$ uma aplicação continuamente diferenciável, tal que a derivada $df(x) \in \mathcal{L}(X, Y)$ é uma aplicação linear invertível para todo $x \in U$. Então $V := f(U)$ é um subconjunto aberto de Y .*

Demonstração. Seja $y_0 \in V$, fixe $x_0 \in U$ com $f(x_0) = y_0$. Então existe, segundo o Teorema 3.1.11, um subconjunto aberto $U_0 \subset U$ com $x_0 \in U_0$ tal que $V_0 := f(U_0)$ é um subconjunto aberto de Y . Como $y_0 = f(x_0) \in f(U_0) = V_0$, existe então um $\delta > 0$ com $B_\delta(y_0) \subset V_0 \subset V$. Isso prova Corolário 3.1.13. \square

Corolário 3.1.14 (Imagem difeomorfo). *Sejam X, Y espaços de Banach e seja $U \subset X$ um conjunto aberto, e seja $f : U \rightarrow Y$ uma aplicação injetiva da classe C^ℓ com $\ell \in \mathbb{N} \cup \{\infty\}$, tal que $df(x) \in \mathcal{L}(X, Y)$ é uma aplicação linear invertível para todo $x \in U$. Então $V := f(U)$ é um subconjunto aberto de Y e $f : U \rightarrow V$ é um C^ℓ -difeomorfismo.*

Demonstração. O subconjunto $V \subset Y$ é aberto, segundo o Corolário 3.1.13, e a aplicação $f : U \rightarrow V$ é bijetiva, segundo a hipótese. Dado $y_0 \in V$, então segundo o Teorema 3.1.11, existe um conjunto aberto $V_0 \subset V$, tal que $y_0 \in V_0$ e $f^{-1}|_{V_0} : V_0 \rightarrow U$ é uma aplicação C^ℓ . Como isto vale para todo $y_0 \in V$ que a inversa f^{-1} é uma aplicação C^ℓ . Com isso Corolário 3.1.14 fica provado. \square

Exemplo 3.1.15. A função seno é suave com $\text{sen}'(0) = 1$. Dali e segundo o Teorema 3.1.11 sua restrição a um intervalo aberto adequado $U_0 \subset \mathbb{R}$ com $0 \in U_0$, é um difeomorfismo suave ao conjunto imagem correspondente. O intervalo máximo é $U_0 = (-\pi/2, \pi/2)$. Como a função seno é injetiva neste intervalo e tem uma derivada positiva, a função $\text{sen} : (-\pi/2, \pi/2) \rightarrow (-1, 1)$ é um difeomorfismo segundo o Corolário 3.1.14.

Exemplo 3.1.16. Seja X um espaço de Banach e $\mathcal{A} := \mathcal{L}(X)$. (Exemplos 1.3.6 e 2.3.13). A aplicação exponencial $\exp : \mathcal{A} \rightarrow \mathcal{A}$ é continuamente diferenciável, veja o Exemplo 2.3.13, com derivada $d\exp(0) = \text{id}$ em $A_0 = 0$. Em Corolário 4.4.4 vamos ver que a aplicação exponencial ainda é suave. Daí segue de Teorema 3.1.11 que existem conjuntos abertos $U, V \subset \mathcal{A}$ com $0 \in U$, tal que $\exp : U \rightarrow V$ é um difeomorfismo suave. A série logaritmo no Exemplo 1.6.2 mostra que podemos escolher $V = \{B \in \mathcal{A} \mid \|B - \mathbb{1}\| < 1\}$. A aplicação inversa então é $\exp^{-1}(B) = \log(B) = -\sum_{k=1}^{\infty} (\mathbb{1} - B)^k / k$ para $B \in V$.

Exercício 3.1.17. Pontos críticos não-degenerados são isolados. (Para a definição veja o Exercício 2.7.17.)

3.2 O Teorema da Função Implícita

Nesta seção trata-se da questão, dado uma equação

$$f(x, y) = 0, \quad (3.2.1)$$

nas quais hipóteses pode-se *resolver localmente para y* , ou seja, existe uma função continuamente diferenciável $y = g(x)$ tal que o conjunto das soluções é *localmente* o gráfico da função g . Então dizemos que a função g é *definido implicitamente* através da equação (3.2.1). Aqui um exemplo simples para já.

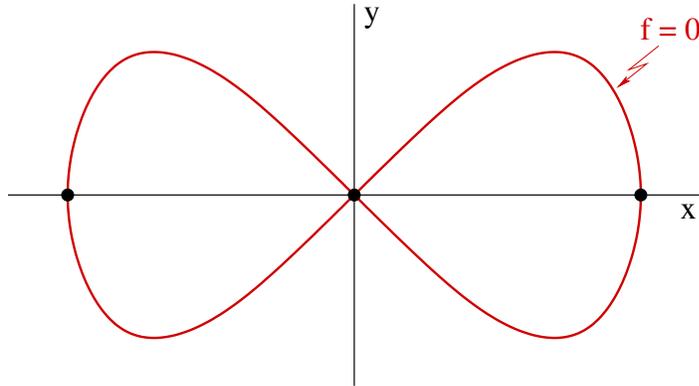


Figura 3.1: Uma função definida implicitamente.

Exemplo 3.2.1. Seja $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ a função

$$f(x, y) := x^2 - x^4 - y^2 \quad (3.2.2)$$

para $(x, y) \in \mathbb{R}^2$. Então toda solução da equação (3.2.1) satisfaz as desigualdades $|x| \leq 1$ e $|y| \leq 1/2$. Além disso, pode-se resolver explicitamente a equação para y , e recebe-se a fórmula

$$y = \pm x\sqrt{1 - x^2} \quad \text{para } -1 \leq x \leq 1. \quad (3.2.3)$$

Esta fórmula representa funções suaves, com ambos sinais para $-1 < x < 0$ e para $0 < x < 1$. Os três pontos excepcionais no conjunto das soluções são $(x, y) = (\pm 1, 0)$ e $(x, y) = (0, 0)$, e isso são exatamente os elementos do \mathbb{R}^2 as quais satisfazem as condições $f(x, y) = 0$ e $\frac{\partial f}{\partial y}(x, y) = 0$ (veja a Figura 3.1).

A situação geral nesta seção refere-se a aplicações continuamente diferenciáveis $f = (f_1, \dots, f_m) : \Omega \rightarrow \mathbb{R}^m$ as quais são definidas num subconjunto aberto $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$. Denotamos de $(x, y) = (x_1, \dots, x_n, y_1, \dots, y_m)$ os elementos do $\mathbb{R}^n \times \mathbb{R}^m$. Então a tarefa é resolver as m equações $f_k(x, y) = 0$, $k = 1, \dots, m$, com respeito às m variáveis y_1, \dots, y_m . É importante que o número de equações coincide com o de variáveis. Nesta situação é útil escrever a matriz jacobiana de f num ponto $(x, y) \in \Omega$ como matriz bloco

$$df(x, y) = \left(\frac{\partial f}{\partial x}(x, y) \mid \frac{\partial f}{\partial y}(x, y) \right) \in \mathbb{R}^{m \times (n+m)}$$

com

$$\frac{\partial f}{\partial x}(x, y) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x, y) & \cdots & \frac{\partial f_1}{\partial x_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x, y) & \cdots & \frac{\partial f_m}{\partial x_n}(x, y) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

$$\frac{\partial f}{\partial y}(x, y) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \cdots & \frac{\partial f_1}{\partial y_m}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(x, y) & \cdots & \frac{\partial f_m}{\partial y_m}(x, y) \end{pmatrix} \in \mathbb{R}^{m \times m}.$$

A condição decisiva para a solubilidade local para y do sistema de equações $f(x, y) = 0$ com respeito y é que o determinante da matriz quadrada $\frac{\partial f}{\partial y}(x, y)$ não é nulo. Ora, a derivada direcional de f no ponto $(x, y) \in \Omega$ na direção de um vetor $(\xi, \eta) \in \mathbb{R}^n \times \mathbb{R}^m$ tem a forma

$$\begin{aligned} df(x, y) \begin{pmatrix} \xi \\ \eta \end{pmatrix} &= \frac{\partial f}{\partial x}(x, y)\xi + \frac{\partial f}{\partial y}(x, y)\eta \\ &= \frac{d}{dt} \Big|_{t=0} f(x + t\xi, y) + \frac{d}{dt} \Big|_{t=0} f(x, y + t\eta). \end{aligned} \quad (3.2.4)$$

A mesma fórmula recebe-se se X, Y, Z são espaços de Banach e $f : \Omega \rightarrow Z$ uma C^1 -aplicação num subconjunto aberto $\Omega \subset X \times Y$. Neste caso define-se

$$d^X f(x, y) := df_y(x) \in \mathcal{L}(X, Z), \quad d^Y f(x, y) := df_x(y) \in \mathcal{L}(Y, Z)$$

para $(x, y) \in \Omega$ e onde $f_y(x) := f(x, y) =: f_x(y)$ define as C^1 -aplicações $f_y : \Omega_y := \{x \in X \mid (x, y) \in \Omega\} \rightarrow Z$ e $f_x : \Omega_x := \{y \in Y \mid (x, y) \in \Omega\} \rightarrow Z$.

Teorema 3.2.2 (Teorema da Função Implícita). *Sejam X, Y, Z espaços de Banach, seja $\Omega \subset X \times Y$ um subconjunto aberto, seja $f : \Omega \rightarrow Z$ uma aplicação C^ℓ com $\ell \in \mathbb{N} \cup \{\infty\}$, e seja dado $(x_0, y_0) \in \Omega$, tal que $f(x_0, y_0) = 0$ e a aplicação linear $d^Y f(x_0, y_0) \in \mathcal{L}(Y, Z)$ é bijetiva. Então existem subconjuntos abertos $U \subset X$ e $V \subset Y$ e uma aplicação $g : U \rightarrow V$ da classe C^ℓ com as seguintes propriedades.*

(i) Vale $(x_0, y_0) \in U \times V \subset \Omega$ e

$$g(x_0) = y_0. \quad (3.2.5)$$

(ii) Para todos os $x \in U$ e todos os $y \in V$ vale a equivalência

$$f(x, y) = 0 \quad \Longleftrightarrow \quad g(x) = y. \quad (3.2.6)$$

(iii) Para todo $x \in U$ a aplicação linear $d^Y f(x, g(x)) \in \mathcal{L}(Y, Z)$ é bijetiva e

$$dg(x) = -d^Y f(x, g(x))^{-1} d^X f(x, g(x)). \quad (3.2.7)$$

Demonstração. Utilizamos as abreviações

$$A := d^X f(x_0, y_0) \in \mathcal{L}(X, Z), \quad B := d^Y f(x_0, y_0) \in \mathcal{L}(Y, Z)$$

e definimos a aplicação $F : \Omega \rightarrow X \times Z$ através de

$$F(x, y) := (x, f(x, y))$$

para $(x, y) \in \Omega$. Então F é uma aplicação C^ℓ e sua derivada no ponto (x_0, y_0) é dada através de

$$\begin{aligned} dF(x_0, y_0)(\xi, \eta) &= \left. \frac{d}{dt} \right|_{t=0} F(x_0 + t\xi, y_0 + t\eta) \\ &= \left. \frac{d}{dt} \right|_{t=0} (x_0 + t\xi, f(x_0 + t\xi, y_0 + t\eta)) \\ &= (\xi, d^X f(x_0, y_0)\xi + d^Y f(x_0, y_0)\eta) \\ &= (\xi, A\xi + B\eta) \end{aligned}$$

para $(\xi, \eta) \in X \times Y$. Como $B : Y \rightarrow Z$ é uma aplicação bijetiva, a aplicação derivada $dF(x_0, y_0) : X \times Y \rightarrow X \times Z$ também é injetiva e tem a inversa

$$dF(x_0, y_0)^{-1}(\xi, \zeta) = (\xi, B^{-1}(\zeta - A\xi))$$

para $\xi \in X$ e $\zeta \in Z$.

Segundo o Teorema 3.1.11 existe um subconjunto aberto $\Omega_0 \subset \Omega$ contendo $(x_0, y_0) \in \Omega_0$ tal que $F(\Omega_0) \subset X \times Z$ é aberto e $F_0 := F|_{\Omega_0} : \Omega_0 \rightarrow F(\Omega_0)$ é um C^ℓ -difeomorfismo. Assim $F_0^{-1} : F(\Omega_0) \rightarrow \Omega_0$ é uma aplicação C^ℓ com

$$(x_0, 0) \in F(\Omega_0), \quad F_0^{-1}(x_0, 0) = (x_0, y_0). \quad (3.2.8)$$

Como Ω_0 é um subconjunto aberto de $X \times Y$, existe um $\varepsilon > 0$ com

$$B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y) \subset \Omega_0. \quad (3.2.9)$$

Como $F_0^{-1} : F(\Omega_0) \rightarrow \Omega_0$ é uma aplicação contínua e $F_0^{-1}(x_0, 0) = (x_0, y_0)$, $F(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y)) = (F_0^{-1})^{-1}(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y)) \subset X \times Z$ é uma vizinhança aberta de $(x_0, 0)$. Daí existe uma constante $0 < \delta \leq \varepsilon$ com

$$B_\delta(x_0; X) \times B_\delta(0; Z) \subset F(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y)) \subset F(\Omega_0). \quad (3.2.10)$$

Ora, definimos aplicações $\iota_X : X \rightarrow X \times Z$ e $\pi_Y : X \times Y \rightarrow Y$ através de $\iota_X(x) := (x, 0)$ para $x \in X$ e $\pi_Y(x, y) := y$ para $(x, y) \in X \times Y$. Então ambas estas aplicações são linear e assim suave.

Ora, se $x \in B_\delta(x_0; X)$, então $\iota_X(x) = (x, 0) \in F(B_\varepsilon(x_0; X) \times B_\varepsilon(y_0; Y))$ segundo (3.2.10) e com isso vale $\pi_Y \circ F_0^{-1} \circ \iota_X(x) \in B_\varepsilon(y_0; Y)$. Portanto encontraremos uma aplicação bem definida $g : U \rightarrow V$ através de

$$U := B_\delta(x_0; X), \quad V := B_\varepsilon(y_0; Y), \quad g(x) := \pi_Y \circ F_0^{-1} \circ \iota_X(x) \quad (3.2.11)$$

para $x \in U$. Isto é uma aplicação C^ℓ segundo o Teorema 2.5.9. Além disso,

$$(x_0, y_0) \in U \times V \subset \Omega_0 \subset \Omega$$

segundo (3.2.9), e $g(x_0) = \pi_Y(F_0^{-1}(x_0, 0)) = \pi_Y(x_0, y_0) = y_0$ segundo (3.2.8). Consequentemente g satisfaz a condição (i). Para a prova de (ii) fixamos dois elementos $x \in U$ e $y \in V$. Então $(x, y), (x, g(x)) \in \Omega_0$ e valem as duas equações $F(x, y) = (x, f(x, y))$ e $F(x, g(x)) = (x, 0)$. Como $F|_{\Omega_0}$ é injetivo, vale então $g(x) = y$ se e somente se $f(x, y) = 0$. Com isso g satisfaz as condições (ii). Para a prova de (iii) fixamos um elemento $x \in U$. Então o par $(x, g(x)) \in \Omega_0$ e daí a aplicação linear $dF(x, g(x)) : X \times Y \rightarrow X \times Z$ é invertível, segundo o Lema 3.1.2. Com

$$dF(x, g(x))(\xi, \eta) = (\xi, d^X f(x, g(x))\xi + d^Y f(x, g(x))\eta)$$

para $(\xi, \eta) \in X \times Y$ segue dali que $d^Y f(x, g(x)) : Y \rightarrow Z$ é invertível. Diferenciamos ora a equação $f(x, g(x)) = 0$, então recebemos

$$0 = d^X f(x, g(x))\xi + d^Y f(x, g(x))dg(x)\xi$$

para todos os $\xi \in X$ e dali segue (3.2.7). Isso prova o Teorema 3.2.2. \square

Exemplo 3.2.3. Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto com $x_0 \in \Omega$, seja $\Lambda \subset \mathbb{R}^N$ um conjunto aberto com $0 \in \Lambda$, e seja $f : \Lambda \times \Omega \rightarrow \mathbb{R}$ uma função da classe C^2 . Para $\lambda \in \Lambda$ definimos a C^2 -função $f_\lambda : \Omega \rightarrow \mathbb{R}$ através de

$$f_\lambda(x) := f(\lambda, x) \quad \text{para } x \in \Omega.$$

Seja $x_0 \in \Omega_0$ um **ponto crítico não-degenerado** de f_0 , ou seja

$$df_0(x_0) = 0, \quad \det(d^2f_0(x_0)) \neq 0. \quad (3.2.12)$$

(Veja também o Exercício 2.7.17 e o Exercício 3.1.17.) Então existem vizinhanças abertas $\Lambda_0 \subset \Lambda$ de $0 \in \mathbb{R}^N$ e $U_0 \subset \Omega$ de $x_0 \in \mathbb{R}^n$, assim como uma C^1 -aplicação $g : \Lambda_0 \rightarrow U_0$ com as seguintes propriedades.

(i) $g(0) = x_0$.

(ii) Para todos os $\lambda \in \Lambda_0$ e todos os $x \in U_0$ vale $x = g(\lambda) \iff df_\lambda(x) = 0$.

(iii) Para $\lambda \in \Lambda_0$ o valor $g(\lambda)$ é um ponto crítico não-degenerado de f_λ e

$$\frac{\partial g}{\partial \lambda_k}(\lambda) = - (d^2f_\lambda(g(\lambda)))^{-1} \frac{\partial \nabla f}{\partial \lambda_k}(\lambda, g(\lambda)), \quad k = 1, \dots, N.$$

(iv) Se f é suave, então g também é suave.

As afirmações (i-iii) seguem do Teorema 3.2.2, onde tem-se que substituir a tuplo $(n, m, x, y, \Omega, X, Y, Z, f)$ por $(N, n, \lambda, x, \Lambda \times \Omega, \mathbb{R}^N, \mathbb{R}^n, \mathbb{R}^n, \nabla f)$. Deixamos como exercício a parte (iv). Resumindo, o exemplo mostra que pontos críticos não-degenerados ficam preservados sob perturbações pequenos.

Exemplo 3.2.4 (Transformação de Legendre). Seja $\Omega \subset \mathbb{R}^n \times \mathbb{R}^n$ um conjunto aberto. Denotamos os elementos de Ω de

$$(x, v) = (x_1, \dots, x_n, v_1, \dots, v_n).$$

Seja $L : \Omega \rightarrow \mathbb{R}$ uma função suave a qual satisfaça a condição

$$\det \left(\left(\frac{\partial^2 L}{\partial v_i \partial v_j}(x, v) \right)_{i,j=1}^n \right) \neq 0 \quad (3.2.13)$$

para todos os $(x, v) \in \Omega$. Isto é a **condição de Legendre**. No seguinte utilizamos para $(x, v) \in \Omega$ a notação

$$\begin{aligned} \frac{\partial L}{\partial x}(x, v) &:= \left(\frac{\partial L}{\partial x_1}(x, v), \dots, \frac{\partial L}{\partial x_n}(x, v) \right) \\ \frac{\partial L}{\partial v}(x, v) &:= \left(\frac{\partial L}{\partial v_1}(x, v), \dots, \frac{\partial L}{\partial v_n}(x, v) \right). \end{aligned}$$

Ora, seja $(x_0, v_0) \in \Omega$ e $y_0 := \frac{\partial L}{\partial v}(x_0, v_0)$. Então segue da condição de Legendre (3.2.13) e do Teorema 3.2.2 que pode-se resolver o sistema de equações

$$y = \frac{\partial L}{\partial v}(x, v) \quad (3.2.14)$$

localmente com respeito a v . Ou seja, existem abertos $U, V, W \subset \mathbb{R}^n$ com

$$x_0 \in U, \quad v_0 \in V, \quad y_0 \in W, \quad U \times V \subset \Omega$$

assim como uma aplicação suave $g : U \times W \rightarrow V$, tal que $g(x_0, y_0) = v_0$ e a seguinte afirmação é satisfeita para todos os $x \in U$, $v \in V$ e $y \in W$:

$$v = g(x, y) \quad \iff \quad \frac{\partial L}{\partial v}(x, v) = y. \quad (3.2.15)$$

Ora definimos a **função hamiltoniana** $H : U \times W \rightarrow \mathbb{R}$ através de

$$H(x, y) := \sum_{j=1}^n y_j g_j(x, y) - L(x, g(x, y)) \quad (3.2.16)$$

para $x \in U$ e $y \in W$. Com ajuda da equação $\frac{\partial L}{\partial v_j}(x, g(x, y)) = y_j$ resultam então para as derivadas parciais da função hamiltoniana as fórmulas

$$\begin{aligned} \frac{\partial H}{\partial x_i}(x, y) &= \frac{\partial}{\partial x_i} \left(\sum_{j=1}^n y_j g_j(x, y) - L(x, g(x, y)) \right) \\ &= \sum_{j=1}^n y_j \frac{\partial g_j}{\partial x_i}(x, y) - \frac{\partial L}{\partial x_i}(x, g(x, y)) - \sum_{j=1}^n \frac{\partial L}{\partial v_j}(x, g(x, y)) \frac{\partial g_j}{\partial x_i}(x, y) \\ &= -\frac{\partial L}{\partial x_i}(x, g(x, y)) \\ \frac{\partial H}{\partial y_i}(x, y) &= \frac{\partial}{\partial y_i} \left(\sum_{j=1}^n y_j g_j(x, y) - L(x, g(x, y)) \right) \\ &= g_i(x, y) + \sum_{j=1}^n y_j \frac{\partial g_j}{\partial y_i}(x, y) - \sum_{j=1}^n \frac{\partial L}{\partial v_j}(x, g(x, y)) \frac{\partial g_j}{\partial y_i}(x, y) \\ &= g_i(x, y). \end{aligned}$$

Em resumo assim encontraremos para $x \in U$ e $y \in W$ as equações

$$\frac{\partial H}{\partial x_i}(x, y) = -\frac{\partial L}{\partial x_i}(x, g(x, y)), \quad \frac{\partial H}{\partial y_i}(x, y) = g_i(x, y). \quad (3.2.17)$$

Ora, seja $I \subset \mathbb{R}$ um intervalo aberto e $x : I \rightarrow U$ uma solução (suave) das equações de Euler–Lagrange

$$\frac{d}{dt} \frac{\partial L}{\partial v_i}(x(t), \dot{x}(t)) = \frac{\partial L}{\partial x_i}(x(t), \dot{x}(t)), \quad i = 1, \dots, n \quad (3.2.18)$$

com $\dot{x}(t) \in V$ e $y(t) \in W$ para todos os $t \in I$, onde $y(t)$ é definido através de

$$y_i(t) := \frac{\partial L}{\partial v_i}(x(t), \dot{x}(t)), \quad i = 1, \dots, n. \quad (3.2.19)$$

Então vale $g_i(x(t), y(t)) = \dot{x}_i(t)$ para $i = 1, \dots, n$ segundo (3.2.15), e dali segue segundo (3.2.17) que a aplicação $(x, y) : I \rightarrow U \times W$ resolve as **equações de Hamilton**

$$\dot{x}_i(t) = \frac{\partial H}{\partial y_i}(x(t), y(t)), \quad \dot{y}_i(t) = -\frac{\partial H}{\partial x_i}(x(t), y(t)) \quad (3.2.20)$$

para $i = 1, \dots, n$ e $t \in I$. Em resumo, isto significa que a transformação de Legendre (3.2.14) converte as equações de Euler–Lagrange (3.2.18) (uma equação diferencial ordinária de segunda ordem em n variáveis) nas equações de Hamilton (3.2.20) (uma equação diferencial ordinária de primeira ordem em $2n$ variáveis) com a função hamiltoniana (3.2.16), também chamado do **Hamiltoniano**. Além disso o Hamiltoniano H tem a significância física de *energia*, a qual fica constante ao longo da solução da equação (3.2.20), e assim também ao longo das soluções de (3.2.18) (Exercício).

Exemplo 3.2.5. Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto e seja $V : \Omega \rightarrow \mathbb{R}$ uma função suave. Então as **equações de Newton**

$$\ddot{x}(t) = -\nabla V(x(t)) \quad (3.2.21)$$

correspondem ao caso especial da equação (3.2.18), onde $L : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}$ é dado através de

$$L(x, v) = \frac{1}{2} \|v\|^2 - V(x) \quad (3.2.22)$$

(energia *cinética* menos *potencial*). Neste caso a equação (3.2.14) tem a forma $y = \frac{\partial L}{\partial v}(x, v) = v$, a função hamiltoniana é

$$H(x, y) = \frac{1}{2} \|y\|^2 + V(x) \quad (3.2.23)$$

(energia *cinética* mais *potencial*), e as equações de Hamilton não fornecem nada novo em comparação a (3.2.21). Porém resulta de (3.2.20) imediatamente o *Teorema da Preservação da Energia*. (Veja também o Exemplo 2.3.8.)

3.3 Subvariedades

A noção de uma subvariedade

Nesta seção será introduzido a noção de subvariedade (de um espaço vetorial normado da *dimensão finita*) e serão discutidos uns exemplos. Este conceito desempenha um papel importante para muitas áreas da matemática e da física. Neste manuscrito será utilizado na Seção 3.4 para a investigação de extremos de funções sujeitos a restrições, assim como no Capítulo 6 sobre a integração de funções em variedades e o Teorema da Divergência de Gauß e no Capítulo 7 sobre a integração de formas diferenciais e o Teorema de Stokes.

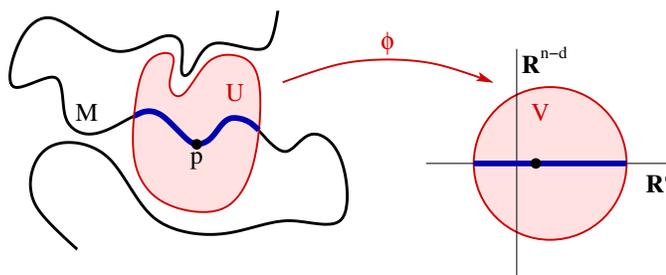


Figura 3.2: Uma subvariedade.

Definição 3.3.1 (Subvariedade). *Sejam d e n números inteiros, tais que $0 \leq d \leq n$, seja $\ell \in \mathbb{N} \cup \{\infty\}$, e seja X um espaço vetorial normado de dimensão n . Um subconjunto $M \subset X$ chama-se de **C^ℓ -subvariedade d -dimensional** de X , se satisfaz as seguintes condições. Para todo ponto $p \in M$ existe um subconjunto aberto $U \subset X$ com $p \in U$, um aberto $V \subset \mathbb{R}^n$, e um C^ℓ -difeomorfismo $\phi : U \rightarrow V$ com*

$$\phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\}). \quad (3.3.1)$$

(Veja a Figura 3.2. No caso de $\ell = \infty$ também falamos de uma **subvariedade suave d -dimensional** de X .) Nesta situação o C^ℓ -difeomorfismo ϕ é chamado de **carta de M** e o conjunto $U \cap M$ de **domínio de carta de M** . Se I é um conjunto de índices e $\phi_i : U_i \rightarrow V_i$ uma carta de M para todo $i \in I$, tal que os domínios de cartas cobrem o todo M , ou seja $M \subset \bigcup_{i \in I} U_i$, então chama-se a coleção $\{\phi_i\}_{i \in I}$ destas cartas de **atlas de M** .

Exemplo 3.3.2. No caso de $d = 0$ um subconjunto $M \subset X$ é uma C^ℓ -subvariedade 0-dimensional de X se e somente se M é um **subconjunto discreto** de X , ou seja, se para todo ponto $p \in M$ existe um subconjunto aberto $U \subset X$ com $U \cap M = \{p\}$.

Exemplo 3.3.3. No caso de $d = n = \dim(X)$ um subconjunto $M \subset X$ é uma C^ℓ -subvariedade n -dimensional de X se e somente se M é um subconjunto aberto de X .

Exemplo 3.3.4. A esfera unitária

$$\mathbb{S}^{n-1} = \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 = 1 \right\}$$

é uma subvariedade suave $(n-1)$ -dimensional do \mathbb{R}^n . Um atlas com $2n$ cartas $\phi_i^\pm : U_i^\pm \rightarrow V^\pm$ é dado através de

$$\begin{aligned} U_i^\pm &:= \left\{ x \in \mathbb{R}^n \mid \pm x_i > 0, \sum_{j \neq i} x_j^2 < 1 \right\} \\ V^\pm &:= \left\{ y \in \mathbb{R}^n \mid \sum_{i=1}^{n-1} y_i^2 < 1, \sqrt{1 - \sum_{i=1}^{n-1} y_i^2} \pm y_n > 0 \right\} \\ \phi_i^\pm(x) &:= \left(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n, x_i \mp \sqrt{1 - \sum_{j \neq i} x_j^2} \right) \end{aligned}$$

para $i = 1, \dots, n$ e $x \in U_i^\pm$.

Exemplo 3.3.5. Sejam X e Y espaços vetoriais normados com $\dim(X) = d$ e $\dim(Y) = m$, seja $\Omega \subset X$ um subconjunto aberto, e seja $g : \Omega \rightarrow Y$ uma C^ℓ -aplicação. Então o conjunto

$$M := \text{graph}(g) = \{(x, y) \in X \times Y \mid x \in \Omega, y = g(x)\}$$

é uma C^ℓ -subvariedade d -dimensional de $X \times Y$. Possui um atlas o qual é composto de uma única carta $\phi : U \rightarrow V$. Para esse fim escolhemos isomorfismos de espaços vetoriais $i : X \rightarrow \mathbb{R}^d$ e $j : Y \rightarrow \mathbb{R}^m$, e definimos

$$U := \Omega \times Y, \quad V := i(\Omega) \times \mathbb{R}^m, \quad \phi(x, y) := (i(x), j(y - g(x)))$$

para $x \in \Omega$ e $y \in Y$.

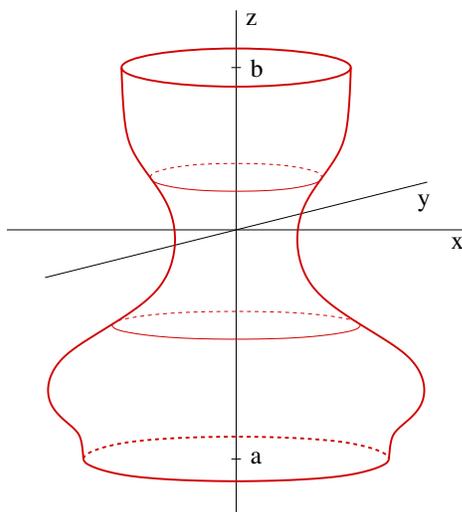


Figura 3.3: Uma superfície de rotação.

Exemplo 3.3.6. Sejam $a < b$ números reais e seja $r : (a, b) \rightarrow (0, \infty)$ uma função suave. Então o conjunto

$$M_r := \{(x, y, z) \in \mathbb{R}^3 \mid a < z < b, x^2 + y^2 = r(z)^2\} \quad (3.3.2)$$

é uma subvariedade suave 2-dimensional do \mathbb{R}^3 . Toda subvariedade desta forma é chamada de **superfície de rotação** (veja a Figura 3.3). **Exercício:** Construa um atlas de M_r .

Exemplo 3.3.7. a) O conjunto

$$S := \{(x, y) \in \mathbb{R}^2 \mid y \neq 0 \Rightarrow 1/y \in \mathbb{N}\}$$

não é uma subvariedade do \mathbb{R}^2 . Para a prova pode-se utilizar a noção conexo no Apêndice B. Se $M \subset X$ é uma subvariedade e $p \in M$, então pode-se escolher uma carta $\phi : U \rightarrow V$ de M com $p \in U$, tal que $V = B_\varepsilon(\phi(p))$ é uma bola no \mathbb{R}^n com centro $\phi(p)$. Então $V \cap (\mathbb{R}^d \times \{0\})$ é um subconjunto conexo do \mathbb{R}^n e assim, segundo (3.3.1), o domínio de carta $U \cap M$ também é conexo. Mas no nosso exemplo o conjunto $U \cap S$ é desconexo para toda vizinhança aberta $U \subset \mathbb{R}^2$ do ponto $(0, 0) \in S$ (Exercício), e dali S não pode ser uma subvariedade do \mathbb{R}^2 . b) Mas $\tilde{S} := S \setminus (\mathbb{R} \times \{0\})$ é uma subvariedade.

Exemplo 3.3.8. O conjunto $S := \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ não é uma subvariedade do \mathbb{R}^2 . Para a prova pode-se utilizar o seguinte teorema.

Valores regulares

Teorema 3.3.9. *Seja X um espaço vetorial normado de dimensão n e $M \subset X$ um subconjunto não-vazio. Seja $\ell \in \mathbb{N} \cup \{\infty\}$ e seja d um número inteiro com $0 \leq d < n$. Então as seguintes afirmações são equivalentes.*

(i) *M é uma C^ℓ -subvariedade d -dimensional de X .*

(ii) *Para todo ponto $p \in M$ existe um subconjunto aberto $U \subset X$ com $p \in U$ e C^ℓ -funções $f_1, \dots, f_{n-d} : U \rightarrow \mathbb{R}$, tal que*

$$U \cap M = \{x \in U \mid f_1(x) = f_2(x) = \dots = f_{n-d}(x) = 0\} \quad (3.3.3)$$

e as funcionais lineares $df_1(x), df_2(x), \dots, df_{n-d}(x) \in \mathcal{L}(X, \mathbb{R})$ são linearmente independente para todo ponto $x \in U \cap M$.

A prova utiliza o seguinte lema da Álgebra Linear.

Lema 3.3.10. *Seja X um espaço vetorial e sejam $\Lambda_1, \dots, \Lambda_m : X \rightarrow \mathbb{R}$ funcionais lineares. Então as seguintes afirmações são equivalentes.*

(i) *As funcionais lineares $\Lambda_1, \dots, \Lambda_m$ são linearmente independente.*

(ii) *Existem vetores $x_1, \dots, x_m \in X$ com $\Lambda_i(x_j) = \delta_{ij}$ para $i, j = 1, \dots, m$.*

(iii) *A aplicação $X \rightarrow \mathbb{R}^m : x \mapsto (\Lambda_1(x), \dots, \Lambda_m(x))$ é surjetiva.*

Demonstração. Provaremos (i) \implies (ii) por indução sobre m . Para $m = 1$ independência linear significa que $\Lambda_1 \neq 0$, e daí existe um vetor $x_1 \in X$ com $\Lambda_1(x_1) = 1$. Ora, seja $m \geq 2$ e suponhamos que a afirmação é válida para $m - 1$. Então provaremos

$$\bigcap_{j \neq i} \ker \Lambda_j \not\subset \ker \Lambda_i \quad \text{para } i = 1, \dots, m. \quad (3.3.4)$$

Pela hipótese da indução existem vetores $\xi_1, \dots, \xi_{m-1} \in X$ com $\Lambda_i(\xi_j) = \delta_{ij}$ para $i, j = 1, \dots, m - 1$. Se a afirmação (3.3.4) não vale para $i = m$, então resulta para todo $x \in X$ que $x - \sum_{i=1}^{m-1} \Lambda_i(x)\xi_i \in \bigcap_{j=1}^{m-1} \ker \Lambda_j \subset \ker \Lambda_m$ e daí vale $\Lambda_m(x) = \sum_{i=1}^m \Lambda_i(x)\Lambda_m(\xi_i)$, de onde segue $\Lambda_m = \sum_{i=1}^{m-1} \Lambda_m(\xi_i)\Lambda_i$, em contradição a (i). Com isso (3.3.4) fica provado para $i = m$. Através de trocar i e m recebemos (3.3.4) para todo i , e dali segue imediatamente parte (ii).

Provaremos (ii) \implies (iii). Sejam $x_1, \dots, x_m \in X$ como na parte (ii) e seja $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$. Então o vetor $x := \sum_{j=1}^m \lambda_j x_j$ satisfaz a condição $\Lambda_i(x) = \sum_{j=1}^m \lambda_j \Lambda_i(x_j) = \sum_{j=1}^m \lambda_j \delta_{ij} = \lambda_i$ para $i = 1, \dots, m$.

Provaremos (iii) \implies (i). Sejam $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ com $\sum_{i=1}^m \lambda_i \Lambda_i = 0$. Segundo (iii) existe então um vetor $x \in X$ com $\Lambda_i(x) = \lambda_i$ para $i = 1, \dots, m$. Segue daí $0 = \sum_{i=1}^m \lambda_i \Lambda_i(x) = \sum_{i=1}^m \lambda_i^2$ e dali $\lambda_i = 0$ para todo i . \square

Demonstração do Teorema 3.3.9. Provaremos (i) \implies (ii). Seja $p \in M$. Segundo (i) existe um aberto $U \subset X$ com $p \in U$, um aberto $V \subset \mathbb{R}^n$, e um C^ℓ -difeomorfismo $\phi = (\phi_1, \dots, \phi_n) : U \rightarrow V$ satisfazendo (3.3.1). Como a aplicação linear $d\phi(x) : X \rightarrow \mathbb{R}^n$ é bijetiva, veja Lema 3.1.2, as funcionais lineares $d\phi_1(x), \dots, d\phi_n(x) : X \rightarrow \mathbb{R}$ são linearmente independente para todo $x \in U$, veja Lema 3.3.10. Além disso, segundo (3.3.1), vale para todo $x \in U$

$$x \in M \cap U \iff \phi(x) \in \mathbb{R}^d \times \{0\} \iff \phi_{d+1}(x) = \dots = \phi_n(x) = 0.$$

Com isto as funções $f_i := \phi_{d+i}$, para $i = 1, \dots, n-d$, satisfazem parte (ii).

Provaremos (ii) \implies (i). Sejam $p \in M$ e $f_1, \dots, f_{n-d} : U \rightarrow \mathbb{R}$ como em (ii). Então existem funcionais lineares $\lambda_1, \dots, \lambda_d : X \rightarrow \mathbb{R}$, tal que

$$\lambda_1, \dots, \lambda_d, df_1(p), \dots, df_{n-d}(p)$$

é uma base do espaço vetorial $\mathcal{L}(X, \mathbb{R})$ de dimensão n . Ora definimos a aplicação $\phi : U \rightarrow \mathbb{R}^n$ para $x \in U$ através de

$$\phi(x) := (\lambda_1(x), \dots, \lambda_d(x), f_1(x), \dots, f_{n-d}(x)).$$

É C^ℓ segundo o Teorema 2.5.9, sua derivada $d\phi(p) : X \rightarrow \mathbb{R}^n$ no ponto p é

$$d\phi(p)\xi = (\lambda_1(\xi), \dots, \lambda_d(\xi), df_1(p)\xi, \dots, df_{n-d}(p)\xi).$$

Como as funcionais lineares $\lambda_1, \dots, \lambda_d, df_1(p), \dots, df_{n-d}(p)$ são linearmente independente, $d\phi(p) : X \rightarrow \mathbb{R}^n$ é bijetiva, segundo o Lema 3.3.10. Segundo o Teorema 3.1.11 existe dali um aberto $U_0 \subset X$ com $p \in U_0$, tal que o conjunto $V_0 := \phi(U_0) \subset \mathbb{R}^n$ é aberto e $\phi_0 := \phi|_{U_0} : U_0 \rightarrow V_0$ um C^ℓ -difeomorfismo. Daí

$$\begin{aligned} \phi_0(U_0 \cap M) &= \{\phi(x) \mid x \in U_0, f_1(x) = \dots = f_{n-d}(x) = 0\} \\ &= V_0 \cap (\mathbb{R}^d \times \{0\}) \end{aligned}$$

e com isso ϕ_0 é uma carta de M . Como $p \in M$ foi escolhido arbitrariamente, isto significa que M é uma C^ℓ -subvariedade d -dimensional de X . \square

Definição 3.3.11 (Valor regular). *Sejam X, Y espaços de Banach, seja $\Omega \subset X$ um conjunto aberto, e seja $f : \Omega \rightarrow Y$ uma aplicação continuamente diferenciável. Um vetor $y \in Y$ chama-se de **valor regular de f** se em todo ponto $x \in \Omega$ com valor $f(x) = y$, primeiro, a derivada $df(x) \in \mathcal{L}(X, Y)$ é uma aplicação sobrejetiva e, segundo, seu núcleo $X_0 := \ker df(x)$ possui um complemento topológico¹ X_1 . (A segunda condição é vazia se $\dim X < \infty$.) Um vetor $y \in Y$ chama-se de **valor singular de f** , se y não é regular.*

¹ Um subespaço linear fechado $X_1 \subset X$, tal que $X_0 \cap X_1 = \{0\}$ e $X_0 \cup X_1$ gera X .

Corolário 3.3.12 (Valores regulares). *Sejam $\ell \in \mathbb{N} \cup \{\infty\}$ e $d, m, n \in \mathbb{Z}$ com $0 \leq d \leq n$ e $m := n - d$. Sejam X e Y espaços vetoriais das dimensões $n = \dim(X)$ e $m = \dim(Y)$, seja $\Omega \subset X$ um conjunto aberto, seja $f : \Omega \rightarrow Y$ uma C^ℓ -aplicação, e seja $y \in Y$ um valor regular de f . Então*

$$M := f^{-1}(y) = \{x \in \Omega \mid f(x) = y\}$$

é uma C^ℓ -subvariedade d -dimensional de X .

Demonstração. No caso de $d = n$ tem-se $Y = \{0\}$ e assim $M = \Omega$ é um subconjunto aberto de X como afirmado. Seja então $d < n$ e daí $m = n - d > 0$. Sejam e_1, \dots, e_m uma base de Y e $f_1, \dots, f_m : \Omega \rightarrow \mathbb{R}$ funções definidas assim

$$f_1(x)e_1 + \dots + f_m(x)e_m := f(x) - y \quad \text{para } x \in \Omega.$$

Estas funções são C^ℓ segundo o Teorema 2.3.2. Além disso, vale

$$M = \{x \in \Omega \mid f(x) = y\} = \{x \in \Omega \mid f_1(x) = \dots = f_m(x) = 0\}.$$

Ora a aplicação linear $X \rightarrow \mathbb{R}^m : \xi \mapsto (df_1(x)\xi, \dots, df_m(x)\xi)$ é surjetiva para todo $x \in M$, como y é um valor regular de f . Dali, segundo o Lema 3.3.10, as funcionais lineares $df_1(x), \dots, df_m(x)$ são linearmente independente para todo $x \in M$. Com isso o Corolário 3.3.12 segue do Teorema 3.3.9. \square

Comentário 3.3.13. (i) Corolário 3.3.12 também fica válido se $y \notin f(\Omega)$, e daí M é o conjunto vazio. Porque o conjunto vazio $M = \emptyset \subset X$ é uma subvariedade d -dimensional para todo d .

(ii) Sejam X, Y espaços vetoriais normados e $\Omega \subset X$ um subconjunto aberto. Seja $f : \Omega \rightarrow Y$ uma C^ℓ -aplicação com $\ell \geq 1$ e $\ell \geq \dim(X) - \dim(Y)$. Então o **Teorema de Sard** diz que *quase todos os* $y \in Y$ são valores regulares de f . Isto significa que o conjunto dos valores singulares de f é um conjunto Lebesgue-nulo, ou seja, que pode ser coberto com um número contável de blocos com volume total arbitrariamente pequeno. Particularmente o conjunto dos valores regulares de f é denso em Y .

(iii) O Teorema de Sard também vale no caso de $\dim(Y) > \dim(X)$. Aqui o conjunto dos valores regulares é o complemento do conjunto das imagens de f : A derivada $df(x) : X \rightarrow Y$ nunca pode ser surjetiva por razões de dimensão. No caso de $\dim(Y) > \dim(X)$ o Teorema de Sard diz portanto que o conjunto imagem $f(\Omega)$ de toda C^1 -aplicação $f : \Omega \rightarrow Y$ é um conjunto Lebesgue-nulo em Y . Isto já segue de Teorema 5.6.7 (iii), como Ω é uma união contável de conjuntos compactos. A prova no caso relevante $\dim(Y) \leq \dim(X)$ é considerável mais difícil. Para funções suaves encontra-se uma prova em [3].

Exemplos

Exemplo 3.3.14. Que a superfície de rotação $M_r \subset \mathbb{R}^3$ é uma subvariedade suave 2-dimensional do \mathbb{R}^3 segue imediatamente do Corolário 3.3.12 com

$$\Omega := \{(x, y, z) \in \mathbb{R}^3 \mid a < z < b\}$$

e a função suave $f : \Omega \rightarrow \mathbb{R}$ definida através de

$$f(x, y, z) := x^2 + y^2 - r(z)^2$$

para $(x, y, z) \in \Omega$. Como $r(z) > 0$ para todos os $(x, y, z) \in \Omega$, o número 0 é um valor regular de f , e daí $M_r = f^{-1}(0)$ é uma subvariedade suave 2-dimensional do \mathbb{R}^3 .

Exemplo 3.3.15. O conjunto $S = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ no Exemplo 3.3.8 não é uma subvariedade, isto pode ser provado através do Teorema 3.3.9. Como S é nem aberto nem discreto o único candidato para a dimensão seria o número $d = 1$. Mas se $f_1 : U \rightarrow \mathbb{R}$ seja uma função C^1 num conjunto aberto $U \subset \mathbb{R}^2$ com $(0, 0) \in U \cap M = \{(x, y) \in U \mid f_1(x, y) = 0\}$, então vale $\frac{\partial f_1}{\partial x}(0, 0) = \frac{\partial f_1}{\partial y}(0, 0) = 0$. O ponto $p = (0, 0) \in S$ lesa daí a condição (ii) no Teorema 3.3.9, e assim S não é uma subvariedade do \mathbb{R}^2 .

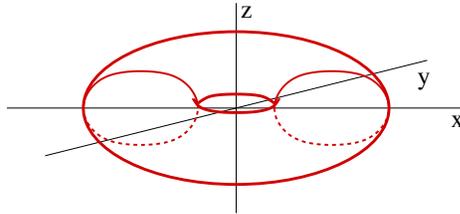


Figura 3.4: Um 2-toro no \mathbb{R}^3 .

Exemplo 3.3.16. Seja $0 < r < 1$ e seja $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ a função suave

$$f(x, y, z) := (x^2 + y^2 - 1 + r^2 - z^2)^2 - 4(x^2 + y^2)(r^2 - z^2).$$

Então 0 é um valor regular de f e dali $M := f^{-1}(0)$ é uma subvariedade suave do \mathbb{R}^3 (veja a Figura 3.4). A aplicação

$$\mathbb{S}^1 \times \mathbb{S}^1 \rightarrow M : (e^{is}, e^{it}) \mapsto \begin{pmatrix} (1 + r \cos(s)) \cos(t) \\ (1 + r \cos(s)) \sin(t) \\ r \sin(s) \end{pmatrix}$$

identifica o 2-toro padrão $\mathbb{S}^1 \times \mathbb{S}^1 \subset \mathbb{C} \times \mathbb{C}$ com M .

Exemplo 3.3.17. Se $M \subset \mathbb{R}^n$ e $M' \subset \mathbb{R}^{n'}$ são subvariedades suaves, então o produto $M \times M'$ é uma subvariedade suave de $\mathbb{R}^n \times \mathbb{R}^{n'} \cong \mathbb{R}^{n+n'}$. Particularmente o n -ésimo produto

$$\mathbb{T}^n := \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$$

é uma subvariedade suave do \mathbb{C}^n . Isto é o **n -toro padrão**.

Exemplo 3.3.18. O grupo ortogonal

$$\mathrm{O}(n) := \{A \in \mathbb{R}^{n \times n} \mid A^T A = \mathbb{1}\}$$

é uma subvariedade suave de $\mathbb{R}^{n \times n}$ e

$$\dim(\mathrm{O}(n)) = \frac{n(n-1)}{2}.$$

Para a prova consideramos os espaços vetoriais normados

$$X := \mathbb{R}^{n \times n}, \quad Y := \mathcal{S}_n := \{S \in \mathbb{R}^{n \times n} \mid S^T = S\}.$$

e a aplicação $f : X \rightarrow Y$ com

$$f(A) := A^T A \quad \text{para } A \in X.$$

Provaremos que a matriz unitária $\mathbb{1} \in Y$ é um valor regular de f . A derivada de f no ponto $A \in X$ é a aplicação linear $df(A) : X \rightarrow Y$ dada através de

$$df(A)\widehat{A} = A^T \widehat{A} + \widehat{A}^T A.$$

Seja $A \in \mathrm{O}(n)$ e $S = S^T \in \mathcal{S}_n$, e defina

$$\widehat{A} := \frac{1}{2}AS \in \mathbb{R}^{n \times n}.$$

Então vale

$$df(A)\widehat{A} = A^T \widehat{A} + \widehat{A}^T A = \frac{1}{2}A^T AS + \frac{1}{2}(AS)^T A = \frac{1}{2}(S + S^T) = S.$$

Isto mostra que a aplicação linear $df(A) : X \rightarrow Y$ é surjetiva para todas as matrizes $A \in \mathrm{O}(n)$. Assim $\mathbb{1} \in Y$ é um valor regular de f , como afirmado, e dali, segundo o Corolário 3.3.12, a pre-imagem $\mathrm{O}(n) = f^{-1}(\mathbb{1})$ é uma subvariedade suave de X de dimensão

$$\dim(\mathrm{O}(n)) = \dim(X) - \dim(Y) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Exemplo 3.3.19. Seja $0 \neq S = S^T \in \mathbb{R}^{n \times n}$ uma matriz simétrica não-nula e (compare Exemplo 2.1.8) seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a função suave

$$f(x) := x^T S x \quad \text{para } x \in \mathbb{R}^n.$$

Então todo número $c \in \mathbb{R} \setminus \{0\}$ é um valor regular de f (Exercício). Por isso

$$Q := f^{-1}(c) = \{x \in \mathbb{R}^n \mid x^T S x = c\}$$

é, para todo número real $c \neq 0$, uma subvariedade suave do \mathbb{R}^n de dimensão $(n - 1)$. No seguinte apresentamos uns casos especiais desta construção.

(a) Com $S = \text{diag}(1/r_1^2, \dots, 1/r_n^2)$ e $c = 1$ recebe-se o elipsoide

$$E(r_1, \dots, r_n) := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n \frac{x_i^2}{r_i^2} = 1 \right\}.$$

(b) Em (a) resulta com $r_i = 1$ a esfera unitária $Q = \mathbb{S}^{n-1}$ no Exemplo 3.3.4.

(c) A hipérbole $Q = \{(x, y) \in \mathbb{R}^2 \mid xy = c\}$ é uma subvariedade suave 1-dimensional de \mathbb{R}^2 para $c \neq 0$, mas não para $c = 0$ (Exemplo 3.3.14).

(d) Com $f(x, y) := \|x - y\|^2$ para $x, y \in \mathbb{R}^3$ e $c = r^2$ com $r > 0$ resulta que o conjunto $Q := \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mid \|x - y\| = r\}$ é, para todo $r > 0$, uma subvariedade suave 5-dimensional de \mathbb{R}^6 .

(e) Com uma modificação do Exemplo em (d) pode-se provar que o conjunto

$$M := \{(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \mid \|x\| = \|x - y\| = \|y - z\| = 1\}$$

é uma subvariedade suave 6-dimensional de \mathbb{R}^9 . Isto é o espaço configuração de um pêndulo triplo.

Exemplo 3.3.20. O conjunto

$$S := \{(x^2, y^2, z^2, yz, zx, xy) \mid x, y, z \in \mathbb{R}, x^2 + y^2 + z^2 = 1\} \quad (3.3.5)$$

é uma subvariedade suave 2-dimensional de \mathbb{R}^6 (Exercício). Este exemplo origina-se de Jakob Steiner (1796–1863). Trata-se aqui de um *mergulho* do *espaço projetivo real* 2-dimensional (o conjunto de todas as retas no \mathbb{R}^3 passando a origem) no \mathbb{R}^6 . Projetando esta subvariedade nas últimas três coordenadas, recebe-se a chamada **superfície dos romanos**

$$\begin{aligned} R &= \{(yz, zx, xy) \mid x, y, z \in \mathbb{R}, x^2 + y^2 + z^2 = 1\} \\ &= \{(\xi, \eta, \zeta) \in \mathbb{R}^3 \mid \eta^2 \zeta^2 + \zeta^2 \xi^2 + \xi^2 \eta^2 = \xi \eta \zeta\}. \end{aligned}$$

Isso não é uma subvariedade do \mathbb{R}^3 (Exercício).

O espaço tangente

Definição 3.3.21 (Vetor tangente).

Seja X um espaço vetorial normado de dimensão finita e seja $M \subset X$ uma C^1 -subvariedade d -dimensional de X . Seja $p \in M$. Chama-se um vetor $v \in X$ de **vetor tangente de M no ponto p** , se exista um $\varepsilon > 0$ e uma C^1 -aplicação $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ com

$$\gamma(0) = p, \quad \dot{\gamma}(0) = v. \quad (3.3.6)$$

(Veja a Figura 3.5.) Chama-se o conjunto

$$T_p M := \{v \in X \mid v \text{ é um vetor tangente de } M \text{ no ponto } p\} \quad (3.3.7)$$

o **espaço tangente de M no ponto p** .

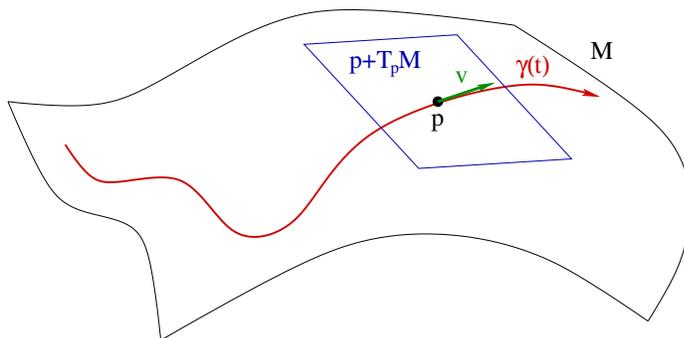


Figura 3.5: O espaço tangente.

Teorema 3.3.22 (O espaço tangente). *Seja X um espaço vetorial normado de dimensão n e seja $M \subset X$ uma C^ℓ -subvariedade d -dimensional de X com $\ell \in \mathbb{N} \cup \{\infty\}$. Seja $p \in M$. Então vale o seguinte.*

(i) *Se $U \subset X$ e $V \subset \mathbb{R}^n$ são conjuntos abertos com $p \in U$ e se $\phi : U \rightarrow V$ é um C^ℓ -difeomorfismo com $\phi(U \cap M) = V \cap (\mathbb{R}^d \times \{0\})$, então vale*

$$T_p M = d\phi(p)^{-1}(\mathbb{R}^d \times \{0\}). \quad (3.3.8)$$

(ii) *Se $U \subset X$ é aberto com $p \in U$ e se $f : U \rightarrow \mathbb{R}^{n-d}$ é uma C^ℓ -aplicação com $U \cap M = f^{-1}(0)$ e tal que $df(p) : X \rightarrow \mathbb{R}^{n-d}$ é sobrejetivo, então vale*

$$T_p M = \ker df(p). \quad (3.3.9)$$

(iii) $T_p M$ é um subespaço linear de X de dimensão d .

(iv) Se $v \in T_p M$, então existe uma C^ℓ -curva $\gamma : \mathbb{R} \rightarrow M$ satisfazendo (3.3.6).

Demonstração. Na situação da parte (i) provaremos para já a inclusão

$$d\phi(p)^{-1}(\mathbb{R}^d \times \{0\}) \subset T_p M. \quad (3.3.10)$$

Seja então dado $v \in d\phi(p)^{-1}(\mathbb{R}^d \times \{0\})$ e seja

$$x := \phi(p) \in V \cap (\mathbb{R}^d \times \{0\}), \quad \xi := d\phi(p)v \in \mathbb{R}^d \times \{0\}. \quad (3.3.11)$$

Como $V \subset \mathbb{R}^n$ é aberto, existe um número $\varepsilon > 0$ tal que $x+t\xi \in V \cap (\mathbb{R}^d \times \{0\})$ para todo $t \in (-\varepsilon, \varepsilon)$. Ora definimos $\gamma : (-\varepsilon, \varepsilon) \rightarrow U \cap M$ através de

$$\gamma(t) := \phi^{-1}(x + t\xi) \quad \text{para } -\varepsilon < t < \varepsilon. \quad (3.3.12)$$

Isso é uma C^ℓ -aplicação com

$$\gamma(0) = \phi^{-1}(x) = p, \quad \dot{\gamma}(0) = d\phi^{-1}(x)\xi = d\phi(p)^{-1}\xi = v.$$

Aqui a penúltima equação segue do Lema 3.1.2 e a última de (3.3.11). Assim $v \in T_p M$ e com isso (3.3.10) fica provado.

Na situação da parte (ii) provaremos a inclusão

$$T_p M \subset \ker df(p). \quad (3.3.13)$$

Ora seja $v \in T_p M$ e $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ uma C^1 -curva satisfazendo (3.3.6). Ora, seja $f : U \rightarrow \mathbb{R}^{n-d}$ como na parte (ii). Como $\gamma(0) = p \in U$ e U é um subconjunto aberto de X , existe um número $0 < \delta \leq \varepsilon$, tal que $\gamma(t) \in U \cap M$ para $-\delta < t < \delta$. Segue dali $f(\gamma(t)) = 0$ para $-\delta < t < \delta$ e assim vale

$$df(p)v = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = 0.$$

Portanto $v \in \ker df(p)$ e assim (3.3.13) fica provado.

Segundo (3.3.10) e (3.3.13) vale $d\phi(p)^{-1}(\mathbb{R}^d \times \{0\}) \subset T_p M \subset \ker df(p)$. Segundo a hipótese $d\phi(p)^{-1}(\mathbb{R}^d \times \{0\})$ e $\ker df(p)$ são ambos subespaços lineares de X de dimensão d e por isso tem que coincidir. Com isso as partes (i), (ii), e (iii) ficam provados.

Provaremos parte (iv). Seja então $v \in T_p M$. Então $d\phi(p)v \in \mathbb{R}^d \times \{0\}$ segundo a parte (i), e daí as equações (3.3.11) e (3.3.12) definem uma C^ℓ -curva $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ a qual satisfaz (3.3.6). Dali segue que a fórmula

$$\beta(t) := \gamma \left(\frac{\varepsilon t}{\sqrt{\varepsilon^2 + t^2}} \right) \quad \text{para } t \in \mathbb{R} \quad (3.3.14)$$

define uma C^ℓ -curva $\beta : \mathbb{R} \rightarrow M$ com $\beta(0) = \gamma(0) = p$, $\dot{\beta}(0) = \dot{\gamma}(0) = v$. Com isso a parte (iv) e o Teorema 3.3.22 ficam provados. \square

Exemplo 3.3.23. No caso de $d = 0$ (Exemplo 3.3.2) o espaço tangente de um conjunto discreto $M \subset X$ no qualquer ponto $p \in M$ é o subespaço $T_p M = \{0\}$.

Exemplo 3.3.24. No caso de $d = n$ (Exemplo 3.3.3) o espaço tangente de um conjunto aberto $M \subset X$ no qualquer ponto $p \in M$ é o subespaço $T_p M = X$.

Exemplo 3.3.25. Em Exemplo 3.3.18 o espaço tangente do grupo ortogonal $O(n)$ num ponto $A \in O(n)$ é dado através de

$$T_A O(n) = \left\{ \widehat{A} \in \mathbb{R}^{n \times n} \mid A^T \widehat{A} + \widehat{A}^T A = 0 \right\}$$

segundo a parte (ii) de Teorema 3.3.22. Para $A = \mathbb{1}$ recebe-se o espaço

$$\mathfrak{o}(n) := T_{\mathbb{1}} O(n) = \left\{ \widehat{A} \in \mathbb{R}^{n \times n} \mid \widehat{A} + \widehat{A}^T = 0 \right\}$$

das matrizes anti-simétricas. Este espaço tangente possui a propriedade que o comutador $[\widehat{A}, \widehat{B}] := \widehat{A}\widehat{B} - \widehat{B}\widehat{A}$ de duas matrizes $\widehat{A}, \widehat{B} \in \mathfrak{o}(n)$ é de novo um elemento de $\mathfrak{o}(n)$. Além disso, vale $\exp(t\widehat{A}) \in O(n)$ para todos os $\widehat{A} \in \mathfrak{o}(n)$ e todos os $t \in \mathbb{R}$ (Exercício).

Exemplo 3.3.26. O espaço tangente da subvariedade $(n - 1)$ -dimensional $Q = \{x \in \mathbb{R}^n \mid x^T S x = c\}$ no Exemplo 3.3.19 é no ponto $x \in Q$ dado por

$$T_x Q = \{\xi \in \mathbb{R}^n \mid \langle Sx, \xi \rangle = 0\} = (Sx)^\perp$$

segundo a parte (ii) de Teorema 3.3.22. Particularmente vale $T_x \mathbb{S}^{n-1} = x^\perp$ para $x \in \mathbb{S}^{n-1}$.

Exemplo 3.3.27. Em Exemplo 3.3.20 o espaço tangente da subvariedade $S \subset \mathbb{R}^6$ no ponto $p = (x^2, y^2, z^2, yz, zx, xy) \in S$ com $x^2 + y^2 + z^2 = 1$ é o subespaço 2-dimensional

$$T_p S = \left\{ \left(\begin{array}{c} 2x\xi \\ 2y\eta \\ 2z\zeta \\ y\zeta + z\eta \\ z\xi + x\zeta \\ x\eta + y\xi \end{array} \right) \mid \xi, \eta, \zeta \in \mathbb{R}, x\xi + y\eta + z\zeta = 0 \right\}. \quad (3.3.15)$$

Aqui mostra-se a inclusão “ \subset ” através de diferenciar curvas em S . A equação segue então do fato que, segundo a parte (iii) de Teorema 3.3.22, ambos lados da equação (3.3.15) são subespaços lineares 2-dimensionais do \mathbb{R}^6 . Com este método pode-se determinar espaços tangente em muitos exemplos num jeito simples.

3.4 Extremas sob restrições

Nesta seção investigamos a questão como pode-se encontrar elementos de um conjunto U as quais minimizam uma função $f : U \rightarrow \mathbb{R}$ sob restrições dado na forma de condições $h_i(x) = 0$. Aqui $U \subset \mathbb{R}^n$ é um conjunto aberto e suponhamos que $f : U \rightarrow \mathbb{R}$ e $h = (h_1, \dots, h_m) : U \rightarrow \mathbb{R}^m$ são diferenciáveis. Além disso, lembramos a notação

$$\nabla f(x) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \in \mathbb{R}^n$$

para o gradiente de f no ponto $x \in U$. De $\|x\|$ denotamos sempre a norma euclidiana de um vetor $x \in \mathbb{R}^n$.

Teorema 3.4.1 (Lagrange). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, sejam $f : U \rightarrow \mathbb{R}$ e $h = (h_1, \dots, h_m) : \mathbb{R}^n \supset U \rightarrow \mathbb{R}^m$ aplicações C^1 , seja $0 \in \mathbb{R}^m$ um valor regular de h , e seja*

$$M := h^{-1}(0) = \{x \in U \mid h(x) = 0\}. \quad (3.4.1)$$

Se um elemento $x_0 \in M$ satisfaça a condição

$$f(x_0) \leq f(x) \quad \text{para todos os } x \in M, \quad (3.4.2)$$

então existem números reais $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ com

$$\nabla f(x_0) = \sum_{i=1}^m \lambda_i \nabla h_i(x_0). \quad (3.4.3)$$

Estes números λ_i são chamados de **multiplicadores de Lagrange**.

Demonstração. Provaremos para já que

$$\nabla f(x_0) \in (T_{x_0}M)^\perp. \quad (3.4.4)$$

Ora, seja $v \in T_{x_0}M$. Então existe, segundo a parte (iv) de Teorema 3.3.22, uma C^1 -curva $\gamma : \mathbb{R} \rightarrow M \subset \mathbb{R}^n$ com $\gamma(0) = x_0$ und $\dot{\gamma}(0) = v$. Dali segue a desigualdade $f(\gamma(0)) \leq f(\gamma(t))$ para todo $t \in \mathbb{R}$ segundo (3.4.2), e daí vale

$$\langle \nabla f(x_0), v \rangle = df(x_0)v = df(\gamma(0))\dot{\gamma}(0) = \left. \frac{d}{dt} \right|_{t=0} f(\gamma(t)) = 0.$$

Aqui o último passo segue de Teorema 2.7.2. Com isso (3.4.4) fica provado.

Ora, seja $d := n - m$. Segundo o Corolário 3.3.12 o subconjunto $M \subset \mathbb{R}^n$ é uma C^1 -subvariedade d -dimensional de \mathbb{R}^n e segunda a parte (ii) de Teorema 3.3.22 seu espaço tangente no ponto $x_0 \in M$ é o subespaço linear d -dimensional

$$\begin{aligned} T_{x_0}M &= \ker dh(x_0) \\ &= \{\xi \in \mathbb{R}^n \mid dh(x_0)\xi = 0\} \\ &= \{\xi \in \mathbb{R}^n \mid \langle \nabla h_i(x_0), \xi \rangle = 0 \text{ para } i = 1, \dots, m\} \\ &= \bigcap_{i=1}^m \nabla h_i(x_0)^\perp. \end{aligned} \quad (3.4.5)$$

Que este subespaço possui dimensão $d = n - m$ segue dos teoremas mencionados, ou também do fato que os vetores $\nabla h_1(x_0), \dots, \nabla h_m(x_0)$ são linearmente independente (Lema 3.3.10). Como todo subespaço linear $V \subset \mathbb{R}^n$ satisfaz a equação $V = V^{\perp\perp}$, recebemos segundo (3.4.5) a equação

$$\begin{aligned} (T_{x_0}M)^\perp &= \left(\bigcap_{i=1}^m \nabla h_i(x_0)^\perp \right)^\perp \\ &= \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x_0) \mid \lambda_1, \dots, \lambda_n \in \mathbb{R} \right\}^{\perp\perp} \\ &= \left\{ \sum_{i=1}^m \lambda_i \nabla h_i(x_0) \mid \lambda_1, \dots, \lambda_n \in \mathbb{R} \right\}. \end{aligned} \quad (3.4.6)$$

Como $\nabla f(x_0)$ é ortogonal a $T_{x_0}M$, segundo (3.4.4), segue dali a afirmação de Teorema 3.4.1. \square

Exemplo 3.4.2. Este exemplo mostra que não pode-se abrir mão da hipótese em Teorema 3.4.1 que 0 é um valor regular de h . Sejam $f, h : \mathbb{R} \rightarrow \mathbb{R}$ as funções

$$f(x) := x, \quad h(x) := x^2$$

para $x \in \mathbb{R}$. Então 0 não é um valor regular de h . Todas as outras hipóteses de Teorema 3.4.1 são satisfeitas com $x_0 = 0 \in M = \{0\}$, porém neste exemplo $\nabla f(x_0) = 1$ não é um múltiplo real de $\nabla h(x_0) = 0$.

Exemplo 3.4.3 (Média geométrica e aritmética).

Seja $U \subset \mathbb{R}^n$ o quadrante positivo

$$U := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i > 0 \text{ para } i = 1, \dots, n\}$$

e sejam $f, h : \mathbb{R}^n \rightarrow \mathbb{R}$ as funções

$$f(x) := x_1 x_2 \cdots x_n, \quad h(x) := x_1 + x_2 + \cdots + x_n - 1 \quad (3.4.7)$$

para $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$. Então $M := U \cap h^{-1}(0)$ é o conjunto

$$M = \left\{ x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid \begin{array}{l} x_i > 0 \text{ para } i = 1, 2, \dots, n, \\ x_1 + x_2 + \cdots + x_n = 1 \end{array} \right\}. \quad (3.4.8)$$

A função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é contínua e o fecho \overline{M} de M é compacto. Dali existe um vetor $a = (a_1, a_2, \dots, a_n) \in \overline{M}$ com

$$f(x) \leq f(a) \quad \text{para todos os } x \in M. \quad (3.4.9)$$

Como f é positivo ao longo de M e anula-se ao longo de $\overline{M} \setminus M$, vale $a \in M$. Segundo o Teorema 3.4.1 existe um número $\lambda \in \mathbb{R}$ com $\nabla f(a) = \lambda \nabla h(a)$. Então vale

$$\frac{a_1 a_2 \cdots a_n}{a_i} = \frac{\partial f}{\partial x_i}(a) = \lambda \frac{\partial h}{\partial x_i}(a) = \lambda \quad (3.4.10)$$

para $i = 1, \dots, n$ e segue dali $a_1 = a_2 = \cdots = a_n = 1/n$. Existe então só um vetor $a \in M$ o qual satisfaz (3.4.10), e este também tem que satisfazer (3.4.9). Com isso temos provado a desigualdade

$$x_1 x_2 \cdots x_n \leq \left(\frac{1}{n}\right)^n \quad (3.4.11)$$

para todos os $x = (x_1, x_2, \dots, x_n) \in M$. Sejam ora x_1, x_2, \dots, x_n número positivos reais. Então $\frac{1}{x_1 + x_2 + \cdots + x_n}(x_1, x_2, \dots, x_n) \in M$ e segundo (3.4.11) vale

$$\frac{x_1 x_2 \cdots x_n}{(x_1 + x_2 + \cdots + x_n)^n} \leq \left(\frac{1}{n}\right)^n.$$

Daí em contrapartida segue $x_1 x_2 \cdots x_n \leq \left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)^n$, e por isso vale

$$\left(\prod_{i=1}^n x_i\right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^n x_i \quad (3.4.12)$$

para toda n -tuplo de números positivos reais x_i . Assim temos mostrado que a *média geométrica* é menor ou igual à *média aritmética* (como com certeza já é conhecido do curso Análise I.)

Exemplo 3.4.4 (Autovalores). Seja $S = S^T \in \mathbb{R}^{n \times n}$ uma matriz simétrica e sejam $f, h : \mathbb{R}^n \rightarrow \mathbb{R}$ as funções suaves

$$f(x) := \frac{1}{2}x^T Sx, \quad h(x) := \frac{1}{2}(\|x\|^2 - 1)$$

para $x \in \mathbb{R}^n$. Então 0 é um valor regular de h e $M := h^{-1}(0) = \mathbb{S}^{n-1}$ é a esfera unitária. Então existe um vetor $v_1 \in \mathbb{S}^{n-1}$ com

$$f(v_1) \leq f(x) \quad \text{para todos os } x \in \mathbb{S}^{n-1}.$$

Pelo Teorema 3.4.1 existe um número real $\lambda_1 \in \mathbb{R}$ com $\nabla f(v_1) = \lambda_1 \nabla h(v_1)$. Com $\nabla f(x) = Sx$ e $\nabla h(x) = x$ esta equação é equivalente a

$$Sv_1 = \lambda_1 v_1. \quad (3.4.13)$$

Com isso temos provado a existência de um autovalor real de S .

Ora consideramos um novo problema de extremo com a mesma função f sujeito a duas restrições dadas através de funções $h, h_1 : \mathbb{R}^n \rightarrow \mathbb{R}$ com

$$h_1(x) = \langle v_1, x \rangle \quad \text{para } x \in \mathbb{R}^n.$$

Então $(0, 0)$ é um valor regular de $(h, h_1) : \mathbb{R}^n \rightarrow \mathbb{R}^2$ e

$$\begin{aligned} M_1 &:= h^{-1}(0) \cap h_1^{-1}(0) \\ &= \mathbb{S}^{n-1} \cap v_1^\perp \\ &= \{x \in \mathbb{S}^{n-1} \mid \langle v_1, x \rangle = 0\}. \end{aligned}$$

Como M_1 é compacto, existe um vetor $v_2 \in M_1$ com

$$f(v_2) \leq f(x) \quad \text{para todos os } x \in M_1.$$

Pelo Teorema 3.4.1 existem dois multiplicadores de Lagrange $\lambda_2, \mu \in \mathbb{R}$ com

$$\nabla f(v_2) = \lambda_2 \nabla h(v_2) + \mu \nabla h_1(v_2).$$

Com $\nabla f(x) = Sx$, $\nabla h(x) = x$, $\nabla h_1(x) = v_1$ segue dali $Sv_2 = \lambda_2 v_2 + \mu v_1$. Como $\langle \lambda_2 v_2 - Sv_2, v_1 \rangle = 0$, segue daí por sua vez $\|\lambda_2 v_2 - Sv_2\|^2 + \mu^2 = 0$. Com isso temos mostrado que v_2 e λ_2 satisfazem as condições

$$Sv_2 = \lambda_2 v_2, \quad \langle v_1, v_2 \rangle = 0. \quad (3.4.14)$$

Continuando este argumento através de indução recebe-se uma base ortogonal v_1, v_2, \dots, v_n do \mathbb{R}^n a qual é composta de autovetores da matriz S . Os autovalores correspondentes $\lambda_1, \lambda_2, \dots, \lambda_n$ são reais e aparecem neste método na ordem $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ (Exercício).

Exemplo 3.4.5 (Distância a uma subvariedade). Seja $M \subset \mathbb{R}^n$ uma subvariedade suave d -dimensional com $0 \leq d < n$. Seja dado $y_0 \in \mathbb{R}^n$. Então existe um elemento $x_0 \in M$ com

$$\|y_0 - x_0\| \leq \|y_0 - x\| \quad \text{para todos os } x \in M \quad (3.4.15)$$

e este elemento x_0 satisfaz a condição

$$y_0 - x_0 \perp T_{x_0}M. \quad (3.4.16)$$

A existência de um tal elemento x_0 segue do fato que a função contínua $f : \mathbb{R}^n \rightarrow \mathbb{R}$ definida através de $f(x) := \frac{1}{2} \|y_0 - x\|^2$ para $x \in \mathbb{R}^n$ assume seu mínimo no subconjunto compacto $M \subset \mathbb{R}^n$. Então a condição (3.4.16) é da forma $\nabla f(x_0) \in (T_{x_0}M)^\perp$ e isto segue da equação (3.4.4) na prova de Teorema 3.4.1. Frequentemente em exemplos concretos pode-se encontrar o vetor procurado $x_0 \in M$, o qual satisfaz a condição (3.4.15), através de investigar todos os vetores as quais satisfazem a condição necessária (3.4.16).

Exercício 3.4.6 (Vizinhanças tubulares). Seja $M \subset \mathbb{R}^n$ como no Exemplo 3.4.5. Para $\varepsilon > 0$ definimos os conjuntos

$$\begin{aligned} W_\varepsilon &:= \{(x, \eta) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \in M, \eta \perp T_x M, \|\eta\| < \varepsilon\} \\ U_\varepsilon &:= \left\{ y \in \mathbb{R}^n \mid \inf_{x \in M} \|y - x\| < \varepsilon \right\} \end{aligned} \quad (3.4.17)$$

e a aplicação $\phi_\varepsilon : W_\varepsilon \rightarrow \mathbb{R}^n$ através de $\phi_\varepsilon(x, \eta) := x + \eta$ para $(x, \eta) \in W_\varepsilon$. Então vale o seguinte.

- (i) W_ε é uma subvariedade suave n -dimensional de $\mathbb{R}^n \times \mathbb{R}^n$.
- (ii) U_ε é um subconjunto aberto de \mathbb{R}^n e vale $\phi_\varepsilon(W_\varepsilon) = U_\varepsilon$.
- (iii) Se $\varepsilon > 0$ é suficientemente pequeno, então $\phi_\varepsilon : W_\varepsilon \rightarrow U_\varepsilon$ é bijetivo e possui uma aplicação inversa suave. Mais detalhado, se $\varepsilon > 0$ é suficientemente pequeno, então primeiro, existe para todo elemento $y \in U_\varepsilon$ exatamente um elemento $\pi(y) \in M$ com

$$\|y - \pi(y)\| = \inf_{x \in M} \|y - x\|$$

segundo, a aplicação $\pi : U_\varepsilon \rightarrow M$ é suave, e terceiro, para todo $y \in U_\varepsilon$ vale que $\phi_\varepsilon^{-1}(y) = (\pi(y), y - \pi(y))$.

- (iv) Para $\varepsilon > 0$ suficientemente pequeno chama-se o conjunto aberto $U_\varepsilon \subset \mathbb{R}^n$ em (3.4.17) uma **vizinhança tubular** de M e a afirmação (iii) é o *Teorema da vizinhança tubular*. Uma prova encontra-se em [3].

Capítulo 4

Campos de Vetores e Fluxos

Este capítulo é uma introdução às equações diferenciais. A primeira seção repete uns conteúdos da Análise I as quais referem-se à existência e unicidade de soluções de um problema de valor inicial. A Seção 4.2 introduz o fluxo de um campo de vetores e mostra que uma solução, a qual existe só num intervalo temporal finito, tem que sair do todo subconjunto compacto da região da definição do campo de vetores. Seção 4.3 mostra a dependência contínua do valor inicial e Seção 4.4 prova a diferenciabilidade do fluxo.

4.1 Existência e unicidade

Seja $U \subset \mathbb{R}^n$ um conjunto aberto e seja $f : U \rightarrow \mathbb{R}^n$ uma aplicação localmente Lipschitz-contínua. Entendemos esta aplicação como um **campo de vetores** o qual associa a todo ponto $x \in U$ um vetor velocidade $f(x) \in \mathbb{R}^n$. Trata-se do problema de entender as soluções da equação diferencial

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0. \quad (4.1.1)$$

Aqui $x_0 \in U$ é o **valor inicial** e as equações em (4.1.1) são chamadas de **problema de valor inicial**. Uma **solução** do problema de valor inicial (4.1.1) é uma aplicação continuamente diferenciável $x : I \rightarrow U$, definida num intervalo aberto $I \subset \mathbb{R}$ com $0 \in I$, a qual assume no momento $t = 0$ o valor x_0 e cuja derivada é dada através de $\dot{x}(t) = f(x(t))$ para $t \in I$. Para já lembramos o que é conhecido do curso Análise I sobre aplicações localmente Lipschitz-contínuas e sobre as soluções de (4.1.1). No seguinte sempre denotamos de $\|x\| := \sqrt{x_1^2 + \cdots + x_n^2}$ a norma euclidiana de um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Definição 4.1.1. *Seja $U \subset \mathbb{R}^n$ aberto. Uma aplicação $f : U \rightarrow \mathbb{R}^n$ chama-se de **localmente Lipschitz-contínua** se para todo $x_0 \in U$ existem dois números $\varepsilon > 0$ e $c > 0$, tal que a bola aberta*

$$B_\varepsilon(x_0) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < \varepsilon\} \subset U$$

é contido em U e ao longo desta bola vale a desigualdade

$$\|f(x) - f(y)\| \leq c \|x - y\|$$

para todos os $x, y \in B_\varepsilon(x_0)$.

Lema 4.1.2. *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $f : U \rightarrow \mathbb{R}^n$ uma aplicação localmente Lipschitz-contínua, e seja $K \subset U$ um subconjunto compacto. Então existe uma constante $c > 0$, tal que todos os $x, y \in K$ satisfazem a desigualdade $\|f(x) - f(y)\| \leq c \|x - y\|$.*

Demonstração. Suponhamos por absurdo que a afirmação seja falsa. Então existem (segundo o axioma da escolha, versão contável) duas sequências $x_i, y_i \in K$, tal que para todos os $i \in \mathbb{N}$ a desigualdade

$$\|f(x_i) - f(y_i)\| > i \|x_i - y_i\| \tag{4.1.2}$$

é satisfeita. Como K é compacto podemos, através de transição a uma subsequência, supor sem perda de generalidade que a sequência $(x_i)_{i \in \mathbb{N}}$ converge. Denotamos o limite de

$$x_0 := \lim_{i \rightarrow \infty} x_i \in K.$$

Como a função $K \rightarrow \mathbb{R} : x \mapsto \|f(x)\|$ é contínua e K é compacto, existe segundo um teorema da Análise I um elemento $\xi \in K$, tal que

$$C := \|f(\xi)\| \geq \|f(x)\| \quad \text{para todo } x \in K. \tag{4.1.3}$$

De (4.1.2) e (4.1.3) segue a desigualdade

$$\|y_i - x_i\| \leq \frac{\|f(y_i) - f(x_i)\|}{i} \leq \frac{\|f(y_i)\| + \|f(x_i)\|}{i} \leq \frac{2C}{i}$$

para todos os $i \in \mathbb{N}$. Portanto vale $\lim_{i \rightarrow \infty} (y_i - x_i) = 0$ e daí a sequência $(y_i)_{i \in \mathbb{N}}$ também converge para x_0 .

Como f é localmente Lipschitz-contínuo, existem números reais $\varepsilon > 0$ e $c > 0$ tais que para todos os $x, y \in \mathbb{R}^n$ vale o seguinte:

$$\begin{aligned} \|x - x_0\| < \varepsilon, \\ \|y - x_0\| < \varepsilon \end{aligned} \quad \implies \quad \begin{aligned} x, y \in U \text{ e} \\ \|f(x) - f(y)\| \leq c \|x - y\|. \end{aligned} \quad (4.1.4)$$

Como as sequências $(x_i)_{i \in \mathbb{N}}$ e $(y_i)_{i \in \mathbb{N}}$ convergem ambas para x_0 , existe um número natural $i_0 \in \mathbb{N}$ tal que todo $i \in \mathbb{N}$ com $i \geq i_0$ satisfaz as desigualdades

$$i \geq c, \quad \|x_i - x_0\| < \varepsilon, \quad \|y_i - x_0\| < \varepsilon.$$

Para $i \in \mathbb{N}$ com $i \geq i_0$ segue daí conforme (4.1.2) a desigualdade

$$\|f(x_i) - f(y_i)\| > i \|x_i - y_i\| \geq c \|x_i - y_i\|.$$

Aquela está na contradição a (4.1.4) e prova assim Lema 4.1.2. \square

Lema 4.1.3. *Seja $U \subset \mathbb{R}^n$ um conjunto aberto e seja $K \subset U$ um conjunto compacto. Então vale o seguinte.*

(i) *Para todo $\varepsilon > 0$ o conjunto*

$$K_\varepsilon := \{y \in \mathbb{R}^n \mid \text{existe um } x \in K, \text{ tal que } \|x - y\| \leq \varepsilon\}$$

é compacto.

(ii) *Existe um $\varepsilon > 0$ com $K_\varepsilon \subset U$.*

Demonstração. Vale $\sup_{y \in K_\varepsilon} \|y\| = \sup_{x \in K} \|x\| + \varepsilon < \infty$ e dali K_ε é limitado. Provaremos ora que K_ε também é fechado. Seja $y_i \in K_\varepsilon$ uma sequência a qual converge para $y \in \mathbb{R}^n$. Então existe uma sequência $x_i \in K$, tal que

$$\|x_i - y_i\| \leq \varepsilon \quad \text{para todos os } i \in \mathbb{N}.$$

Como K é compacto podemos, através de transição a uma subsequência, supor sem perda de generalidade que a sequência x_i converge para um elemento $x \in K$. Disso segue que

$$\|x - y\| = \lim_{i \rightarrow \infty} \|x_i - y_i\| \leq \varepsilon.$$

Portanto vale $y \in K_\varepsilon$. Com isso é mostrado que K_ε é fechado e limitado. Segundo o Teorema de Heine–Borel com isso K_ε é compacto.

Provaremos parte (ii) indireto e suponhamos que $K_\varepsilon \not\subset U$ para todo $\varepsilon > 0$. Então existe segundo o axioma da escolha, versão contável, uma sequência

$$y_i \in \mathbb{R}^n \setminus U$$

tal que $y_i \in K_{1/i}$ para todo $i \in \mathbb{N}$. Então existe, de novo segundo o axioma da escolha, versão contável, uma sequência $x_i \in K$, tal que

$$\|x_i - y_i\| \leq \frac{1}{i}$$

para todo $i \in \mathbb{N}$. Como K é compacto podemos, através de transição a uma subsequência, supor sem perda de generalidade que a sequência x_i converge para um elemento $x \in K$. Dali segue

$$x = \lim_{i \rightarrow \infty} x_i = \lim_{i \rightarrow \infty} y_i \notin U,$$

porque $\mathbb{R}^n \setminus U$ é fechado. Portanto $K \not\subset U$ em contraposição à nossa hipótese. Com isso Lema 4.1.3 fica provado. \square

O Teorema da Barreira (Teorema 2.4.1) diz que toda aplicação continuamente diferenciável $f : U \rightarrow \mathbb{R}^n$ é localmente Lipschitz-contínua.

O seguinte teorema da Análise I garante a existência e unicidade de soluções do problema de valor inicial (4.1.1) sob a hipótese que o campo de vetores $f : U \rightarrow \mathbb{R}^n$ é localmente Lipschitz-contínuo.

Teorema 4.1.4 (Existência e unicidade). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto, seja $f : U \rightarrow \mathbb{R}^n$ uma aplicação localmente Lipschitz-contínua, e seja $K \subset U$ um subconjunto compacto. Então vale o seguinte.*

(i) *Existe um intervalo aberto $I \subset \mathbb{R}$ com $0 \in I$, tal que o problema de valor inicial (4.1.1) possui uma solução $x : I \rightarrow U$ para todo $x_0 \in K$.*

(ii) *Se $I \subset \mathbb{R}$ é um intervalo aberto com $0 \in I$ e se $x, y : I \rightarrow U$ são duas soluções de (4.1.1) com $x_0 \in U$, então vale $x(t) = y(t)$ para todos os $x \in I$.*

Demonstração. Veja a página 139. \square

Sejam f e x_0 como em Teorema 4.1.4, sejam $I, J \subset \mathbb{R}$ dois intervalos abertos com $0 \in I \cap J$, e sejam $x : I \rightarrow U$ e $y : J \rightarrow U$ dois soluções de (4.1.1). Então vale $x(t) = y(t)$ para todo $t \in I \cap J$ segundo a parte (ii) de Teorema 4.1.4. Disso segue que a equação (4.1.1) também possui uma solução no intervalo $I \cup J$, a qual é dada através de $x(t)$ para $t \in I$ e através de $y(t)$ para $t \in J$. Isso lida ao conceito do *intervalo máximo de existência* em Definição 4.2.1.

Lema 4.1.5. *Sejam $f : U \rightarrow \mathbb{R}^n$ e $x_0 \in U$ como em Teorema 4.1.4, seja $I \subset \mathbb{R}$ um intervalo aberto com $0 \in I$, e seja $x : I \rightarrow U$ uma aplicação contínua. Então as seguintes afirmações são equivalentes.*

- (i) *A aplicação $x : I \rightarrow U$ é uma solução de (4.1.1).*
(ii) *A aplicação $x : I \rightarrow U$ satisfaz a equação integral*

$$x(t) = x_0 + \int_{t_0}^t f(x(s)) ds \quad \text{para todos os } t \in I. \quad (4.1.5)$$

Demonstração. Isto segue imediato do Teorema Fundamental do Cálculo, respectivamente das partes (v) e (vi) de Lema 1.7.3. \square

Demonstração do Teorema 4.1.4. Como $K \subset U$ é um subconjunto compacto, existe segundo o Lema 4.1.3 um número $\varepsilon > 0$, tal que

$$K_\varepsilon = \{x \in \mathbb{R}^n \mid \text{existe um } x_0 \in K \text{ com } \|x - x_0\| \leq \varepsilon\} \subset U. \quad (4.1.6)$$

Como, segundo o Lema 4.1.3, K_ε é um subconjunto compacto de U , existe segundo o Lema 4.1.2 uma constante $c > 0$, tal que

$$\|f(x) - f(y)\| \leq c \|x - y\| \quad \text{para todos os } x, y \in K_\varepsilon. \quad (4.1.7)$$

Além disso, a função $K \rightarrow \mathbb{R} : x \mapsto \|f(x)\|$ é contínua e daí, segundo um teorema do curso Análise I, limitada. Defina

$$M := \sup_{x \in K} \|f(x)\| + c\varepsilon \quad (4.1.8)$$

e escolha $\delta > 0$ tão pequeno que

$$\delta M < \varepsilon, \quad \delta c < 1. \quad (4.1.9)$$

Seja dado $x_0 \in K$. Provaremos em cinco passos a existência e unicidade de uma solução $x : [-\delta, \delta] \rightarrow U$ do problema de valor inicial (4.1.1).

Passo 1. *Seja $x \in \mathbb{R}^n$ com $\|x - x_0\| \leq \varepsilon$. Então vale*

$$x \in U, \quad \|f(x)\| \leq M.$$

Segundo (4.1.6) vale $x \in U$, e segundo (4.1.7) e (4.1.8) vale

$$\begin{aligned} \|f(x)\| &\leq \|f(x_0)\| + \|f(x) - f(x_0)\| \\ &\leq \|f(x_0)\| + c \|x - x_0\| \\ &\leq \|f(x_0)\| + c\varepsilon \\ &= M. \end{aligned}$$

Com isso Passo 1 fica provado.

Passo 2. *Seja $x : [-\delta, \delta] \rightarrow U$ uma solução de (4.1.1). Então vale*

$$\|x(t) - x_0\| < \varepsilon$$

para todos os $t \in [-\delta, \delta]$.

Provaremos Passo 2 através de um argumento de contradição e suponhamos que exista um número real $t \in [0, \delta]$ com $\|x(t) - x_0\| \geq \varepsilon$. Seja

$$\tau := \inf \{t \in [0, \delta] \mid \|x(t) - x_0\| \geq \varepsilon\}.$$

Então vale $0 < \tau \leq \delta$ e

$$\|x(t) - x_0\| < \varepsilon \quad \text{para todos os } t \in [0, \tau), \quad (4.1.10)$$

$$\|x(\tau) - x_0\| = \varepsilon. \quad (4.1.11)$$

Disso segue que

$$\|x(\tau) - x_0\| = \left\| \int_0^\tau f(x(t)) dt \right\| \leq \int_0^\tau \|f(x(t))\| dt \leq \tau M < \varepsilon.$$

Aqui o primeiro passo segue de Lema 4.1.5, o segundo da parte (iii) de Lema 1.7.3, o terceiro de (4.1.10) e passo 1, e o último passo segue de (4.1.9). Esta desigualdade está em contradição a (4.1.11) e com isso é mostrado que $\|x(t) - x_0\| < \varepsilon$ para todo $t \in [0, \delta]$. O mesmo argumento pode ser utilizado para $t \in [-\delta, 0]$ e com isso Passo 2 fica provado.

Passo 3. *Seja*

$$\mathcal{X} := \left\{ x : [-\delta, \delta] \rightarrow \mathbb{R}^n \left| \begin{array}{l} x \text{ é contínuo e} \\ \|x(t) - x_0\| \leq \varepsilon \\ \text{para todos os } t \in [-\delta, \delta] \end{array} \right. \right\} \quad (4.1.12)$$

e defina $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ através de

$$d(x, y) := \|x - y\|_\infty := \sup_{|t| \leq \delta} \|x(t) - y(t)\| \quad (4.1.13)$$

para $x, y \in \mathcal{X}$. Então (\mathcal{X}, d) é um espaço métrico completo não-vazio.

O espaço é não-vazio, porque a função constante $x(t) = x_0$ é um elemento de \mathcal{X} . Que a aplicação $d : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ é uma função distância, segue direto das definições. Que o espaço métrico (\mathcal{X}, d) é completo, segue dali que toda sequência de Cauchy de números reais converge e que, segundo um teorema da Análise I, o limite de uma sequência uniformemente convergente de funções contínuas é mesmo uma função contínua de novo. Com isso Passo 3 fica provado.

Passo 4. Seja \mathcal{X} dado através de equação (4.1.12) em Passo 3 e seja $x \in \mathcal{X}$. Defina a função $\mathcal{F}(x) : [-\delta, \delta] \rightarrow \mathbb{R}^n$ através de

$$(\mathcal{F}(x))(t) := x_0 + \int_0^t f(x(s)) ds \quad (4.1.14)$$

para $-\delta \leq t \leq \delta$. Então vale $\mathcal{F}(x) \in \mathcal{X}$.

Para $-\delta \leq s \leq \delta$ vale $\|x(s) - x_0\| \leq \varepsilon$ e daí $x(s) \in U$ e $\|f(x(s))\| \leq M$ segundo o Passo 1. Disso segue que a função $\mathcal{F}(x)$ em (4.1.14) é bem definido e satisfaz para $0 \leq t \leq \delta$ a desigualdade

$$\begin{aligned} \|(\mathcal{F}(x))(t) - x_0\| &= \left\| \int_0^t f(x(s)) ds \right\| \\ &\leq \int_0^t \|f(x(s))\| ds \\ &\leq tM \\ &< \varepsilon. \end{aligned}$$

Aqui o segundo passo segue de parte (iii) de Lema 1.7.3, e o último passo de (4.1.9). Igualmente vale $\|(\mathcal{F}(x))(t) - x_0\| < \varepsilon$ para $-\delta \leq t \leq 0$. Demais $\mathcal{F}(x)$ é contínuo e daí vale $\mathcal{F}(x) \in \mathcal{X}$. Com isso Passo 4 fica provado.

Passo 5. Seja (\mathcal{X}, d) o espaço métrico completo não-vazio em Passo 3 e seja $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ a aplicação em Passo 4. Então vale

$$\|\mathcal{F}(x) - \mathcal{F}(y)\|_\infty \leq \delta c \|x - y\|_\infty$$

para todos os $x, y \in \mathcal{X}$ e dali \mathcal{F} é uma contração.

Sejam $x, y \in \mathcal{X}$. Então vale

$$\begin{aligned} \|(\mathcal{F}(x))(t) - (\mathcal{F}(y))(t)\| &= \left\| \int_0^t (f(x(s)) - f(y(s))) ds \right\| \\ &\leq \int_0^t \|f(x(s)) - f(y(s))\| ds \\ &\leq \int_0^t c \|x(s) - y(s)\| ds \\ &\leq \delta c \|x - y\|_\infty \end{aligned}$$

para $0 \leq t \leq \delta$. Argumenta-se igualmente para $-\delta \leq t \leq 0$ e dali segue a desigualdade $\|\mathcal{F}(x) - \mathcal{F}(y)\|_\infty \leq \delta c \|x - y\|_\infty$. Como $\delta c < 1$, segue daí que $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ é uma contração. Com isso Passo 5 fica provado.

Segundo o Passo 5 e o Teorema E.0.7 do Ponto Fixo de Banach a contração \mathcal{F} possui um ponto fixo único

$$x = \mathcal{F}(x) \in \mathcal{X}.$$

Este ponto fixo é, segundo (4.1.6), uma função contínua

$$x : [-\delta, \delta] \rightarrow U,$$

a qual satisfaz, segundo a definição de \mathcal{F} , a equação

$$x(t) = (\mathcal{F}(x))(t) = x_0 + \int_0^t f(x(s)) ds$$

para todo $t \in [-\delta, \delta]$. Isto é a equação (4.1.5). Portanto segue de Lema 4.1.5 que x é uma solução de (4.1.1). Vice versa, segue de Passo 2 e Lema 4.1.5 que toda solução $x : [-\delta, \delta] \rightarrow U$ de (4.1.1) é um ponto fixo de \mathcal{F} . Com isso a existência e a unicidade de uma solução de (4.1.1) no intervalo $[-\delta, \delta]$ ficam provadas para todo valor inicial $x_0 \in K$.

Ora, seja dado $x_0 \in U$, seja $I \subset \mathbb{R}$ um intervalo aberto com $0 \in I$, e sejam $x, y : I \rightarrow U$ duas soluções de (4.1.1). Suponhamos que exista um elemento $t \in I$ com

$$x(t) \neq y(t).$$

Então vale $t \neq 0$. No caso de $t > 0$ definimos

$$t_0 := \sup \{t \in I \mid t > 0 \text{ e } x(s) = y(s) \text{ para todos os } s \in [0, t]\}.$$

Então vale $t_0 \in I$ e $t_0 > 0$ e $x(s) = y(s)$ para $0 \leq s \leq t_0$. Mas ora segue da existência já provada com o valor inicial $x(t_0)$, aplicado às soluções $x(t - t_0)$ e $y(t - t_0)$, que existe um número $\delta > 0$ com $t_0 - \delta, t_0 + \delta \in I$ e $x(t_0 + t) = y(t_0 + t)$ para $-\delta \leq t \leq \delta$. Isto está na contradição à definição de t_0 . Igualmente recebe-se uma contradição no caso de $t < 0$ e com isso Teorema 4.1.4 fica provado. \square

No Teorema 4.1.4 também vale a afirmação de existência em parte (i) ainda se o campo de vetores $f : U \rightarrow \mathbb{R}^n$ só é contínuo, mas não Lipschitz-contínuo. Com efeito esta afirmação mais geral exige uma outra prova. Além disso, na afirmação de unicidade em parte (ii) não pode-se abrir mão da Lipschitz-continuidade, como mostra o seguinte exemplo.

Exemplo 4.1.6. Seja $n = 1$ e seja $f : \mathbb{R} \rightarrow \mathbb{R}$ o campo de vetores contínuo

$$f(x) := \sqrt{|x|} \quad \text{para } x \in \mathbb{R}. \quad (4.1.15)$$

Então as funções

$$x(t) := 0, \quad y(t) := \begin{cases} t^2/4, & \text{se } t \geq 0, \\ 0, & \text{se } t \leq 0, \end{cases}$$

são duas soluções diferentes do problema de valor inicial (4.1.1) com $x_0 = 0$.

Exemplo 4.1.7. Para $n = 1$ consideramos o problema de valor inicial

$$\dot{x} := -x^2 \quad x(0) = x_0. \quad (4.1.16)$$

Se $x_0 = 0$, então $x(t) = 0$ é a solução única de (4.1.16) e ela existe em todo o \mathbb{R} . Se $x_0 \neq 0$, então toda solução $x : I \rightarrow \mathbb{R}$ de (4.1.16) é em todo ponto não-nulo, assim tem uma derivada negativa, e assim é decrescente segundo o Teorema do Valor Médio. Disso resulta que a função $I \rightarrow \mathbb{R} : t \mapsto x(t)$ possui uma função inversa $x \mapsto t(x)$ diferenciável. Como $\dot{x}(t) = -x(t)^2$ para todo t , a função inversa tem a derivada $dt/dx = -1/x^2$. Segundo o Teorema Fundamental do Cálculo vale

$$t = t(x) - t(x_0) = - \int_{x_0}^x \frac{d\xi}{\xi^2} = \frac{1}{x} - \frac{1}{x_0}.$$

Dali a solução de (4.1.16) no caso de $x_0 > 0$ é dada através da fórmula

$$x(t) = \frac{x_0}{1 + tx_0}, \quad -\frac{1}{x_0} < t < \infty. \quad (4.1.17)$$

Esta solução não pode-se estender a um intervalo maior.

Exemplo 4.1.8. Seja $A \in \mathbb{R}^{n \times n}$ e $x_0 \in \mathbb{R}^n$. Então a solução única do problema de valor inicial

$$\dot{x} = Ax, \quad x(0) = x_0 \quad (4.1.18)$$

é dada através da matriz exponencial, ou seja

$$x(t) = e^{tA}x_0 = \sum_{k=0}^{\infty} \frac{t^k A^k x_0}{k!} \quad \text{para } t \in \mathbb{R}. \quad (4.1.19)$$

4.2 O fluxo de um campo de vetores

Daqui em diante suponha-se que são dados um subconjunto aberto $U \subset \mathbb{R}^n$ e um campo de vetores localmente Lipschitz-contínuo $f : U \rightarrow \mathbb{R}^n$.

Definição 4.2.1. *Seja $x_0 \in U$. O intervalo máximo de existência de x_0 (para equação (4.1.1)) é o conjunto*

$$I(x_0) := \bigcup \left\{ I \subset \mathbb{R} \mid \begin{array}{l} I \text{ é um intervalo aberto com } 0 \in I \\ \text{e o problema de valor inicial (4.1.1)} \\ \text{possui uma solução } x : I \rightarrow U \end{array} \right\}. \quad (4.2.1)$$

Segundo o Teorema 4.1.4, o subconjunto $I(x_0) \subset \mathbb{R}$ é um intervalo aberto contendo zero $0 \in I(x_0)$, o problema de valor inicial (4.1.1) possui uma solução $x : I(x_0) \rightarrow U$, esta solução é única, e não existe nenhuma solução de (4.1.1) num intervalo estritamente maior.

Definição 4.2.2. *Seja*

$$\Omega := \{(t, x_0) \in \mathbb{R} \times U \mid t \in I(x_0)\}. \quad (4.2.2)$$

O fluxo de f é a aplicação

$$\phi : \Omega \rightarrow U,$$

a qual associa a todo par $(t, x_0) \in \Omega$ a solução única de (4.1.1) no momento temporal t , ou seja $\phi(t, x_0) := x(t)$, onde $x : I(x_0) \rightarrow U$ é a única solução de (4.1.1). Em outras palavras, $\phi : \Omega \rightarrow U$ é a única aplicação a qual é parcialmente diferenciável com respeito a t e satisfaz as equações

$$\partial_t \phi(t, x) = f(\phi(t, x)), \quad \phi(0, x_0) = x_0 \quad (4.2.3)$$

para todos os $(t, x) \in \Omega$ e todos os $x_0 \in U$.

Exemplo 4.2.3. *Seja $A \in \mathbb{R}^{n \times n}$ e seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ o campo de vetores linear*

$$f(x) := Ax$$

para $x \in \mathbb{R}^n$. Então $\Omega = \mathbb{R} \times \mathbb{R}^n$. Além disso, segundo o Exemplo 4.1.8, o fluxo $\phi : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ de f é a aplicação

$$\phi(t, x) = e^{tA}x$$

para $t \in \mathbb{R}$ e $x \in \mathbb{R}^n$.

Exemplo 4.2.4. Seja $n = 1$, $U = \mathbb{R}$, e seja $f : \mathbb{R} \rightarrow \mathbb{R}$ o campo de vetores

$$f(x) := -x^2 \quad \text{para } x \in \mathbb{R}.$$

(Veja o Exemplo 4.1.7.) Para $x_0 \in \mathbb{R}$ a curva

$$x(t) := \frac{x_0}{1 + tx_0}$$

é a solução única de (4.1.1) no intervalo

$$I(x_0) = \begin{cases} (-x_0^{-1}, \infty), & \text{para } x_0 > 0, \\ \mathbb{R}, & \text{para } x_0 = 0, \\ (-\infty, -x_0^{-1}), & \text{para } x_0 < 0. \end{cases}$$

Daí as equações

$$\Omega = \{(t, x) \in \mathbb{R}^2 \mid tx > -1\}, \quad \phi(t, x) = \frac{x}{1 + tx}$$

representam o fluxo de f .

Lema 4.2.5. *Seja $x_0 \in U$ e $t_0 \in I(x_0)$. Então vale*

$$I(\phi(t_0, x_0)) = I(x_0) - t_0 \tag{4.2.4}$$

e

$$\phi(s, \phi(t_0, x_0)) = \phi(t_0 + s, x_0) \tag{4.2.5}$$

para todos os $s \in I(\phi(t_0, x_0))$.

Demonstração. Seja

$$y_0 := \phi(t_0, x_0)$$

e

$$J := I(x_0) - t_0 = \{t - t_0 \mid t \in I(x_0)\} = \{s \in \mathbb{R} \mid t_0 + s \in I(x_0)\}.$$

Então $J \subset \mathbb{R}$ é um intervalo aberto e $0 \in J$. Defina $y : J \rightarrow \mathbb{R}$ através de

$$y(s) := \phi(t_0 + s, x_0) \quad \text{para } s \in J.$$

Então y é continuamente diferenciável e vale $\dot{y}(s) = f(y(s))$ para todo $s \in J$ e $y(0) = y_0$. Portanto $J \subset I(y_0)$ e $y(s) = \phi(s, y_0)$ para todo $s \in J$. Com isso temos provado a inclusão

$$I(x_0) - t_0 \subset I(\phi(t_0, x_0))$$

e a equação (4.2.5) para todos os $s \in I(x_0) - t_0$.

Resta mostrar que vale a inclusão inversa, ou seja

$$I(\phi(t_0, x_0)) \subset I(x_0) - t_0.$$

Neste respeito notamos que, segundo o que foi mostrado até agora, vale o seguinte:

$$-t_0 \in I(x_0) - t_0 \subset I(y_0), \quad \phi(-t_0, y_0) = \phi(0, x_0) = x_0.$$

Por isso resulta da primeira parte da prova, com (t_0, x_0) substituído de $(-t_0, y_0)$, a inclusão $I(y_0) - (-t_0) \subset I(\phi(-t_0, y_0))$. Ou seja,

$$I(\phi(t_0, x_0)) + t_0 \subset I(x_0)$$

e dali $I(\phi(t_0, x_0)) \subset I(x_0) - t_0$. Com isso Lema 4.2.5 fica provado. \square

Teorema 4.2.6. *Seja $x_0 \in U$ um valor inicial com*

$$I(x_0) \cap [0, \infty) = [0, b), \quad 0 < b < \infty.$$

Então existe para todo conjunto compacto $K \subset U$ um número $0 \leq t_K < b$ com

$$t_K < t < b \quad \implies \quad \phi(t, x_0) \notin K. \quad (4.2.6)$$

Em outras palavras, se a solução do problema de valor inicial (4.1.1) existe só num intervalo temporal finito, então ele tem que sair de todo subconjunto compacto de U .

Demonstração. Segundo o Teorema 4.1.4 existe um número $\delta > 0$ com

$$(-\delta, \delta) \subset I(x) \quad \text{para todos os } x \in K.$$

Aqui podemos escolher $\delta > 0$ tão pequeno que vale $\delta \leq b$. Então o número

$$t_K := b - \delta$$

satisfaz a condição (4.2.6). Isto é t um número real com

$$b - \delta < t < b,$$

então vale segundo o Lema 4.2.5

$$I(\phi(t, x_0)) = I(x_0) - t.$$

Dali segue

$$I(\phi(t, x_0)) \cap [0, \infty) = [0, b - t).$$

Como $b - t < \delta$, vale $(-\delta, \delta) \not\subset I(\phi(t, x_0))$ e daí $\phi(t, x_0)$ não pode ser um elemento de K . Com isso Teorema 4.2.6 fica provado. \square

4.3 Continuidade do fluxo

Nosso objetivo é mostrar que o fluxo de um campo de vetores localmente Lipschitz-contínuo também é uma aplicação localmente Lipschitz-contínua.

Teorema 4.3.1. *Seja $U \subset \mathbb{R}^n$ aberto, seja $f : U \rightarrow \mathbb{R}^n$ um campo de vetores localmente Lipschitz-contínuo, e seja $\phi : \Omega \rightarrow U$ o fluxo de f . Então Ω , definido em (4.2.2), é aberto em $\mathbb{R} \times U$ e ϕ é localmente Lipschitz-contínuo.*

Demonstração. Veja a página 148. \square

A prova resta no **Lema de Gronwall**.

Lema 4.3.2 (Gronwall). *Sejam $I \subset \mathbb{R}$ um intervalo contendo 0, $A, B \geq 0$, e seja $g : I \rightarrow [0, \infty)$ uma função contínua a qual satisfaça a desigualdade*

$$g(t) \leq A + B \left| \int_0^t g(s) ds \right| \quad (4.3.1)$$

para todos os $t \in I$. Então vale

$$g(t) \leq Ae^{B|t|} \quad (4.3.2)$$

para todos os $t \in I$.

Demonstração. Defina a função $G : I \cap [0, \infty) \rightarrow [0, \infty)$ através de

$$G(t) := A + B \int_0^t g(s) ds \quad \text{para } t \in I \text{ com } t \geq 0.$$

Ela satisfaz a desigualdade $g(t) \leq G(t)$ e dali

$$\dot{G}(t) = Bg(t) \leq BG(t)$$

para todos os $t \in I$ com $t \geq 0$. Disso segue

$$\frac{d}{dt}(e^{-Bt}G(t)) = e^{-Bt}(\dot{G}(t) - BG(t)) \leq 0$$

e daí

$$e^{-Bt}G(t) \leq G(0) = A$$

para todos os $t \in I$ com $t \geq 0$. Portanto vale

$$g(t) \leq G(t) \leq Ae^{Bt}$$

para todos os $t \in I$ com $t \geq 0$. No caso de $t \leq 0$ resulta a desigualdade desejada no que substitui-se a função g pela função $t \mapsto g(-t)$. Com isso Lema 4.3.2 fica provado. \square

Demonstração do Teorema 4.3.1. A prova tem três passos. O primeiro passo utiliza a notação de Lema 4.1.3.

Passo 1. *Seja $K \subset U$ compacto e seja $\varepsilon > 0$ tal que $K_\varepsilon \subset U$. Então existe um $\delta > 0$ com $[-\delta, \delta] \times K \subset \Omega$ e $\phi([-\delta, \delta] \times K) \subset K_\varepsilon$.*

Segundo a hipótese $U_\varepsilon := \bigcup_{x \in K} B_\varepsilon(x)$ é um subconjunto aberto de U . Segundo a parte (i) de Teorema 4.1.4, com U_ε em vez de U , existe então um número $\delta > 0$ com $[-\delta, \delta] \subset I(x_0)$ e $\phi(t, x_0) \in U_\varepsilon$ para todos os $x_0 \in K$ e todos os $t \in [-\delta, \delta]$. Com isso Passo 1 fica provado.

Passo 2. *Sejam K, ε, δ como no Passo 1. Então ϕ é Lipschitz-contínuo no conjunto $[-\delta, \delta] \times K \subset \Omega$.*

Como f é localmente Lipschitz-contínuo e, segundo o Lema 4.1.3, $K_\varepsilon \subset U$ é compacto, a restrição $f|_{K_\varepsilon}$ é Lipschitz-contínuo (Lema 4.1.2), e isso significa

$$c := \sup_{\substack{x, y \in K_\varepsilon \\ x \neq y}} \frac{\|f(x) - f(y)\|}{\|x - y\|} < \infty, \quad M := \sup_{x \in K_\varepsilon} \|f(x)\| < \infty. \quad (4.3.3)$$

Sejam dados $x_0, y_0 \in K$, e sejam $x, y : [-\delta, \delta] \rightarrow U$ as soluções correspondentes de (4.1.1), ou seja $x(t) := \phi(t, x_0)$ e $y(t) := \phi(t, y_0)$ para $t \in [-\delta, \delta]$. Então vale $x(t), y(t) \in K_\varepsilon$ para $t \in [-\delta, \delta]$ segundo o Passo 1 e daí

$$\|x(s) - x(t)\| = \left\| \int_s^t f(x(r)) dr \right\| \leq \int_s^t \|f(x(r))\| dr \leq M|s - t|$$

para todos os $s, t \in [-\delta, \delta]$ com $s < t$. Demais vale para todos os $t \in [-\delta, \delta]$

$$\begin{aligned} \|x(t) - y(t)\| &= \left\| x_0 - y_0 + \int_0^t (f(x(s)) - f(y(s))) ds \right\| \\ &\leq \|x_0 - y_0\| + \left\| \int_0^t (f(x(s)) - f(y(s))) ds \right\| \\ &\leq \|x_0 - y_0\| + \left| \int_0^t \|f(x(s)) - f(y(s))\| ds \right| \\ &\leq \|x_0 - y_0\| + c \left| \int_0^t \|x(s) - y(s)\| ds \right|. \end{aligned}$$

Disso segue $\|x(t) - y(t)\| \leq e^{c|t|} \|x_0 - y_0\|$, segundo o Lema 4.3.2, e daí

$$\begin{aligned} \|\phi(s, x_0) - \phi(t, y_0)\| &\leq \|x(s) - x(t)\| + \|x(t) - y(t)\| \\ &\leq M|s - t| + e^{c\delta} \|x_0 - y_0\| \end{aligned}$$

para todos os $x_0, y_0 \in K$ e todos os $s, t \in [-\delta, \delta]$. Isso prova Passo 2.

Passo 3. O conjunto $\Omega \subset \mathbb{R} \times U$, definido em (4.2.2), é aberto e a aplicação $\phi : \Omega \rightarrow U$ é localmente Lipschitz-contínua.

Seja $(t_0, x_0) \in \Omega$ com $t_0 \geq 0$ e defina

$$\Gamma := \{\phi(t, x_0) \mid 0 \leq t \leq t_0\}.$$

Isto é um subconjunto compacto de U . Dali existe, segundo o Lema 4.1.3, uma constante $\varepsilon > 0$, tal que

$$K := \Gamma_\varepsilon = \{y \in \mathbb{R}^n \mid \text{existe um } x \in \Gamma, \text{ tal que } \|x - y\| \leq \varepsilon\} \subset U.$$

Além disso, segundo o Lema 4.1.3, o conjunto K é compacto e o conjunto

$$V := \Gamma_\varepsilon^\circ = \{y \in \mathbb{R}^n \mid \text{existe um } x \in \Gamma, \text{ tal que } \|x - y\| < \varepsilon\} = \bigcup_{x \in \Gamma} B_\varepsilon(x)$$

é aberto. Segundo o Passo 1 e o Passo 2 existe uma constante $\delta > 0$, tal que $[-\delta, \delta] \times K \subset \Omega$ e ϕ é Lipschitz-contínuo em $[-\delta, \delta] \times K$. Escolha $N \in \mathbb{N}$ tão grande que

$$\tau := \frac{t_0}{N} < \delta.$$

Então $[0, \tau] \times V \subset \Omega$, e definimos a aplicação $\phi_\tau : V \rightarrow U$ através de

$$\phi_\tau(y) := \phi(\tau, y) \quad \text{para } y \in V.$$

Esta aplicação é contínua segundo o Passo 2. Definimos ora uma sequência finita de conjuntos abertos $V_0 \supset V_1 \supset \cdots \supset V_N$ através de

$$\begin{aligned} V_0 &:= V, \\ V_1 &:= \phi_\tau^{-1}(V_0) = \{y \in V \mid \phi_\tau(y) \in V\} \\ V_2 &:= \phi_\tau^{-1}(V_1) = \{y \in V \mid \phi_\tau(y) \in V, \phi_\tau(\phi_\tau(y)) \in V\} \\ &\vdots \\ V_N &:= \phi_\tau^{-1}(V_{N-1}) = \{y \in V \mid \phi_\tau^k(y) \in V \text{ para } k = 1, \dots, N\}. \end{aligned}$$

Aqui denotamos de $\phi_\tau^k := \phi_\tau \circ \cdots \circ \phi_\tau$ a composição (k vezes) da aplicação $\phi_\tau : V \rightarrow U$ com se mesma. A região da sua definição é exatamente o conjunto V_{k-1} . Os conjuntos $V_k \subset U$ são abertos, como $\phi_\tau : V \rightarrow U$ é contínuo. Além disso, segue de Lema 4.2.5 a igualdade $\phi_\tau^k(x_0) = \phi(k\tau, x_0) \in \Gamma \subset V$ para $k = 1, \dots, N$ e dali vale $x_0 \in V_k$ para $k = 0, 1, \dots, N$.

Também segue de Lema 4.2.5 que para todos os $x \in V_N$ vale o seguinte:

$$t_0 = N\tau \in I(x), \quad \phi(t_0, x) \in V.$$

Disso segue que $[-\delta, \delta] \subset I(\phi(t_0, x))$ para todos os $x \in V_N$ e daí, segundo o Lema 4.2.5, vale

$$[t_0 - \delta, t_0 + \delta] \times V_N \subset \Omega.$$

Como ϕ é Lipschitz-contínuo em $[-\delta, \delta] \times K$, existe uma constante $c > 0$, tal que todos os $s, s' \in [-\delta, \delta]$ e todos os $y, y' \in K$ satisfazem a desigualdade

$$\|\phi(s, y) - \phi(s', y')\| \leq c|s - s'| + c\|y - y'\|. \quad (4.3.4)$$

Para $x, x' \in V_N$ e $t, t' \in (t_0 - \delta, t_0 + \delta)$ segue dali a desigualdade

$$\begin{aligned} \|\phi(t, x) - \phi(t', x')\| &= \|\phi(t - t_0, \phi_\tau^N(x)) - \phi(t' - t_0, \phi_\tau^N(x'))\| \\ &\leq c|t - t'| + c\|\phi_\tau^N(x) - \phi_\tau^N(x')\| \\ &\leq c|t - t'| + c^{N+1}\|x - x'\|. \end{aligned}$$

Aqui a última desigualdade segue através de aplicar a desigualdade (4.3.4) N vezes. Com isso temos mostrado que existe para todo elemento $(t_0, x_0) \in \Omega$ com $t_0 \geq 0$ um conjunto aberto $W = (t_0 - \delta, t_0 + \delta) \times V_N \subset \mathbb{R} \times \mathbb{R}^n$ tal que $(t_0, x_0) \in W \subset \Omega$ e tal que a restrição $\phi|_W : W \rightarrow U$ é Lipschitz-contínua. Para $t_0 < 0$ segue a mesma afirmação no que substitui-se f por $-f$ e $\phi(t, x)$ por $\phi(-t, x)$. Com isso Passo 3 e Teorema 4.3.1 ficam provados. \square

Comentário 4.3.3. Para todo $t \in \mathbb{R}$ o conjunto

$$U_t := \{x_0 \in U \mid t \in I(x_0)\}$$

é aberto segundo o Teorema 4.3.1. Além disso, segue de Lema 4.2.5 que para todo $x \in U_t$ vale que $-t \in I(x) - t = I(\phi(t, x))$ e dali que $\phi(t, x) \in U_{-t}$. Frequentemente é útil usar para o fluxo a notação

$$\phi_t : U_t \rightarrow U_{-t}, \quad \phi_t(x) := \phi(t, x).$$

Segundo o Teorema 4.3.1 a aplicação $\phi_t : U_t \rightarrow U_{-t}$ é um homeomorfismo com a aplicação inversa

$$\phi_t^{-1} = \phi_{-t}.$$

Além disso, segundo o Lema 4.2.5, vale $U_0 = U$, $\phi_0 = \text{id} : U \rightarrow U$, e

$$\phi_{s+t} = \phi_s \circ \phi_t \quad \text{em } U_{s+t} \cap \phi_t^{-1}(U_s)$$

para todos os $s, t \in \mathbb{R}$. Esta notação simplifica-se consideravelmente se, para todo $x \in U$, vale $I(x) = \mathbb{R}$. Então $U_t = U$ para todos os $t \in \mathbb{R}$.

4.4 Diferenciabilidade do fluxo

Teorema 4.4.1. *Seja $U \subset \mathbb{R}^n$ aberto e seja $f : U \rightarrow \mathbb{R}^n$ um campo de vetores de classe C^ℓ com $\ell \in \mathbb{N} \cup \{\infty\}$. Então o fluxo $\phi : \Omega \rightarrow U$ de f é C^ℓ .*

Demonstração. Veja a página 153. \square

Antes de provar este teorema formulamos primeiramente um lema o qual pressupõe a diferenciabilidade contínua do fluxo. Este lema fornece uma fórmula para a derivada de ϕ e esta fórmula será utilizada então para provar Teorema 4.4.1.

Lema 4.4.2. *Seja $U \subset \mathbb{R}^n$ aberto, seja $f : U \rightarrow \mathbb{R}^n$ um campo de vetores continuamente diferenciável, e seja o fluxo $\phi : \Omega \rightarrow U$ de f uma aplicação continuamente diferenciável. Sejam dados $x_0 \in U$ e $\xi_0 \in \mathbb{R}^n$. Seja $I := I(x_0)$ e defina as aplicações $x : I \rightarrow U$ e $\xi : I \rightarrow \mathbb{R}^n$ através de*

$$x(t) := \phi_t(x_0), \quad \xi(t) := d\phi_t(x_0)\xi_0 \quad (4.4.1)$$

para $t \in I$. Então x e ξ são continuamente diferenciável e vale

$$\begin{aligned} \dot{x}(t) &= f(x(t)), & x(0) &= x_0 \\ \dot{\xi}(t) &= df(x(t))\xi(t), & \xi(0) &= \xi_0 \end{aligned} \quad (4.4.2)$$

para todos os $t \in I$.

Demonstração. A derivada parcial $\partial\phi/\partial t = f \circ \phi : \Omega \rightarrow \mathbb{R}^n$ existe segundo a definição do fluxo e é continuamente diferenciável por hipótese. Além disso, vale segundo a regra da cadeia

$$\frac{\partial}{\partial x_i} \frac{\partial \phi}{\partial t}(t, x) = df(\phi(t, x)) \frac{\partial \phi}{\partial x_i}(t, x) \quad \text{para todos os } (t, x) \in \Omega.$$

Segundo o Teorema de Schwarz (sobre a comutatividade das segundas derivadas) segue daí que a aplicação $\partial\phi/\partial x_i : \Omega \rightarrow \mathbb{R}^n$ é parcialmente diferenciável com respeito a t e satisfaz a equação

$$\frac{\partial}{\partial t} \frac{\partial \phi}{\partial x_i}(t, x) = \frac{\partial}{\partial x_i} \frac{\partial \phi}{\partial t}(t, x) = df(\phi(t, x)) \frac{\partial \phi}{\partial x_i}(t, x)$$

para $\forall (t, x) \in \Omega$. Então a função $\xi(t) = d\phi_t(x_0)\xi_0 = \sum_{i=1}^n \frac{\partial \phi}{\partial x_i}(t, x_0)\xi_{0i}$, onde $\xi_0 = (\xi_{01}, \dots, \xi_{0n}) \in \mathbb{R}^n$, é continuamente diferenciável com derivada

$$\dot{\xi}(t) = \sum_{i=1}^n \frac{\partial}{\partial t} \frac{\partial \phi}{\partial x_i}(t, x_0)\xi_{0i} = \sum_{i=1}^n df(\phi(t, x_0)) \frac{\partial \phi}{\partial x_i}(t, x_0)\xi_{0i} = df(x(t))\xi(t).$$

Com isso Lema 4.4.2 fica provado. \square

Lema 4.4.3. *Seja $I \subset \mathbb{R}$ um intervalo aberto com $0 \in I$ e $A : I \rightarrow \mathbb{R}^{n \times n}$ uma aplicação contínua. Então o problema de valor inicial*

$$\dot{\xi}(t) = A(t)\xi(t), \quad \xi(0) = \xi_0 \quad (4.4.3)$$

possui para todo valor inicial $\xi_0 \in \mathbb{R}^n$ uma solução única $\xi : I \rightarrow \mathbb{R}^n$ continuamente diferenciável.

Demonstração. Seja $T \in I$ com $T > 0$ e defina

$$c := 2 \sup_{0 \leq t \leq T} \|A(t)\|, \quad \|A(t)\| := \sup_{\xi \in \mathbb{R}^n \setminus \{0\}} \frac{\|A(t)\xi\|}{\|\xi\|}.$$

Seja $\mathcal{X} := C([0, T], \mathbb{R}^n)$ o espaço das aplicações contínuas $\xi : [0, T] \rightarrow \mathbb{R}^n$. Isto é um espaço de Banach com a função norma

$$\|\xi\|_c := \sup_{0 \leq t \leq T} e^{-ct} \|\xi(t)\| \quad \text{para } \xi \in \mathcal{X}.$$

Defina a aplicação $\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$ através de

$$\mathcal{F}(\xi)(t) := \xi_0 + \int_0^t A(s)\xi(s) ds \quad \text{para } 0 \leq t \leq T.$$

Provaremos a desigualdade

$$\|\mathcal{F}(\xi) - \mathcal{F}(\eta)\|_c \leq \frac{1}{2} \|\xi - \eta\|_c \quad \text{para todos os } \xi, \eta \in \mathcal{X}. \quad (4.4.4)$$

Sejam então dados $\xi, \eta \in \mathcal{X}$ e seja $t \in [0, T]$. Então vale

$$\begin{aligned} \|\mathcal{F}(\xi)(t) - \mathcal{F}(\eta)(t)\| &= \left\| \int_0^t A(s)(\xi(s) - \eta(s)) ds \right\| \\ &\leq \int_0^t \|A(s)\| \|\xi(s) - \eta(s)\| ds \\ &\leq \frac{1}{2} \int_0^t c \|\xi(s) - \eta(s)\| ds \\ &= \frac{1}{2} \int_0^t c e^{cs} e^{-cs} \|\xi(s) - \eta(s)\| ds \\ &\leq \frac{1}{2} \int_0^t c e^{cs} ds \|\xi - \eta\|_c \\ &\leq \frac{1}{2} e^{ct} \|\xi - \eta\|_c \end{aligned}$$

para todos os $t \in [0, T]$. Disso segue a desigualdade (4.4.4) no que multiplica-se ambos lados com e^{-ct} e forma-se o supremo sobre todo $t \in [0, T]$.

Da desigualdade (4.4.4) segue, segundo o Teorema do Ponto Fixo de Banach (Teorema E.0.7), que a aplicação

$$\mathcal{F} : \mathcal{X} \rightarrow \mathcal{X}$$

possui um ponto fixo único

$$\xi = \mathcal{F}(\xi).$$

Este ponto fixo é a única solução $\xi : [0, T] \rightarrow \mathbb{R}^n$ do problema de valor inicial (4.4.3) no intervalo $[0, T]$. Como uma tal solução única existe em todo subintervalo compacto $[0, T] \subset I$, também existe em $I \cap [0, \infty)$. Então segue a existência de uma solução única no intervalo I inteiro através de substituir a função $A : I \rightarrow \mathbb{R}^{n \times n}$ pela função $-I \rightarrow \mathbb{R}^{n \times n} : t \mapsto -A(-t)$. Com isso Lema 4.4.3 fica provado. \square

Lema 4.4.2 mostra que a linearização $\xi(t) := d\phi_t(x_0)\xi_0$ do fluxo, caso exista, solve a equação diferencial linearizada (4.4.2). Lema 4.4.3 mostra que a segunda equação em (4.4.2) sempre possui uma solução $\xi : I \rightarrow \mathbb{R}^n$ no intervalo de existência $I = I(x_0)$ inteiro. A tarefa na prova de Teorema 4.4.1 ora consiste em mostrar que esta solução realmente é dada pela derivada de ϕ_t por meio de (4.4.1).

Demonstração do Teorema 4.4.1. Primeiro vamos provar a afirmação para $\ell = 1$. Seja então $f : U \rightarrow \mathbb{R}^n$ um campo de vetores continuamente diferenciável. Resta mostrar que o fluxo $\phi : \Omega \rightarrow U$ de f também é continuamente diferenciável. Seja $x_0 \in U$ e $I := I(x_0) \subset \mathbb{R}$. Defina as aplicações

$$x : I \rightarrow U, \quad A : I \rightarrow \mathbb{R}^{n \times n}$$

através de

$$x(t) := \phi_t(x_0), \quad A(t) := df(x(t)) \quad (4.4.5)$$

para $t \in I$. Segundo o Lema 4.4.3 existe uma única aplicação continuamente diferenciável $\Phi : I \rightarrow \mathbb{R}^{n \times n}$, tal que

$$\dot{\Phi}(t) := A(t)\Phi(t), \quad \Phi(0) = \mathbb{1} \quad (4.4.6)$$

para todos os $t \in I$. Provaremos para já que para todo $t \in I$ a aplicação $\phi_t : U_t \rightarrow U_{-t}$ é diferenciável no ponto $x_0 \in U_t$ com a matriz jacobiana

$$d\phi_t(x_0) = \Phi(t) \quad \text{para } t \in I. \quad (4.4.7)$$

Seja $T \in I$ com $T > 0$. Então o seguinte deve ser mostrado.

Afirmção 1: Para todo $\varepsilon > 0$ existe um $\delta > 0$, tal que para todos os $t \in [0, T]$ e todos os $\xi_0 \in \mathbb{R}^n$ com $\|\xi_0\| < \delta$ vale $x_0 + \xi_0 \in U_t$ e

$$\|\phi_t(x_0 + \xi_0) - \phi_t(x_0) - \Phi(t)\xi_0\| \leq \varepsilon \|\xi_0\|. \quad (4.4.8)$$

Demonstração da Afirmção 1. Seja dado um número real $\varepsilon > 0$. Determinamos a constante $\delta > 0$ nas seguintes quatro passos.

(a) Segundo o Teorema 4.3.1 existe uma constante $r > 0$, tal que

$$[0, T] \times \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\} \subset \Omega.$$

(b) Segundo o Teorema 4.3.1 existe uma constante $c > 0$, tal que

$$\|\phi_t(x_0 + \xi_0) - \phi_t(x_0)\| \leq c \|\xi_0\|, \quad \|df(x(t))\| \leq c$$

para todos os $t \in [0, T]$ e todos os $\xi_0 \in \mathbb{R}^n$ com $\|\xi_0\| \leq r$.

(c) Existe uma constante $\rho > 0$, tal que para todos os $t \in [0, T]$ e todos os $h \in \mathbb{R}^n$ com $\|h\| < \rho$ vale $x(t) + h \in U$ e

$$\|df(x(t) + h) - df(x(t))\| \leq e^{-cT} \varepsilon.$$

Aqui usamos o fato que uma aplicação contínua é uniformemente contínua em todo subconjunto compacto da sua região de definição.

(d) Escolha $\delta > 0$, tal que $\delta \leq r$ e $c\delta \leq \rho$.

Provaremos que Afirmção 1 vale com este δ . Para $t \in [0, T]$ e $h \in \mathbb{R}^n$ com $\|h\| < \rho$ seja

$$R(t, h) := f(x(t) + h) - f(x(t)) - df(x(t))h. \quad (4.4.9)$$

Então vale

$$R(t, h) = \int_0^1 (df(x(t) + sh) - df(x(t)))h \, ds$$

e daí, segundo (c),

$$\|R(t, h)\| \leq \int_0^1 \|df(x(t) + sh) - df(x(t))\| \|h\| \, ds \leq e^{-cT} \varepsilon \|h\| \quad (4.4.10)$$

para todos os $t \in [0, T]$ e todos os $h \in \mathbb{R}^n$ com $\|h\| < \rho$.

Ora, seja $\xi_0 \in \mathbb{R}^n$ com $\|\xi_0\| < \delta$ e defina $\eta : [0, T] \rightarrow \mathbb{R}^n$ através de

$$\eta(t) := \phi_t(x_0 + \xi_0) - \phi_t(x_0) - \Phi(t)\xi_0 \quad \text{para } 0 \leq t \leq T. \quad (4.4.11)$$

Então $\eta(0) = 0$ e

$$\begin{aligned} \dot{\eta}(t) &= f(\phi_t(x_0 + \xi_0)) - f(\phi_t(x_0)) - A(t)\Phi(t)\xi_0 \\ &= f(\phi_t(x_0 + \xi_0)) - f(\phi_t(x_0)) - A(t)(\phi_t(x_0 + \xi_0) - \phi_t(x_0)) + A(t)\eta(t) \\ &= R(t, \phi_t(x_0 + \xi_0) - \phi_t(x_0)) + A(t)\eta(t). \end{aligned}$$

Segundo (a), (b), a (d) vale $x_0 + \xi_0 \in U_t$ e

$$\|\phi_t(x_0 + \xi_0) - \phi_t(x_0)\| \leq c\|\xi_0\| < c\delta \leq \rho \quad \text{para } 0 \leq t \leq T.$$

Segundo (4.4.10) segue dali a desigualdade

$$\begin{aligned} \|R(t, \phi_t(x_0 + \xi_0) - \phi_t(x_0))\| &\leq e^{-cT}\varepsilon\|\phi_t(x_0 + \xi_0) - \phi_t(x_0)\| \\ &\leq ce^{-cT}\varepsilon\|\xi_0\| \end{aligned}$$

para $0 \leq t \leq T$. Disso, em contrapartida, segue

$$\begin{aligned} \|\dot{\eta}(t)\| &\leq \|R(t, \phi_t(x_0 + \xi_0) - \phi_t(x_0))\| + \|A(t)\eta(t)\| \\ &\leq ce^{-cT}\varepsilon\|\xi_0\| + c\|\eta(t)\| \\ &\leq ce^{-cT}\varepsilon\|\xi_0\| + c\int_0^t \|\dot{\eta}(s)\| ds \end{aligned}$$

para $0 \leq t \leq T$. Segundo o Lema 4.3.2 de Gronwall vale então a desigualdade

$$\|\dot{\eta}(t)\| \leq ce^{-cT}\varepsilon\|\xi_0\| e^{ct}$$

e daí segue

$$\begin{aligned} \|\eta(t)\| &\leq \int_0^t \|\dot{\eta}(s)\| ds \\ &\leq e^{-cT}\varepsilon\|\xi_0\| \int_0^t ce^{cs} ds \\ &\leq \varepsilon\|\xi_0\| \end{aligned}$$

para $0 \leq t \leq T$. Como $\eta(t)$ foi definido através de (4.4.11), a Afirmação 1 fica provado com isso. Assim temos mostrado que ϕ_t é diferenciável para $t > 0$ e a prova para $t < 0$ é análoga.

Em seguida provaremos que a aplicação $\Omega \rightarrow \mathbb{R}^{n \times n} : (t, x) \mapsto d\phi_t(x)$ é contínua. Seja então $x_0 \in U$ e seja $T \in I(x_0)$ com $T > 0$. Defina as aplicações $A, \Phi : [0, T] \rightarrow \mathbb{R}^{n \times n}$ através de

$$A(t) := df(\phi_t(x_0)), \quad \Phi(t) := d\phi_t(x_0)$$

para $0 \leq t \leq T$.

Afirmação 2: Para todo $\varepsilon > 0$ existe um $\delta > 0$, tal que para todo $t \in [0, T]$ e todo $y_0 \in \mathbb{R}^n$ com $\|x_0 - y_0\| < \delta$ vale $y_0 \in U_t$ e

$$\|d\phi_t(x_0) - d\phi_t(y_0)\| < \varepsilon. \quad (4.4.12)$$

Demonstração da Afirmação 2. Seja dada uma constante $0 < \varepsilon \leq 1$. Determinamos a constante $\delta > 0$ nas seguintes três passos.

(a) Escolha $c > 0$, tal que

$$\|A(t)\| + 1 \leq c, \quad \|\Phi(t)\| \leq c \quad \text{para } 0 \leq t \leq T.$$

(b) Existe uma constante $\rho > 0$, tal que para todos os $t \in [0, T]$ e todos os $y \in \mathbb{R}^n$ com $\|y - \phi_t(x_0)\| < \rho$ vale $y \in U$ e

$$\|df(y) - df(\phi_t(x_0))\| < e^{-cT}\varepsilon.$$

(c) Segundo o Teorema 4.3.1 existe uma constante $\delta > 0$ tal que para todos os $t \in [0, T]$ e todos os $y_0 \in \mathbb{R}^n$ com $\|x_0 - y_0\| < \delta$ vale que $y_0 \in U_t$ e

$$\|\phi_t(x_0) - \phi_t(y_0)\| < \rho.$$

Seja ora $y_0 \in \mathbb{R}^n$ com $\|x_0 - y_0\| < \delta$ e defina $B, \Psi : [0, T] \rightarrow \mathbb{R}^{n \times n}$ através de

$$B(t) := df(\phi_t(y_0)), \quad \Psi(t) := d\phi_t(y_0)$$

para $0 \leq t \leq T$. Segundo do que já foi provado Ψ é continuamente diferenciável e $\dot{\Psi}(t) = B(t)\Psi(t)$ para $0 \leq t \leq T$ e $\Psi(0) = \mathbb{1}$. Além disso, segundo (c), vale $\|\phi_t(x_0) - \phi_t(y_0)\| < \rho$ para $0 \leq t \leq T$. Dali, segundo (b), segue

$$\|A(t) - B(t)\| \leq e^{-cT}\varepsilon \leq 1 \quad \text{para } 0 \leq t \leq T$$

e daí por sua vez, conforme (a), segue

$$\|B(t)\| \leq \|A(t)\| + 1 \leq c \quad \text{para } 0 \leq t \leq T.$$

Destas duas desigualdades segue

$$\begin{aligned}
\|\dot{\Phi}(t) - \dot{\Psi}(t)\| &= \|A(t)\Phi(t) - B(t)\Psi(t)\| \\
&\leq \|A(t)\Phi(t) - B(t)\Phi(t)\| + \|B(t)\Phi(t) - B(t)\Psi(t)\| \\
&< ce^{-cT}\varepsilon + c\|\Phi(t) - \Psi(t)\| \\
&\leq ce^{-cT}\varepsilon + c\int_0^T \|\dot{\Phi}(s) - \dot{\Psi}(s)\| ds
\end{aligned}$$

para $0 \leq t \leq T$. Segundo o Lema 4.3.2 de Gronwall segue daí a desigualdade $\|\dot{\Phi}(t) - \dot{\Psi}(t)\| < e^{-cT}ce^{ct}\varepsilon$ e dali resulta através de integração a desigualdade $\|\Phi(t) - \Psi(t)\| < \varepsilon$ para $0 \leq t \leq T$. Com isso a Afirmação 2 fica provada.

Temos mostrado então que a aplicação $(t, x) \mapsto d\phi_t(x)$ é contínua para $t \geq 0$ e a prova para $t \leq 0$ é análoga. Com isso todas as derivadas parciais $\partial_t \phi = f \circ \phi$ e $\partial \phi / \partial x_i$ para $i = 1, \dots, n$ são contínuas, e assim $\phi : \Omega \rightarrow U$ é continuamente diferenciável.

Com isso temos provado Teorema 4.4.1 para $\ell = 1$. Ora, seja $\ell \in \mathbb{N}$ e suponhamos que o teorema vale para este ℓ (e qualquer campo de vetores de classe C^ℓ em qualquer subconjunto aberto de um espaço euclidiano de qualquer dimensão). Seja $f : U \rightarrow \mathbb{R}^n$ um campo de vetores de classe $C^{\ell+1}$ e seja $\phi : \Omega \rightarrow \mathbb{R}^n$ o fluxo de f . Então ϕ é uma C^ℓ -aplicação, conforme hipótese da indução, e temos que mostrar que ϕ é uma aplicação de classe $C^{\ell+1}$. Para esse fim definimos o conjunto aberto

$$\tilde{U} := U \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$$

e a aplicação $\tilde{f} : \tilde{U} \rightarrow \mathbb{R}^n \times \mathbb{R}^n$ assim

$$\tilde{f}(x, \xi) := (f(x), df(x)\xi) \quad \text{para } x \in U \text{ e } \xi \in \mathbb{R}^n.$$

Além disso, seja $\tilde{\Omega} := \Omega \times \mathbb{R}^n$ e seja a aplicação $\tilde{\phi} : \tilde{\Omega} \rightarrow \tilde{U}$ definida através de

$$\tilde{\phi}(t, x, \xi) := (\phi_t(x), d\phi_t(x)\xi) \quad \text{para } (t, x) \in \Omega \text{ e } \xi \in \mathbb{R}^n,$$

Então \tilde{f} é um campo de vetores C^ℓ e $\tilde{\phi}$ é o fluxo de \tilde{f} como temos provado acima. Conforme hipótese da indução $\tilde{\phi}$ é então uma aplicação C^ℓ . Disso segue que todas as derivadas parciais $\partial_t \phi = f \circ \phi$ e $\partial \phi / \partial x_i$ para $i = 1, \dots, n$ são aplicações C^ℓ . Com isso fica provado que ϕ é uma aplicação $C^{\ell+1}$.

Este argumento de indução mostra que a afirmação do teorema vale para todo $\ell \in \mathbb{N}$. Que também vale para $\ell = \infty$ segue do fato que uma função é suave se e somente se é ℓ vezes continuamente diferenciável para todo $\ell \in \mathbb{N}$. Com isso Teorema 4.4.1 fica provado. \square

No caso de $\ell = \infty$ Teorema 4.4.1 diz que o fluxo $\phi : \Omega \rightarrow U$ de um campo de vetores $f : U \rightarrow \mathbb{R}^n$ também é uma aplicação suave, e particularmente, que a aplicação $\phi_t : U_t \rightarrow U_{-t}$ em Comentário 4.3.3 é um difeomorfismo suave para todo $t \in \mathbb{R}$. Isto fornece, em muitas situações, métodos úteis para a construção de difeomorfismos suaves, no que construí-se primeiramente um campo de vetores suave com as propriedades desejadas, e em seguida definem-se os difeomorfismos mediante as soluções da equação diferencial. Além disso, podemos utilizar Teorema 4.4.1 para mostrar que a aplicação exponencial é suave.

Corolário 4.4.4. *A aplicação exponencial $\exp : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ é suave.*

Demonstração. Seja $f : \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ o campo de vetores definido através de

$$f(A, B) := (0, AB) \quad \text{para } A, B \in \mathbb{R}^{n \times n}.$$

O fluxo de f é a aplicação $\phi : \mathbb{R} \times \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}$ definida através de

$$\phi(t, A, B) := (A, e^{At}B) \quad \text{para } t \in \mathbb{R} \text{ e } A, B \in \mathbb{R}^{n \times n}.$$

Como f é suave segue de Teorema 4.4.1 que ϕ é suave. Por isso a aplicação

$$\mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n} : A \mapsto \phi(1, A, \mathbb{1}) = (A, e^A)$$

também é suave. A aplicação exponencial $\exp : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ é a composição desta aplicação com a projeção no segundo fator e, por isso, também é suave. Com isso Corolário 4.4.4 fica provado. \square

Capítulo 5

Integrais Múltiplas

O objetivo deste capítulo é definir a integral de Riemann de uma função de várias variáveis sobre um conjunto compacto Jordan-mensurável, bem como deduzir as propriedades fundamentais da integral de Riemann.

5.1 A integral de Riemann

Para a definição da integral de Riemann é necessário a introdução de umas noções básicas, como aquela de uma partição em dimensões superiores.

Partições

Sejam $a = (a_1, \dots, a_n)$ e $b = (b_1, \dots, b_n)$ dois vetores reais com $a_i < b_i$ para todo $i = 1, \dots, n$. O conjunto

$$\begin{aligned} Q &:= Q(a, b) \\ &:= \{x \in \mathbb{R}^n \mid a_i < x_i < b_i \text{ para } i = 1, \dots, n\} \\ &= (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n) \end{aligned} \tag{5.1.1}$$

chama-se de **bloco aberto (eixo-paralelo)**. O **volume de Q** é o número

$$\text{Vol}_n(Q) := \prod_{i=1}^n (b_i - a_i). \tag{5.1.2}$$

Importa sublinhar que assim o volume é definido até agora **só** para blocos abertos eixo-paralelos. Se $Q \subset \mathbb{R}^n$ é definido através de (5.1.1), então o fecho

de Q (com respeito á norma euclidiana no \mathbb{R}^n) é o conjunto

$$\begin{aligned}\bar{Q} &= \{x \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \text{ para } i = 1, \dots, n\} \\ &= [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]\end{aligned}\tag{5.1.3}$$

(Exercício: prova esta fórmula.) Todo tal conjunto é chamado de **bloco fechado (eixo-paralelo)**. Frequentemente também é útil definir o **volume de \bar{Q}** através do lado direito da equação (5.1.2). Mais geral, definimos

$$\text{Vol}_n(B) := \prod_{i=1}^n (b_i - a_i) \quad \text{para } Q \subset B \subset \bar{Q}.$$

Definição 5.1.1. *Seja $B \subset \mathbb{R}^n$ um subconjunto compacto. Uma partição de B é um conjunto finito $P = \{P_1, \dots, P_k\}$ de blocos abertos eixo-paralelos, tal que*

$$B = \bigcup_{i=1}^k \bar{P}_i, \quad P_i \cap P_j = \emptyset \quad \text{para } i \neq j.$$

O conjunto das partições de B denota-se de $\mathcal{P}(B)$.

Um exemplo de uma partição no caso de $n = 2$ é visualizado na Figura 5.1. Nota-se que não todo subconjunto do \mathbb{R}^n possui uma partição. Se $B \subset \mathbb{R}^n$ possui uma partição, então B é necessariamente compacto e não-vazio, e ainda para subconjuntos compactos não-vazios do \mathbb{R}^n a existência de uma partição é uma propriedade *rara*.

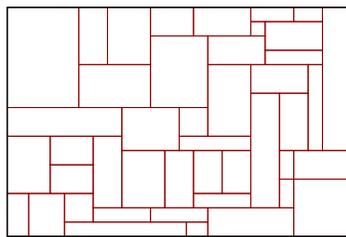


Figura 5.1: Uma partição de um retângulo.

Exemplo 5.1.2. No caso de $n = 1$ pode-se representar uma partição de um intervalo compacto $B = [a, b]$ através de uma sequência finita da forma $a = t_0 < t_1 < t_2 < \cdots < t_k = b$ com $P_i = (t_{i-1}, t_i)$ (veja [5]).

Exercício 5.1.3. O disco unitário fechado no \mathbb{R}^2 não possui nenhuma partição. Um triângulo fechado no \mathbb{R}^2 não possui nenhuma partição.

Exercício 5.1.4. Um subconjunto $B \subset \mathbb{R}^n$ possui uma partição se e somente se é uma união finita de blocos eixo-paralelos fechados. Um tal conjunto chama-se **prédio de blocos** (Definição 5.6.1).

Exercício 5.1.5. Sejam $K \subset U \subset \mathbb{R}^n$, tal que U é aberto e K compacto. Então existe um subconjunto $B \subset \mathbb{R}^n$ com $\mathcal{P}(B) \neq \emptyset$ e $K \subset B \subset U$.

Lema 5.1.6. *Seja $B \subset \mathbb{R}^n$ um subconjunto compacto e $P = \{P_1, \dots, P_k\}$ e $Q = \{Q_1, \dots, Q_\ell\}$ duas partições de B . Então também o conjunto*

$$P \wedge Q := \{P_i \cap Q_j \mid i = 1, \dots, k, j = 1, \dots, \ell, P_i \cap Q_j \neq \emptyset\} \quad (5.1.4)$$

é uma partição de B . O conjunto $\{P_i \cap Q_j \mid i = 1, \dots, k, P_i \cap Q_j \neq \emptyset\}$ é para todo j uma partição de $\overline{Q_j}$, e $\{P_i \cap Q_j \mid j = 1, \dots, \ell, P_i \cap Q_j \neq \emptyset\}$ é para todo i uma partição de $\overline{P_i}$.

Demonstração. Segue das definições que o conjunto $P_i \cap Q_j$, caso não-vazio, é um bloco eixo-paralelo aberto. Provaremos o seguinte.

Afirmção: *Para todo $x \in \overline{P_i}$ existe um j tal que $x \in \overline{P_i \cap Q_j}$.*

Pela hipótese existe uma sequência $x_\nu \in P_i$ a qual converge para um x . Como $P_i \subset B$, vale $x_\nu \in \overline{Q_1} \cup \dots \cup \overline{Q_\ell}$ para todos os ν . Portanto existe (pelo menos um) $j \in \{1, \dots, \ell\}$, tal que $\overline{Q_j}$ contem um número infinito de membros da sequência x_ν . Através de transição a uma sequência parcial podemos supor que vale $x_\nu \in \overline{Q_j}$ para todos os ν . Escolha uma sequência $0 < \varepsilon_\nu < 1/\nu$ com

$$B_{\varepsilon_\nu}(x_\nu) \subset P_i$$

para todos os ν . Então também vale $B_{\varepsilon_\nu}(x_\nu) \cap Q_j \neq \emptyset$ para todos os ν . Segundo o axioma da escolha, versão contável, existe então uma sequência $y_\nu \in Q_j$, tal que vale $\|x_\nu - y_\nu\| < \varepsilon_\nu$ para todo ν . Então vale $y_\nu \in P_i \cap Q_j$ e

$$x = \lim_{\nu \rightarrow \infty} x_\nu = \lim_{\nu \rightarrow \infty} y_\nu.$$

Portanto $x \in \overline{P_i \cap Q_j}$ e com isso a afirmação fica provada.

A afirmação diz que

$$\overline{P}_i = \bigcup_{j=1}^{\ell} \overline{P_i \cap Q_j}$$

para todos os i . Igualmente vale $\overline{Q}_j = \bigcup_{i=1}^k \overline{P_i \cap Q_j}$ e

$$B = \bigcup_{i=1}^k \overline{P}_i = \bigcup_{i=1}^k \bigcup_{j=1}^{\ell} \overline{P_i \cap Q_j}.$$

Com isso Lema 5.1.6 fica provado. \square

Lema 5.1.7. *Seja $Q = Q(a, b) \subset \mathbb{R}^n$ um bloco eixo-paralelo aberto e $P = (P_1, \dots, P_k)$ uma partição de $B := \overline{Q}$. Então vale*

$$\text{Vol}_n(Q) = \sum_{i=1}^k \text{Vol}_n(P_i).$$

Demonstração. A demonstração resta numa ideia de John von Neumann (quem estudou de 1921 até 1923 na ETH Zurique). Para todo $\varepsilon > 0$ consideramos a grade

$$\Lambda_\varepsilon := \varepsilon\mathbb{Z}^n$$

e investigamos o número dos pontos da grade as quais pertencem ao bloco aberto Q , respectivamente ao bloco fechado \overline{Q} . Agora vale a seguinte desigualdade

$$\prod_{i=1}^n (b_i - a_i - \varepsilon) \leq \varepsilon^n \#(\Lambda_\varepsilon \cap Q) \leq \varepsilon^n \#(\Lambda_\varepsilon \cap \overline{Q}) \leq \prod_{i=1}^n (b_i - a_i + \varepsilon). \quad (5.1.5)$$

Para a prova de (5.1.5) consideramos para já o caso $n = 1$. Sejam então $a < b$ dois números reais e seja $\varepsilon > 0$. Escolha $\ell, m \in \mathbb{Z}$, tal que

$$(\ell - 1)\varepsilon < a \leq \ell\varepsilon, \quad m\varepsilon \leq b < (m + 1)\varepsilon.$$

Então vale $\varepsilon\mathbb{Z} \cap [a, b] = \{j\varepsilon \mid \ell \leq j \leq m\}$ e daí

$$\#(\varepsilon\mathbb{Z} \cap [a, b]) = m - \ell + 1 \leq \frac{b - a + \varepsilon}{\varepsilon}.$$

Ora escolha $\ell, m \in \mathbb{Z}$, tal que

$$(\ell - 1)\varepsilon \leq a < \ell\varepsilon, \quad m\varepsilon < b \leq (m + 1)\varepsilon.$$

Então vale $\varepsilon\mathbb{Z} \cap (a, b) = \{j\varepsilon \mid \ell \leq j \leq m\}$ e daí

$$\#(\varepsilon\mathbb{Z} \cap (a, b)) = m - \ell + 1 = (m + 1) - (\ell - 1) - 1 \geq \frac{b - a - \varepsilon}{\varepsilon}.$$

Dali segue a desigualdade

$$b - a - \varepsilon \leq \varepsilon\#(\varepsilon\mathbb{Z} \cap (a, b)) \leq \varepsilon\#(\varepsilon\mathbb{Z} \cap [a, b]) \leq b - a + \varepsilon.$$

Substituimos ora a, b por a_i, b_i e formamos o produto sobre $i = 1, \dots, n$, então resulta a desigualdade (5.1.5).

De (5.1.5) segue através do limite $\varepsilon \rightarrow 0$ que

$$\text{Vol}_n(Q) = \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap Q) = \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap \overline{Q}). \quad (5.1.6)$$

Ora, obviamente vale

$$\sum_{i=1}^k \#(\Lambda_\varepsilon \cap P_i) \leq \#(\Lambda_\varepsilon \cap Q) \leq \sum_{i=1}^k \#(\Lambda_\varepsilon \cap \overline{P}_i).$$

Multiplicamos esta desigualdade com ε^n e formamos o limite $\varepsilon \rightarrow 0$, então resulta sob uso triplo de (5.1.6) que

$$\begin{aligned} \sum_i \text{Vol}_n(P_i) &= \lim_{\varepsilon \rightarrow 0} \sum_i \varepsilon^n \#(\Lambda_\varepsilon \cap P_i) \\ &\leq \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap Q) = \text{Vol}_n(Q) \\ &\leq \lim_{\varepsilon \rightarrow 0} \sum_i \varepsilon^n \#(\Lambda_\varepsilon \cap \overline{P}_i) \\ &= \sum_i \text{Vol}_n(P_i). \end{aligned}$$

Dali segue imediato a afirmação de Lema 5.1.7. □

Somas superiores e somas inferiores

Definição 5.1.8. *Seja $B \subset \mathbb{R}^n$ um conjunto compacto e $P = \{P_1, \dots, P_k\}$ uma partição de B . Seja $f : B \rightarrow \mathbb{R}$ uma função limitada. A **soma superior de f e P** é o número*

$$\overline{S}(f, P) := \sum_{i=1}^k \left(\sup_{\overline{P}_i} f \right) \text{Vol}_n(P_i). \quad (5.1.7)$$

A **soma inferior de f e P** é o número

$$\underline{S}(f, P) := \sum_{i=1}^k \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i). \quad (5.1.8)$$

Lema 5.1.9. *Seja $B \subset \mathbb{R}^n$ compacto, seja $f : B \rightarrow \mathbb{R}$ uma função limitada, e sejam*

$$P = \{P_1, \dots, P_k\}, \quad Q = \{Q_1, \dots, Q_\ell\}$$

partições de B . Então vale

$$\underline{S}(f, P) \leq \underline{S}(f, P \wedge Q) \leq \overline{S}(f, P \wedge Q) \leq \overline{S}(f, Q).$$

Demonstração. Segue da Definição 5.1.8 assim como de Lema 5.1.6 e de Lema 5.1.7 que

$$\begin{aligned} \underline{S}(f, P) &= \sum_i \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \\ &= \sum_i \sum_j \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\leq \sum_i \sum_j \left(\inf_{\overline{P}_i \cap Q_j} f \right) \text{Vol}_n(P_i \cap Q_j) = \underline{S}(f, P \wedge Q) \\ &\leq \sum_i \sum_j \left(\sup_{\overline{P}_i \cap Q_j} f \right) \text{Vol}_n(P_i \cap Q_j) = \overline{S}(f, P \wedge Q) \\ &\leq \sum_j \sum_i \left(\sup_{\overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &= \sum_j \left(\sup_{\overline{Q}_j} f \right) \text{Vol}_n(Q_j) \\ &= \overline{S}(f, Q). \end{aligned}$$

Com isso Lema 5.1.9 fica provado. □

Definição da integral de Riemann

Segundo o Lema 5.1.9, todo subconjunto compacto $B \subset \mathbb{R}^n$ com $\mathcal{P}(B) \neq \emptyset$ e toda função limitada $f : B \rightarrow \mathbb{R}$ satisfaz a desigualdade

$$\sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) \leq \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q). \quad (5.1.9)$$

Definição 5.1.10. *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$. Uma função limitada $f : B \rightarrow \mathbb{R}$ chama-se de **Riemann-integrável**, se*

$$\sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) = \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q).$$

Se $f : B \rightarrow \mathbb{R}$ é Riemann-integrável, então o número real

$$\int_B f(x) dx_1 \cdots dx_n := \sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) = \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q) \quad (5.1.10)$$

é chamado de **integral (de Riemann) de f sobre B** .

Exemplo 5.1.11. Seja $B \subset \mathbb{R}^n$ um compacto que possui uma partição. Como primeiro exemplo consideramos uma função constante $f : B \rightarrow \mathbb{R}$. Seja então $c \in \mathbb{R}$, tal que $f(x) = c$ para todos os $x \in B$. Se $P = \{P_1, \dots, P_k\}$ é uma partição de B , então vale $\inf_{\overline{P}_i} f = \sup_{\overline{P}_i} f = c$ para todos os i e daí

$$\underline{S}(f, P) = \overline{S}(f, P) = c \sum_{i=1}^k \text{Vol}_n(P_i).$$

Dali segue, segundo o Lema 5.1.9, que todos os $P, Q \in \mathcal{P}(B)$ satisfazem as desigualdades

$$\overline{S}(f, P) = \underline{S}(f, P) \leq \overline{S}(f, Q) = \underline{S}(f, Q) \leq \overline{S}(f, P).$$

Portanto as somas superiores e inferiores de f são independente da escolha da partição. Segue disso que a função constante $f \equiv c$ é Riemann-integrável e que vale a equação

$$\int_B c dx_1 \cdots dx_n = c \sum_{i=1}^k \text{Vol}_n(P_i)$$

para toda partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$.

Definição 5.1.12 (Volume de um prédio de blocos). *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$. A integral da função constante $f \equiv 1$ sobre B chama-se de **volume (n -dimensional) de B** e denota-se de*

$$\text{Vol}_n(B) := \int_B 1 \, dx_1 \cdots dx_n. \quad (5.1.11)$$

Segundo o Exemplo 5.1.11 vale a fórmula $\text{Vol}_n(B) = \sum_{i=1}^k \text{Vol}_n(P_i)$ para toda partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$.

Lema 5.1.13. *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$ e seja $f : B \rightarrow \mathbb{R}$ uma função limitada. As seguintes afirmações são equivalente.*

(i) *f é Riemann-integrável.*

(ii) *Para todo $\varepsilon > 0$ existe uma partição $P \in \mathcal{P}(B)$ com*

$$\overline{S}(f, P) - \underline{S}(f, P) < \varepsilon.$$

Demonstração. Provaremos para já que (ii) implica (i). Seja então (ii) satisfeito. Então vale a desigualdade

$$\inf_{Q \in \mathcal{P}(P)} \overline{S}(f, Q) < \sup_{P \in \mathcal{P}(P)} \underline{S}(f, P) + \varepsilon$$

para todo $\varepsilon > 0$. Segue daí

$$\inf_{Q \in \mathcal{P}(P)} \overline{S}(f, Q) \leq \sup_{P \in \mathcal{P}(P)} \underline{S}(f, P).$$

Dali segue da desigualdade (5.1.9) que f é Riemann-integrável.

Provaremos para já que (i) implica (ii). Seja então $f : B \rightarrow \mathbb{R}$ uma função Riemann-integrável e seja

$$c := \int_B f(x) dx_1 \cdots dx_n.$$

Seja $\varepsilon > 0$. Então existem segundo a definição da integral Riemann duas partições $P, Q \in \mathcal{P}(B)$, tal que

$$\underline{S}(f, P) > c - \frac{\varepsilon}{2}, \quad \overline{S}(f, Q) < c + \frac{\varepsilon}{2}.$$

Daí segue, segundo o Lema 5.1.9, que

$$\overline{S}(f, P \wedge Q) - \underline{S}(f, P \wedge Q) \leq \overline{S}(f, Q) - \underline{S}(f, P) < \varepsilon.$$

Com isso Lema 5.1.13 fica provado. \square

Exercício 5.1.14. Toda função contínua $f : B \rightarrow \mathbb{R}$ num subconjunto compacto $B \subset \mathbb{R}^n$ com $\mathcal{P}(B) \neq \emptyset$ é Riemann-integrável. **Dica:** Lema 5.1.13.

5.2 Conjuntos Jordan-nulos e continuidade

Segundo o exercício anterior toda função contínua $f : B \rightarrow \mathbb{R}$ num subconjunto compacto $B \subset \mathbb{R}^n$ com $\mathcal{P}(B) \neq \emptyset$ é Riemann-integrável. Nesta seção vamos provar uma afirmação bem mais geral a qual já garante a integrabilidade de Riemann de uma função limitada $f : B \rightarrow \mathbb{R}$ se ela é contínua fora de um chamado assim *conjunto Jordan-nulo*.

Definição 5.2.1. *Chama-se um subconjunto $A \subset \mathbb{R}^n$ de conjunto Jordan-nulo se para todo $\varepsilon > 0$ existe um número finito de blocos abertos W_1, \dots, W_N no \mathbb{R}^n cobrindo $A \subset \bigcup_{\nu=1}^N W_\nu$ e com soma de volumes $\sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon$.*

Teorema 5.2.2. *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$, seja $A \subset B$ um conjunto Jordan-nulo, e seja $f : B \rightarrow \mathbb{R}$ uma função limitada.*

(i) *Se f é contínua em $B \setminus A$, então f é Riemann-integrável.*

(ii) *Se $f(x) = 0$ para todos os $x \in B \setminus A$, então $\int_B f = 0$.*

Demonstração. Veja a página 171. □

Segue direto de Definição 5.2.1 que todo conjunto Jordan-nulo é limitado, que todo subconjunto de um conjunto Jordan-nulo também é um conjunto Jordan-nulo, e que toda união finita de conjuntos Jordan-nulos igualmente é um conjunto Jordan-nulo.

Lema 5.2.3. *Se $A \subset \mathbb{R}^n$ é um conjunto Jordan-nulo, então $\mathring{A} = \emptyset$ e o fecho \bar{A} também é um conjunto Jordan-nulo.*

Demonstração. Se $\mathring{A} \neq \emptyset$, então existe um cubo eixo-paralelo aberto $Q \subset A$ com $\delta := \text{Vol}_n(Q) > 0$. Ora se $W_1, \dots, W_N \subset \mathbb{R}^n$ são blocos eixo-paralelos abertos com $A \subset \bigcup_{\nu=1}^N W_\nu$, então segue do argumento de von Neumann na prova de Lema 5.1.7 a desigualdade

$$\delta = \lim_{\varepsilon \rightarrow 0} \varepsilon^n \#(\Lambda_\varepsilon \cap Q) \leq \lim_{\varepsilon \rightarrow 0} \sum_{\nu=1}^N \varepsilon^n \#(\Lambda_\varepsilon \cap W_\nu) = \sum_{\nu=1}^N \text{Vol}_n(W_\nu),$$

em contradição à hipótese que A é um conjunto Jordan-nulo.

Ora, seja dado $\varepsilon > 0$ e sejam $W_1, \dots, W_N \subset \mathbb{R}^n$ blocos eixo-paralelos abertos com $A \subset \bigcup_{\nu=1}^N W_\nu$ e $\sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \frac{\varepsilon}{2}$. Para $\nu = 1, \dots, N$ seja W'_ν um bloco eixo-paralelo aberto com $\bar{W}_\nu \subset W'_\nu$ e $\text{Vol}_n(W'_\nu) \leq 2\text{Vol}_n(W_\nu)$. Então vale $\bar{A} \subset \bigcup_{\nu=1}^N \bar{W}_\nu \subset \bigcup_{\nu=1}^N W'_\nu$ e $\sum_{\nu=1}^N \text{Vol}_n(W'_\nu) < \varepsilon$. Com isso Lema 5.2.3 fica provado. □

Exemplo 5.2.4 (Conjuntos Jordan-nulos).

(i) Se $B \subset \mathbb{R}^{n-1}$ é um conjunto limitado e $c \in \mathbb{R}$, então o conjunto

$$A := B \times \{c\} \subset \mathbb{R}^{n-1} \times \mathbb{R} = \mathbb{R}^n$$

é um conjunto Jordan-nulo. **Prova:** Escolha $R > 0$ e $\delta > 0$ com

$$B \subset (-R, R)^{n-1}, \quad \delta < \frac{\varepsilon}{2^n R^{n-1}}.$$

Então o conjunto $W := (-R, R)^{n-1} \times (c - \delta, c + \delta)$ é um bloco aberto contendo A e vale $\text{Vol}_n(W) = (2R)^{n-1} 2\delta < \varepsilon$.

(ii) Seja $Q \subset \mathbb{R}^n$ um bloco eixo-paralelo aberto. Então ∂Q é um conjunto Jordan-nulo segundo (i).

(iii) Um subconjunto não-vazio aberto do \mathbb{R}^n não é um conjunto Jordan-nulo segundo o Lema 5.2.3.

(iv) O conjunto limitado contável $A := \mathbb{Q} \cap [0, 1] \subset \mathbb{R}$ não é um conjunto Jordan-nulo, porque $\overline{A} = [0, 1]$ não é um conjunto Jordan-nulo segundo (iii).

(v) O conjunto Cantor padrão $K \subset [0, 1]$ é um conjunto Jordan-nulo.

(vi) Pode-se modificar a construção do conjunto Cantor tal que receba-se um conjunto compacto $K' \subset [0, 1]$ o qual não é um conjunto Jordan-nulo. Neste caso $U := [0, 1] \setminus K'$ é um subconjunto aberto de \mathbb{R} cujo bordo $\partial U = K'$ não é um conjunto Jordan-nulo.

Exemplo 5.2.5 (Integrabilidade de Riemann e monotonia).

(i) Seja $f : [0, 1]^2 \rightarrow \mathbb{R}$ uma função a qual é monótona em ambas variáveis, ou seja, para todos os $x, y, x', y' \in \mathbb{R}$ vale

$$x \leq x' \text{ e } y \leq y' \implies f(x, y) \leq f(x', y'). \quad (5.2.1)$$

Então f é Riemann-integrável: Se $P = \{P_{ij} \mid i, j = 1, \dots, N\} \in \mathcal{P}([0, 1]^2)$ com $P_{ij} := (\frac{i-1}{N}, \frac{i}{N}) \times (\frac{j-1}{N}, \frac{j}{N})$, então vale $\overline{S}(f, P) - \underline{S}(f, P) = \frac{f(1,1) - f(0,0)}{N}$.

(ii) Seja $\mathbb{N} \rightarrow \mathbb{Q} \cap [0, 1] : i \mapsto q_i$ uma aplicação bijetiva e definida a função $f : [0, 1]^2 \rightarrow \mathbb{R}$ através de $f(x, 0) := f(0, y) := 0$ para $x, y \in [0, 1]$ e

$$f(x, y) := \sum_{i \in \mathbb{N}, q_i \leq x} 2^{-i} + \sum_{j \in \mathbb{N}, q_j \leq y} 2^{-j} \quad (5.2.2)$$

para $x, y > 0$. Esta função satisfaz (5.2.1) e daí é Riemann-integrável. Demais f é contínua no ponto $(x, y) \in [0, 1]^2$ se e somente se x e y ambos são irracionais. Dali o conjunto das descontinuidades de f não é um conjunto Jordan-nulo.

Definição 5.2.6. *Seja $B \subset \mathbb{R}^n$ um conjunto compacto e $P = \{P_1, \dots, P_k\}$ uma partição de B . Então todo bloco aberto P_i é um produto de intervalos abertos*

$$P_i =: (a_{i1}, b_{i1}) \times (a_{i2}, b_{i2}) \times \cdots \times (a_{in}, b_{in})$$

com $a_{ij} < b_{ij}$. Para $i = 1, \dots, k$ chama-se o número

$$\delta(P_i) := \max \{b_{i1} - a_{i1}, b_{i2} - a_{i2}, \dots, b_{in} - a_{in}\}$$

de comprimento do lado máximo de P_i . O número

$$\delta(P) := \max_{i=1, \dots, k} \delta(P_i) = \max_{\substack{i=1, \dots, k \\ j=1, \dots, n}} (b_{ij} - a_{ij})$$

chama-se de **fineza da partição P** .

Exercício 5.2.7. *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$. Então existe para todo $\delta_0 > 0$ uma partição $P \in \mathcal{P}(B)$ com $\delta(P) < \delta_0$.*

O seguinte lema é da significância fundamental para o entendimento de conjuntos Jordan-nulos.

Lema 5.2.8 (O lema dos conjuntos nulos). *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$, seja $A \subset \mathbb{R}^n$ um conjunto Jordan-nulo, e seja $\varepsilon > 0$. Então existe um número $\delta_0 > 0$, tal que para todo $P \in \mathcal{P}(B)$ vale*

$$\delta(P) < \delta_0 \quad \implies \quad \sum_{\overline{P_i} \cap A \neq \emptyset} \text{Vol}_n(P_i) < \varepsilon.$$

Demonstração. Seja dado $\varepsilon > 0$. Segundo o Lema 5.2.3 o fecho \overline{A} é um conjunto Jordan-nulo. Daí existem blocos abertos $W_1, \dots, W_N \subset \mathbb{R}^n$, tal que

$$\overline{A} \subset \bigcup_{\nu=1}^N W_\nu, \quad \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon. \quad (5.2.3)$$

Passo 1. *Existe um $\rho > 0$, tal que para todo $a \in \overline{A}$ existe um $\nu \in \{1, \dots, N\}$ com $B_\rho(a) \subset W_\nu$.*

Para todo $a \in \overline{A}$ existe um número $\nu = \nu(a) \in \{1, \dots, N\}$ tal que $a \in W_\nu$. Como W_ν é um conjunto aberto, existe além disso para todo $a \in A$ um número $\rho(a) > 0$ com $B_{2\rho(a)}(a) \subset W_{\nu(a)}$. Então os conjuntos $B_{\rho(a)}(a)$, $a \in \overline{A}$

formam uma cobertura aberta do conjunto \bar{A} (Seção C.1). O fecho \bar{A} também é limitado e daí, segundo o Teorema de Heine–Borel, compacto sequencial. Segundo o Teorema C.1.2 existem dali um número finito de elementos $a_1, \dots, a_m \in \bar{A}$ com $\bar{A} \subset \bigcup_{i=1}^m B_{\rho(a_i)}(a_i)$. Para $i = 1, \dots, m$ introduzimos ora as abreviações $\rho_i := \rho(a_i)$ e $\nu_i := \nu(a_i)$. Seja $\rho := \min\{\rho_1, \dots, \rho_m\} > 0$. Então vale $B_{\rho_i+\rho}(a_i) \subset B_{2\rho_i}(a_i) \subset W_{\nu_i}$ para $i = 1, \dots, m$. Ora, seja $a \in \bar{A}$. Então existe $i \in \{1, \dots, m\}$ com $a \in B_{\rho_i}(a_i)$ e daí $B_\rho(a) \subset B_{\rho+\rho_i}(a_i) \subset W_{\nu_i}$. Com isso Passo 1 fica provado.

Passo 2. *Seja $\rho > 0$ como no Passo 1 e $\delta_0 := \rho/\sqrt{n}$. Seja $P = \{P_1, \dots, P_k\}$ uma partição de B com $\delta(P) < \delta_0$ e seja $j \in \{1, \dots, k\}$. Então vale*

$$\bar{P}_j \cap A \neq \emptyset \quad \implies \quad \exists \nu \in \{1, \dots, N\} \text{ com } \bar{P}_j \subset W_\nu.$$

Seja $a \in \bar{P}_j \cap A$. Segundo o Passo 1 existe $\nu \in \{1, \dots, N\}$ com $B_\rho(a) \subset W_\nu$. Como $\delta(P_j) < \delta_0$, todo $x \in \bar{P}_j$ satisfaz a desigualdade

$$\|x - a\| = \sqrt{\sum_{i=1}^n (x_i - a_i)^2} \leq \sqrt{n} \delta(P_j) < \sqrt{n} \delta_0 = \rho.$$

Consequentemente vale $\bar{P}_j \subset B_\rho(a) \subset W_\nu$. Com isso Passo 1 fica provado.

Passo 3. *A afirmação de Lema 5.2.8 vale com $\delta_0 = \rho/\sqrt{n}$.*

Seja $P = \{P_1, \dots, P_k\}$ uma partição de B com $\delta(P) < \delta_0$. Defina

$$J_A := \{j \in \{1, \dots, k\} \mid \bar{P}_j \cap A \neq \emptyset\}$$

e, para $\nu = 1, \dots, N$,

$$J_\nu := \{j \in \{1, \dots, k\} \mid \bar{P}_j \subset W_\nu\}.$$

Então vale segundo o Passo 2 que $J_A \subset \bigcup_{\nu=1}^N J_\nu$. Dali segue

$$\sum_{j \in J_A} \text{Vol}_n(P_j) \leq \sum_{\nu=1}^N \left(\sum_{j \in J_\nu} \text{Vol}_n(P_j) \right) \leq \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon.$$

Aqui a segunda desigualdade segue de Lema 5.1.7 e a última de (5.2.3). Com isso Lema 5.2.8 fica provado. \square

Demonstração de Teorema 5.2.2. A prova tem dois passos.

Passo 1. Se $f(x) = 0$ para $x \in B \setminus A$, então $f \in \mathcal{R}(B)$ e

$$\int_B f = 0.$$

Suponhamos sem perda de generalidade que f não anule-se identicamente. Portanto $\|f\| = \sup_B |f| \neq 0$. Seja ora dado $\varepsilon > 0$. Segundo o Lema 5.2.8 existe uma partição

$$P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$$

tal que

$$\sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) < \frac{\varepsilon}{4\|f\|}. \quad (5.2.4)$$

Então vale

$$\overline{S}(f, P) = \sum_{\overline{P}_i \cap A \neq \emptyset} \left(\sup_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \leq \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \leq \frac{\varepsilon}{4}$$

e igualmente

$$\underline{S}(f, P) = \sum_{\overline{P}_i \cap A \neq \emptyset} \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \geq -\|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \geq -\frac{\varepsilon}{4}.$$

Com isso o Passo 1 segue de Teorema 5.3.1.

Passo 2. Se f é contínuo em $B \setminus A$, então $f \in \mathcal{R}(B)$.

Seja dado $\varepsilon > 0$. Suponhamos que f não se anule identicamente e escolhemos uma partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$, tal que vale (5.2.4). Considere o conjunto compacto

$$K := \bigcup_{\overline{P}_i \cap A = \emptyset} \overline{P}_i.$$

A restrição $f|_K : K \rightarrow \mathbb{R}$ é contínuo. Segundo um Teorema de Análise I a restrição $f|_K$ é contínuo uniformemente. Seja

$$V := \sum_{i=1}^k \text{Vol}_n(P_i).$$

Então existe uma constante $\delta_0 > 0$ com

$$x, y \in K, \quad \max_{\nu=1, \dots, n} |x_\nu - y_\nu| < \delta_0 \quad \implies \quad |f(x) - f(y)| \leq \frac{\varepsilon}{2V}.$$

Seja $Q = \{Q_1, \dots, Q_\ell\} \in \mathcal{P}(B)$ uma partição com $\delta(Q) < \delta_0$. Então vale

$$|f(x) - f(y)| \leq \frac{\varepsilon}{2V}$$

para todos os $j = 1, \dots, \ell$ e todos os $x, y \in Q_j \cap K$, e dali

$$\overline{P}_i \cap A = \emptyset \quad \implies \quad \sup_{P_i \cap Q_j} f - \inf_{P_i \cap Q_j} f \leq \frac{\varepsilon}{2V}$$

para todos os i, j . Disso segue que

$$\begin{aligned} & \overline{S}(f, P \wedge Q) - \underline{S}(f, P \wedge Q) \\ &= \sum_i \sum_j \left(\sup_{P_i \cap Q_j} f - \inf_{P_i \cap Q_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\leq \sum_{\overline{P}_i \cap A = \emptyset} \sum_j \left(\sup_{P_i \cap Q_j} f - \inf_{P_i \cap Q_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\quad + 2 \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \sum_j \text{Vol}_n(P_i \cap Q_j) \\ &\leq \frac{\varepsilon}{2V} \sum_{\overline{P}_i \cap A = \emptyset} \sum_j \text{Vol}_n(P_i \cap Q_j) \\ &\quad + 2 \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \sum_j \text{Vol}_n(P_i \cap Q_j) \\ &= \frac{\varepsilon}{2V} \sum_{\overline{P}_i \cap A = \emptyset} \text{Vol}_n(P_i) + 2 \|f\| \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \\ &< \frac{\varepsilon}{2V} \sum_i \text{Vol}_n(P_i) + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Aqui a penúltima desigualdade segue de (5.2.4). Segundo o Lema 5.1.13 ora f é Riemann-integrável. Com isso Passo 2 e Teorema 5.2.2 ficam provados. \square

5.3 Somas de Riemann

O primeiro objetivo desta seção é a prova do seguinte critério importante para a Riemann-integrabilidade. O teorema mostra que uma função limitada é Riemann-integrável se e somente se as chamadas assim *somas de Riemann* convergem. Este teorema será usado para deduzir umas propriedades fundamentais da integral Riemann.

Teorema 5.3.1 (Somas de Riemann). *Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$, seja $c \in \mathbb{R}$, e seja $f : B \rightarrow \mathbb{R}$ uma função limitada. Então as seguintes afirmações são equivalente.*

(i) f é Riemann-integrável e

$$c = \int_B f(x) dx_1 \cdots dx_n.$$

(ii) Para todo $\varepsilon > 0$ existe um $P \in \mathcal{P}(B)$, tal que

$$c - \varepsilon < \underline{S}(f, P) \leq \overline{S}(f, P) < c + \varepsilon. \quad (5.3.1)$$

(iii) Para todo $\varepsilon > 0$ existe um $\delta_0 > 0$, tal que para todo $P \in \mathcal{P}(B)$ vale

$$\delta(P) < \delta_0 \quad \implies \quad c - \varepsilon < \underline{S}(f, P) \leq \overline{S}(f, P) < c + \varepsilon. \quad (5.3.2)$$

(iv) Para todo $\varepsilon > 0$ existe um número real $\delta_0 > 0$ tal que para toda partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ e todos os $x_1, \dots, x_k \in \mathbb{R}^n$ vale

$$\begin{aligned} \delta(P) < \delta_0 \\ x_i \in \overline{P}_i \forall i \end{aligned} \quad \implies \quad \left| c - \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \right| < \varepsilon. \quad (5.3.3)$$

Demonstração. Veja a página 175. □

Comentário 5.3.2. Os números $\sum_{i=1}^k f(x_i) \text{Vol}_n(P_i)$ na parte (iv) de Teorema 5.3.1 são chamados de **somas de Riemann** de f e o critério na parte (iv) de Teorema 5.3.1 pode ser sumariado na fórmula

$$\int_B f(x) dx_1 \cdots dx_n = \lim_{\substack{\delta(P) \rightarrow 0 \\ P=(P_1, \dots, P_k) \in \mathcal{P}(B) \\ x_1 \in \overline{P}_1, \dots, x_k \in \overline{P}_k}} \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i). \quad (5.3.4)$$

As somas de Riemann convergem portanto para a integral para $\delta(P) \rightarrow 0$.

Para a prova de Teorema 5.3.1 precisamos o seguinte lema.

Lema 5.3.3. *Seja $B \subset \mathbb{R}^n$ um subconjunto compacto com $\mathcal{P}(B) \neq \emptyset$, seja $Q = (Q_1, \dots, Q_\ell) \in \mathcal{P}(B)$, seja $f : B \rightarrow \mathbb{R}$ uma função limitada e $\varepsilon > 0$. Então existe um $\delta_0 > 0$, tal que para todos os $P \in \mathcal{P}(B)$ vale*

$$\delta(P) < \delta_0 \quad \Longrightarrow \quad \begin{aligned} \underline{S}(f, P) &\geq \underline{S}(f, Q) - \varepsilon, \\ \overline{S}(f, P) &\leq \overline{S}(f, Q) + \varepsilon. \end{aligned} \quad (5.3.5)$$

Demonstração. Seja $M := \sup_{x \in B} |f(x)|$ e $A := \bigcup_{j=1}^{\ell} \partial Q_j$. Então A é um conjunto Jordan-nulo (Exemplo 5.2.4). Segundo o Lema 5.2.8 existe daí um número $\delta_0 > 0$, tal que para toda partição $P = \{P_1, \dots, P_k\}$ de B vale

$$\delta(P) < \delta_0 \quad \Longrightarrow \quad \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) < \frac{\varepsilon}{2M}. \quad (5.3.6)$$

Seja ora $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ com $\delta(P) < \delta_0$. Então vale

$$\begin{aligned} \underline{S}(f, P) &\leq \underline{S}(f, P \wedge Q) \\ &= \sum_{i=1}^k \sum_{j=1}^{\ell} \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &= \sum_{\overline{P}_i \cap A = \emptyset} \sum_{j=1}^{\ell} \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\quad + \sum_{\overline{P}_i \cap A \neq \emptyset} \sum_{j=1}^{\ell} \left(\inf_{\overline{P}_i \cap \overline{Q}_j} f \right) \text{Vol}_n(P_i \cap Q_j) \\ &\leq \sum_{\overline{P}_i \cap A = \emptyset} \left(\inf_{\overline{P}_i} f \right) \text{Vol}_n(P_i) \\ &\quad + M \sum_{\overline{P}_i \cap A \neq \emptyset} \left(\sum_{j=1}^{\ell} \text{Vol}_n(P_i \cap Q_j) \right) \\ &\leq \underline{S}(f, P) + 2M \sum_{\overline{P}_i \cap A \neq \emptyset} \text{Vol}_n(P_i) \\ &< \underline{S}(f, P) + \varepsilon. \end{aligned}$$

A desigualdade $\overline{S}(f, P) > \overline{S}(f, Q) - \varepsilon$ mostra-se igualmente. Com isso Lema 5.3.3 fica provado. \square

Demonstração do Teorema 5.3.1. Provaremos primeiro que (iii) e (iv) são equivalentes. A implicação (iii) \implies (iv) segue imediatamente da desigualdade

$$\underline{S}(f, P) \leq \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \leq \overline{S}(f, P)$$

para toda partição $P = \{P_1, \dots, P_k\}$ de B e todos os $x_i \in \overline{P}_i$. Para mostrar o reverso (iv) \implies (iii), fixamos um $\varepsilon > 0$ e escolhemos $\delta_0 > 0$ tal que vale (5.3.3). Seja $P = (P_1, \dots, P_k) \in \mathcal{P}(B)$ com $\delta(P) < \delta_0$. Então vale

$$\underline{S}(f, P) = \inf_{x_i \in \overline{P}_i} \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i), \quad \overline{S}(f, P) = \sup_{x_i \in \overline{P}_i} \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i),$$

e segundo (5.3.3) recebemos daí a desigualdade

$$c - \varepsilon \leq \underline{S}(f, P) \leq \overline{S}(f, P) \leq c + \varepsilon.$$

Com isso temos mostrado que (iii) e (iv) são equivalentes.

Provaremos (i) \implies (iii). Seja então $f : B \rightarrow \mathbb{R}$ Riemann-integrável, seja

$$c = \int_B f(x) dx_1 \cdots dx_n = \sup_{P \in \mathcal{P}(B)} \underline{S}(f, P) = \inf_{Q \in \mathcal{P}(B)} \overline{S}(f, Q),$$

e seja $\varepsilon > 0$. Segundo o Lema 5.1.13 existe uma partição $Q \in \mathcal{P}(B)$ com

$$\overline{S}(f, Q) - \underline{S}(f, Q) < \frac{\varepsilon}{2}.$$

Como $\underline{S}(f, Q) \leq c \leq \overline{S}(f, Q)$, segue daí as desigualdades

$$c - \frac{\varepsilon}{2} < \underline{S}(f, Q) \leq c \leq \overline{S}(f, Q) < c + \frac{\varepsilon}{2}.$$

Segundo o Lema 5.3.3 existe um $\delta_0 > 0$ tal que toda partição $P \in \mathcal{P}(B)$ com $\delta(P) < \delta_0$ satisfaz as desigualdades

$$\underline{S}(f, P) \geq \underline{S}(f, Q) - \frac{\varepsilon}{2} > c - \varepsilon, \quad \overline{S}(f, P) \leq \overline{S}(f, Q) + \frac{\varepsilon}{2} < c + \varepsilon.$$

Portanto f satisfaz a condição (iii).

A implicação (iii) \implies (ii) segue da existência de uma partição P de B com $\delta(P) < \delta_0$ (Exercício 5.2.7). A implicação (ii) \implies (i) segue diretamente da definição da integral Riemann. Com isso Teorema 5.3.1 fica provado. \square

O espaço de Banach das funções integráveis

Seja $B \subset \mathbb{R}^n$ um subconjunto compacto com $\mathcal{P}(B) \neq \emptyset$, ou seja B possui uma partição (Definição 5.1.1). Denotamos a conjunto das funções Riemann-integráveis em B (Definição 5.1.10) de

$$\mathcal{R}(B) := \{f : B \rightarrow \mathbb{R} \mid f \text{ é Riemann-integrável}\}$$

e usamos às vezes a abreviação $\int_B f := \int_B f(x) dx_1 \cdots dx_n$ para a integral. Lembramos o fato que o **volume do conjunto** B é definido através de

$$\text{Vol}_n(B) := \int_B 1 = \overline{S}(1, P) = \underline{S}(1, P) = \sum_{i=1}^k \text{Vol}_n(P_i) \quad (5.3.7)$$

para $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$. Este número é independente de P , como temos visto no Exemplo 5.1.11. A norma do supremo de uma função limitada $f : B \rightarrow \mathbb{R}$ é denotada de

$$\|f\| := \sup_{x \in B} |f(x)|. \quad (5.3.8)$$

O seguinte teorema engloba umas propriedades fundamentais da integral Riemann. Particularmente $\mathcal{R}(B)$ é um espaço de Banach, veja (i) e (v), e a aplicação $\mathcal{R}(B) \rightarrow \mathbb{R} : f \mapsto \int_B f$ é um funcional linear limitado, veja (i) e (vi).

Teorema 5.3.4 (Propriedades do integral Riemann). *Seja $B \subset \mathbb{R}^n$ um subconjunto compacto com $\mathcal{P}(B) \neq \emptyset$. Então vale o seguinte.*

(i) *Se $f, g : B \rightarrow \mathbb{R}$ são Riemann-integrável e $\lambda \in \mathbb{R}$, então as funções $f + g : B \rightarrow \mathbb{R}$ e $\lambda f : B \rightarrow \mathbb{R}$ também são Riemann-integráveis e vale*

$$\int_B (f + g) = \int_B f + \int_B g, \quad \int_B \lambda f = \lambda \int_B f.$$

(ii) *Se $f, g : B \rightarrow \mathbb{R}$ são Riemann-integráveis então as funções fg , $\max\{f, g\}$, $\min\{f, g\}$, $|f|$ também são Riemann-integráveis.*

(iii) *Se $f, g \in \mathcal{R}(B)$ e vale $f(x) \leq g(x)$ para todo $x \in B$, então $\int_B f \leq \int_B g$.*

(iv) *Para todo $f \in \mathcal{R}(B)$ vale $|\int_B f| \leq \int_B |f| \leq \|f\| \text{Vol}_n(B)$.*

(v) *Se $f_\nu \in \mathcal{R}(B)$, $\nu \in \mathbb{N}$, é uma sequência de funções Riemann-integráveis, a qual converge uniformemente para $f : B \rightarrow \mathbb{R}$, então também f é Riemann-integrável, e vale $\int_B f = \lim_{\nu \rightarrow \infty} \int_B f_\nu$.*

(vi) *$\mathcal{R}(B)$ é um espaço de Banach sob a norma do supremo e o funcional linear $\mathcal{R}(B) \rightarrow \mathbb{R} : f \mapsto \int_B f$ é contínuo com respeito a esta norma.*

Demonstração. Provaremos (i). Sejam $f, g : B \rightarrow \mathbb{R}$ Riemann-integráveis e

$$a := \int_B f, \quad b := \int_B g.$$

Provaremos que a função $h := f + g : B \rightarrow \mathbb{R}$ e $c := a + b \in \mathbb{R}$ satisfazem a condição (iv) no Teorema 5.3.1. Seja dado $\varepsilon > 0$. Então existe, segundo o Teorema 5.3.1, um $\delta_0 > 0$, tal que toda partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ com $\delta(P) < \delta_0$ e todos os vetores $x_i \in \overline{P}_i$ satisfazem as desigualdades

$$\left| a - \sum_{i=1}^k f(x_i) \text{Vol}_n(P_i) \right| < \frac{\varepsilon}{2}, \quad \left| b - \sum_{i=1}^k g(x_i) \text{Vol}_n(P_i) \right| < \frac{\varepsilon}{2}.$$

Daí segue segundo a desigualdade triangular

$$\left| a + b - \sum_{i=1}^k (f(x_i) + g(x_i)) \text{Vol}_n(P_i) \right| < \varepsilon,$$

também para toda partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$ com $\delta(P) < \delta_0$ e todos os vetores $x_i \in \overline{P}_i$. Com isso temos mostrado que $h := f + g$ e $c := a + b$ satisfazem a condição (iv) no Teorema 5.3.1. Portanto $f + g$ é Riemann-integrável e

$$\int_B (f + g) = a + b = \int_B f + \int_B g.$$

Do mesmo modo mostra-se que para todo $f \in \mathcal{R}(B)$ e todo número real λ , a função $\lambda f : B \rightarrow \mathbb{R}$ é Riemann-integrável e $\int_B \lambda f = \lambda \int_B f$. Com isso (i) fica provado.

Provaremos (ii). Seja $P = \{P_1, \dots, P_k\}$ uma partição de B . Então vale para todo i e todos os $x, y \in \overline{P}_i$

$$\begin{aligned} f(x)g(x) - f(y)g(y) &= f(x)(g(x) - g(y)) + (f(x) - f(y))g(y) \\ &\leq \|f\| (g(x) - g(y)) + \|g\| (f(x) - f(y)) \\ &\leq \|f\| \left(\sup_{\overline{P}_i} g - \inf_{\overline{P}_i} g \right) + \|g\| \left(\sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f \right). \end{aligned}$$

Se formamos o supremo sobre $x, y \in \overline{P}_i$, então resulta para $i = 1, \dots, k$ a desigualdade

$$\sup_{\overline{P}_i} fg - \inf_{\overline{P}_i} fg \leq \|f\| \left(\sup_{\overline{P}_i} g - \inf_{\overline{P}_i} g \right) + \|g\| \left(\sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f \right).$$

Dali segue

$$\overline{S}(fg, P) - \underline{S}(fg, P) \leq \|f\| (\overline{S}(g, P) - \underline{S}(g, P)) + \|g\| (\overline{S}(f, P) - \underline{S}(f, P)).$$

Ora provaremos que fg é Riemann-integrável. Para isso suponhamos sem perda de generalidade que, nem f , nem g , anulam-se identicamente, consequentemente $\|f\| \neq 0$ e $\|g\| \neq 0$. Seja ora dado $\varepsilon > 0$. Como as funções f e g são Riemann-integráveis, existem duas partições $P, Q \in \mathcal{P}(B)$, segundo o Lema 5.1.13, com

$$\overline{S}(f, P) - \underline{S}(f, P) \leq \frac{\varepsilon}{2\|g\|}, \quad \overline{S}(g, Q) - \underline{S}(g, Q) \leq \frac{\varepsilon}{2\|f\|}.$$

Sob uso de Lema 5.1.9 segue dali a cadeia de desigualdades

$$\begin{aligned} & \overline{S}(fg, P \wedge Q) - \underline{S}(fg, P \wedge Q) \\ & \leq \|f\| (\overline{S}(g, P \wedge Q) - \underline{S}(g, P \wedge Q)) + \|g\| (\overline{S}(f, P \wedge Q) - \underline{S}(f, P \wedge Q)) \\ & \leq \|f\| (\overline{S}(g, Q) - \underline{S}(g, Q)) + \|g\| (\overline{S}(f, P) - \underline{S}(f, P)) \\ & < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Daí segue segundo o Lema 5.1.13 que fg é Riemann-integrável. Provaremos ora que $|f|$ é Riemann-integrável. Para este efeito escolhemos de novo uma partição $P = \{P_1, \dots, P_k\}$ de B . Então vale para todo i e todo $x, y \in \overline{P}_i$

$$|f(x)| - |f(y)| \leq |f(x) - f(y)| \leq \sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f.$$

Dali segue $\sup_{\overline{P}_i} |f| - \inf_{\overline{P}_i} |f| \leq \sup_{\overline{P}_i} f - \inf_{\overline{P}_i} f$ para todos os i e daí

$$\overline{S}(|f|, P) - \underline{S}(|f|, P) \leq \overline{S}(f, P) - \underline{S}(f, P).$$

Consequentemente segue de Lema 5.1.13 que $|f|$ é Riemann-integrável. Ora vale

$$\max\{f, g\} + \min\{f, g\} = f + g, \quad \max\{f, g\} - \min\{f, g\} = |f - g|.$$

Estas duas funções são Riemann-integráveis conforme o que já foi mostrado. Portanto, segue de (i), que as funções $\max\{f, g\}$ e $\min\{f, g\}$ também são Riemann-integráveis. Com isso (ii) fica provado.

Provaremos (iii). Se $f, g \in \mathcal{R}(B)$ e $f \leq g$, então vale $\underline{S}(f, P) \leq \underline{S}(g, P)$ para toda partição $P \in \mathcal{P}(B)$ e daí segue imediatamente que $\int_B f \leq \int_B g$.

Provaremos (iv). Se $f \in \mathcal{R}(B)$, então as funções $-|f|$ e $|f|$ são Riemann-integráveis segundo (i) e (ii). Além disso, vale $-|f| \leq f \leq |f|$. Portanto segue de (iii), que

$$-\int_B |f| \leq \int_B f \leq \int_B |f|.$$

Isso é equivalente a $|\int_B f| \leq \int_B |f|$. Demais a integral de uma função constante $g(x) = c$ é dada, segundo o Exemplo 5.1.11, através de $\int_B c = c \text{Vol}_n(B)$. Com $c = \|f\| \geq |f|$ segue dali segundo (iii) a desigualdade

$$\|f\| \text{Vol}_n(B) = \int_B \|f\| \geq \int_B |f|.$$

Com isso (iv) fica provado.

Provaremos (v). Seja ora $f_\nu : B \rightarrow \mathbb{R}$ uma sequência de funções Riemann-integráveis, as quais convergem uniformemente para $f : B \rightarrow \mathbb{R}$. Da definição de convergência uniforme segue que f é limitado e que a sequência f_ν converge para f com respeito à norma do supremo. Particularmente $(f_\nu)_{\nu \in \mathbb{N}}$ é então uma sequência de Cauchy com respeito à norma do supremo. Ora definimos uma sequência $(c_\nu)_{\nu \in \mathbb{N}}$ através de

$$c_\nu := \int_B f_\nu, \quad \nu \in \mathbb{N}.$$

Conforme (iv) vale para todos os $\nu, \nu' \in \mathbb{N}$ a desigualdade

$$|c_\nu - c_{\nu'}| \leq \int_B |f_\nu - f_{\nu'}| \leq \|f_\nu - f_{\nu'}\| \text{Vol}_n(B). \quad (5.3.9)$$

Dali segue que $(c_\nu)_{\nu \in \mathbb{N}}$ é uma sequência de Cauchy em \mathbb{R} . Como toda sequência de Cauchy de números reais converge, também existe o limite

$$c := \lim_{\nu \rightarrow \infty} c_\nu.$$

Formamos ora em (5.3.9) o limite $\nu' \rightarrow \infty$, então recebemos a desigualdade

$$|c_\nu - c| \leq \|f_\nu - f\| \text{Vol}_n(B). \quad (5.3.10)$$

Provaremos que f é Riemann-integrável e $\int_B f = c$. Seja dado $\varepsilon > 0$. Como f_ν converge para f na norma do supremo, existe um número $\nu \in \mathbb{N}$, tal que

$$\|f_\nu - f\| < \frac{\varepsilon}{3\text{Vol}_n(B)}, \quad |c_\nu - c| < \frac{\varepsilon}{3}. \quad (5.3.11)$$

Aqui a segunda desigualdade segue da primeira e (5.3.10). Como $c_\nu = \int_B f_\nu$, existe segundo 5.3.1 uma partição $P \in \mathcal{P}(B)$, tal que

$$c_\nu - \frac{\varepsilon}{3} < \underline{S}(f_\nu, P) \leq \overline{S}(f_\nu, P) \leq c_\nu + \frac{\varepsilon}{3}. \quad (5.3.12)$$

Além disso, segue da definição de soma superior e soma inferior que

$$\begin{aligned} |\overline{S}(f_\nu, P) - \overline{S}(f, P)| &\leq \|f - f_\nu\| \text{Vol}_n(B) < \frac{\varepsilon}{3} \\ |\underline{S}(f_\nu, P) - \underline{S}(f, P)| &\leq \|f - f_\nu\| \text{Vol}_n(B) < \frac{\varepsilon}{3}. \end{aligned} \quad (5.3.13)$$

Combinamos as desigualdades (5.3.11), (5.3.12), e (5.3.13), então recebemos

$$c - \varepsilon < c_\nu - \frac{2\varepsilon}{3} < \underline{S}(f_\nu, P) - \frac{\varepsilon}{3} < \underline{S}(f, P)$$

e

$$\overline{S}(f, P) < \overline{S}(f_\nu, P) + \frac{\varepsilon}{3} < c_\nu + \frac{2\varepsilon}{3} < c + \varepsilon.$$

Portanto segue de Teorema 5.3.1 que f é Riemann-integrável e $\int_B f = c$. Com isso (v) fica provado.

Provaremos (vi). Seja $f_\nu \in \mathcal{R}(B)$ uma seqüência de Cauchy com respeito à norma do supremo. A desigualdade

$$|f_\nu(x) - f_{\nu'}(x)| \leq \|f_\nu - f_{\nu'}\|$$

para todos os $x \in B$ e todos os $\nu, \nu' \in \mathbb{N}$ mostra que a seqüência $(f_\nu(x))_{\nu \in \mathbb{N}}$ em \mathbb{R} para todo $x \in B$ é uma seqüência de Cauchy. Consequentemente segue do axioma de completude dos números reais que a seqüência $(f_\nu(x))_{\nu \in \mathbb{N}}$ converge para todo $x \in B$. Denotamos o limite de

$$f(x) := \lim_{\nu \rightarrow \infty} f_\nu(x).$$

Isto define uma função $f : B \rightarrow \mathbb{R}$. Provaremos que a seqüência de funções f_ν converge uniformemente para f . Seja dado $\varepsilon > 0$. Como f_ν é uma seqüência

de Cauchy com respeito à norma do supremo, existe um número natural $\nu_0 \in \mathbb{N}$, tal que para todos os $\nu, \nu' \in \mathbb{N}$ vale que

$$\nu, \nu' \geq \nu_0 \quad \implies \quad \|f_\nu - f_{\nu'}\| \leq \varepsilon.$$

Dali segue $|f_\nu(x) - f_{\nu'}(x)| \leq \varepsilon$ para todos os $x \in B$ e todos os $\nu, \nu' \geq \nu_0$. Formamos ora o limite $\nu' \rightarrow \infty$, então resulta a desigualdade

$$|f_\nu(x) - f(x)| = \lim_{\nu' \rightarrow \infty} |f_\nu(x) - f_{\nu'}(x)| \leq \varepsilon$$

para todo $\nu \in \mathbb{N}$ com $\nu \geq \nu_0$ e todo $x \in B$. Formamos o supremo sobre $x \in B$, então recebemos a desigualdade

$$\|f_\nu - f\| \leq \varepsilon$$

para todos os $\nu \in \mathbb{N}$ com $\nu \geq \nu_0$. Em outras palavras, a sequência f_ν converge para f na norma do supremo, e isso é equivalente com a afirmação que a sequência f_ν converge uniformemente para f . Ora segue de (iv) que f é Riemann-integrável. Portanto temos mostrado que $\mathcal{R}(B)$ munido da norma do supremo é um espaço vetorial normado completo e assim um espaço de Banach. A continuidade do funcional linear

$$\mathcal{R}(B) \rightarrow \mathbb{R} : f \mapsto \int_B f$$

segue imediatamente de (iv). Com isso Teorema 5.3.4 fica provado. \square

Seja $B \subset \mathbb{R}^n$ um conjunto com $\mathcal{P}(B) \neq \emptyset$ e seja

$$\mathcal{B}(B) := \{f : B \rightarrow \mathbb{R} \mid \exists c > 0 \forall x \in B \ |f(x)| \leq c\}$$

o conjunto das funções limitadas. Isto é um espaço de Banach sob a norma do supremo

$$\|f\| = \sup_{x \in B} |f(x)|.$$

O conjunto $\mathcal{R}(B)$ das funções Riemann-integráveis é, segundo a parte (iv) de Teorema 5.3.4, um subespaço linear fechado de $\mathcal{B}(B)$ e assim mesmo um espaço de Banach. Além disso, segundo o Teorema 5.2.2, o conjunto

$$\mathcal{C}(B) = \{f : B \rightarrow \mathbb{R} \mid f \text{ é contínuo}\}$$

de todas as funções contínuas é um subespaço linear fechado de $\mathcal{B}(B)$ e assim também um espaço de Banach sob a norma do supremo (pois que o limite de uma sequência uniformemente convergente de funções contínuas é contínuo). Com isso temos inclusões de espaços de Banach com a norma do supremo

$$\mathcal{C}(B) \subset \mathcal{R}(B) \subset \mathcal{B}(B).$$

5.4 Integrais iteradas

O **Teorema de Fubini** tem um significado central para a integral Riemann. Será provado primeiro e subsequente utilizado para o cálculo do volume do símplice unitário.

Teorema 5.4.1 (Fubini). *Dado $p, q \in \mathbb{N}$ e $n := p+q$. Dado $A \subset \mathbb{R}^p$, $B \subset \mathbb{R}^q$ conjuntos compactos com $\mathcal{P}(A) \neq \emptyset$ e $\mathcal{P}(B) \neq \emptyset$. Então $\mathcal{P}(A \times B) \neq \emptyset$. Seja $f : A \times B \rightarrow \mathbb{R}$ uma função Riemann-integrável e para todo $x \in A$ seja a função $B \rightarrow \mathbb{R} : y \mapsto f_x(y) := f(x, y)$ Riemann-integrável. Então a função $A \rightarrow \mathbb{R} : x \mapsto \int_B f_x$ é Riemann-integrável e vale*

$$\int_{A \times B} f = \int_A \left(\int_B f(x, y) dy_1 \cdots dy_q \right) dx_1 \cdots dx_p. \quad (5.4.1)$$

Demonstração. Defina $F : A \rightarrow \mathbb{R}$ e $c \in \mathbb{R}$ através de

$$F(x) := \int_B f_x = \int_B f(x, y) dy_1 \cdots dy_q, \quad c := \int_{A \times B} f.$$

Provaremos em dois passos que $F \in \mathcal{R}(A)$ e $\int_A F = c$.

Passo 1. *Sejam $P = \{P_1, \dots, P_k\} \in \mathcal{P}(A)$ e $Q = \{Q_1, \dots, Q_\ell\} \in \mathcal{P}(B)$ duas partições. Então o conjunto*

$$P \times Q := \{P_i \times Q_j \mid i = 1, \dots, k, j = 1, \dots, \ell\}$$

é uma partição de $A \times B$ com $\delta(P \times Q) = \max \{\delta(P), \delta(Q)\}$ e vale

$$\underline{S}(f, P \times Q) \leq \underline{S}(F, P) \leq \overline{S}(F, P) \leq \overline{S}(f, P \times Q).$$

Para todo $i \in \{1, \dots, k\}$ e todo $j \in \{1, \dots, \ell\}$ o conjunto $P_i \times Q_j$ é um bloco aberto em $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^q$ com fecho $\overline{P_i \times Q_j} = \overline{P_i} \times \overline{Q_j}$, comprimento de lado máximo

$$\delta(P_i \times Q_j) = \max \{\delta(P_i), \delta(Q_j)\}$$

e volume n -dimensional

$$\text{Vol}_n(P_i \times Q_j) = \text{Vol}_p(P_i) \text{Vol}_q(Q_j).$$

Conseqüentemente os blocos abertos $P_i \times Q_j$ formam uma partição $P \times Q$ de $A \times B$ da fineza $\delta(P \times Q) = \max\{\delta(P), \delta(Q)\}$. Além disso, vale para todo $i \in \{1, \dots, k\}$ e $x \in \overline{P}_i$ a desigualdade

$$\begin{aligned} F(x) &\leq \overline{S}(f_x, Q) \\ &= \sum_{j=1}^{\ell} \left(\sup_{y \in \overline{Q}_j} f(x, y) \right) \text{Vol}_q(Q_j) \\ &\leq \sum_{j=1}^{\ell} \left(\sup_{P_i \times Q_j} f \right) \text{Vol}_q(Q_j). \end{aligned}$$

Daí segue

$$\left(\sup_{\overline{P}_i} F \right) \text{Vol}_p(P_i) \leq \sum_{j=1}^{\ell} \left(\sup_{P_i \times Q_j} f \right) \text{Vol}_n(P_i \times Q_j).$$

Formamos ora a soma sobre todos os i , então resulta

$$\overline{S}(F, P) \leq \overline{S}(f, P \times Q).$$

Igualmente mostra-se a desigualdade

$$\underline{S}(F, P) \geq \underline{S}(f, P \times Q).$$

Com isso o Passo 1 fica provado.

Passo 2. Para todo $\varepsilon > 0$ existe uma partição $P \in \mathcal{P}(A)$, tal que

$$c - \varepsilon < \underline{S}(F, P) \leq \overline{S}(F, P) < c + \varepsilon.$$

Seja dado $\varepsilon > 0$. Segundo o Teorema 5.3.1 existe uma $\delta_0 > 0$, tal que toda partição $R \in \mathcal{P}(A \times B)$ com $\delta(R) < \delta_0$ satisfaz a desigualdade

$$c - \varepsilon < \underline{S}(f, R) \leq \overline{S}(f, R) < c + \varepsilon.$$

Escolha uma partição $P \in \mathcal{P}(A)$ da fineza $\delta(P) < \delta_0$ e mais uma partição $Q \in \mathcal{P}(B)$ da fineza $\delta(Q) < \delta_0$. Segundo o Passo 1 o produto $P \times Q$ é uma partição de $A \times B$ da fineza $\delta(P \times Q) < \delta_0$ e vale

$$c - \varepsilon < \underline{S}(f, P \times Q) \leq \underline{S}(F, P) \leq \overline{S}(F, P) \leq \overline{S}(f, P \times Q) < c + \varepsilon.$$

Com isso o Passo 2 fica provado.

Segundo o Passo 2 e o Teorema 5.3.1 a função $F : A \rightarrow \mathbb{R}$ é Riemann-integrável e sua integral é $\int_A F = c$. Com isso Teorema 5.4.1 fica provado. \square

Exemplo 5.4.2. Este exemplo mostra que não segue da integrabilidade de f no Teorema 5.4.1 a integrabilidade de f_x para todo $x \in A$. Seja

$$A = B = [0, 1]$$

e seja $K \subset [0, 1]$ o conjunto de Cantor. Então K é um conjunto Jordan-nulo em \mathbb{R} e $K \times [0, 1]$ um conjunto Jordan-nulo em \mathbb{R}^2 (Exercício). Ora definimos a função $f : A \times B \rightarrow \mathbb{R}$ através de

$$f(x, y) := \begin{cases} 1, & \text{para } x \in K, y \in [0, 1] \cap \mathbb{Q}, \\ 0, & \text{caso contrário.} \end{cases}$$

Esta função é contínua em $(A \times B) \setminus (K \times B)$ e dali é Riemann-integrável segundo o Teorema 5.2.2. Porém $f_x : B \rightarrow \mathbb{R}$ é Riemann-integrável só para $x \in A \setminus K$, mas não para $x \in K$.

Exemplo 5.4.3. Este exemplo mostra que da integrabilidade de f_x para todo $x \in A$ não segue a integrabilidade de f . Seja $A = B = [0, 1]$ e defina $f : A \times B \rightarrow \mathbb{R}$ através de

$$f(x, y) := \begin{cases} 1, & \text{para } x \in [0, 1] \cap \mathbb{Q}, \\ 0, & \text{para } x \in [0, 1] \setminus \mathbb{Q}. \end{cases}$$

Então $f_x : B \rightarrow \mathbb{R}$ é constante para todo $x \in [0, 1]$ e por causa disso Riemann-integrável. No entanto $\underline{S}(f, P) = 0$ e $\overline{S}(f, P) = 1$ para toda partição P de $[0, 1]^2$. Consequentemente f não é Riemann-integrável.

Exemplo 5.4.4. Sejam $A \subset \mathbb{R}^p$ e $B \subset \mathbb{R}^q$ conjuntos compactos com $\mathcal{P}(A) \neq \emptyset$ e $\mathcal{P}(B) \neq \emptyset$. Sejam $f : A \rightarrow \mathbb{R}$ e $g : B \rightarrow \mathbb{R}$ Riemann-integrável. Defina $h : A \times B \rightarrow \mathbb{R}$ através de

$$h(x, y) := f(x)g(y).$$

Então $h : A \times B \rightarrow \mathbb{R}$ é Riemann-integrável (Exercício). Além disso a função $h_x := h(x, \cdot) = f(x)g$ é Riemann-integrável para todo $x \in A$. Portanto segue de Teorema 5.4.1 e Teorema 5.3.4 (i) a fórmula

$$\int_{A \times B} h = \left(\int_A f \right) \left(\int_B g \right).$$

Exemplo 5.4.5. Para $n \in \mathbb{N}$ e $m \in \mathbb{N}_0$ seja $f_{n,m} : [0, 1]^n \rightarrow \mathbb{R}$ a função

$$f_{n,m}(x) := \begin{cases} (1 - x_1 - \cdots - x_n)^m, & \text{se } x_1 + \cdots + x_n \leq 1, \\ 0, & \text{se } x_1 + \cdots + x_n > 1. \end{cases} \quad (5.4.2)$$

Então $f_{n,m}$ é Riemann-integrável e

$$\int_{[0,1]^n} f_{n,m} = \frac{1}{(m+1)(m+2)\cdots(m+n)} = \frac{m!}{(m+n)!}. \quad (5.4.3)$$

Á primeira $f_{n,m}$ é contínua para $m \geq 1$, e $f_{n,0}$ é contínua no complemento do conjunto $A := \{x \in [0, 1]^n \mid x_1 + \cdots + x_n = 1\}$. Que trata-se de A por um conjunto Jordan-nulo, segue do fato que A é o gráfico de uma função contínua (veja também o Exemplo 5.6.10). Segundo o Teorema 5.2.2 as funções $f_{n,m}$ são daí Riemann-integráveis. Verificamos ora a equação (5.4.3) através de indução sobre n .

Á primeira vale para todo número inteiro $m \geq 0$ a equação

$$\int_0^1 f_{1,m}(x) = \int_0^1 (1-x)^m dx = \int_0^1 t^m dt = \frac{1}{m+1}.$$

Assim (5.4.3) é mostrado para $n = 1$. Suponhamos ora que $n \geq 2$ e a equação (5.4.3) vale para $n - 1$. Segundo o Teorema 5.4.1, e com Δ^{n-1} como em (5.4.4) na próxima página, recebemos então a equação

$$\begin{aligned} \int_{[0,1]^n} f_{n,m} &= \int_{[0,1]^{n-1}} \left(\int_0^1 f_{n,m}(x) dx_n \right) dx_1, \dots, dx_{n-1} \\ &= \int_{\Delta^{n-1}} \left(\int_0^{1-\sum_{i=1}^{n-1} x_i} \left(1 - \sum_{i=1}^n x_i \right)^m dx_n \right) dx_1 \cdots dx_{n-1} \\ &= \int_{\Delta^{n-1}} \frac{(1 - x_1 - \cdots - x_{n-1})^{m+1}}{m+1} dx_1 \cdots dx_{n-1} \\ &= \frac{1}{m+1} \int_{[0,1]^{n-1}} f_{n-1,m+1}(x) dx_1 \cdots dx_{n-1} \\ &= \frac{1}{(m+1)(m+2)\cdots(m+n)}. \end{aligned}$$

O último passo segue da hipótese da indução. Além disso $\int_{\Delta^{n-1}} f$ denota a integral sobre $[0, 1]^{n-1}$ da função a qual coincide em Δ^{n-1} com f e fora de Δ^{n-1} anula-se.

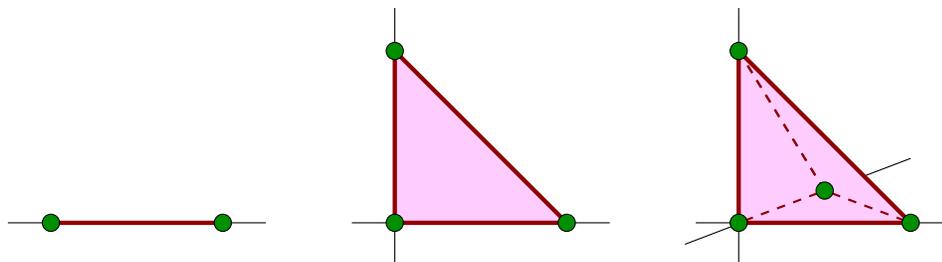


Figura 5.2: O símplice unitário para $n = 1, 2, 3$.

Para $m = 0$ pode-se interpretar geometricamente a fórmula (5.4.3) como segue. O **símplice unitário** no \mathbb{R}^n é o conjunto

$$\Delta^n := \left\{ x \in \mathbb{R}^n \mid x_i \geq 0 \forall i, \sum_{i=1}^n x_i \leq 1 \right\}. \quad (5.4.4)$$

A função $f_{n,0} : [0, 1]^n \rightarrow \mathbb{R}$ é a **função característica** deste conjunto, ou seja

$$f_{n,0}(x) = \begin{cases} 1, & \text{para } x \in \Delta^n, \\ 0, & \text{para } x \notin \Delta^n. \end{cases}$$

Portanto, nós podemos também interpretar a integral

$$\int_{[0,1]^n} f_{n,0}(x) dx_1 \cdots dx_n = \frac{1}{n!}$$

como integral da função constante $f(x) = 1$ no símplice unitário Δ^n e a integral dela como o **volume do símplice unitário**. Isto pode-se escrever na forma

$$\text{Vol}_n(\Delta^n) = \int_{\Delta^n} 1 dx_1 \cdots dx_n = \frac{1}{n!}.$$

Portanto o símplice unitário no \mathbb{R}^n tem o volume $\frac{1}{n!}$. Porém aqui tem-se que dar atenção no fato que até agora temos definida a integral de uma função só sobre conjuntos as quais possuem uma partição. Mas o conjunto Δ^n não possui uma partição. Por isso, faz sentido estender a noção da integral um pouco. Isto será efetuado na seguinte seção.

5.5 Conjuntos Jordan-mensuráveis

Até agora temos definido meramente o volume de tais subconjuntos do \mathbb{R}^n as quais possuem partições (Definição 5.1.12). Isto é uma condição muito restritiva, e esta seção vai tratar-se de estender dramaticamente a classe dos conjuntos cujos volumes podemos definir.

Definição 5.5.1. *Um subconjunto limitado $B \subset \mathbb{R}^n$ chama-se de **Jordan-mensurável**, se seu bordo ∂B é um conjunto Jordan-nulo.*

Definição 5.5.2 (Integração sobre conjuntos Jordan-mensurável). *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável e $Q \subset \mathbb{R}^n$ um cubo eixo-paralelo fechado o qual contém B . Seja $f : Q \rightarrow \mathbb{R}$ uma função Riemann-integrável. O número*

$$\int_B f := \int_Q 1_B(x)f(x) dx_1 \cdots dx_n \quad (5.5.1)$$

chama-se de **integral de f sobre B** . Aqui denotamos de

$$1_B : \mathbb{R}^n \rightarrow \mathbb{R}, \quad 1_B(x) := \begin{cases} 1, & \text{para } x \in B, \\ 0, & \text{para } x \in \mathbb{R}^n \setminus B, \end{cases} \quad (5.5.2)$$

a **função característica de B** . O número real

$$\mu_n(B) := \text{Vol}_n(B) := \int_B 1 = \int_Q 1_B(x) dx_1 \cdots dx_n \quad (5.5.3)$$

é chamado da **medida Jordan de B ou o volume de B** .

Comentário 5.5.3. (i) A função $1_B : Q \rightarrow \mathbb{R}$ é contínua em $Q \setminus \partial B$. Como ∂B é um conjunto Jordan-nulo 1_B é Riemann-integrável segundo 5.2.2. Conforme Teorema 5.3.4 o produto $1_B f : Q \rightarrow \mathbb{R}$ é Riemann-integrável.

(ii) A integral $\int_Q 1_B f$ em (5.5.1) é independente da escolha do bloco eixo-paralelo fechado Q contendo B .

(iii) Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-nulo. Então $\partial B = \overline{B}$ é um conjunto Jordan-nulo segundo o Lema 5.2.3. Por isso B é Jordan-mensurável e vale $\text{Vol}_n(B) = 0$ segundo o Teorema 5.2.2.

(iv) Se $Q \subset \mathbb{R}^n$ é um bloco eixo-paralelo fechado, então ∂Q é um conjunto Jordan-nulo segundo o Exemplo 5.2.4 (ii). Daí Q é Jordan-mensurável. Além disso, neste caso a definição do volume na Definição 5.5.2 coincide, conforme Exemplo 5.1.11, com aquela em (5.1.2). Para a prova desta coincidência também pode-se utilizar Teorema 5.4.1.

(v) Seja $B \subset \mathbb{R}^n$ um conjunto compacto com $\mathcal{P}(B) \neq \emptyset$. Seja $P = \{P_1, \dots, P_k\}$ uma partição de B . Então vale $\partial B \subset \bigcup_i \partial P_i$. Daí B é Jordan-mensurável e tem o volume $\text{Vol}_n(B) = \int_B 1 = \sum_{i=1}^k \text{Vol}_n(P_i)$ segundo o Exemplo 5.1.11. Portanto a definição do volume de B em Definição 5.5.2 coincide com aquela na Definição 5.1.12.

(vi) Segundo o Exemplo 5.4.5 o n -símplice padrão $\Delta^n \subset \mathbb{R}^n$ é Jordan-mensurável e tem o volume $\text{Vol}_n(\Delta^n) = \frac{1}{n!}$.

(vii) O conjunto $B := [0, 1] \cap \mathbb{Q}$ não é Jordan-mensurável, como $\partial B = [0, 1]$.

Lema 5.5.4. *Sejam $A, B \subset \mathbb{R}^n$ dois conjuntos Jordan-mensuráveis, $Q \subset \mathbb{R}^n$ um bloco eixo-paralelo fechado com $A, B \subset Q$, e $f : Q \rightarrow \mathbb{R}$ uma função Riemann-integrável. Então vale o seguinte.*

(i) *Os conjuntos $A \cap B$, $A \cup B$, $A \setminus B$, $B \setminus A$ são Jordan-mensuráveis.*

(ii) *Vale*

$$\int_{A \cup B} f + \int_{A \cap B} f = \int_A f + \int_B f \quad (5.5.4)$$

e $\text{Vol}_n(A \cup B) + \text{Vol}_n(A \cap B) = \text{Vol}_n(A) + \text{Vol}_n(B)$.

(iii) *Se $A \subset B$, então vale $\text{Vol}_n(A) \leq \text{Vol}_n(B)$.*

Demonstração. Parte (i) segue do fato que cada um dos conjuntos $\partial(A \cup B)$, $\partial(A \cap B)$, $\partial(A \setminus B)$, $\partial(B \setminus A)$ é contido no conjunto Jordan-nulo $\partial A \cup \partial B$ e assim é mesmo um conjunto Jordan-nulo. Parte (ii) segue de Teorema 5.3.4 (i) e o fato que $1_{A \cup B} + 1_{A \cap B} = 1_A + 1_B$. Parte (iii) segue de Teorema 5.3.4 (iii) e da desigualdade $1_A \leq 1_B$ para $A \subset B$. \square

Definição 5.5.5. *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável. Uma função $f : B \rightarrow \mathbb{R}$ chama-se de **Riemann-integrável**, se a função estendida*

$$\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \tilde{f}(x) := \begin{cases} f(x), & \text{para } x \in B, \\ 0, & \text{para } x \in \mathbb{R}^n \setminus B, \end{cases}$$

é Riemann-integrável sobre todo bloco fechado $Q \subset \mathbb{R}^n$ contendo B . Neste caso chama-se o número

$$\int_B f := \int_Q \tilde{f}(x) dx_1 \cdots dx_n$$

da integral de f sobre B . Segundo o Lema 5.5.4 a integral é independente da escolha do bloco eixo-paralelo fechado no qual f é estendido. O conjunto das funções Riemann-integráveis $f : B \rightarrow \mathbb{R}$ denota-se de $\mathcal{R}(B)$.

Comentário 5.5.6. Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável. Então toda função contínua limitada $f : B \rightarrow \mathbb{R}$ é Riemann-integrável. Isto segue de Teorema 5.2.2 e o fato que a função estendida $\tilde{f} : Q \rightarrow \mathbb{R}$ na Definição 5.5.5 é contínua em $Q \setminus \partial B$.

Com esta definição temos estendido a integral Riemann a funções definida em conjuntos Jordan-mensuráveis $B \subset \mathbb{R}^n$ arbitrários. Particularmente toda função contínua limitada $f : B \rightarrow \mathbb{R}$ é Riemann-integrável segundo o Teorema 5.2.2 e Definição 5.5.5. Além disso, todas as afirmações de Teorema 5.3.4 mantem-se válidas para conjuntos Jordan-mensuráveis B . Ainda por cima pode-se generalizar o Teorema de Fubini da seguinte forma,

Corolário 5.5.7 (Fubini). *Sejam $p, q \in \mathbb{N}$ e $n := p + q$, e sejam $P \subset \mathbb{R}^p$ e $Q \subset \mathbb{R}^q$ blocos eixo-paralelos fechados.*

(i) *Seja $B \subset P \times Q$ um conjunto Jordan-mensurável tal que para todo $x \in P$ o conjunto $B_x := \{y \in \mathbb{R}^q \mid (x, y) \in B\} \subset Q$ é Jordan-mensurável. Então a função $P \rightarrow \mathbb{R} : x \mapsto \text{Vol}_q(B_x)$ é Riemann-integrável e vale*

$$\text{Vol}_n(B) = \int_P \text{Vol}_q(B_x) dx_1 \cdots dx_p.$$

(ii) *Seja $B \subset P \times Q$ como em (i). Seja $f : B \rightarrow \mathbb{R}$ Riemann-integrável e para todo $x \in P$ seja a função $B_x \rightarrow \mathbb{R} : y \mapsto f_x(y) := f(x, y)$ Riemann-integrável. Então a função $P \rightarrow \mathbb{R} : x \mapsto \int_{B_x} f_x$ é Riemann-integrável e vale*

$$\int_B f = \int_P \left(\int_{B_x} f_x \right) dx_1 \cdots dx_p.$$

Demonstração. Provaremos (ii). Defina $\tilde{f} : P \times Q \rightarrow \mathbb{R}$ através de $\tilde{f}(x, y) := f(x, y)$ para $(x, y) \in B$ e $\tilde{f}(x, y) := 0$ para $(x, y) \in (P \times Q) \setminus B$. Para $x \in P$ defina $\tilde{f}_x : Q \rightarrow \mathbb{R}$ através de $\tilde{f}_x(y) := f_x(y)$ para $y \in B_x$ e $\tilde{f}_x(y) := 0$ para $y \in Q \setminus B_x$. Então vale $\tilde{f}_x(y) = \tilde{f}(x, y)$ para todos os $x \in P$ e $y \in Q$. Segundo a hipótese e definição 5.5.5 as funções $\tilde{f} : P \times Q \rightarrow \mathbb{R}$ e $\tilde{f}_x : Q \rightarrow \mathbb{R}$ (para $x \in P$) são Riemann-integráveis, e vale

$$\int_B f = \int_{P \times Q} \tilde{f}, \quad \int_{B_x} f_x = \int_Q \tilde{f}_x.$$

Portanto (ii) segue de Teorema 5.4.1. A afirmação (i) segue de (ii) com $f \equiv 1$. Com isso Corolário 5.5.7 fica provado. \square

Exemplo 5.5.8. Este exemplo mostra que da mensurabilidade Jordan de B em Corolário 5.5.7 não segue a mensurabilidade Jordan de B_x para todo $x \in P$. Seja $P = Q = [0, 1]$, seja $K \subset [0, 1]$ o conjunto Cantor, e seja

$$B := \{(x, y) \in [0, 1]^2 \mid x \in K \implies y \in \mathbb{Q}\}.$$

Então $\partial B = K \times [0, 1]$, por isso B é Jordan-mensurável (Exemplo 5.4.2). Mas $B_x \subset [0, 1]$ é Jordan-mensurável só para $x \in [0, 1] \setminus K$, não para $x \in K$.

Exemplo 5.5.9. Este exemplo mostra que da mensurabilidade Jordan dos conjuntos B_x para todos os $x \in P$ não segue a mensurabilidade Jordan do conjunto B . Seja $P = Q = [0, 1]$ e $B := ([0, 1] \cap \mathbb{Q}) \times [0, 1]$. Então para todo $x \in [0, 1]$ o conjunto B_x é, ou vazio, ou igual $[0, 1]$, e daí Jordan-mensurável. Porém vale $\partial B = [0, 1]^2$ e dali B não é Jordan-mensurável.

Exemplo 5.5.10. O conjunto compacto $B := \{(s, t) \in \mathbb{R}^2 \mid 0 \leq s \leq t \leq 1\}$ é o triângulo localizado no quadrado unitário em \mathbb{R}^2 acima da diagonal. Este conjunto é Jordan-mensurável e toda função contínua $f : B \rightarrow \mathbb{R}$ é Riemann-integrável segundo o Teorema 5.2.2. Segundo o Corolário 5.5.7 toda função Riemann-integrável $f : B \rightarrow \mathbb{R}$ satisfaz a equação

$$\int_B f = \int_0^1 \left(\int_0^t f(s, t) ds \right) dt = \int_0^1 \left(\int_s^1 f(s, t) dt \right) ds.$$

Exemplo 5.5.11. Seja $B \subset \mathbb{R}^n$ Jordan-mensurável e $h > 0$. O **cilindro sobre B com altura h** é o conjunto $Z_B := B \times [0, h]$. Este conjunto é Jordan-mensurável e tem o volume $\text{Vol}_{n+1}(Z_B) = h \text{Vol}_n(B)$.

Exemplo 5.5.12. Seja $B \subset \mathbb{R}^n$ Jordan-mensurável e $\lambda > 0$. Então o conjunto $\lambda B := \{\lambda x \mid x \in B\}$ é Jordan-mensurável e $\text{Vol}_n(\lambda B) = \lambda^n \text{Vol}_n(B)$. A mensurabilidade Jordan de λB segue do fato que $\partial(\lambda B) = \lambda \partial B$. A fórmula para o volume segue por um lado da fórmula da transformação de variáveis em Teorema 5.7.11, mas também pode ser deduzida diretamente das definições.

Exemplo 5.5.13. Seja $B \subset \mathbb{R}^n$ Jordan-mensurável e $h > 0$. O conjunto

$$K_B := \{(x, y) \mid 0 \leq y \leq h, x \in (1 - y/h) B\}$$

é o **cone B de altura h** . Este conjunto é Jordan-mensurável, como vamos ver na próxima seção (Exemplo 5.6.12). Se aceita-se a mensurabilidade Jordan do cone K_B , então resulta conforme Corolário 5.5.7 (i) e Exemplo 5.5.12 a fórmula

$$\text{Vol}_{n+1}(K_B) = \int_0^h \left(1 - \frac{y}{h}\right)^n \text{Vol}_n(B) dy = \frac{h}{n+1} \text{Vol}_n(B) = \frac{\text{Vol}_{n+1}(Z_B)}{n+1}.$$

5.6 Conjuntos Jordan-mensuráveis – mais

Como já temos visto ao longo de vários exemplos, é uma questão importante como pode-se reconhecer conjuntos Jordan-mensuráveis e conjuntos Jordan-nulos. Com isso ocupa-se esta seção.

Caracterização de conjuntos Jordan-mensuráveis

Definição 5.6.1 (Edifícios de blocos). Um subconjunto compacto $B \subset \mathbb{R}^n$ chama-se de **edifício de blocos** se, ou é vazio, ou pode ser escrito como união de um número finito de blocos eixo-paralelo fechados (com volume positivo). Ou seja, um subconjunto não-vazio $B \subset \mathbb{R}^n$ é um edifício de blocos se e somente se $\mathcal{P}(B) \neq \emptyset$ (veja o Exercício 5.1.4).

Teorema 5.6.2 (Caracterização de conjuntos Jordan-mensuráveis).

Seja $B \subset \mathbb{R}^n$ um conjunto limitado e $b \geq 0$ um número real. Seja $Q \subset \mathbb{R}^n$ um bloco eixo-paralelo fechado o qual contém B . Então as seguintes afirmações são equivalente.

- (i) B é Jordan-mensurável e tem o volume $\text{Vol}_n(B) = b$.
- (ii) A função característica $1_B : Q \rightarrow \mathbb{R}$ é Riemann-integrável e $\int_Q 1_B = b$.
- (iii) Para todo $\varepsilon > 0$ existem edifícios de blocos $B_0, B_1 \subset \mathbb{R}^n$ com

$$B_0 \subset B \subset B_1, \quad b - \varepsilon < \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) < b + \varepsilon.$$

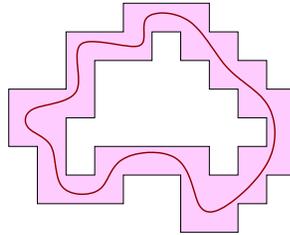


Figura 5.3: Conjuntos Jordan-mensuráveis.

Comentário 5.6.3. Pode-se escolher os edifícios de blocos B_0, B_1 em Teorema 5.6.2, parte (iii), tal que todos os seus pontos extremais são vetores (cantos) da forma $x = (2^{-m}k_1, \dots, 2^{-m}k_n)$ com $m \in \mathbb{N}$ e $k_1, \dots, k_n \in \mathbb{Z}$.

Demonstração de Teorema 5.6.2. A implicação (i) \implies (ii) segue imediatamente da definição da medida de Jordan e do Teorema 5.2.2.

Provaremos (ii) \implies (iii). Seja $Q \subset \mathbb{R}^n$ um bloco eixo-paralelo fechado o qual contem B . Pela hipótese a função característica $1_B : Q \rightarrow \mathbb{R}$ é Riemann-integrável e $\int_Q 1_B = b$. Seja $\varepsilon > 0$. Segundo o Teorema 5.3.1 existe uma partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(Q)$ com

$$b - \varepsilon < \underline{S}(1_B, P) \leq \overline{S}(1_B, P) < b + \varepsilon.$$

Defina

$$J_0 := \{j \in \{1, \dots, k\} \mid \overline{P}_j \subset B\}, \quad J_1 := \{j \in \{1, \dots, k\} \mid \overline{P}_j \cap B \neq \emptyset\}.$$

Então os conjuntos $B_0 := \bigcup_{j \in J_0} \overline{P}_j$ e $B_1 := \bigcup_{j \in J_1} \overline{P}_j$ são edifícios de blocos com $B_0 \subset B \subset B_1$ e

$$b - \varepsilon < \underline{S}(1_B, P) = \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) = \overline{S}(1_B, P) < b + \varepsilon.$$

Com isso fica provado que (ii) implica (iii).

Provaremos (iii) \implies (i). Seja dado $\varepsilon > 0$. Segundo (iii) existem dois edifícios de blocos $B_0, B_1 \subset \mathbb{R}^n$ com $B_0 \subset B \subset B_1$ e

$$b - \frac{\varepsilon}{4} < \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) < b + \frac{\varepsilon}{4}.$$

Defina

$$A := B_1 \setminus \overset{\circ}{B}_0 = \overline{B_1} \setminus \overline{B_0}.$$

Então $\partial B \subset A$. Como B_0 é um edifício de bloco, coincide o bordo do interior $\overset{\circ}{B}_0$ com o borde de B_0 . Daí $\overset{\circ}{B}_0$ é Jordan-mensurável e vale $\text{Vol}_n(\overset{\circ}{B}_0) = \text{Vol}_n(B_0) = \text{Vol}_n(B_0)$. Segundo o Lema 5.5.4 o conjunto A é Jordan-mensurável e vale

$$\int_{B_1} 1_A = \text{Vol}_n(A) = \text{Vol}_n(B_1) - \text{Vol}_n(B_0) < \frac{\varepsilon}{2}.$$

Portanto existe uma partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B_1)$ com soma superior $\overline{S}(1_A, P) < \varepsilon/2$. Seja $J := \{j \in \{1, \dots, k\} \mid \overline{P}_j \cap A \neq \emptyset\}$. Para $j \in J$ seja $W_j \subset \mathbb{R}^n$ um bloco aberto com $\overline{P}_j \subset W_j$ e $\text{Vol}_n(W_j) < 2\text{Vol}_n(P_j)$. Então vale $\partial B \subset \bigcup_{j \in J} W_j$ e

$$\sum_{j \in J} \text{Vol}_n(W_j) < 2 \sum_{j \in J} \text{Vol}_n(P_j) = 2\overline{S}(1_A, P) < \varepsilon.$$

Com isso temos mostrado que ∂B é um conjunto Jordan-nulo. Então B é Jordan-mensurável. Ora a equação $\text{Vol}_n(B) = b$ segue imediato de (iii). Com isso Teorema 5.6.2 fica provado. \square

Corolário 5.6.4. *Um subconjunto limitado $B \subset \mathbb{R}^n$ é um conjunto Jordan-nulo se e somente se é Jordan-mensurável e de volume nulo $\text{Vol}_n(B) = 0$.*

Demonstração. Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável com volume nulo $\text{Vol}_n(B) = 0$. Seja $\varepsilon > 0$. Então existe segundo o Teorema 5.6.2 um edifício de blocos $B_1 \subset \mathbb{R}^n$ com $B \subset B_1$ e $\text{Vol}_n(B_1) < \varepsilon/2$. Escolha uma partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B_1)$. Então vale $B \subset B_1 = \bigcup_{i=1}^k \bar{P}_i$ e $\sum_{i=1}^k \text{Vol}_n(P_i) = \text{Vol}_n(B_1) < \varepsilon/2$. Escolha blocos eixo-paralelos abertos W_i com $\bar{P}_i \subset W_i$ e $\text{Vol}_n(W_i) \leq 2\text{Vol}_n(P_i)$. Então $B \subset \bigcup_i W_i$ e $\sum_i \text{Vol}_n(W_i) < \varepsilon$. Então B é um conjunto Jordan-nulo.

Se B é um conjunto Jordan-nulo, então $\overset{\circ}{B} = \emptyset$ e \bar{B} é um conjunto Jordan-nulo segundo o Lema 5.2.3. Segue dali que $\partial B = \bar{B} \setminus \overset{\circ}{B} = \bar{B}$ é um conjunto Jordan-nulo. Além disso, segue de parte (ii) de Teorema 5.2.2 que o volume é nulo $\text{Vol}_n(B) = \int_Q 1_B = 0$ (para todo cubo fechado Q o qual contem B). Com isso Corolário 5.6.4 fica provado. \square

Corolário 5.6.5. *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável e $f : B \rightarrow \mathbb{R}$ uma função Riemann-integrável, tal que $f(x) \geq 0$ para todo $x \in B$. Então*

$$C := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid x \in B, 0 \leq y \leq f(x)\}$$

é um conjunto Jordan-mensurável e o seu volume é

$$\text{Vol}_{n+1}(C) = \int_B f.$$

Demonstração. Seja $Q \subset \mathbb{R}^n$ um bloco eixo-paralelo fechado contendo B . Defina $\tilde{f} : Q \rightarrow \mathbb{R}$ através de $\tilde{f}(x) := f(x)$ para $x \in B$ e de $\tilde{f}(x) := 0$ para $x \in Q \setminus B$. Então $\tilde{f} : Q \rightarrow \mathbb{R}$ é Riemann-integrável e $c := \int_Q \tilde{f} = \int_B f$. Segundo Teorema 5.3.1 existe uma partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(Q)$ com

$$c - \varepsilon < \underline{S}(\tilde{f}, P) \leq \bar{S}(\tilde{f}, P) < c + \frac{\varepsilon}{2}.$$

Defina

$$C_0 := \bigcup_{\inf_{\bar{P}_i} f > 0} \bar{P}_i \times [0, \inf_{\bar{P}_i} f], \quad C_1 := \bigcup_{i=1}^k \bar{P}_i \times \left[0, \sup_{\bar{P}_i} f + \frac{\varepsilon}{2\text{Vol}_n(Q)}\right].$$

Então C_0 e C_1 são dois edifício de blocos em \mathbb{R}^{n+1} com $C_0 \subset C \subset C_1$ e

$$c - \varepsilon < \underline{S}(\tilde{f}, P) = \text{Vol}_{n+1}(C_0) \leq \text{Vol}_{n+1}(C_1) = \bar{S}(\tilde{f}, P) + \frac{\varepsilon}{2} < c + \varepsilon.$$

Como $\varepsilon > 0$ foi escolhido arbitrariamente, segue de Teorema 5.6.2 que C é Jordan-mensurável e $\text{Vol}_{n+1}(C) = c$. Isso prova Corolário 5.6.5. \square

Corolário 5.6.6. *Sejam $p, q \in \mathbb{N}$ e $n := p + q$. Sejam $A \subset \mathbb{R}^p$ e $B \subset \mathbb{R}^q$ Jordan-mensurável. Então $A \times B \subset \mathbb{R}^n$ é Jordan-mensurável e vale*

$$\text{Vol}_n(A \times B) = \text{Vol}_p(A)\text{Vol}_q(B).$$

Demonstração. Segue diretamente das definições que o produto de um conjunto Jordan-nulo em \mathbb{R}^p com um subconjunto limitado de \mathbb{R}^q é um conjunto Jordan-nulo em \mathbb{R}^n . Neste caso segue a afirmação de Corolário 5.6.4. Portanto podemos supor que nem A , nem B , é um conjunto Jordan-nulo. Então vale

$$a := \text{Vol}_p(A) > 0, \quad b := \text{Vol}_q(B) > 0.$$

Dado $\varepsilon > 0$, suponhamos sem perda de generalidade que $\varepsilon < \min\{a, b, 1\}$. Segundo o Teorema 5.6.2 existem edifícios de blocos $A_0, A_1 \subset \mathbb{R}^p$ e $B_0, B_1 \subset \mathbb{R}^q$ com

$$A_0 \subset A \subset A_1, \quad B_0 \subset B \subset B_1$$

e

$$\begin{aligned} a - \frac{\varepsilon}{2b} < \text{Vol}_p(A_0) \leq \text{Vol}_p(A_1) < a + \frac{\varepsilon}{2(b+1)} \\ b - \frac{\varepsilon}{2a} < \text{Vol}_q(B_0) \leq \text{Vol}_q(B_1) < b + \frac{\varepsilon}{2(a+1)}. \end{aligned}$$

Então $A_0 \times B_0$ e $A_1 \times B_1$ são edifícios de blocos em \mathbb{R}^n com

$$A_0 \times B_0 \subset A \times B \subset A_1 \times B_1$$

e vale

$$\begin{aligned} ab - \varepsilon &< \left(a - \frac{\varepsilon}{2b}\right) \left(b - \frac{\varepsilon}{2a}\right) \\ &< \text{Vol}_p(A_0)\text{Vol}_q(B_0) \\ &= \text{Vol}_n(A_0 \times B_0) \\ &\leq \text{Vol}_n(A_1 \times B_1) \\ &= \text{Vol}_p(A_1)\text{Vol}_q(B_1) \\ &< \left(a + \frac{\varepsilon}{2(b+1)}\right) \left(b + \frac{\varepsilon}{2(a+1)}\right) \\ &\leq ab + \varepsilon. \end{aligned}$$

Portanto, segundo o Teorema 5.6.2, o produto $A \times B$ é Jordan-mensurável com volume $\text{Vol}_n(A \times B) = ab$. Com isso Corolário 5.6.6 fica provado. \square

Conjuntos Jordan-nulos e Lipschitz-continuidade

Parte (iii) do seguinte teorema pode ser considerada como caso especial do **Teorema de Sard** (veja o Comentário 3.3.13).

Teorema 5.6.7 (Conjuntos Jordan-nulos). *Sejam $p, q, n \in \mathbb{N}$.*

(i) *Seja $A \subset \mathbb{R}^p \times \mathbb{R}^q$ um conjunto compacto e seja*

$$A^y := \{x \in \mathbb{R}^p \mid (x, y) \in A\}$$

um conjunto Jordan-nulo em \mathbb{R}^p para todo $y \in \mathbb{R}^q$. Então A é um conjunto Jordan-nulo.

(ii) *Seja $U \subset \mathbb{R}^n$ um conjunto aberto, seja $A \subset U$ um conjunto Jordan-nulo compacto, e seja $f : U \rightarrow \mathbb{R}^n$ uma aplicação localmente Lipschitz-continua. Então $f(A) \subset \mathbb{R}^n$ também é um conjunto Jordan-nulo.*

(iii) *Seja $V \subset \mathbb{R}^d$ um conjunto aberto, seja $B \subset V$ um subconjunto compacto, e seja $g : V \rightarrow \mathbb{R}^n$ uma aplicação localmente Lipschitz-continua com $d < n$. Então $g(B)$ é um conjunto Jordan-nulo.*

Demonstração. Provaremos (i). Sejam $E \subset \mathbb{R}^p$ e $F \subset \mathbb{R}^q$ dois blocos eixo-paralelos fechados, tal que

$$A \subset \overset{\circ}{E} \times \overset{\circ}{F}$$

e seja $V := \text{Vol}_q(F) > 0$. Seja dado $\varepsilon > 0$ e defina

$$B := \{y \in F \mid A^y \neq \emptyset\}.$$

Então B é um subconjunto compacto de $\overset{\circ}{F}$ e para todo $y \in B$ existe uma partição

$$P(y) = \{P_1(y), \dots, P_{k(y)}(y)\} \in \mathcal{P}(E)$$

tal que

$$\sum_{i \in I(y)} \text{Vol}_p(P_i(y)) < \frac{\varepsilon}{V}, \quad I(y) := \left\{ i \in \{1, \dots, k(y)\} \mid \overline{P_i(y)} \cap A^y \neq \emptyset \right\}.$$

Para $y \in B$ defina

$$J(y) := \{1, \dots, k(y)\} \setminus I(y) = \left\{ i \in \{1, \dots, k(y)\} \mid \overline{P_i(y)} \cap A^y = \emptyset \right\}$$

e

$$U(y) := \left\{ y' \in \mathbb{R}^q \mid A^{y'} \cap \overline{P_i(y)} = \emptyset \text{ para todos os } i \in J(y) \right\}.$$

Então $U(y)$ é, para todo $y \in B$, um conjunto aberto contendo y . Por isso existe para todo $y \in B$ um $r(y) > 0$ com $B_{2r(y)}(y) \subset U(y) \cap \overset{\circ}{F}$.

Com esta construção as bolas abertas $B_{r(y)}(y)$ com $y \in B$ formam uma cobertura aberta do conjunto compacto B . Existem dali, segundo o Teorema C.1.2, um número finito de elementos $y_1, \dots, y_N \in B$ com

$$B \subset \bigcup_{\nu=1}^N B_{r(y_\nu)}(y_\nu).$$

Seja ora $Q = \{Q_1, \dots, Q_\ell\} \in \mathcal{P}(F)$ uma partição com a fineza

$$\delta(Q) < \min_{\nu=1, \dots, N} \frac{r(y_\nu)}{\sqrt{q}}.$$

Então existe para todo $j \in \{1, \dots, \ell\}$ com $\overline{Q}_j \cap B \neq \emptyset$ um $\nu \in \{1, \dots, N\}$ com $\overline{Q}_j \subset U(y_\nu)$. (Pois se $\eta \in \overline{Q}_j \cap B$, então existe um $\nu \in \{1, \dots, N\}$ com $\eta \in B_{r(y_\nu)}(y_\nu)$ e segue dali que $\overline{Q}_j \subset B_{r(y_\nu)}(\eta) \subset B_{2r(y_\nu)}(y_\nu) \subset U(y_\nu)$.) Ora, seja

$$J := \{j \in \{1, \dots, \ell\} \mid \overline{Q}_j \cap B \neq \emptyset\}$$

e escolha uma aplicação

$$J \rightarrow \{1, \dots, N\} : j \mapsto \nu(j)$$

tal que $\overline{Q}_j \subset U(y_{\nu(j)})$ para todos os $j \in J$. Então vale

$$A^y \subset \bigcup_{i \in I(y_{\nu(j)})} \overline{P_i(y_{\nu(j)})} \quad \text{para todos os } j \in J \text{ e todos os } y \in \overline{Q}_j$$

e daí

$$A \subset \bigcup_{j \in J} \bigcup_{i \in I(y_{\nu(j)})} \overline{P_i(y_{\nu(j)})} \times \overline{Q}_j.$$

Portanto o conjunto A é contido na união de um número finito de blocos fechados com volume total

$$\begin{aligned} \sum_{j \in J} \sum_{i \in I(y_{\nu(j)})} \text{Vol}_{p+q}(P_i(y_{\nu(j)}) \times Q_j) &= \sum_{j \in J} \sum_{i \in I(y_{\nu(j)})} \text{Vol}_p(P_i(y_{\nu(j)})) \text{Vol}_q(Q_j) \\ &< \frac{\varepsilon}{V} \sum_{j \in J} \text{Vol}_q(Q_j) \\ &\leq \varepsilon. \end{aligned}$$

Por isso A é um conjunto Jordan-nulo.

Provaremos (ii). Como f é localmente Lipschitz-contínua, a restrição de f a qualquer um subconjunto compacto de U é Lipschitz-contínua. Assim $f|_A : A \rightarrow \mathbb{R}^n$ é Lipschitz-contínua, ou seja, existe um $c > 0$, tal que para todos os $x, y \in \mathbb{R}^n$ vale

$$x, y \in A \quad \implies \quad \|f(x) - f(y)\| \leq c \|x - y\|.$$

Aqui utilizamos a norma euclidiana. Um **cubo fechado** em \mathbb{R}^n é um subconjunto da forma

$$Q = [a_1, a_1 + s] \times [a_2, a_2 + s] \times \cdots \times [a_n, a_n + s]$$

com $a_1, \dots, a_n \in \mathbb{R}$ e $s > 0$. O número s chama-se de **comprimento de canto do cubo** Q . Seja dado $\varepsilon > 0$. Como A é um conjunto Jordan-nulo, existem um número finito de cubos fechados $Q_1, \dots, Q_\ell \subset \mathbb{R}^n$, tal que

$$A \subset \bigcup_{j=1}^{\ell} Q_j, \quad \sum_{j=1}^{\ell} \text{Vol}_n(Q_j) \leq \frac{\varepsilon}{(3c\sqrt{n})^n}.$$

Seja $s_j > 0$ o comprimento de canto do cubo Q_j . Então $\text{Vol}_n(Q_j) = s_j^n$ e vale $\|x - y\| \leq \sqrt{n}s_j$ para todos os $x, y \in Q_j$. Segue dali

$$\|f(x) - f(y)\| \leq c\sqrt{n}s_j \quad \text{para todos os } x, y \in Q_j \cap A.$$

Daí existem cubos abertos $W_j \subset \mathbb{R}^n$, $j = 1, \dots, \ell$, do comprimento de lado $t_j := 3c\sqrt{n}s_j$ tal que $f(Q_j \cap A) \subset W_j$ para todo j . Como $A \subset \bigcup_j Q_j$, segue daí $f(A) \subset \bigcup_j W_j$ e

$$\sum_j \text{Vol}_n(W_j) = \sum_j (3c\sqrt{n}s_j)^n = (3c\sqrt{n})^n \sum_j \text{Vol}_n(Q_j) < \varepsilon.$$

Portanto $f(A)$ é um conjunto Jordan-nulo e com isso (ii) fica provado.

Provaremos (iii). Seja $V \subset \mathbb{R}^d$ um conjunto aberto, seja $B \subset V$ compacto, e seja $g : V \rightarrow \mathbb{R}^n$ uma aplicação localmente Lipschitz-contínua. Defina os conjuntos $A \subset U \subset \mathbb{R}^n$ e a função $f : U \rightarrow \mathbb{R}^n$ através de

$$U := V \times \mathbb{R}^{n-d}, \quad A := B \times \{0\}, \quad f(x_1, \dots, x_n) := g(x_1, \dots, x_d).$$

Então U é aberto, o subconjunto $A \subset U$ é um conjunto Jordan-nulo compacto, e $f : U \rightarrow \mathbb{R}^n$ é localmente Lipschitz-contínuo. Portanto $g(B) = f(A)$ é um conjunto Jordan-nulo segundo a parte (ii). Com isso Teorema 5.6.7 fica provado. \square

Exemplos

Exemplo 5.6.8. A hipótese que A é compacto, não pode ser omitida em parte (i) de Teorema 5.6.7. O conjunto

$$A := \{(x, y) = 2^{-m}(k, \ell) \in \mathbb{R}^2 \mid m \in \mathbb{N}, k, \ell \in \{1, 3, \dots, 2^m - 1\}\}$$

satisfaz $\bar{A} = \partial A = [0, 1]^2$ e dali não é um conjunto Jordan-nulo, ainda que o conjunto $A^y \subset [0, 1]$ contem, para todo o $y \in \mathbb{R}$, só um número finito de elementos e por isso é um conjunto Jordan-nulo.

Exemplo 5.6.9. A hipótese que g é localmente Lipschitz-contínuo, não pode ser omitida em parte (iii) de Teorema 5.6.7. Se $g : \mathbb{R} \rightarrow \mathbb{R}^2$ é uma aplicação contínua com $g([0, 1]) = [0, 1]^2$ (*curva enchendo espaço*), então $g([0, 1])$ não é um conjunto Jordan-nulo.

Exemplo 5.6.10. Seja $E \subset \mathbb{R}^n$ um subespaço linear com

$$d = \dim E < n$$

e seja $a \in \mathbb{R}^n$. Então todo subconjunto compacto

$$K \subset a + E$$

do subespaço afino $a + E$ é um conjunto Jordan-nulo em \mathbb{R}^n .

Demonstração: Seja e_1, \dots, e_d uma base de E e seja $g : \mathbb{R}^d \rightarrow \mathbb{R}^n$ a aplicação $g(x_1, \dots, x_d) := a + \sum_{i=1}^d x_i e_i$. Esta aplicação é suave e tem o conjunto imagem $g(\mathbb{R}^d) = a + E$. Como $B := g^{-1}(K)$ é compacto, o conjunto $K = g(B)$ é Jordan-nulo segundo a parte (iii) de Teorema 5.6.7.

Exemplo 5.6.11. O n -símplice $\Delta^n = \{x \in \mathbb{R}^n \mid x_i \geq 0, \sum_i x_i \leq 1\}$ é Jordan-mensurável segundo o Exemplo 5.6.10. (Veja também o Exemplo 5.4.5.)

Exemplo 5.6.12. Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável, seja $h > 0$, e seja

$$K_B := \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid 0 \leq y \leq h, x \in (1 - \frac{y}{h})B\}$$

o cone sobre B como em Exemplo 5.5.13. Então

$$\partial K_B = K_{\partial B} \cup (\bar{B} \times \{0\})$$

é um conjunto Jordan-nulo segundo a parte (i) de Teorema 5.6.7. Portanto K_B é Jordan-mensurável.

Exemplo 5.6.13. Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade cuja dimensão satisfaz $d = \dim M < n$ e seja $A \subset M$ um subconjunto compacto. Então A é um conjunto Jordan-nulo.

Demonstração: Para todo $x \in A$ existem dois subconjuntos abertos $U_x, V_x \subset \mathbb{R}^n$ e um C^1 -difeomorfismo $\phi_x : U_x \rightarrow V_x$ com

$$x \in U_x, \quad \phi_x(U_x \cap M) = V_x \cap (\mathbb{R}^d \times \{0\}).$$

Escolha vizinhanças compactas $K_x \subset U_x$ de x , uma para todo $x \in A$. Então para todo $x \in A$ o conjunto $L_x := \phi_x(K_x \cap M)$ é compacto e, segundo Exemplo 5.6.10, um conjunto Jordan-nulo. Dali $K_x \cap M = \phi_x^{-1}(L_x)$ também é um conjunto Jordan-nulo segundo (iii) em Teorema 5.6.7. Os conjuntos $\overset{\circ}{K}_x$, para $x \in A$, formam uma cobertura aberta de A . Como A é compacto, existe segundo o Teorema C.1.2 um número finito de pontos $x_1, \dots, x_\ell \in A$ com $A \subset \bigcup_{j=1}^{\ell} \overset{\circ}{K}_{x_j}$. Portanto A é a união finita dos conjuntos Jordan-nulos $K_{x_j} \cap A$ para $j = 1, \dots, \ell$ e por isso também é um conjunto Jordan-nulo.

Exemplo 5.6.14. Uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ chama-se de **própria**, se pre-imagens de compactos são compacto: Para todo subconjunto compacto $K \subset \mathbb{R}$ o conjunto $f^{-1}(K) = \{x \in \mathbb{R}^n \mid f(x) \in K\}$ é compacto. A isso equivalente é o critério que para toda sequência $(x_\nu)_{\nu \in \mathbb{N}}$ de vetores em \mathbb{R}^n vale

$$\sup_{\nu} |f(x_\nu)| < \infty \implies (x_\nu) \text{ possui uma subsequência convergente.} \quad (5.6.1)$$

Ora, seja $f : \mathbb{R}^n \rightarrow [0, \infty)$ continuamente diferenciável e próprio. Seja $c > 0$ um valor regular de f . Então o conjunto

$$B := \{x \in \mathbb{R}^n \mid f(x) \leq c\}$$

é compacto. Seu bordo é o conjunto compacto $\partial B = \{x \in \mathbb{R}^n \mid f(x) = c\}$. Como c é um valor regular de f , o bordo ∂B é uma C^1 -subvariedade $(n-1)$ -dimensional de \mathbb{R}^n (Corolário 3.3.12) e por isso é, segundo o Exemplo 5.6.13, um conjunto Jordan-nulo. Portanto B é Jordan-mensurável.

Exemplo 5.6.15. A bola unitária

$$B^n := \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

é Jordan-mensurável segundo o Exemplo 5.6.14 com $f(x) = \|x\|^2$, $c = 1$. Para todo número natural $k \in \mathbb{N}$ vamos provar as fórmulas

$$\text{Vol}_{2k}(B^{2k}) = \frac{\pi^k}{k!}, \quad \text{Vol}_{2k+1}(B^{2k+1}) = \frac{2^{k+1}\pi^k}{3 \cdot 5 \cdots (2k+1)}. \quad (5.6.2)$$

Para isso consideramos a semi-bola superior em B^{n+1} . O corte por esta semi-bola na altura $t \in [0, 1]$ é a bola do raio $r = \sqrt{1-t^2}$ no \mathbb{R}^n . Este é Jordan-mensurável segundo o Exemplo 5.5.12 com $\text{Vol}_n(rB^n) = r^n \text{Vol}_n(B^n)$. Dali segue do Teorema de Fubini (Corolário 5.5.7) a fórmula

$$\text{Vol}_{n+1}(B^{n+1}) = 2 \int_0^1 (1-t^2)^{n/2} dt \cdot \text{Vol}_n(B^n). \quad (5.6.3)$$

O fator

$$c_n := \int_0^1 (1-t^2)^{n/2} dt, \quad n \in \mathbb{N}_0 = \mathbb{N} \cup \{0\} \quad (5.6.4)$$

pode ser calculado indutivamente como segue. Vale $c_0 = 1$ e, com a substituição $t = \text{sen}(\theta)$ para $0 \leq \theta \leq \pi/2$,

$$c_1 = \int_0^1 \sqrt{1-t^2} dt = \int_0^{\pi/2} \cos^2(\theta) d\theta = \frac{\pi}{4}. \quad (5.6.5)$$

Como $\text{Vol}_1(B^1) = 2$, segue daí $\text{Vol}_2(B^2) = 2c_1 \text{Vol}_1(B^1) = \pi$ segundo (5.6.3). Além disso, vale $\frac{d}{dt}t(1-t^2)^{n/2} = (n+1)(1-t^2)^{n/2} - n(1-t^2)^{(n-2)/2}$. Dali segue através de integração a equação $0 = (n+1)c_n - nc_{n-2}$ e dali

$$c_n = \frac{n}{n+1} c_{n-2}$$

para todos os $n \in \mathbb{N}$ com $n \geq 2$. Com $c_0 = 1$ e $c_1 = \pi/4$ recebemos dali

$$c_{2k} = \frac{2k}{2k+1} \frac{2k-2}{2k-1} \cdots \frac{2}{3}, \quad c_{2k+1} = \frac{2k+1}{2k+2} \frac{2k-1}{2k} \cdots \frac{3}{4} \frac{\pi}{4}$$

para $k \in \mathbb{N}$. Estas fórmulas em contrapartida podem ser escritas na forma

$$c_{2k} = \frac{2^k k!}{3 \cdot 5 \cdot 7 \cdots (2k+1)}, \quad c_{2k+1} = \frac{3 \cdot 5 \cdot 7 \cdots (2k+1)}{2^{k+2} (k+1)!} \pi.$$

Segundo (5.6.3) e (5.6.4) seguem daí as equações

$$\begin{aligned} \frac{\text{Vol}_{2k+2}(B^{2k+2})}{\text{Vol}_{2k}(B^{2k})} &= 4c_{2k}c_{2k+1} = \frac{\pi}{k+1} \\ \frac{\text{Vol}_{2k+1}(B^{2k+1})}{\text{Vol}_{2k-1}(B^{2k-1})} &= 4c_{2k-1}c_{2k} = \frac{2\pi}{2k+1}. \end{aligned}$$

Com $\text{Vol}_1(B^1) = 2$ e $\text{Vol}_2(B^2) = \pi$ segue disso (5.6.2).

5.7 A fórmula da transformação de variáveis

A fórmula da transformação de variáveis (Teorema 5.7.11) é a extensão natural da regra de substituição a funções de várias variáveis. Ela é de importância fundamental para a integral de Riemann e a demonstração dela é a demonstração mais operosa do capítulo presente. Isto é em contraste grave ao caso $n = 1$, onde pode-se deduzir a regra de substituição facilmente do Teorema Fundamental de Cálculo. A seguinte subseção estende a noção de uma partição, no que blocos são substituídos por conjuntos Jordan-mensuráveis. Precisa-se a convergência das somas de Riemann, correspondentemente generalizadas, na demonstração da fórmula da transformação de variáveis.

Somas de Riemann generalizadas

Definição 5.7.1 (Jordan-partição). *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável. Uma **Jordan-partição de B** é um conjunto $Z = \{B_1, \dots, B_k\}$ de conjuntos Jordan-mensuráveis compactos $B_i \subset B$ tal que $B = \bigcup_{i=1}^k B_i$ e para todos os $i, j \in \{1, \dots, k\}$ com $i \neq j$ a interseção $B_i \cap B_j$ é um conjunto Jordan-nulo. O conjunto de todas as partições de B denota-se de $\mathcal{P}_{\text{Jord}}(B)$.*

Comentário 5.7.2. *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável e seja $f \in \mathcal{R}(B)$. Seja $Z = \{B_1, \dots, B_k\}$ uma Jordan-partição de B . Então vale*

$$\int_B f = \sum_{i=1}^k \int_{B_i} f, \quad \text{Vol}_n(B) = \sum_{i=1}^k \text{Vol}_n(B_i).$$

A primeira equação segue de Lema 5.5.4 através de indução, e a segunda equação segue da primeira com $f \equiv 1$.

Ora utilizamos a norma $\|x\|_\infty := \max_i |x_i|$ para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Denotamos de $d_\infty : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $(x, y) \mapsto \|x - y\|_\infty$, a distância associada. O **diâmetro de um subconjunto $B \subset \mathbb{R}^n$** com respeito desta métrica é o número $\delta(B) := \sup_{x, y \in B} d_\infty(x, y)$. O diâmetro de B é finito se e somente se B é limitado. A **distância de dois subconjuntos não-vazios $A, B \subset \mathbb{R}^n$** é definida através de $d_\infty(A, B) := \inf_{x \in A, y \in B} d_\infty(x, y)$. Esta distância pode ser plenamente nulo sem os conjuntos A e B se intersectam, ainda menos coincidem.

Definição 5.7.3 (Fineza). *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável. A **fineza** de uma Jordan-partição $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ é o número $\delta(Z) := \max_j \delta(B_j)$ (o diâmetro máximo dos B_j com respeito d_∞).*

Lema 5.7.4. *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável e $\Gamma \subset B$ um conjunto Jordan-nulo. Seja dado $\varepsilon > 0$. Então existe um número $\delta_0 > 0$, tal que para toda Jordan-partição $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ vale*

$$\delta(Z) < \delta_0 \quad \implies \quad \sum_{d_\infty(B_i, \Gamma) \leq \delta_0} \text{Vol}_n(B_i) < \varepsilon.$$

Demonstração. Escolha blocos abertos $W_1, \dots, W_N \subset \mathbb{R}^n$, tal que

$$\bar{\Gamma} \subset \bigcup_{\nu=1}^N W_\nu, \quad \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon.$$

Pode-se mostrar como em Passo 1 na prova de Lema 5.2.8 que existe um número $\rho > 0$, tal que para todo ponto $x \in \bar{\Gamma}$ a bola de raio ρ com centro x (com respeito à métrica d_∞) é contido por inteiro num dos conjuntos W_1, \dots, W_N , ou seja $\exists \rho > 0 \forall x \in \bar{\Gamma} \exists \nu \in \{1, \dots, N\} \forall \xi \in \mathbb{R}^n$:

$$d_\infty(x, \xi) < \rho \quad \implies \quad \xi \in W_\nu. \quad (5.7.1)$$

Seja ora $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ com $\delta(Z) < \rho/2$. Defina

$$\begin{aligned} J_\Gamma &:= \{j \in \{1, \dots, k\} \mid d_\infty(B_j, \bar{\Gamma}) \leq \rho/2\}, \\ J_\nu &:= \{j \in \{1, \dots, k\} \mid B_j \subset W_\nu\}, \quad \nu = 1, \dots, N. \end{aligned}$$

Então vale

$$J_\Gamma \subset \bigcup_{\nu=1}^N J_\nu. \quad (5.7.2)$$

Isto é seja $j \in J_\Gamma$. Então existem dois elementos $x \in \bar{\Gamma}$ e $y \in B_j$, tal que $d_\infty(x, y) \leq \rho/2$. (Aqui utilizamos o fato que $\bar{\Gamma}$ e B_j são conjuntos compactos.) Para este x existe ora um $\nu \in \{1, \dots, N\}$, tal que vale (5.7.1). Disso segue $B_j \subset W_\nu$, pois que para $\xi \in B_j$ vale a desigualdade $d_\infty(x, \xi) \leq d_\infty(x, y) + d_\infty(y, \xi) < \rho/2 + \delta(B_j) < \rho$ e dali $\xi \in W_\nu$. Disso segue $j \in J_\nu$. Com isso (5.7.2) fica provado. De (5.7.2) segue a desigualdade

$$\begin{aligned} \sum_{j \in J_\Gamma} \text{Vol}_n(B_j) &\leq \sum_{\nu=1}^N \sum_{j \in J_\nu} \text{Vol}_n(B_j) = \sum_{\nu=1}^N \text{Vol}_n \left(\bigcup_{j \in J_\nu} B_j \right) \\ &\leq \sum_{\nu=1}^N \text{Vol}_n(W_\nu) < \varepsilon. \end{aligned}$$

Com isso o Lema 5.7.4 fica provado. \square

Seja ora $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável e $Z = \{B_1, \dots, B_k\}$ uma Jordan-partição de B . Seja $f : B \rightarrow \mathbb{R}$ uma função limitada. Exato como na Definição 5.1.8 definimos a **soma superior** e a **soma inferior** de f e Z através de

$$\bar{S}(f, Z) := \sum_{j=1}^k (\sup_{B_j} f) \text{Vol}_n(B_j), \quad \underline{S}(f, Z) := \sum_{j=1}^k (\inf_{B_j} f) \text{Vol}_n(B_j).$$

Do mesmo modo como em Lema 5.1.9 prova-se a desigualdade

$$\sup_{Z \in \mathcal{P}_{\text{Jord}}(B)} \underline{S}(f, Z) \leq \inf_{Z \in \mathcal{P}_{\text{Jord}}(B)} \bar{S}(f, Z). \quad (5.7.3)$$

O seguinte teorema mostra que uma função limitada $f : B \rightarrow \mathbb{R}$ é Riemann-integrável se e somente se reina igualdade em (5.7.3).

Teorema 5.7.5. *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável, $f : B \rightarrow \mathbb{R}$ uma função limitada, e $c \in \mathbb{R}$. Então as seguintes afirmações são equivalente.*

(i) *f é Riemann-integrável e $c = \int_B f$.*

(ii) *Para todo $\varepsilon > 0$ existe um $Z \in \mathcal{P}_{\text{Jord}}(B)$, tal que vale*

$$c - \varepsilon < \underline{S}(f, Z) \leq \bar{S}(f, Z) < c + \varepsilon.$$

(iii) *Para todo $\varepsilon > 0$ existe um $\delta_0 > 0$ tal que para todo $Z \in \mathcal{P}_{\text{Jord}}(B)$ vale*

$$\delta(Z) < \delta_0 \quad \implies \quad c - \varepsilon < \underline{S}(f, Z) \leq \bar{S}(f, Z) < c + \varepsilon. \quad (5.7.4)$$

(iv) *Para todo $\varepsilon > 0$ existe um número $\delta_0 > 0$ tal que para toda Jordan-partição $P = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ e todos os $x_1, \dots, x_k \in \mathbb{R}^n$ vale*

$$\delta(Z) < \delta_0 \quad \implies \quad \left| c - \sum_{i=1}^k f(x_i) \text{Vol}_n(B_i) \right| < \varepsilon. \quad (5.7.5)$$

Os números $S_k = \sum_{i=1}^k f(x_i) \text{Vol}_n(B_i)$ chamam-se de **somas de Riemann (generalizadas)** de f .

Demonstração. Veja a página 205. □

Para a prova de Teorema 5.7.5 precisamos o seguinte Lema o qual compare as somas superiores e inferiores de f e Z com as somas superiores e inferiores de Definição 5.1.8.

Lema 5.7.6. *Sejam B , f , c como em Teorema 5.3.1 e seja $Q \subset \mathbb{R}^n$ um bloco eixo-paralelo fechado o qual contem B . Defina $\tilde{f} : Q \rightarrow \mathbb{R}$ através de*

$$\tilde{f}(x) := \begin{cases} f(x), & \text{para } x \in B, \\ 0, & \text{para } x \in Q \setminus B. \end{cases}$$

Seja $P = \{P_1, \dots, P_k\}$ uma partição de Q e defina Z_P e $\varepsilon_P > 0$ através de

$$Z_P := \{B_1, \dots, B_k\}, \quad B_i := \bar{P}_i \cap B,$$

e

$$\varepsilon_P := \sum_{\bar{P}_i \cap \partial B \neq \emptyset} \text{Vol}_n(P_i).$$

Então $Z_P \in \mathcal{P}_{\text{Jord}}(B)$ e vale

$$\left| \bar{S}(\tilde{f}, P) - \bar{S}(f, Z) \right| \leq 2 \|f\| \varepsilon_P, \quad \left| \underline{S}(\tilde{f}, P) - \underline{S}(f, Z) \right| \leq 2 \|f\| \varepsilon_P. \quad (5.7.6)$$

Aqui denota $\|f\| := \sup_B |f|$ a norma do supremo de f .

Demonstração. Defina os conjuntos de índices

$$I_0 := \{i \in \{1, \dots, k\} \mid \bar{P}_i \subset \overset{\circ}{B}\}, \quad I_1 := \{i \in \{1, \dots, k\} \mid \bar{P}_i \cap \partial B \neq \emptyset\}.$$

Então vale

$$\begin{aligned} \bar{S}(\tilde{f}, P) &= \sum_{i=1}^k (\sup_{\bar{P}_i} \tilde{f}) \text{Vol}_n(P_i) \\ &= \sum_{i \in I_0} (\sup_{B_i} f) \text{Vol}_n(B_i) + \sum_{i \in I_1} (\sup_{\bar{P}_i} \tilde{f}) \text{Vol}_n(P_i) \\ &= \sum_{i=1}^k (\sup_{B_i} f) \text{Vol}_n(B_i) - \sum_{i \in I_1} (\sup_{B_i} f) \text{Vol}_n(B_i) + \sum_{i \in I_1} (\sup_{\bar{P}_i} \tilde{f}) \text{Vol}_n(P_i) \\ &= \bar{S}(f, Z_P) - \sum_{i \in I_1} (\sup_{B_i} f) \text{Vol}_n(B_i) + \sum_{i \in I_1} (\sup_{\bar{P}_i} \tilde{f}) \text{Vol}_n(P_i). \end{aligned}$$

O módulo dos últimos dois somandos é menor como $\varepsilon_P \|f\|$. Dali segue a primeira desigualdade em (5.7.6). A segunda desigualdade prova-se igualmente. Com isso o Lema 5.7.6 fica provado. \square

Demonstração do Teorema 5.7.5. Provaremos (i) \implies (ii). Suponhamos sem perda de generalidade que f não anule-se identicamente. Sejam Q e \tilde{f} como em Lema 5.7.6. Escolha uma partição $P \in \mathcal{P}(Q)$, tal que

$$\varepsilon_P < \frac{\varepsilon}{4\|f\|}, \quad c - \frac{\varepsilon}{2} < \underline{S}(\tilde{f}, P) \leq \overline{S}(\tilde{f}, P) < c + \frac{\varepsilon}{2}.$$

A existência de uma tal partição segue de Lema 5.2.8 e Teorema 5.3.1. Segundo o Lema 5.7.6 segue daí

$$\overline{S}(f, Z_P) \leq \overline{S}(\tilde{f}, P) + 2\|f\|\varepsilon_P < c + \frac{\varepsilon}{2} + 2\|f\|\varepsilon_P < c + \varepsilon$$

$$\underline{S}(f, Z_P) \geq \underline{S}(\tilde{f}, P) - 2\|f\|\varepsilon_P > c - \frac{\varepsilon}{2} - 2\|f\|\varepsilon_P > c - \varepsilon.$$

Com isso fica provado que (i) implica (ii).

Provaremos (ii) \implies (iii). Seja dado $\varepsilon > 0$. Segundo (ii) existe uma Jordan-partição $Y = \{A_1, \dots, A_k\} \in \mathcal{P}_{\text{Jord}}(B)$ com

$$c - \varepsilon/2 < \underline{S}(f, Y) \leq \overline{S}(f, Y) < c + \varepsilon/2.$$

Então

$$\Gamma := \bigcup_{i=1}^k \partial A_i$$

é um conjunto Jordan-nulo. Então existe segundo o Lema 5.7.4 um $\delta_0 > 0$, tal que para todo $Z = \{B_1, \dots, B_\ell\} \in \mathcal{P}_{\text{Jord}}(B)$ vale

$$\delta(Z) < \delta_0 \quad \implies \quad \sum_{d_\infty(B_j, \Gamma) \leq \delta_0} \text{Vol}_n(B_j) < \frac{\varepsilon}{4\|f\|}.$$

Afirmção. *Seja $Z = \{B_1, \dots, B_\ell\} \in \mathcal{P}_{\text{Jord}}(B)$ com $\delta(Z) < \delta_0$ e seja $j \in \{1, \dots, \ell\}$ com $d_\infty(B_j, \Gamma) > \delta_0$. Então existe um $i \in \{1, \dots, k\}$ com*

$$B_j \subset A_i.$$

Como $B_j \subset B = \bigcup_i A_i$, existe um índice $i \in \{1, \dots, k\}$ com $A_i \cap B_j \neq \emptyset$. Seja $a \in A_i \cap B_j$. Então $d_\infty(a, \partial A_i) > \delta_0$ e dali

$$B_j \subset \overline{B}_{\delta_0}(a; d_\infty) = \{x \in \mathbb{R}^n \mid d_\infty(a, x) \leq \delta_0\} \subset A_i.$$

Com isso a afirmação fica provada.

Ora, seja $Z = \{B_1, \dots, B_\ell\} \in \mathcal{P}_{\text{Jord}}(B)$ com $\delta(Z) < \delta_0$. Defina

$$\begin{aligned} J_0 &:= \{j \in \{1, \dots, \ell\} \mid d_\infty(B_j, \Gamma) \leq \delta_0\} \\ J_1 &:= \{j \in \{1, \dots, \ell\} \mid d_\infty(B_j, \Gamma) > \delta_0\}. \end{aligned}$$

Além disso, seja

$$Y \wedge Z := \{A_i \cap B_j \mid i = 1, \dots, k, j = 1, \dots, \ell\}.$$

Então $Y \wedge Z \in \mathcal{P}_{\text{Jord}}(B)$ e

$$\begin{aligned} c - \frac{\varepsilon}{2} &< \underline{S}(f, Y) \\ &\leq \underline{S}(f, Y \wedge Z) \\ &= \sum_{i=1}^k \sum_{j=1}^{\ell} \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) \\ &= \sum_{j \in J_0} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) + \sum_{j \in J_1} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) \\ &= \sum_{j \in J_0} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) + \sum_{j \in J_1} \left(\inf_{B_j} f \right) \text{Vol}_n(B_j) \\ &= \sum_{j \in J_0} \sum_{i=1}^k \left(\inf_{A_i \cap B_j} f \right) \text{Vol}_n(A_i \cap B_j) - \sum_{j \in J_0} \left(\inf_{B_j} f \right) \text{Vol}_n(B_j) + \underline{S}(f, Z) \\ &\leq 2 \|f\| \sum_{j \in J_0} \text{Vol}_n(B_j) + \underline{S}(f, Z) \\ &< \frac{\varepsilon}{2} + \underline{S}(f, Z). \end{aligned}$$

Aqui no quinto passo temos utilizado a afirmação. Com isso vale

$$\underline{S}(f, Z) > c - \varepsilon$$

e igualmente

$$\overline{S}(f, Z) < c + \varepsilon$$

para todo $Z \in \mathcal{P}_{\text{Jord}}(B)$ com $\delta(Z) < \delta_0$. Com isso temos mostrado que (ii) implica (iii).

Provaremos (iii) \implies (i). Sejam Z e \tilde{f} como em Lema 5.7.6. Podemos supor sem perda de generalidade que f não anule-se identicamente. Seja dado $\varepsilon > 0$. Segundo (iii) existe uma constante $\delta_0 > 0$, tal que para todo $Z \in \mathcal{P}_{\text{Jord}}(B)$ vale

$$\delta(Z) < \delta_0 \quad \implies \quad c - \varepsilon/2 < \underline{S}(f, Z) \leq \overline{S}(f, Z) < c + \varepsilon/2.$$

Segundo o Lema 5.2.8 podemos escolher δ_0 tão pequeno, tal que para todos os $P \in \mathcal{P}(Q)$ vale

$$\delta(P) < \delta_0 \quad \implies \quad \varepsilon_P < \frac{\varepsilon}{4\|f\|}.$$

Ora, seja $P \in \mathcal{P}(Q)$ com $\delta(P) < \delta_0$. Seja $Z_P \in \mathcal{P}_{\text{Jord}}(B)$ como em Lema 5.7.6. Então vale $\delta(Z_P) < \delta_0$ e disso segue

$$\begin{aligned} c - \varepsilon &< c - \varepsilon/2 - 2\|f\|\varepsilon_P \\ &< \underline{S}(f, Z_P) - 2\|f\|\varepsilon_P \\ &\leq \underline{S}(\tilde{f}, P) \\ &\leq \overline{S}(\tilde{f}, P) \\ &\leq \overline{S}(f, Z_P) + 2\|f\|\varepsilon_P \\ &< c + \varepsilon/2 + 2\|f\|\varepsilon_P \\ &< c + \varepsilon. \end{aligned}$$

Como $\varepsilon > 0$ foi escolhido arbitrariamente, segue de Teorema 5.3.1 que a função $\tilde{f} : Q \rightarrow \mathbb{R}$ é Riemann-integrável e $\int_Q \tilde{f} = c$. Disso segue, segundo a definição, que a função $f : B \rightarrow \mathbb{R}$ é Riemann-integrável e $\int_B f = c$.

Com isso temos mostrado que as afirmações (i), (ii) e (iii) são equivalente uma à outra. A equivalência das afirmações (iii) e (iv) prova-se exatamente como em Teorema 5.3.1. Com isso Teorema 5.7.5 fica provado. \square

Exercício 5.7.7. (i) Todo subconjunto conexo limitado de \mathbb{R} é Jordan-mensurável.

(ii) Existe um subconjunto compacto $K \subset \mathbb{R}$ o qual não é Jordan-mensurável. **Dica:** Modifique a construção do conjunto de Cantor.

(iii) Existe um subconjunto conexo compacto de \mathbb{R} o qual não é Jordan-mensurável.

Transformações lineares

Teorema 5.7.8. *Seja $B \subset \mathbb{R}^n$ Jordan-mensurável e $\Phi \in \mathbb{R}^{n \times n}$, $v \in \mathbb{R}^n$. Então o conjunto $\Phi B + v := \{\Phi x + v \mid x \in B\}$ é Jordan-mensurável e tem a medida Jordan $\text{Vol}_n(\Phi B + v) = |\det(\Phi)| \text{Vol}_n(B)$.*

Demonstração. Veja a página 209. □

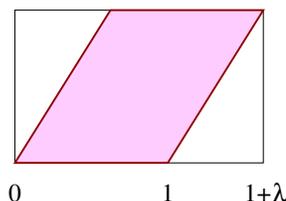


Figura 5.4: O cisalhamento de um quadrado.

Exemplo 5.7.9. Seja $\lambda > 0$ e

$$B := [0, 1]^2, \quad \Phi := \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}.$$

Então $A := \Phi B = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq 1, \lambda y \leq x \leq \lambda y + 1\}$. (Veja a Figura 5.4.) Para $y \in [0, 1]$ seja $A^y := \{x \in \mathbb{R} \mid (x, y) \in A\} = [\lambda y, \lambda y + 1]$. Então A^y é Jordan-mensurável e $\text{Vol}_1(A^y) = 1$ para todo $y \in [0, 1]$. Demais $\partial A = \Phi \partial B$ é, segundo o Teorema 5.6.7, um conjunto Jordan-nulo e por isso A é Jordan-mensurável. Portanto vale segundo o Corolário 5.5.7

$$\text{Vol}_2(B\Phi) = \int_0^1 \text{Vol}_1(A^y) dy = 1 = |\det(\Phi)| \text{Vol}_2(B).$$

Exemplo 5.7.10. Sejam a_1, \dots, a_n números reais positivos. O elipsoide

$$E(a_1, \dots, a_n) := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^n \frac{x_i^2}{a_i^2} \leq 1 \right\}$$

é a imagem da bola unitária $B := \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1\}$ sob a matriz diagonal $\Phi := \text{diag}(a_1, \dots, a_n)$. Portanto $E(a_1, \dots, a_n)$ é, segundo o Teorema 5.7.8, Jordan-mensurável e tem o volume

$$\text{Vol}_n(E(a_1, \dots, a_n)) = a_1 \cdots a_n \text{Vol}_n(B).$$

Aqui temos utilizado a fórmula $\det(\Phi) = a_1 \cdots a_n$. Posteriormente vamos deduzir uma fórmula para o volume da bola unitária (Exemplo 5.6.15).

Demonstração do Teorema 5.7.8. A prova tem sete passos.

Passo 1. *Seja $\Phi \in \mathbb{R}^{n \times n}$ com $\det(\Phi) = 0$ e $v \in \mathbb{R}^n$. Seja $B \subset \mathbb{R}^n$ um subconjunto limitado. Então $\Phi B + v$ é um conjunto Jordan-nulo.*

O conjunto imagem $E := \text{im } \Phi = \{\Phi x \mid x \in \mathbb{R}^n\}$ de Φ é um subespaço linear de \mathbb{R}^n de dimensão $d := \dim E < n$ e $\Phi(B)$ é um subconjunto limitado de E . Portanto $\Phi(B)$ é, segundo o Exemplo 5.6.10, um conjunto Jordan-nulo.

Passo 2. *Seja $\Phi \in \mathbb{R}^{n \times n}$ e $v \in \mathbb{R}^n$. Se $B \subset \mathbb{R}^n$ é um conjunto Jordan-mensurável, então $\Phi B + v$ também é Jordan-mensurável.*

Se $\det(\Phi) = 0$, então segue a Jordan-mensurabilidade de $\Phi B + v$ de Passo 1. Se $\det(\Phi) \neq 0$, então a aplicação $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $\phi(x) := \Phi x + v$ para $x \in \mathbb{R}^n$ é um difeomorfismo. E por isso o bordo de $\Phi B + v = \phi(B)$ é o conjunto $\partial(\Phi B + v) = \phi(\partial B) = \Phi \partial B + v$. Isto é um conjunto Jordan-nulo segundo o Teorema 5.6.7. Daí $\Phi B + v$ é Jordan-mensurável.

Passo 3. *Seja $B \subset \mathbb{R}^n$ um conjunto Jordan-mensurável e $v \in \mathbb{R}^n$ e $s \geq 0$. Então vale $\text{Vol}_n(sB + v) = s^n \text{Vol}_n(B)$.*

Se B seja um bloco aberto ou fechado, então a afirmação segue imediatamente das definições. Se B seja um edifício de blocos, escolha uma partição $P = \{P_1, \dots, P_k\} \in \mathcal{P}(B)$. Então $\text{Vol}_n(B) = \sum_{i=1}^k \text{Vol}_n(P_i)$ segundo o Comentário 5.5.3 (v). Demais

$$sP + v := \{sP_1 + v, \dots, sP_k + v\}$$

é uma partição de $sB + v$. Disso segue a afirmação para edifícios de blocos.

Seja ora $B \subset \mathbb{R}^n$ qualquer conjunto Jordan-mensurável e $b := \text{Vol}_n(B)$. Se $s = 0$, então $sB + v = \{v\}$ é um conjunto Jordan-nulo e satisfaz consequentemente a afirmação de Passo 3. Seja então $s > 0$ e escolha uma constante $\varepsilon > 0$. Então existem, segundo o Teorema 5.6.2, dois edifícios de blocos $B_0, B_1 \subset \mathbb{R}^n$, tal que

$$B_0 \subset B \subset B_1, \quad b - \frac{\varepsilon}{s^n} < \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) < b + \frac{\varepsilon}{s^n}.$$

Disso segue $sB_0 + v \subset sB + v \subset sB_1 + v$ e

$$s^n b - \varepsilon < \text{Vol}_n(sB_0 + v) \leq \text{Vol}_n(sB + v) \leq \text{Vol}_n(sB_1 + v) < s^n b + \varepsilon.$$

Como $\varepsilon > 0$ foi escolhido arbitrariamente a afirmação segue de Teorema 5.6.2.

Passo 4. Seja $\Phi \in \mathbb{R}^{n \times n}$ e $v \in \mathbb{R}^n$. Seja $W_0 := [0, 1]^n$ o cubo unitário fechado. Defina o número $\lambda(\Phi) \geq 0$ através de

$$\lambda(\Phi) := \text{Vol}_n(\Phi W_0).$$

Então vale

$$\text{Vol}_n(\Phi B + v) = \lambda(\Phi) \text{Vol}_n(B) \quad (5.7.7)$$

para todo conjunto Jordan-mensurável $B \subset \mathbb{R}^n$.

Seja $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ e $s > 0$ e

$$B := [a_1, a_1 + s] \times \cdots \times [a_n, a_n + s] = sW_0 + a.$$

Então $\text{Vol}_n(B) = s^n$ e dali segue de Passo 3 que

$$\text{Vol}_n(\Phi B) = \text{Vol}_n(s\Phi W_0 + sa) = s^n \text{Vol}_n(\Phi W_0) = \lambda(\Phi) \text{Vol}_n(B).$$

Com isso o Passo 4 fica provado para cubos fechados.

Ora, seja $B \subset \mathbb{R}^n$ um edifício de blocos o qual possui uma partição $P = \{P_1, \dots, P_k\}$, tal que \bar{P}_i é um cubo fechado para todo i . Então os conjuntos $\Phi \bar{P}_i + v$, $i = 1, \dots, k$, formam uma Jordan-partição de $\Phi B + v$. Conforme o que foi provado até agora e Comentário 5.7.2 vale daí que

$$\text{Vol}_n(\Phi B + v) = \sum_{i=1}^k \text{Vol}_n(\Phi \bar{P}_i + v) = \lambda(\Phi) \sum_{i=1}^k \text{Vol}_n(P_i) = \lambda(\Phi) \text{Vol}_n(B).$$

Com isso Passo 4 fica provado para todo edifício de blocos o qual possui uma partição de cubos.

Seja ora $B \subset \mathbb{R}^n$ qualquer conjunto Jordan-mensurável e $b := \text{Vol}_n(B)$. Seja $\varepsilon > 0$. Então existem, segundo o Teorema 5.6.2, dois edifícios de blocos $B_0, B_1 \subset \mathbb{R}^n$, tal que $B_0 \subset B \subset B_1$ e

$$b - \varepsilon \leq \text{Vol}_n(B_0) \leq \text{Vol}_n(B_1) \leq b + \varepsilon.$$

Segundo o Comentário 5.6.3 pode-se escolher tais B_0 e B_1 as quais possuem partições de cubos. Com isso segue do que foi provado até agora que

$$\lambda(\Phi)(b - \varepsilon) \leq \text{Vol}_n(\Phi B_0 + v) \leq \text{Vol}_n(\Phi B_1 + v) \leq \lambda(\Phi)(b + \varepsilon).$$

Disso segue $\lambda(\Phi)(b - \varepsilon) \leq \text{Vol}_n(\Phi B + v) \leq \lambda(\Phi)(b + \varepsilon)$ para todo $\varepsilon > 0$ e dali $\text{Vol}_n(\Phi B + v) = \lambda(\Phi)b$. Com isso o Passo 4 fica provado.

Passo 5. A função $\lambda : \mathbb{R}^{n \times n} \rightarrow [0, \infty)$ definida em Passo 4 tem propriedades como segue.

- (a) Se $\Phi \in \mathbb{R}^{n \times n}$ com $\det(\Phi) = 0$, então vale $\lambda(\Phi) = 0$.
- (b) Para todo $s \geq 0$ vale $\lambda(s\mathbb{1}) = s^n$.
- (c) Para todo $\Phi, \Psi \in \mathbb{R}^{n \times n}$ vale $\lambda(\Phi\Psi) = \lambda(\Phi)\lambda(\Psi)$.

Se $\det(\Phi) = 0$, então ΦW_0 é um conjunto Jordan-nulo segundo o Passo 1. Portanto vale segundo o Corolário 5.6.4

$$\lambda(\Phi) = \text{Vol}_n(\Phi W_0) = 0.$$

Se $s \geq 0$, então vale segundo o Passo 3

$$\lambda(s\mathbb{1}) = \text{Vol}_n(sW_0) = s^n.$$

Se $\Phi, \Psi \in \mathbb{R}^{n \times n}$, então vale segundo o Passo 4

$$\lambda(\Phi\Psi) = \text{Vol}_n(\Phi\Psi W_0) = \lambda(\Phi)\text{Vol}_n(\Psi W_0) = \lambda(\Phi)\lambda(\Psi).$$

Com isso o Passo 5 fica provado.

Passo 6. Para toda matriz elementar $\Phi \in \mathbb{R}^{n \times n}$ vale $\lambda(\Phi) = |\det(\Phi)|$.

Uma *matriz elementar*, ou é uma matriz de permutação, ou uma matriz diagonal, ou um cisalhamento. Uma matriz de permutação $\Phi \in \mathbb{R}^{n \times n}$ tem entradas 0 e 1, com exato um 1 na toda linha e toda coluna. Uma tal matriz representa uma aplicação $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ da forma

$$\Phi x = (x_{\sigma(1)}, \dots, x_{\sigma(n)}), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n,$$

onde $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ é uma permutação (logo uma aplicação bijetiva). Tal aplicação tem o determinante $\det(\Phi) = \pm 1$ e leva o cubo unitário W_0 bijetivamente para si mesmo. Por isso vale

$$\lambda(\Phi) = \text{Vol}_n(\Phi W_0) = \text{Vol}_n(W_0) = 1 = |\det(\Phi)|.$$

Se $\Phi = \text{diag}(\lambda_1, \dots, \lambda_n)$ é uma matriz diagonal, então $\det(\Phi) = \lambda_1 \cdots \lambda_n$ e

$$\Phi W_0 = I_1 \times \cdots \times I_n, \quad I_i := \begin{cases} [0, \lambda_i], & \text{se } \lambda_i \geq 0, \\ [\lambda_i, 0], & \text{se } \lambda_i < 0. \end{cases}$$

O intervalo I_i tem então o comprimento $|\lambda_i|$ e disso segue

$$\lambda(\Phi) = \text{Vol}_n(\Phi W_0) = \prod_{i=1}^n |\lambda_i| = |\det(\Phi)|.$$

Um cisalhamento é uma matriz da forma

$$\Phi = \begin{pmatrix} 1 & \lambda & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

com $\lambda > 0$. Esta matriz tem o determinante 1 e a imagem do cubo unitário é o conjunto

$$\Phi W_0 = \{(x_1 + \lambda x_2, x_2, \dots, x_n) \mid 0 \leq x_i \leq 1\} = P \times Q$$

com

$$P := \{(x_1 + \lambda x_2, x_2) \mid 0 \leq x_i \leq 1\}, \quad Q := [0, 1]^{n-2}.$$

Segundo o Exemplo 5.7.9 vale $\text{Vol}_2(P) = 1$. Portanto segue de Corolário 5.6.6 a equação

$$\lambda(\Phi) = \text{Vol}_n(P \times Q) = \text{Vol}_2(P) \text{Vol}_{n-2}(Q) = 1 = |\det(\Phi)|.$$

Com isso o Passo 6 fica provado.

Passo 7. Para toda matriz $\Phi \in \mathbb{R}^{n \times n}$ vale $\lambda(\Phi) = |\det(\Phi)|$.

Segundo um teorema da Álgebra Linear pode-se escrever toda matriz $\Phi \in \mathbb{R}^{n \times n}$ como produto

$$\Phi = \Phi_1 \Phi_2 \cdots \Phi_\ell$$

de matrizes elementares. Como o determinante de um produto de matrizes quadradas é igual ao produto dos determinantes (também segundo um teorema da Álgebra Linear), segue do Passo 5 e Passo 6 a equação

$$\lambda(\Phi) = \prod_{i=1}^{\ell} \lambda(\Phi_i) = \prod_{i=1}^{\ell} |\det(\Phi_i)| = |\det(\Phi)|.$$

Com isso Passo 7 fica provado.

A afirmação de Teorema 5.7.8 segue direto do Passo 2, Passo 4 e Passo 7.

□

Transformações não-lineares

Segue o teorema principal desta seção.

Teorema 5.7.11 (Fórmula da transformação de variáveis). *Seja $U \subset \mathbb{R}^n$ um aberto e $\phi : U \rightarrow \mathbb{R}^n$ uma aplicação continuamente diferenciável. Seja $A \subset U$ um conjunto Jordan-mensurável compacto e seja $N \subset A$ um conjunto Jordan-nulo compacto. Suponhamos que a restrição $\phi|_{A \setminus N} : A \setminus N \rightarrow \mathbb{R}^n$ seja injetiva e satisfaça a condição*

$$\det(d\phi(x)) \neq 0 \quad \text{para todos os } x \in A \setminus N. \quad (5.7.8)$$

Então o conjunto $B := \phi(A) \subset \mathbb{R}^n$ é Jordan-mensurável. Se $f : B \rightarrow \mathbb{R}$ seja Riemann-integrável, então a composição $f \circ \phi : A \rightarrow \mathbb{R}$ também é Riemann-integrável e vale a equação

$$\int_B f(y) dy_1 \cdots dy_n = \int_A f(\phi(x)) |\det(d\phi(x))| dx_1 \cdots dx_n. \quad (5.7.9)$$

Demonstração. Veja a página 216. □

Lema 5.7.12. *Seja $U \subset \mathbb{R}^n$ um conjunto aberto, seja $A \subset U$ um conjunto Jordan-mensurável compacto, seja $\phi : U \rightarrow \mathbb{R}^n$ uma aplicação continuamente diferenciável, e seja $N := \{x \in A \mid \det(d\phi(x)) = 0\}$ um conjunto Jordan-nulo. Então $B := \phi(A)$ é Jordan-mensurável e vale $\partial B \subset \phi(\partial A \cup N)$.*

Demonstração. Primeiro B é a imagem contínua de um conjunto compacto e por isso compacto e assim também fechado. Demais

$$A \setminus (\partial A \cup N) = \overset{\circ}{A} \setminus N$$

é um subconjunto aberto de U e segundo a hipótese vale $\det(d\phi(x)) \neq 0$ para todos os $x \in A \setminus (\partial A \cup N)$. Pelo Corolário 3.1.13 o conjunto $\phi(A \setminus (\partial A \cup N))$ é dali aberto e disso segue que $\phi(A \setminus (\partial A \cup N)) \subset \overset{\circ}{B}$. Com isso recebemos

$$\partial B = B \setminus \overset{\circ}{B} \subset \phi(A) \setminus \phi(A \setminus (\partial A \cup N)) \subset \phi(\partial A \cup N).$$

Segundo o Teorema 5.6.7 o conjunto $\phi(\partial A \cup N)$ é Jordan-nulo. Então ∂B é um conjunto Jordan-nulo e com isso o Lema 5.7.12 fica provado. □

Comentário 5.7.13. A afirmação de Lema 5.7.12 vale também caso o conjunto $N := \{x \in A \mid \det(d\phi(x)) = 0\}$ não seria Jordan-nulo. Mas sem esta hipótese a prova torna-se mais operosa. Resta no fato que $\phi(N)$ ainda é um conjunto Jordan-nulo ainda se N mesmo não for um conjunto Jordan-nulo. Isto é um caso especial do **Teorema de Sard** (veja o Comentário 3.3.13), e dali segue a afirmação como em Lema 5.7.12.

Lema 5.7.14. *Seja $U \subset \mathbb{R}^n$ um conjunto aberto e $K \subset U$ um subconjunto compacto. Seja $\phi : U \rightarrow \mathbb{R}^n$ uma aplicação continuamente diferenciável, tal que $\det(d\phi(x)) \neq 0$ para todos os $x \in K$. Então existe para todo $\varepsilon > 0$ uma constante $\delta > 0$ tal que vale o seguinte. Se $0 < s < \delta$ e $a \in \mathbb{R}^n$ com*

$$W := [a_1 - s, a_1 + s] \times \cdots \times [a_n - s, a_n + s] \subset K \quad (5.7.10)$$

então vale

$$\left| \text{Vol}_n(\phi(W)) - |\det(d\phi(a))| \text{Vol}_n(W) \right| \leq \varepsilon \text{Vol}_n(W). \quad (5.7.11)$$

Demonstração. Lembramos primeiro as definições das normas

$$\|x\|_\infty := \max_{i=1, \dots, n} |x_i|, \quad \|\Phi\|_\infty := \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|\Phi x\|_\infty}{\|x\|_\infty} \quad (5.7.12)$$

para $x \in \mathbb{R}^n$ e $\Phi \in \mathbb{R}^{n \times n}$. O conjunto $W \subset \mathbb{R}^n$ em (5.7.10) é a bola fechada $W = \overline{B}_s(a; d_\infty)$ de raio s com centro a com respeito à função distância $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} : (x, y) \mapsto d_\infty(x, y) := \|x - y\|_\infty$. Pela hipótese as aplicações

$$K \rightarrow \mathbb{R} : x \mapsto \|d\phi(x)^{-1}\|_\infty, \quad K \rightarrow \mathbb{R} : x \mapsto |\det(d\phi(x))|$$

são contínuas. Como K é compacto, estas funções são contínuas. Por isso existe uma constante $c > 0$, tal que

$$\|d\phi(x)^{-1}\|_\infty \leq c, \quad |\det(d\phi(x))| \leq c \quad \text{para todos os } x \in K. \quad (5.7.13)$$

Seja dado $\varepsilon > 0$. Escolha $\rho > 0$ tão pequeno que

$$1 - \frac{\varepsilon}{c} < (1 - \rho)^n < (1 + \rho)^n < 1 + \frac{\varepsilon}{c}. \quad (5.7.14)$$

Escolha $\delta > 0$ tão pequeno que para todos os $x, y \in \mathbb{R}^n$ vale

$$x, y \in K, \quad \|x - y\|_\infty < \delta \quad \implies \quad \|d\phi(x) - d\phi(y)\| < \frac{\rho}{c}. \quad (5.7.15)$$

Uma tal constante existe, porque a aplicação $d\phi : U \rightarrow \mathbb{R}^{n \times n}$ é uniformemente contínua no conjunto compacto $K \subset U$. Provaremos que a afirmação de Lema 5.7.14 é satisfeita com esta constante δ .

Sejam então dados $a \in \mathbb{R}^n$ e $0 < s < \delta$, tal que

$$W := \overline{B}_s(a; d_\infty) \subset K.$$

Então vale $\|a - x\| < \delta$ para todos os $x \in W$ e dali segue, segundo (5.7.15) com $\Phi := d\phi(a) \in \mathbb{R}^{n \times n}$, a desigualdade

$$\|d\phi(x) - \Phi\| < \frac{\rho}{c} \leq \frac{\rho}{\|\Phi^{-1}\|_\infty} \quad \text{para todos os } x \in W.$$

Aqui a segunda desigualdade segue de (5.7.13) com $x = a$. Consideramos ora a aplicação

$$\psi : U \rightarrow \mathbb{R}^n, \quad \psi(x) := \Phi^{-1}\phi(x).$$

Então vale $d\psi(x) = \Phi^{-1}d\phi(x)$ e

$$\|d\psi(x) - \mathbb{1}\|_\infty = \|\Phi^{-1}(d\phi(x) - \Phi)\|_\infty \leq \|\Phi^{-1}\|_\infty \|d\phi(x) - \Phi\|_\infty < \rho$$

para todos os $x \in W$. Segundo o Lema 3.1.12, segue daí

$$(1 - \rho)\Phi(W - a) \subset \phi(W) - \phi(a) \subset (1 + \rho)\Phi(W + a).$$

Segundo o Lema 5.7.12, o conjunto $\phi(W) - \phi(a)$ é Jordan-mensurável e, segundo o Teorema 5.7.8, vale

$$\begin{aligned} (1 - \rho)^n |\det(\Phi)| \text{Vol}_n(W) &= \text{Vol}_n((1 - \rho)\Phi(W - a)) \\ &\leq \text{Vol}_n(\phi(W)) \\ &\leq \text{Vol}_n((1 + \rho)\Phi(W + a)) \\ &= (1 + \rho)^n |\det(\Phi)| \text{Vol}_n(W). \end{aligned}$$

Disso segue a desigualdade

$$\begin{aligned} &\left| \text{Vol}_n(\phi(W)) - |\det(\Phi)| \text{Vol}_n(W) \right| \\ &\leq \max \{1 - (1 - \rho)^n, (1 + \rho)^n - 1\} |\det(\Phi)| \text{Vol}_n(W) \\ &\leq \max \{1 - (1 - \rho)^n, (1 + \rho)^n - 1\} c \text{Vol}_n(W) \\ &\leq \varepsilon \text{Vol}_n(W). \end{aligned}$$

Aqui a segunda desigualdade segue de (5.7.13) e a terceira de (5.7.14). Com isso o Lema 5.7.14 fica provado. \square

Demonstração do Teorema 5.7.11. A demonstração tem três passos.

Passo 1. Se $N = \emptyset$ e $A \subset U$ é um edifício de blocos o qual possui uma partição de cubos, então vale

$$\text{Vol}_n(\phi(A)) = \int_A |\det(d\phi(x))| dx_1 \cdots dx_n. \quad (5.7.16)$$

Seja dado $\varepsilon > 0$ e escolha $\delta > 0$, tal que a afirmação de Lema 5.7.14 vale para $K = A$. A função $A \rightarrow \mathbb{R} : x \mapsto |\det(d\phi(x))|$ é contínua e por isso Riemann-integrável (veja o Comentário 5.5.6). Segundo o Teorema 5.3.1 existe então uma partição $P = \{P_1, \dots, P_k\}$ de A , tal que \bar{P}_i é um cubo fechado para todo i , e qual satisfaz para todo $x_i \in \bar{P}_i$ a desigualdade

$$\left| \int_A |\det(d\phi(x))| dx_1 \cdots dx_n - \sum_{i=1}^k |\det(d\phi(x_i))| \text{Vol}_n(P_i) \right| < \varepsilon. \quad (5.7.17)$$

Mais além podemos escolher P , tal que $\delta(P) < \delta$. Seja ora $x_i \in P_i$ o centro de P_i . Então vale, segundo a escolha de δ , a desigualdade

$$\left| \text{Vol}_n(\phi(P_i)) - |\det(d\phi(x_i))| \text{Vol}_n(P_i) \right| < \varepsilon \text{Vol}_n(P_i) \quad (5.7.18)$$

para $i = 1, \dots, k$. Ora $\bigcup_i \partial P_i$ é um conjunto Jordan-nulo. Com isso $\phi(\bigcup_i \partial P_i)$ também é um conjunto Jordan-nulo, segundo o Teorema 5.6.7. Segundo o Lema 5.7.12 com $N = \emptyset$, vale além disso que $\partial\phi(P_i) \subset \phi(\partial P_i)$ para todo i . Com isso $\bigcup_i \partial\phi(P_i)$ é um conjunto Jordan-nulo e daí vale $\text{Vol}_n(\phi(A)) = \sum_{i=1}^k \text{Vol}_n(\phi(P_i))$ segundo o Lema 5.5.4. Combinamos isto com as desigualdades (5.7.17) e (5.7.18), então resulta

$$\begin{aligned} \left| \text{Vol}_n(\phi(A)) - \int_A |\det(d\phi)| \right| &= \left| \sum_{i=1}^k \text{Vol}_n(\phi(P_i)) - \int_A |\det(d\phi)| \right| \\ &\leq \sum_{i=1}^k \left| \text{Vol}_n(\phi(P_i)) - |\det(d\phi(x_i))| \text{Vol}_n(P_i) \right| \\ &\quad + \left| \sum_{i=1}^k |\det(d\phi(x_i))| \text{Vol}_n(P_i) - \int_A |\det(d\phi)| \right| \\ &\leq (\text{Vol}_n(A) + 1) \varepsilon. \end{aligned}$$

Como $\varepsilon > 0$ foi escolhido arbitrariamente, segue dali o Passo 1.

Passo 2. Nas hipóteses de Teorema 5.7.11 vale (5.7.16).

Seja dado $\varepsilon > 0$. Escolha $c > 0$ com

$$\|\phi(x) - \phi(x')\|_\infty \leq c \|x - x'\|_\infty \quad \text{para todos os } x, x' \in A. \quad (5.7.19)$$

Uma tal constante $c > 0$ existe segundo o teorema da barreira (Teorema 2.4.1). Escolha uma constante $C > 0$ com

$$|\det(d\phi(x))| \leq C \quad \text{para todos os } x \in A. \quad (5.7.20)$$

Escolha dois edifícios de blocos $A_0, A_1 \subset \mathbb{R}^n$ as quais possuem partições de cubos e satisfazem as seguintes condições

$$\begin{aligned} A_0 &\subset A \setminus N, & \text{Vol}_n(A) - \varepsilon &< \text{Vol}_n(A_0), \\ A \setminus A_0 &\subset A_1, & \text{Vol}_n(A_1) &< \text{Vol}_n(A \setminus A_0) + \varepsilon < 2\varepsilon. \end{aligned} \quad (5.7.21)$$

Tais edifícios de blocos existem segundo o Teorema 5.6.2.

Escolha uma partição $P = \{P_1, \dots, P_k\}$ de A_1 . Então \bar{P}_i é um cubo fechado para todo i . Se s_i é o comprimento lateral de P_i , então $\|x - x'\|_\infty \leq s_i$ para todos os $x, x' \in \bar{P}_i$. Disso segue, segundo (5.7.19), a desigualdade

$$\|\phi(x) - \phi(x')\|_\infty \leq cs_i \quad \text{para todos os } x, x' \in \bar{P}_i \cap A.$$

Dali existe, para todo i , um cubo fechado $Q_i \subset \mathbb{R}^n$ com comprimento lateral $2cs_i$, tal que $\phi(\bar{P}_i \cap A) \subset Q_i$ e $\text{Vol}_n(Q_i) = (2cs_i)^n$. Além disso, vale

$$\phi(A) \setminus \phi(A_0) \subset \phi(A \setminus A_0) \subset \phi\left(\bigcup_i \bar{P}_i \cap A\right) \subset \bigcup_i Q_i$$

e daí

$$\text{Vol}_n(\phi(A) \setminus \phi(A_0)) \leq \sum_{i=1}^k \text{Vol}_n(Q_i) = (2c)^n \sum_{i=1}^k s_i^n = (2c)^n \text{Vol}_n(A_1) < 2(2c)^n \varepsilon.$$

Além disso, vale segundo o Passo 1, bem como (5.7.20) e (5.7.21), que

$$\text{Vol}_n(\phi(A_0)) = \int_{A_0} |\det(d\phi)|, \quad \int_{A \setminus A_0} |\det(d\phi)| \leq C \text{Vol}_n(A \setminus A_0) \leq C\varepsilon.$$

Dali segue

$$\begin{aligned} \left| \text{Vol}_n(\phi(A)) - \int_A |\det(d\phi)| \right| &\leq \text{Vol}_n(\phi(A) \setminus \phi(A_0)) + \int_{A \setminus A_0} |\det(d\phi)| \\ &\leq (2(2c)^n + C)\varepsilon. \end{aligned}$$

Como $\varepsilon > 0$ foi escolhido arbitrariamente, segue daí a afirmação de Passo 2.

Passo 3. *Provaremos Teorema 5.7.11.*

Sejam $N \subset A \subset U$, $\phi : U \rightarrow \mathbb{R}^n$ e $f : B = \phi(A) \rightarrow \mathbb{R}$ como nas hipóteses de Teorema 5.7.11. A norma do supremo de f é denotada assim

$$\|f\| := \sup_{y \in B} |f(y)|.$$

Além disso, utilizamos a notação (5.7.12).

Seja dado $\varepsilon > 0$. Como $A \subset U$ é compacto e $\phi : U \rightarrow \mathbb{R}^n$ é continuamente diferenciável, a restrição $\phi|_A : A \rightarrow \mathbb{R}^n$ é Lipschitz-contínuo segundo o teorema da barreira (Teorema 2.4.1). Por isso existe uma constante $c > 0$, tal que

$$\|\phi(x) - \phi(x')\|_\infty \leq c \|x - x'\|_\infty \quad \text{para todos os } x, x' \in A. \quad (5.7.22)$$

Como $f : B \rightarrow \mathbb{R}$ é Riemann-integrável, existe segundo o Teorema 5.7.5 um $\delta_0 > 0$, tal que para toda Jordan-partição $Z = \{B_1, \dots, B_k\} \in \mathcal{P}_{\text{Jord}}(B)$ e todos os vetores $y_1, \dots, y_k \in \mathbb{R}^n$ vale o seguinte:

$$\delta(Z) < c\delta_0, \quad y_i \in B_i \forall i \implies \left| \int_B f - \sum_{i=1}^k f(y_i) \text{Vol}_n(B_i) \right| < \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)}. \quad (5.7.23)$$

Como a função $A \rightarrow \mathbb{R} : x \mapsto \det(d\phi(x))$ é uniformemente contínua podemos escolher δ_0 tão pequeno que todos os $x, x' \in A$ com $\|x - x'\|_\infty < \delta_0$ satisfazem a desigualdade

$$|\det(d\phi(x)) - \det(d\phi(x'))| < \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)}. \quad (5.7.24)$$

Seja ora $Y = \{A_1, \dots, A_k\} \in \mathcal{P}_{\text{Jord}}(A)$ uma Jordan-partição de A com

$$\delta(Y) = \max_{i=1, \dots, k} \delta(A_i) < \delta_0.$$

Como $\phi|_{A \setminus N}$ é injetivo, vale $\phi(A_i) \cap \phi(A_j) \subset \phi((A_i \cap A_j) \cup N)$. Segundo o Teorema 5.6.7, segue dali para $i \neq j$ que $\phi(A_i) \cap \phi(A_j)$ é um conjunto Jordan-nulo. Com isso o conjunto

$$Z := \{B_1, \dots, B_k\}, \quad B_i := \phi(A_i)$$

é uma Jordan-partição de $B = \phi(A)$. Além disso, vale $\|x - x'\|_\infty < \delta_0$ para todo i e todos os $x, x' \in A_i$. Segundo (5.7.22) segue dali $\|y - y'\|_\infty < c\delta_0$ para todo i e todos os $y, y' \in B_i$. Ou seja, a Jordan-partição $Z \in \mathcal{P}_{\text{Jord}}(B)$ satisfaz a desigualdade $\delta(Z) < c\delta_0$. Então Z satisfaz a hipótese de (5.7.23).

Com $x_i \in A_i$ e $y_i := \phi(x_i)$ resulta

$$\begin{aligned}
& \left| \int_B f - \sum_{i=1}^k f(\phi(x_i)) |\det(d\phi(x_i))| \text{Vol}_n(A_i) \right| \\
& \leq \left| \int_B f - \sum_{i=1}^k f(y_i) \text{Vol}_n(B_i) \right| \\
& \quad + \|f\| \sum_{i=1}^k |\text{Vol}_n(\phi(A_i)) - |\det(d\phi(x_i))| \text{Vol}_n(A_i)| \\
& < \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)} + \|f\| \sum_{i=1}^k \int_{A_i} \| |d\phi(x)| - |d\phi(x_i)| \| dx \\
& \leq \frac{\varepsilon}{1 + \|f\| \text{Vol}_n(A)} + \sum_{i=1}^k \frac{\varepsilon \|f\| \text{Vol}_n(A_i)}{1 + \|f\| \text{Vol}_n(A)} = \varepsilon.
\end{aligned}$$

Aqui a segunda desigualdade segue de (5.7.23) e a terceira de (5.7.24). Como o número $\varepsilon > 0$ foi escolhido arbitrariamente, segue a equação (5.7.9) de Teorema 5.7.5. Com isso o Teorema 5.7.11 fica provado. \square

Exemplo 5.7.15. A transformação de Jacobi $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ é definida através de

$$\phi(u, v) := (u - uv, uv)$$

para $(u, v) \in \mathbb{R}^2$. Denotamos $x := u - uv$ e $y := uv$, então resulta para a aplicação inversa a fórmula $u = x + y$ e $v = y/(x + y)$. Ou seja, a transformação ϕ leva o conjunto aberto $U := \{(u, v) \in \mathbb{R}^2 \mid u \neq 0\}$ difeomorfo ao conjunto $V := \{(x, y) \in \mathbb{R}^2 \mid x + y \neq 0\}$ e tem a derivada

$$d\phi(u, v) = \begin{pmatrix} 1 - v & -u \\ v & u \end{pmatrix}, \quad \det(d\phi(u, v)) = u.$$

Por isso as hipóteses de Teorema 5.7.11 com os conjuntos $A_R := [0, R] \times [0, 1]$, $N := \{0\} \times [0, 1]$, e $\Delta_R := \phi(A_R) = \{(x, y) \in \mathbb{R}^2 \mid x, y \geq 0, x + y \leq R\}$ são satisfeitas, e dali segue a equação

$$\int_{\Delta_R} f(x, y) dx dy = \int_0^R \int_0^1 f(u - uv, uv) u dv du \quad (5.7.25)$$

para toda função contínua $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Em equação (5.7.25) também é interessante considerar o valor limite para $R \rightarrow \infty$, caso exista. Tais valores limites são o tema da seguinte seção.

5.8 Integrais impróprias

Definição 5.8.1. *Seja $U \subset \mathbb{R}^n$ aberto. Uma função $f : U \rightarrow \mathbb{R}$ chama-se de localmente Riemann-integrável, se a restrição $f|_B : B \rightarrow \mathbb{R}$ é Riemann-integrável para todo conjunto Jordan-mensurável compacto $B \subset U$. Usamos as notações*

$$\begin{aligned}\mathcal{B}(U) &:= \{B \subset U \mid B \text{ é compacto e Jordan-mensurável}\} \\ \mathcal{R}_{\text{loc}}(U) &:= \{f : U \rightarrow \mathbb{R} \mid f \text{ é localmente Riemann-integrável}\}.\end{aligned}$$

Definição 5.8.2. *Seja $U \subset \mathbb{R}^n$ aberto. Uma função $f : U \rightarrow \mathbb{R}$ chama-se de Riemann-integrável (imprópria), se é localmente Riemann-integrável e*

$$\sup_{B \in \mathcal{B}(U)} \int_B |f(x)| dx < \infty. \quad (5.8.1)$$

Exercício 5.8.3. *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto. Então existe uma seqüência $B_i \in \mathcal{B}(U)$, tal que*

$$B_i \subset \overset{\circ}{B}_{i+1} \text{ para todos os } i \in \mathbb{N}, \quad U = \bigcup_{i=1}^{\infty} B_i. \quad (5.8.2)$$

Teorema 5.8.4 (A integral imprópria).

Seja $U \subset \mathbb{R}^n$ aberto e seja $f : U \rightarrow \mathbb{R}$ uma função impropriamente Riemann-integrável. Então existe exatamente um número real $c \in \mathbb{R}$, o qual satisfaz as seguintes duas condições.

- (i) *Se a seqüência $B_i \in \mathcal{B}(U)$ satisfaça (5.8.2), então $c = \lim_{i \rightarrow \infty} \int_{B_i} f(x) dx$.*
- (ii) *Para todo $\varepsilon > 0$ existe um $B_0 \in \mathcal{B}(U)$ tal que todo conjunto compacto $B \in \mathcal{B}(U)$ com $B_0 \subset B \subset U$ satisfaz a desigualdade $|c - \int_B f(x) dx| < \varepsilon$.*

Demonstração. Veja a página 221. □

Definição 5.8.5 (A integral imprópria). *Sejam U e f como em Teorema 5.8.4. Então chama-se o número $c \in \mathbb{R}$ em Teorema 5.8.4 a integral Riemann (imprópria) de f sobre U e denota-se de $\int_U f$. Ou seja, vale*

$$\int_U f := \int_U f(x) dx := \lim_{i \rightarrow \infty} \int_{B_i} f(x) dx \quad (5.8.3)$$

para toda seqüência $B_i \in \mathcal{B}(U)$ a qual satisfaz as condições em (5.8.2).

Demonstração do Teorema 5.8.4. A prova tem cinco passos.

Passo 1. Seja $B_i \in \mathcal{B}(U)$ uma seqüência de subconjuntos Jordan-mensuráveis compactos de U a qual satisfaça (5.8.2). Seja $B \subset U$ um subconjunto compacto qualquer. Então existe um número $i \in \mathbb{N}$, tal que $B \subset B_i$.

Argumentamos indireto e suponhamos que $B \not\subset B_i$ para todos os $i \in \mathbb{N}$. Então existe segundo o axioma da escolha, versão contável, uma seqüência $(x_i)_{i \in \mathbb{N}}$, tal que

$$x_i \in B \setminus B_i \quad \text{para todos os } i \in \mathbb{N}. \quad (5.8.4)$$

Como B é compacto existe uma subseqüência $(x_{i_\nu})_{\nu \in \mathbb{N}}$ com limite

$$x_0 := \lim_{\nu \rightarrow \infty} x_{i_\nu} \in B.$$

Como B é um subconjunto de U , vale $x_0 \in U$ e, como $U = \bigcup_{i=1}^{\infty} B_i$ segundo (5.8.2), existe consequentemente um número $i_0 \in \mathbb{N}$ com

$$x_0 \in B_{i_0} \subset \overset{\circ}{B}_{i_0+1}.$$

Como a seqüência $(x_{i_\nu})_{\nu \in \mathbb{N}}$ converge para x_0 e o conjunto $\overset{\circ}{B}_{i_0+1}$ é aberto, existe um número $\nu_0 \in \mathbb{N}$, tal que para todos os $\nu \in \mathbb{N}$ vale o seguinte:

$$\nu \geq \nu_0 \quad \implies \quad x_\nu \in \overset{\circ}{B}_{i_0+1}.$$

Se escolhermos o número $\nu \in \mathbb{N}$ tão grande que $\nu \geq \nu_0$ e $i_\nu > i_0$, então vale

$$x_{i_\nu} \in \overset{\circ}{B}_{i_0+1} \subset B_{i_0+1} \subset B_{i_\nu}.$$

Isto está na contradição a (5.8.4) e com isso o Passo 1 fica provado.

Passo 2. Seja $B_i \in \mathcal{B}(U)$ uma seqüência como em Passo 1 a qual satisfaça (5.8.2). Então vale $\lim_{i \rightarrow \infty} \int_{B_i} |f| = \sup_{B \in \mathcal{B}(U)} \int_K |f|$.

Seja $C := \sup_{B \in \mathcal{B}(U)} \int_K |f|$ e seja dada uma constante $\varepsilon > 0$. Então existe um conjunto $B \in \mathcal{B}(U)$ com

$$\int_B |f| > C - \varepsilon.$$

Segundo o Passo 1 existe um número $i_0 \in \mathbb{N}$ com $B \subset B_{i_0}$. Se $i \in \mathbb{N}$ com $i \geq i_0$ então vale $B \subset B_{i_0} \subset B_i$ e daí

$$C - \varepsilon < \int_B |f| \leq \int_{B_i} |f| \leq C.$$

Com isso o Passo 2 fica provado.

Passo 3. Seja $B_i \in \mathcal{B}(U)$ uma sequência como em Passo 1 a qual satisfaça (5.8.2). Então $c_i := \int_{B_i} f(x) dx$ é uma sequência de Cauchy.

Seja dado $\varepsilon > 0$. Então existe segundo o Passo 2 um número natural $i_0 \in \mathbb{N}$ com $\int_{B_{i_0}} |f| > C - \varepsilon$. Ora, sejam dados $i, j \in \mathbb{N}$ com $j \geq i \geq i_0$. Então vale

$$|c_i - c_j| = \left| \int_{B_j \setminus B_i} f \right| \leq \int_{B_j \setminus B_i} |f| = \int_{B_j} |f| - \int_{B_i} |f| \leq C - \int_{B_{i_0}} |f| < \varepsilon.$$

Com isso o Passo 3 fica provado.

Passo 4. Seja $B_i \in \mathcal{B}(U)$ uma sequência como em Passo 1 a qual satisfaça (5.8.2). Então existe o limite

$$c := \lim_{i \rightarrow \infty} \int_{B_i} f(x) dx \quad (5.8.5)$$

e satisfaz a condição (ii) em Teorema 5.8.4.

O limite existe segundo o Passo 3. Sejam dados $C := \sup_{B \in \mathcal{B}(U)} \int_B |f|$ e $\varepsilon > 0$. Então existe, segundo (5.8.5) e Passo 2, um número $i_0 \in \mathbb{N}$ com

$$\left| c - \int_{B_{i_0}} f \right| < \frac{\varepsilon}{2}, \quad \int_{B_{i_0}} |f| > C - \frac{\varepsilon}{2}.$$

Seja $B \in \mathcal{B}(U)$ com $B_{i_0} \subset B$. Então vale

$$\left| \int_{B \setminus B_{i_0}} f \right| \leq \int_{B \setminus B_{i_0}} |f| = \int_B |f| - \int_{B_{i_0}} |f| \leq C - \int_{B_{i_0}} |f| < \frac{\varepsilon}{2}$$

e daí

$$\left| c - \int_B f \right| = \left| c - \int_{B_{i_0}} f - \int_{B \setminus B_{i_0}} f \right| \leq \left| c - \int_{B_{i_0}} f \right| + \left| \int_{B \setminus B_{i_0}} f \right| < \varepsilon.$$

Com isso o Passo 4 fica provado.

Passo 5. Seja $c \in \mathbb{R}$ uma constante satisfazendo a condição (ii) em Teorema 5.8.4. Então c também satisfaz condição (i) em Teorema 5.8.4.

Como em Passo 1 seja $B_i \in \mathcal{B}(U)$ uma sequência de subconjuntos Jordan-mensuráveis compactos de U a qual satisfaz (5.8.2), e seja $\varepsilon > 0$. Então existe, segundo (ii), um $B_0 \in \mathcal{B}(U)$, tal que todo $B \in \mathcal{B}(U)$ com $B_0 \subset B$ satisfaz a desigualdade $|c - \int_B f| < \varepsilon$. Ora existe segundo o Passo 1 um $i_0 \in \mathbb{N}$ com $B_0 \subset B_{i_0}$. Se $i \in \mathbb{N}$ com $i \geq i_0$, então vale $B_0 \subset B_{i_0} \subset B_i$, e consequentemente $|c - \int_{B_i} f| < \varepsilon$. Com isso temos mostrado que $c = \lim_{i \rightarrow \infty} \int_{B_i} f$. Com isso o Passo 5 e o Teorema 5.8.4 ficam provados. \square

Exemplos

Exemplo 5.8.6. Dado $p, q > 0$, a função $f : (0, \infty)^2 \rightarrow \mathbb{R}$ é dada através de

$$f(x, y) := x^{p-1}y^{q-1}e^{-x-y}$$

para $x, y > 0$. Para $R > 1$ o subconjunto $K_R := [R^{-1}, R]^2$ de $U := (0, \infty)^2$ é compacto e segundo o Teorema de Fubini (Teorema 5.4.1) vale

$$\int_{K_R} f(x, y) dx dy = \left(\int_{1/R} x^{p-1} e^{-x} dx \right) \left(\int_{1/R} y^{q-1} e^{-y} dy \right).$$

Segue daí, com a transição ao limite $R \rightarrow \infty$, que a função $f : U \rightarrow \mathbb{R}$ é Riemann-integrável imprópria e que sua integral é dada assim

$$\int_U f(x, y) dx dy = \Gamma(p)\Gamma(q). \quad (5.8.6)$$

(Aqui $\Gamma(p) := \int_0^\infty x^{p-1} e^{-x} dx$ é a função-Gamma.) Por outro lado também pode-se escrever, segundo o Teorema 5.8.4, a integral de f como valor limite para $R \rightarrow \infty$ das integrais sobre os conjuntos $\Delta_R = \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, x + y < R\}$. Segundo o Exemplo 5.7.15 segue dali

$$\begin{aligned} \int_U f(x, y) dx dy &= \lim_{R \rightarrow \infty} \int_{\Delta_R} f(x, y) dx dy \\ &= \lim_{R \rightarrow \infty} \int_0^R \int_0^1 f(u - uv, uv) u dv du \\ &= \lim_{R \rightarrow \infty} \int_0^R \int_0^1 (u - uv)^{p-1} (uv)^{q-1} u e^{-u} dv du \\ &= \lim_{R \rightarrow \infty} \int_0^R \int_0^1 (1 - v)^{p-1} v^{q-1} u^{p+q-1} e^{-u} dv du \\ &= \int_0^1 (1 - v)^{p-1} v^{q-1} dv \lim_{R \rightarrow \infty} \int_0^R u^{p+q-1} e^{-u} du \\ &= \int_0^1 (1 - v)^{p-1} v^{q-1} dv \Gamma(p + q). \end{aligned}$$

Disso resulta, segundo (5.8.6), a fórmula

$$\int_0^1 (1 - t)^{p-1} t^{q-1} dt = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}. \quad (5.8.7)$$

Exemplo 5.8.7. Repetimos o argumento de Exemplo 5.7.15 para $R = 1$ com os conjuntos abertos $(0, 1)^2$ e

$$\Delta := \{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, x + y < 1\} \quad (5.8.8)$$

e com a função impropriamente Riemann-integrável $f : \Delta \rightarrow \mathbb{R}$, a qual é dada através de $f(x, y) = x^{p-1}y^{q-1}$ com $p, q > 0$. Então resulta para a integral imprópria a fórmula

$$\begin{aligned} \int_{\Delta} x^{p-1}y^{q-1} dx dy &= \int_0^1 \int_0^1 (u - uv)^{p-1}(uv)^{q-1}u dv du \\ &= \left(\int_0^1 (1 - v)^{p-1}v^{q-1} dv \right) \left(\int_0^1 u^{p+q-1} du \right) \\ &= \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \frac{1}{p+q} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q+1)}. \end{aligned} \quad (5.8.9)$$

Aqui segue o primeiro passo da equação (5.7.25) em Exemplo 5.7.15 e o terceiro passo da equação (5.8.7) em Exemplo 5.8.6.

Exemplo 5.8.8. Sejam $a, b, \alpha, \beta, p, q > 0$. Provaremos a fórmula

$$\begin{aligned} \int_{\Delta_{a,b}^{\alpha,\beta}} x^{p-1}y^{q-1} dx dy &= \frac{a^p b^q}{\alpha\beta} \frac{\Gamma\left(\frac{p}{\alpha}\right)\Gamma\left(\frac{q}{\beta}\right)}{\Gamma\left(\frac{p}{\alpha} + \frac{q}{\beta} + 1\right)} \\ \Delta_{a,b}^{\alpha,\beta} &:= \left\{ (x, y) \in \mathbb{R}^2 \mid x > 0, y > 0, \left(\frac{x}{a}\right)^{\alpha} + \left(\frac{y}{b}\right)^{\beta} < 1 \right\}. \end{aligned} \quad (5.8.10)$$

A aplicação $\Delta \rightarrow \Delta_{a,b}^{\alpha,\beta} : (\xi, \eta) \mapsto \phi(\xi, \eta) := (a\xi^{1/\alpha}, b\eta^{1/\beta})$ é um difeomorfismo com a derivada

$$d\phi(\xi, \eta) = \begin{pmatrix} \frac{a}{\alpha}\xi^{\frac{1}{\alpha}-1} & 0 \\ 0 & \frac{b}{\beta}\eta^{\frac{1}{\beta}-1} \end{pmatrix}.$$

Então segue da fórmula da transformação de variáveis em Teorema 5.7.11, com ajuda das transições aos limites para integrais impróprias, a equação

$$\int_{\Delta_{a,b}^{\alpha,\beta}} f(x, y) dx = \frac{ab}{\alpha\beta} \int_{\Delta} f(a\xi^{1/\alpha}, b\eta^{1/\beta}) \xi^{\frac{1}{\alpha}-1} \eta^{\frac{1}{\beta}-1} d\xi d\eta.$$

Com $f(x, y) = x^{p-1}y^{q-1}$ encontraremos

$$\int_{\Delta_{a,b}^{\alpha,\beta}} x^{p-1}y^{q-1} dx = \frac{a^p b^q}{\alpha\beta} \int_{\Delta} \xi^{\frac{p}{\alpha}-1} \eta^{\frac{q}{\beta}-1} d\xi d\eta.$$

Com isso segue (5.8.10) de equação (5.8.9) em Exemplo 5.8.7.

Exemplo 5.8.9. Seja $\|\cdot\|$ a norma euclidiana no \mathbb{R}^n e seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a função $f(x) := e^{-\|x\|^2} = e^{-x_1^2 - \dots - x_n^2}$ para $x \in \mathbb{R}^n$. Esta função é impropriamente Riemann-integrável. Provaremos a fórmula

$$\int_{\mathbb{R}^n} e^{-\|x\|^2} dx = \pi^{n/2}. \quad (5.8.11)$$

Primeiro vale, segundo do Teorema de Fubini, para todo $n \in \mathbb{N}$ a equação

$$\int_{\mathbb{R}^n} e^{-\|x\|^2} dx_1 \cdots dx_n = \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^n. \quad (5.8.12)$$

(Aplique Teorema 5.4.1 à integral sobre $[-R, R]^n$ e considere o valor limite $R \rightarrow \infty$.) Seja $\phi : (0, \infty) \times (-\pi, \pi) \rightarrow \mathbb{R}^2 \setminus ((-\infty, 0] \times \{0\})$ o difeomorfismo definido através de $\phi(r, \theta) := (r \cos(\theta), r \sin(\theta))$ (coordenadas polares para $n = 2$). Então vale $\det(d\phi(r, \theta)) = r$ e daí

$$\int_{\mathbb{R}^2} f(x, y) dx dy = \int_0^{\infty} \int_{-\pi}^{\pi} f(r \cos(\theta), r \sin(\theta)) r d\theta dr$$

para toda função $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ impropriamente Riemann-integrável. Para a função $f(x, y) := e^{-x^2 - y^2}$ encontraremos com ajuda de (5.8.12) a fórmula

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 = \int_0^{\infty} \int_{-\pi}^{\pi} e^{-r^2} r d\theta dr = 2\pi \int_0^{\infty} e^{-r^2} r dr = \pi.$$

Aqui segue o último passo da fórmula de substituição para integrais numa variável com $s := r^2$. Disso resulta a fórmula $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$ e daí segue a equação (5.8.11) de (5.8.12).

Exemplo 5.8.10. Seja $A = A^T \in \mathbb{R}^{n \times n}$ uma matriz simétrica positiva definida. Então vale

$$\int_{\mathbb{R}^n} e^{-x^T A x} dx = \frac{\pi^{n/2}}{\sqrt{\det(A)}}. \quad (5.8.13)$$

Para a prova utilizamos o fato, conhecido da álgebra linear, que existe uma (única) matriz simétrica positiva definida $B = B^T \in \mathbb{R}^{n \times n}$ com $B^2 = A$. Seja $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ o difeomorfismo $\phi(x) := Bx$. Então resulta de Exemplo 5.8.9 junto com a fórmula da transformação em Teorema 5.7.11 a equação

$$\pi^{n/2} = \int_{\mathbb{R}^n} e^{-\|y\|^2} dy = \int_{\mathbb{R}^n} e^{-\|Bx\|^2} \det(B) dx = \int_{\mathbb{R}^n} e^{-x^T A x} dx \sqrt{\det(A)}.$$

5.9 Divergência

Nesta última seção do capítulo vamos relacionar a fórmula da transformação de variáveis em Teorema 5.7.11 com o fluxo de um campo de vetores continuamente diferenciável em Capítulo 4. Para isto começamos com a seguinte observação. Sejam $U, V \subset \mathbb{R}^n$ dois conjuntos abertos e seja $\phi : U \rightarrow V$ um C^1 -difeomorfismo. Se B é um subconjunto compacto Jordan-mensurável de U , então $\phi(B)$ é um subconjunto compacto Jordan-mensurável de V . Que $\phi(B)$ é compacto, segue da continuidade de ϕ . E que o bordo $\partial\phi(B) = \phi(\partial B)$ é um conjunto Jordan-nulo, segue da parte (ii) de Teorema 5.6.7. Com isso $\phi(B)$ é Jordan-mensurável, como afirmado, e tem, segundo da fórmula da transformação de variáveis (5.7.9) em Teorema 5.7.11 com $f \equiv 1$, o volume

$$\text{Vol}_n(\phi(B)) = \int_B |\det(d\phi(x))| dx. \quad (5.9.1)$$

Definição 5.9.1. *Sejam $U, V \subset \mathbb{R}^n$ conjuntos abertos. Diz-se que um C^1 -difeomorfismo $\phi : U \rightarrow V$ **preserva volume** se todo subconjunto Jordan-mensurável compacto $B \subset U$ satisfaz a equação $\text{Vol}_n(\phi(B)) = \text{Vol}_n(B)$.*

Lema 5.9.2. *Sejam $U, V \subset \mathbb{R}^n$ conjuntos abertos e seja $\phi : U \rightarrow V$ um C^1 -difeomorfismo. Então as seguintes afirmações são equivalentes.*

- (i) ϕ preserva volume.
- (ii) Vale $|\det(d\phi(x))| = 1$ para todos os $x \in U$.

Demonstração. A implicação (ii) \implies (i) segue da equação (5.9.1), a qual em contrapartida segue direto da fórmula da transformação (5.7.9) em Teorema 5.7.11. Para a prova da implicação reversa suponhamos que ϕ preserve volume e escolhemos um elemento $a = (a_1, \dots, a_n) \in U$. Para $\varepsilon > 0$ seja

$$Q_\varepsilon := [a_1 - \varepsilon, a_1 + \varepsilon] \times \cdots \times [a_n - \varepsilon, a_n + \varepsilon].$$

Este conjunto é Jordan-mensurável e tem o volume $(2\varepsilon)^n$. Para $\varepsilon > 0$ suficientemente pequeno vale $Q_\varepsilon \subset U$ e dali, segundo (i) e (5.9.1),

$$1 = \frac{\text{Vol}_n(\phi(Q_\varepsilon))}{\text{Vol}_n(Q_\varepsilon)} = \frac{\text{Vol}_n(\phi(Q_\varepsilon))}{(2\varepsilon)^n} = \frac{1}{(2\varepsilon)^n} \int_{Q_\varepsilon} |\det(d\phi(x))| dx.$$

Com a transição ao limite $\varepsilon \rightarrow 0$ segue daí

$$|\det(d\phi(a))| = \lim_{\varepsilon \rightarrow 0} \frac{1}{(2\varepsilon)^n} \int_{Q_\varepsilon} |\det(d\phi(x))| dx = 1.$$

Portanto ϕ satisfaz a condição (ii) e com isso o Lema 5.9.2 fica provado. \square

Exemplo 5.9.3. Seja $\phi : \mathbb{R}^2 \rightarrow \mathbb{R} \times (0, \infty)$ o difeomorfismo definido por

$$\phi(x, y) := (e^{-y}x, e^y)$$

para $(x, y) \in \mathbb{R}^2$. A aplicação inversa é dada através de

$$\phi^{-1}(\xi, \eta) = (\xi\eta, \log(\eta))$$

para $\xi \in \mathbb{R}$ e $\eta > 0$. A matriz jacobiana

$$d\phi(x, y) = \begin{pmatrix} e^{-y} & -e^{-y}x \\ 0 & e^y \end{pmatrix}$$

tem determinante 1 para todos os $(x, y) \in \mathbb{R}^2$ e dali ϕ preserva volume.

Ora nós dedicamos à questão nas quais hipóteses o fluxo de um campo de vetores continuamente diferenciável é composto de difeomorfismos preservando volume. Seja então $U \subset \mathbb{R}^n$ um conjunto aberto, seja $f : U \rightarrow \mathbb{R}^n$ um campo de vetores continuamente diferenciável, e seja $\phi : \Omega \rightarrow U$ o fluxo deste campo de vetores. Isto significa que $\Omega \subset \mathbb{R} \times U$ é o conjunto de todos os pares (t, x_0) com $x_0 \in U$ e $t \in I(x_0)$, onde $I(x_0) \subset \mathbb{R}$ é o intervalo máximo de existência para as soluções do problema de valor inicial

$$\dot{x} = f(x), \quad x(0) = x_0, \quad (5.9.2)$$

e que a função $I(x_0) \rightarrow U : t \mapsto \phi(t, x_0)$ para todo $x_0 \in U$ é exatamente a única solução deste problema de valor inicial (veja Definição 4.2.2). Segundo o Teorema 4.3.1 o subconjunto $\Omega \subset \mathbb{R} \times U$ é aberto e segundo o Teorema 4.4.1 a aplicação $\phi : \Omega \rightarrow U$ é continuamente diferenciável. Com

$$U_t := \{x \in U \mid t \in I(x)\}$$

e $\phi_t(x) := \phi(t, x)$ para $x \in U_t$ segue dali que a aplicação

$$\phi_t : U_t \rightarrow U_{-t}$$

é, para todo $t \in \mathbb{R}$, um C^1 -difeomorfismo entre subconjuntos abertos do \mathbb{R}^n (Comentário 4.3.3). Ora a questão é nas quais hipóteses ϕ_t preserva volume. Para isso lembramos que para $x_0 \in U$ e $t \in I(x_0)$ a matriz jacobina $\Phi(t) := d\phi_t(x_0)$ é a solução do problema de valor inicial

$$\dot{\Phi}(t) = A(t)\Phi(t), \quad \Phi(0) = \mathbb{1}, \quad (5.9.3)$$

com $A(t) := df(\phi_t(x_0))$ (Lema 4.4.2).

Lema 5.9.4. *Seja $I \subset \mathbb{R}$ um intervalo aberto com $0 \in I$, seja $A : I \rightarrow \mathbb{R}^{n \times n}$ uma função contínua, e seja $\Phi : I \rightarrow \text{GL}(n, \mathbb{R})$ a solução de (5.9.3). Então vale a identidade*

$$\frac{d}{dt} \det(\Phi(t)) = \text{tr}(A(t)) \det(\Phi(t)) \quad (5.9.4)$$

para todos os $t \in I$.

Demonstração. A função $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ é um polinômio homogêneo do grau n em n^2 variáveis e por isso é suave (Teorema D.3.3). Seja

$$\rho := \det|_{\text{GL}(n, \mathbb{R})} : \text{GL}(n, \mathbb{R}) \rightarrow \mathbb{R}$$

a restrição do determinante ao grupo linear geral (um subconjunto aberto de $\mathbb{R}^{n \times n}$). Provaremos que a derivada de ρ no ponto $\Phi \in \text{GL}(n, \mathbb{R})$ é a aplicação linear $d\rho(\Phi) : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ dada através de

$$d\rho(\Phi)\widehat{\Phi} = \text{tr}(\Phi^{-1}\widehat{\Phi}) \det(\Phi) \quad (5.9.5)$$

para $\widehat{\Phi} \in \mathbb{R}^{n \times n}$. Com esse fim consideramos para já o caso $\Phi = \mathbb{1}$. Então vale para $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ a equação $\frac{d}{dt}|_{t=0} e^{tA} = A$ (Teorema 1.5.2) e dali

$$\begin{aligned} d\rho(\mathbb{1})A &= \left. \frac{d}{dt} \right|_{t=0} \rho(e^{tA}) \\ &= \left. \frac{d}{dt} \right|_{t=0} \det(e^{tA}e_1, \dots, e^{tA}e_n) \\ &= \sum_{i=1}^n \det(e_1, \dots, e_{i-1}, Ae_i, e_{i+1}, \dots, e_n) \\ &= \sum_{i=1}^n \det(e_1, \dots, e_{i-1}, \sum_{j=1}^n a_{ji}e_j, e_{i+1}, \dots, e_n) \\ &= \sum_{i=1}^n a_{ii} \det(e_1, \dots, e_n) \\ &= \text{tr}(A). \end{aligned}$$

Aqui e_1, \dots, e_n denota a base padrão do \mathbb{R}^n e temos usado o fato que a aplicação determinante $\Phi \mapsto \det(\Phi)$ é linear em cada uma coluna de Φ e se anula logo que duas colunas de Φ coincidem. Com isso a equação (5.9.5) para $\Phi = \mathbb{1}$ fica provada.

Ora, seja $\Phi \in \text{GL}(n, \mathbb{R})$ e $\widehat{\Phi} \in \mathbb{R}^{n \times n}$ e defina

$$A := \Phi^{-1}\widehat{\Phi}.$$

Então vale

$$\left. \frac{d}{dt} \right|_{t=0} \Phi e^{tA} = \Phi A = \widehat{\Phi}$$

e dali

$$\begin{aligned} d\rho(\Phi)\widehat{\Phi} &= \left. \frac{d}{dt} \right|_{t=0} \rho(\Phi e^{tA}) \\ &= \left. \frac{d}{dt} \right|_{t=0} \rho(\Phi)\rho(e^{tA}) \\ &= \rho(\Phi)d\rho(\mathbb{1})A \\ &= \det(\Phi)\text{tr}(A). \end{aligned}$$

Com isso a equação (5.9.5) fica provada, e dali segue para a solução

$$I \rightarrow \text{GL}(n, \mathbb{R}) : t \mapsto \Phi(t)$$

da equação (5.9.3), segundo a regra da cadeia, que

$$\begin{aligned} \frac{d}{dt} \det(\Phi(t)) &= d\rho(\Phi(t))\dot{\Phi}(t) \\ &= \text{tr}(\Phi(t)^{-1}\dot{\Phi}(t)) \det(\Phi(t)) \\ &= \text{tr}(A(t)) \det(\Phi(t)). \end{aligned}$$

Com isso o Lema 5.9.4 fica provado. \square

Definição 5.9.5 (Divergência). *Seja $U \subset \mathbb{R}^n$ um subconjunto aberto e seja $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ um campo de vetores continuamente diferenciável. A **divergência** de f é a função $\text{div}(f) : U \rightarrow \mathbb{R}$ definida através de*

$$(\text{div}(f))(x) := \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) = \text{tr}(df(x)) \quad (5.9.6)$$

para $x \in U$. Diz-se que o campo de vetores f tem **divergência nula** se a divergência de f anula-se em todo o conjunto U completo, ou seja, se $\text{tr}(df(x)) = 0$ para todos os $x \in U$.

Teorema 5.9.6. *Seja $U \subset \mathbb{R}^n$ um conjunto aberto, seja $f : U \rightarrow \mathbb{R}^n$ um campo de vetores continuamente diferenciável, e seja $\phi_t : U_t \rightarrow U_{-t}$ o fluxo de f . Então as seguintes afirmações são equivalentes.*

- (i) *O campo de vetores f tem divergência nula.*
- (ii) *O C^1 -difeomorfismo $\phi_t : U_t \rightarrow U_{-t}$ preserva volume para todos os $t \in \mathbb{R}$.*

Demonstração. Lema 5.9.4 mostra que o determinante da matriz jacobiana de ϕ_t satisfaz a equação diferencial

$$\begin{aligned} \frac{d}{dt} \det(d\phi_t(x_0)) &= \operatorname{tr}(df(\phi_t(x_0))) \det(d\phi_t(x_0)), \\ \det(d\phi_0(x_0)) &= 1. \end{aligned} \quad (5.9.7)$$

De isto segue que o campo de vetores f tem divergência nula exatamente se o fluxo satisfaz a condição $\det(d\phi_t(x_0)) = 1$ para todos os $x_0 \in U$ e todos os $t \in I(x_0)$. Como o determinante de $d\phi_t(x_0)$ nunca pode ser negativo, esta condição é equivalente, segundo o Lema 5.9.2, que $\phi_t : U_t \rightarrow U_{-t}$ preserva volume para todo $t \in \mathbb{R}$. Com isso o Teorema 5.9.6 fica provado. \square

Exemplo 5.9.7. Exemplos simples de campos de vetores no \mathbb{R}^2 com divergência nula são, primeiro, o campo de vetores

$$f(x, y) := (x, -y) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

com uma **singularidade** (um ponto onde um campo de vetores anula-se) hiperbólica na origem e com o fluxo

$$\phi_t(x, y) = (e^t x, e^{-t} y)$$

e, segundo, o campo de vetores que vira um vetor (x, y) por 90 graus

$$f(x, y) := (-y, x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

com uma singularidade elíptica na origem e com o fluxo

$$\phi_t(x, y) = (\cos(t)x - \operatorname{sen}(t)y, \operatorname{sen}(t)x + \cos(t)y).$$

Capítulo 6

Integração ao longo de Subvariedades

O objetivo deste capítulo é definir a integral de Riemann de uma função ao longo de uma subvariedade d -dimensional $M \subset \mathbb{R}^n$ e provar o Teorema de Gauß. O capítulo começa com uma introdução ao volume d -dimensional e à integral sobre um domínio de carta (Seção 6.1). Sequentemente será definida a integral sobre M inteiro (Seção 6.2), discutidos uns exemplos (Seção 6.3), e provado o Teorema de Gauß (Seção 6.4).

6.1 O volume d -dimensional

O comprimento de uma curva

Esta seção começa com a introdução da noção de *comprimento de uma curva* no \mathbb{R}^n . Trabalhamos sempre com o produto euclidiano

$$\langle x, y \rangle := \sum_{i=1}^n x_i y_i, \quad \|x\| := \sqrt{\langle x, x \rangle}$$

para $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ e $y = (y_1, \dots, y_n) \in \mathbb{R}^n$. Seja $I = [a, b]$ im intervalo compacto. Uma **partição** de I é uma $(N + 1)$ -tuplo de números reais $P = (t_0, t_1, \dots, t_N)$ com $N \in \mathbb{N}$ e

$$a = t_0 < t_1 < \dots < t_{N-1} < t_N = b. \quad (6.1.1)$$

(Veja [5].) O conjunto de todas as partições de I é denotado de $\mathcal{P}(I)$.

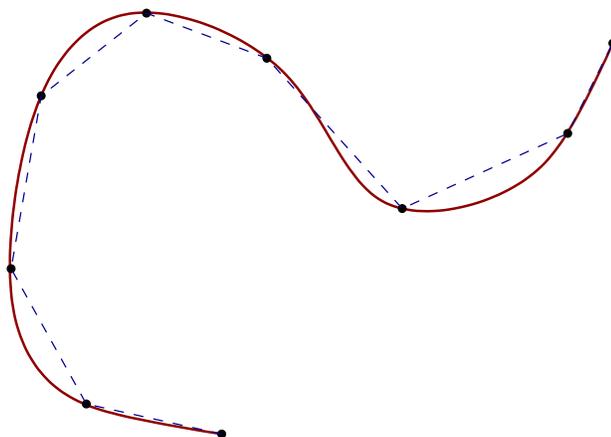


Figura 6.1: O comprimento de uma curva.

Definição 6.1.1 (O comprimento).

Sejam $a < b$ números reais, seja $I := [a, b]$, e seja $\gamma : I \rightarrow \mathbb{R}^n$ uma aplicação contínua. Para uma partição $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ definimos

$$L(\gamma, P) := \sum_{i=1}^N \|\gamma(t_i) - \gamma(t_{i-1})\|. \quad (6.1.2)$$

O comprimento de γ é o número

$$L(\gamma) := \sup_{P \in \mathcal{P}(I)} L(\gamma, P) \in [0, \infty].$$

A curva γ chama-se de **rectificável** se $L(\gamma) < \infty$.

Exemplo 6.1.2. Seja $\gamma : [0, 1] \rightarrow \mathbb{R}$ a função a qual se anula no conjunto de Cantor

$$K := \bigcap_{n=1}^{\infty} K_n, \quad K_n := \bigcup_{a_1, \dots, a_n \in \{0,1\}} \left[2 \sum_{i=1}^n \frac{a_i}{3^i}, 2 \sum_{i=1}^n \frac{a_i}{3^i} + \frac{1}{3^n} \right]$$

e é dada, para $\forall n \in \mathbb{N}$ e $\forall a_1, \dots, a_n \in \{0, 1\}$ com $a_n = 1$, através da fórmula

$$\gamma(t) := \frac{3^n}{2^{n-2}} \min \left\{ \left(t + \frac{1}{3^n} - 2 \sum_{i=1}^n \frac{a_i}{3^i} \right), \left(2 \sum_{i=1}^n \frac{a_i}{3^i} - t \right) \right\}$$

para $2 \sum_{i=1}^n \frac{a_i}{3^i} - \frac{1}{3^n} \leq t \leq 2 \sum_{i=1}^n \frac{a_i}{3^i}$. Esta curva $\gamma : [0, 1] \rightarrow \mathbb{R}$ é contínua, mas não rectificável (veja a Figura 6.2).

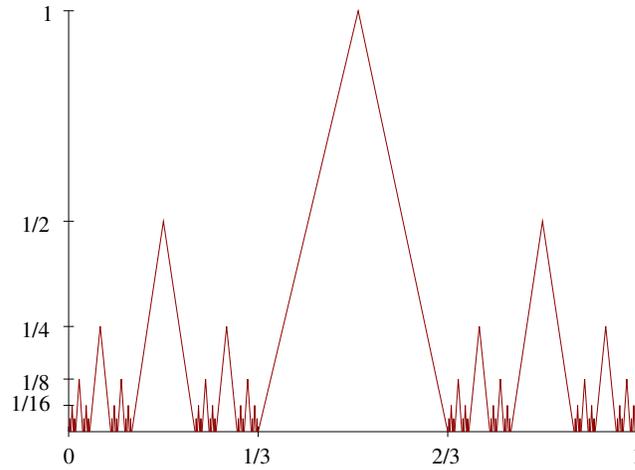


Figura 6.2: Uma curva não rectificável.

Lema 6.1.3. *Seja $I = [a, b]$ um intervalo compacto e seja $\gamma : I \rightarrow \mathbb{R}^n$ uma aplicação continuamente diferenciável. Então γ é rectificável e vale*

$$L(\gamma) = \int_a^b \|\dot{\gamma}(t)\| dt. \quad (6.1.3)$$

Demonstração. A demonstração tem dois passos.

Passo 1. *A curva γ é rectificável e vale $L(\gamma) \leq \int_a^b \|\dot{\gamma}(t)\| dt$.*

Seja $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ uma partição. Então vale, segundo a parte (iii) de Lema 1.7.3, para todo $i \in \{1, \dots, N\}$ a desigualdade

$$\|\gamma(t_i) - \gamma(t_{i-1})\| = \left\| \int_{t_{i-1}}^{t_i} \dot{\gamma}(t) dt \right\| \leq \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t)\| dt$$

e disso segue

$$L(\gamma, P) = \sum_{i=1}^N \|\gamma(t_i) - \gamma(t_{i-1})\| \leq \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t)\| dt = \int_a^b \|\dot{\gamma}(t)\| dt.$$

Ora, formamos o supremo sobre todos os $P \in \mathcal{P}(I)$, então resulta

$$L(\gamma) = \sup_{P \in \mathcal{P}(I)} L(\gamma, P) \leq \int_a^b \|\dot{\gamma}(t)\| dt < \infty.$$

Com isso o Passo 1 fica provado.

Passo 2. O comprimento de γ é dado através de (6.1.3).

Seja dado $\varepsilon > 0$. Como $\dot{\gamma} : I \rightarrow \mathbb{R}^n$ é uniformemente contínuo, existe uma constante $\delta > 0$, tal que para todos os $s, t \in I$ vale:

$$|s - t| < \delta \quad \implies \quad \|\dot{\gamma}(s) - \dot{\gamma}(t)\| < \frac{\varepsilon}{b - a}.$$

Ora, seja $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ uma partição a qual satisfaça a desigualdade

$$0 < t_{i+1} - t_i < \delta$$

para $i = 1, \dots, N$. Então recebemos para todos os $i \in \{1, \dots, N\}$ e todos os tempos $t \in [t_{i-1}, t_i]$ a desigualdade

$$\begin{aligned} \left\| \dot{\gamma}(t) - \frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} \right\| &= \left\| \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} (\dot{\gamma}(t) - \dot{\gamma}(s)) \, ds \right\| \\ &\leq \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t) - \dot{\gamma}(s)\| \, ds \\ &\leq \frac{\varepsilon}{b - a}. \end{aligned}$$

Disso segue a desigualdade

$$\|\dot{\gamma}(t)\| \leq \frac{\|\gamma(t_i) - \gamma(t_{i-1})\|}{t_i - t_{i-1}} + \frac{\varepsilon}{b - a}$$

para $i = 1, \dots, N$ e $t_{i-1} \leq t \leq t_i$. Integrando esta desigualdade sobre o intervalo $[t_{i-1}, t_i]$ e formando a soma sobre $i = 1, \dots, N$, então recebemos

$$\begin{aligned} \int_a^b \|\dot{\gamma}(t)\| \, dt &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \|\dot{\gamma}(t)\| \, dt \\ &\leq \sum_{i=1}^N \left(\|\gamma(t_i) - \gamma(t_{i-1})\| + \varepsilon \frac{t_i - t_{i-1}}{b - a} \right) \\ &= L(\gamma, P) + \varepsilon \\ &\leq L(\gamma) + \varepsilon. \end{aligned}$$

Então vale $\int_a^b \|\dot{\gamma}(t)\| \, dt \leq L(\gamma) + \varepsilon$ para todos os $\varepsilon > 0$ e daí

$$\int_a^b \|\dot{\gamma}(t)\| \, dt \leq L(\gamma).$$

Segundo o Passo 1 segue dali $\int_a^b \|\dot{\gamma}(t)\| \, dt = L(\gamma)$. Com isso o Passo 2 e Lema 6.1.3 ficam provados. \square

Lema 6.1.4. *Sejam $I = [a, b]$ e $I' = [a', b']$ dois intervalos compactos. Seja $\phi : I' \rightarrow I$ um homeomorfismo e $\gamma : I \rightarrow \mathbb{R}^n$ uma aplicação contínua. Então*

$$L(\gamma \circ \phi) = L(\gamma). \quad (6.1.4)$$

Demonstração. Seja $P = (t_0, t_1, \dots, t_N) \in \mathcal{P}(I)$ uma partição a qual satisfaça (6.1.1). Definimos $\phi^*P := (\phi^{-1}(t_0), \phi^{-1}(t_1), \dots, \phi^{-1}(t_N))$ caso ϕ cresça estritamente monótono e $\phi^*P := (\phi^{-1}(t_N), \phi^{-1}(t_{N-1}), \dots, \phi^{-1}(t_0))$ caso ϕ decresça estritamente monótono. Assim temos $\phi^*P \in \mathcal{P}(I')$ e vale a igualdade $L(\gamma \circ \phi, \phi^*P) = L(\gamma, P)$. Como toda partição $P' \in \mathcal{P}(I')$ é da forma $P' = \phi^*P$ para um $P \in \mathcal{P}(I)$, segue daí

$$L(\gamma \circ \phi) = \sup_{P \in \mathcal{P}(I)} L(\gamma \circ \phi, \phi^*P) = \sup_{P \in \mathcal{P}(I)} L(\gamma, P) = L(\gamma).$$

Com isso Lema 6.1.4 fica provado. \square

O volume d -dimensional de um paralelepípedo

Sejam $a = (a_1, \dots, a_n)$ e $b = (b_1, \dots, b_n)$ dois vetores linearmente independente no \mathbb{R}^n . Eles geram um paralelepípedo, ou seja, o conjunto

$$P(a, b) := \{sa + tb \mid s, t \in \mathbb{R}, 0 \leq s, t \leq 1\},$$

e sua conteúdo de área é o número

$$v_2(P(a, b)) := \sqrt{\|a\|^2 \|b\|^2 - \langle a, b \rangle^2}. \quad (6.1.5)$$

Esta fórmula fica válida se a e b sejam linearmente dependente. Neste caso vale $\langle a, b \rangle = \pm \|a\| \|b\|$ e o conteúdo de área é $v_2(P(a, b)) = 0$. No geral consideramos, para um número real $0 \leq d \leq n$ e d vetores $a_1, \dots, a_d \in \mathbb{R}^n$, o paralelepípedo correspondente

$$P(a_1, \dots, a_d) := \left\{ \sum_{i=1}^d t_i a_i \mid t_i \in \mathbb{R}, 0 \leq t_i \leq 1 \text{ para } i = 1, \dots, d \right\} \quad (6.1.6)$$

e buscamos por uma fórmula para seu volume d -dimensional. Em outras palavras, buscamos por uma função $v_d : (\mathbb{R}^n)^d \rightarrow [0, \infty)$ a qual associa com toda d -tuplo de vetores $(a_1, \dots, a_d) \in (\mathbb{R}^n)^d$ o volume d -dimensional do paralelepípedo $P(a_1, \dots, a_d)$. Esta é caracterizada pelos axiomas (V1-3) no seguinte lema.

Lema 6.1.5. *Sejam $d, n \in \mathbb{N}$ com $1 \leq d \leq n$. Então existe exatamente uma função $v_d : (\mathbb{R}^n)^d \rightarrow [0, \infty)$ a qual satisfaz as seguintes três condições.*

(V1) *Para todos os $a_1, \dots, a_d \in \mathbb{R}^n$ e todos os $\lambda_1, \dots, \lambda_d \in \mathbb{R}$ vale*

$$v_d(\lambda_1 a_1, \dots, \lambda_d a_d) = \left(\prod_{i=1}^d |\lambda_i| \right) v_d(a_1, \dots, a_d). \quad (6.1.7)$$

(V2) *Para todos os $a_1, \dots, a_d \in \mathbb{R}^n$ e todos os $i, j \in \{1, \dots, d\}$ com $i \neq j$ vale*

$$v_d(a_1, \dots, a_{i-1}, a_i + a_j, a_{i+1}, \dots, a_d) = v_d(a_1, \dots, a_d). \quad (6.1.8)$$

(V3) *Para todos os $a_1, \dots, a_d \in \mathbb{R}^n$ vale*

$$\langle a_i, a_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{caso } i = j, \\ 0, & \text{caso } i \neq j \end{cases} \implies v_d(a_1, \dots, a_d) = 1. \quad (6.1.9)$$

A aplicação v_d é dada pela fórmula

$$v_d(a_1, \dots, a_d) = \sqrt{\det(A^T A)} \quad (6.1.10)$$

na qual $A := (a_1 \cdots a_d) \in \mathbb{R}^{n \times d}$ é a matriz com as colunas a_1, \dots, a_d , ou seja $A^T A = (\langle a_i, a_j \rangle)_{i,j=1}^d \in \mathbb{R}^{d \times d}$.

Demonstração. Para $d = 1$ a função $v_1 : \mathbb{R}^n \rightarrow [0, \infty)$ definida $v_1(a) := \|a\|$, para $a \in \mathbb{R}^n$, satisfaz os axiomas (V1) e (V3), as quais determinam v_1 unívoco. Neste caso o axioma (V2) é sem significância. Para $d = 2$ a fórmula (6.1.10) coincide com o conteúdo de área (6.1.5). No caso geral segue diretamente das propriedades do determinante que a função $v_d : (\mathbb{R}^n)^d \rightarrow [0, \infty)$ definida através de (6.1.10) satisfaz o axioma (V1) de mudar a escala, o axioma (V2) do cisalhamento, e o axioma (V3) da normalização, as quais determinam v_d unívoco. \square

Definição 6.1.6. *Para $a_1, \dots, a_d \in \mathbb{R}^n$ chama-se o número $v_d(a_1, \dots, a_d)$ em Lema 6.1.5 o **volume d -dimensional** do paralelepípedo $P(a_1, \dots, a_d)$ em (6.1.6) e usa-se o símbolo*

$$\text{Vol}_d(P(a_1, \dots, a_d)) := v_d(a_1, \dots, a_d) \quad (6.1.11)$$

para o volume d -dimensional.

Lema 6.1.7. *Seja $n \geq 2$ um número inteiro e sejam $a_1, \dots, a_{n-1} \in \mathbb{R}^n$. Defina $A = (a_1 \cdots a_{n-1}) \in \mathbb{R}^{n \times (n-1)}$ e para $k = 1, \dots, n$ seja $A_k \in \mathbb{R}^{(n-1) \times (n-1)}$ a matriz quadrada a qual resulta de A através de remover a k -ésima linha. Então vale*

$$v_d(a_1, \dots, a_{n-1}) = \sqrt{\sum_{k=1}^n \alpha_k^2}, \quad \alpha_k := (-1)^{k-1} \det(A_k). \quad (6.1.12)$$

Demonstração. Seja $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$. Então, segundo a parte (iv) de Teorema D.3.5, vale para todos os $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ a equação

$$\det(x, a_1, \dots, a_{n-1}) = \sum_{k=1}^n (-1)^{k-1} x_k \det(A_k) = \sum_{k=1}^n x_k \alpha_k = \langle x, \alpha \rangle.$$

Com $x = a_i$ segue dali $\langle a_i, \alpha \rangle = 0$ para $i = 1, \dots, n$. Daí por sua vez segue que a matriz $B := (\alpha \ a_1 \cdots a_{n-1}) \in \mathbb{R}^{n \times n}$ satisfaz as equações

$$B^T B = \begin{pmatrix} \|\alpha\|^2 & 0 \\ 0 & A^T A \end{pmatrix}, \quad \det(B) = \|\alpha\|^2$$

e conseqüentemente vale $\|\alpha\|^4 = \det(B)^2 = \det(B^T B) = \|\alpha\|^2 \det(A^T A)$. Disso segue $\det(A^T A) = \|\alpha\|^2$ e dali, segundo (6.1.9) em Lema 6.1.5,

$$v_d(a_1, \dots, a_{n-1}) = \sqrt{\det(A^T A)} = \|\alpha\| = \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}.$$

Com isso o Lema 6.1.7 fica provado. \square

Lema 6.1.8. *Seja $A \in \mathbb{R}^{n \times d}$ e seja $Q = [0, \lambda_1] \times \cdots \times [0, \lambda_d]$ com $\lambda_i > 0$. Então o paralelepípedo $AQ := \{Ax \mid x \in Q\}$ tem volume d -dimensional*

$$\text{Vol}_d(AQ) = \sqrt{\det(A^T A)} \cdot \text{Vol}_d(Q). \quad (6.1.13)$$

Demonstração. Sejam $a_1, \dots, a_d \in \mathbb{R}^n$ as colunas da matriz $A \in \mathbb{R}^{n \times d}$. Então

$$AQ = \left\{ \sum_{i=1}^d t_i a_i \mid 0 \leq t_i \leq \lambda_i \text{ para } i = 1, \dots, d \right\} = P(\lambda_1 a_1, \dots, \lambda_d a_d)$$

e por isso vale

$$\text{Vol}_d(AQ) = v_d(\lambda_1 a_1, \dots, \lambda_d a_d) = \lambda_1 \cdots \lambda_d \cdot v_d(a_1, \dots, a_d) = \text{Vol}_d(Q) \sqrt{\det(A^T A)}.$$

Com o isso Lema 6.1.8 fica provado. \square

Integração sobre domínios de carta

O próximo objetivo é integrar uma função real contínua numa subvariedade de um espaço euclidiano sobre o domínio de uma carta. Por isso precisamos as seguintes noções.

Definição 6.1.9 (Domínio de carta). *Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional, sejam $U, W \subset \mathbb{R}^n$ subconjuntos abertos, e seja*

$$\phi : U \rightarrow W$$

um C^1 -difeomorfismo com

$$\phi(U \cap M) = W \cap (\mathbb{R}^d \times \{0\})$$

ou seja ϕ é uma carta em M . Nesta situação chama-se o conjunto $U \cap M$ **domínio de carta**. Portanto todo tal domínio de uma carta é um subconjunto aberto de M com respeito à topologia relativa (Seção B.1). A aplicação

$$\psi : V \rightarrow U \cap M$$

definida através de

$$\begin{aligned} V &:= \{x = (x_1, \dots, x_d) \in \mathbb{R}^d \mid (x_1, \dots, x_d, 0, \dots, 0) \in W\} \\ \psi(x_1, \dots, x_d) &:= \phi^{-1}(x_1, \dots, x_d, 0, \dots, 0) \end{aligned} \quad (6.1.14)$$

para $(x_1, \dots, x_d) \in \Omega$, é chamada de **C^1 -parametrização** do domínio de carta $U \cap M$. Assim toda tal C^1 -parametrização é um homeomorfismo de um subconjunto aberto $V \subset \mathbb{R}^d$ a um domínio de carta $U \cap M$ (com $U \subset \mathbb{R}^n$ aberto), é continuamente diferenciável como aplicação de V a \mathbb{R}^n , e a aplicação inversa $\psi^{-1} : U \cap M \rightarrow V$ pode ser estendida a uma aplicação continuamente diferenciável de U a \mathbb{R}^d .

Seja $\psi : V \rightarrow U \cap M$ uma C^1 -parametrização de um domínio de carta numa C^1 -subvariedade d -dimensional $M \subset \mathbb{R}^n$ como em Definição 6.1.9. Para funções adequadas $f : M \rightarrow \mathbb{R}$ e subconjuntos adequados $B \subset U \cap M$ queremos definir ora a integral de f sobre B . Vamos denotar com o símbolo

$$\int_B f dS$$

esta integral.

Primeiro consideramos um conjunto da forma $B = \psi(A) \subset U \cap M$, onde o subconjunto $A \subset \Omega$ é um edifício de blocos (Definição 5.6.1). Então A possui uma partição segundo o Exercício 5.1.4. Conforme Definição 5.1.1 existem então blocos eixo-paralelos fechados $Q_1, \dots, Q_\ell \subset \Omega$ com

$$A = \bigcup_{i=1}^{\ell} Q_i, \quad \overset{\circ}{Q}_i \cap \overset{\circ}{Q}_j = \emptyset \text{ para } i \neq j.$$

A seguinte consideração heurística motiva a definição da integral de uma função contínua $f : M \rightarrow \mathbb{R}$ sobre B . Se escolha-se os Q_i de perímetro muito pequeno, e se x_i é o centro do bloco Q_i , então o conjunto imagem $\psi(Q_i)$ devia só pouco do conjunto imagem $\psi(x_i) + d\psi(x_i)(Q_i - x_i)$ do mesmo bloco sob a aplicação $x \mapsto \psi(x_i) + d\psi(x_i)(x - x_i)$, porque esta aplicação afim é conforme a definição da derivada uma aproximação *boa* para ψ perto de x_i . Consequentemente o volume do paralelepípedo $d\psi(x_i)Q_i$ é uma aproximação boa para o (ainda a ser definido) volume do conjunto imagem real $\psi(Q_i)$. Este valor aproximativo tem a forma

$$\begin{aligned} \text{Vol}_d(\psi(Q_i)) &\sim \text{Vol}_d(\psi(x_i) + d\psi(x_i)(Q_i - x_i)) \\ &= \text{Vol}_d(d\psi(x_i)Q_i) \\ &= \sqrt{\det(d\psi(x_i)^T d\psi(x_i))} \text{Vol}_d(Q_i). \end{aligned}$$

Aqui segue a última igualdade de Lema 6.1.8. Substituímos ora o (ainda a ser definido) volume $\text{Vol}_d(\psi(Q_i))$ por este valor aproximativo, então resulta para a integral buscada conforme o princípio “*A integral é aproximadamente a soma dos produtos de área da base vezes altura sobre as regiões de uma partição fina*” a fórmula heurística

$$\begin{aligned} \int_B f dS &\sim \sum_{i=1}^{\ell} f(\psi(x_i)) \text{Vol}_d(\psi(Q_i)) \\ &\sim \sum_{i=1}^{\ell} f(\psi(x_i)) \sqrt{\det(d\psi(x_i)^T d\psi(x_i))} \text{Vol}_d(Q_i) \\ &\sim \int_A f(\psi(x)) \sqrt{\det(d\psi(x)^T d\psi(x))} dx_1 \cdots dx_d. \end{aligned}$$

Aqui tem-se que entender o último termo via a convergência das somas de Riemann em Teorema 5.3.1. Exatamente este último termo vamos usar ora para a definição da integral.

Definição 6.1.10 (Integral sobre um domínio de carta). *Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional, seja $U \subset \mathbb{R}^n$ um conjunto aberto, tal que $U \cap M$ é um domínio de carta, seja $V \subset \mathbb{R}^d$ um conjunto aberto, seja*

$$\psi : V \rightarrow U \cap M$$

uma C^1 -parametrização do domínio de carta $U \cap M$, seja $A \subset \Omega$ um conjunto Jordan-mensurável (Definição 5.5.1), seja

$$B := \psi(A),$$

e seja $f : M \rightarrow \mathbb{R}$ uma função com a propriedade que a composição

$$f \circ \psi : A \rightarrow \mathbb{R}$$

é Riemann-integrável. Seja $g^\psi : \Omega \rightarrow \mathbb{R}$ a função contínua

$$g^\psi(x) := \sqrt{\det(d\psi(x)^T d\psi(x))} \quad \text{para } x \in \Omega. \quad (6.1.15)$$

Segundo o Teorema 5.2.2 e o Teorema 5.3.4 a função

$$(f \circ \psi)g^\psi : A \rightarrow \mathbb{R}$$

*é Riemann-integrável (Definição 5.5.2). Nestas hipóteses definimos a **integral de f sobre B** através de*

$$\int_B f \, dS := \int_A f(\psi(x))g^\psi(x) \, dx_1 \cdots dx_d \quad (6.1.16)$$

*e o **volume d -dimensional de B** através de*

$$\text{Vol}_d(B) := \int_B 1 \, dS = \int_A g^\psi(x) \, dx_1 \cdots dx_d. \quad (6.1.17)$$

O seguinte lema justifica a notação $\int_B f \, dS$ para a integral da função $f : M \rightarrow \mathbb{R}$ sobre um subconjunto compacto $B \subset M$. Porque esta notação sugere que a integral realmente só depende de B e f . Olhando exclusivamente à definição este fato porém não é visível imediato. Pois para a definição foi escolhido um domínio de carta $U \cap M$ com $B \subset U \cap M$ assim como uma parametrização $\psi : V \rightarrow U \cap M$ deste domínio de carta. Ora, resta mostrar que o lado direito da equação (6.1.16) não depende de estas *datas adicionais* V, U, ψ , as quais foram utilizados para a definição.

Lema 6.1.11. *A integral (6.1.16) em Definição 6.1.10 é independente da escolha de V, U, ψ com $B \subset U \cap M$.*

Demonstração. Seja $U' \subset \mathbb{R}^n$ um conjunto aberto, tal que $U' \cap M$ é um domínio de carta com $B \subset U' \cap M$. Mesmo seja $V' \subset \mathbb{R}^d$ um conjunto aberto e

$$\psi' : V' \rightarrow U' \cap M$$

uma C^1 -parametrização do domínio de carta $U' \cap M$. Então

$$\tilde{V} := \psi^{-1}(U \cap U' \cap M), \quad \tilde{V}' := (\psi')^{-1}(U \cap U' \cap M)$$

são subconjuntos abertos de \mathbb{R}^d e as derivadas

$$\psi : \tilde{V} \rightarrow U \cap U' \cap M, \quad \psi' : \tilde{V}' \rightarrow U \cap U' \cap M$$

são homeomorfismos. Além disso, a aplicação

$$\phi := \psi^{-1} \circ \psi' : \tilde{V}' \rightarrow \tilde{V}$$

é um C^1 -difeomorfismo, segundo a Definição 6.1.9.

Ora segue de Lema 5.7.12 que o conjunto

$$A' := (\psi')^{-1}(B) = (\psi \circ \phi)^{-1}(B) = \phi^{-1}(A) \subset \tilde{\Omega}'$$

é Jordan-mensurável. Como $\psi \circ \phi = \psi'$, segue da regra da cadeia a fórmula

$$d\psi'(x') = d\psi(\phi(x'))d\phi(x') \quad \text{para todos os } x' \in \tilde{V}'$$

e daí segue

$$g^{\psi'}(x') = g^{\psi}(\phi(x')) |\det(d\phi(x'))| \quad \text{para todos os } x' \in \tilde{V}'.$$

Portanto resulta da fórmula da transformação em Teorema 5.7.11 a equação

$$\begin{aligned} \int_{A'} (f \circ \psi') g^{\psi'} &= \int_{A'} (f \circ \psi \circ \phi) (g^{\psi \circ \phi}) |\det(d\phi)| \\ &= \int_{\phi(A')} (f \circ \psi) g^{\psi} \\ &= \int_A (f \circ \psi) g^{\psi}. \end{aligned}$$

Com isso o Lema 6.1.11 fica provado. \square

6.2 Definição da integral

Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional. Nosso objetivo é definir a integral de uma função contínua $f : M \rightarrow \mathbb{R}$ sobre um subconjunto compacto de M o qual não é contido em só *um* domínio de carta. Isto requer o conceito de um conjunto Jordan-mensurável $B \subset M$, ou seja, de um subconjunto o qual é Jordan-mensurável *relativo a M* , o que em contrapartida significa que seu bordo $\partial_M B$ é um conjunto Jordan-nulo d -dimensional com respeito à topologia relativa de M . Além disso, precisamos o conceito de uma *partição da unidade*, qual permite representar a função f como soma de funções cada uma das quais possui valores não-nulos só em um único domínio de carta.

Subconjuntos Jordan-mensuráveis de M

Definição 6.2.1. *Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional. Um subconjunto compacto $N \subset M$ é chamado de **conjunto Jordan-nulo d -dimensional**, ou também **conjunto Jordan-nulo com respeito a M** , se para toda C^1 -parametrização*

$$\psi : V \rightarrow U \cap M$$

de um domínio de carta $U \cap M$ (com $U \subset \mathbb{R}^n$ e $V \subset \mathbb{R}^d$ aberto) e todo conjunto compacto $K \subset V$, o conjunto

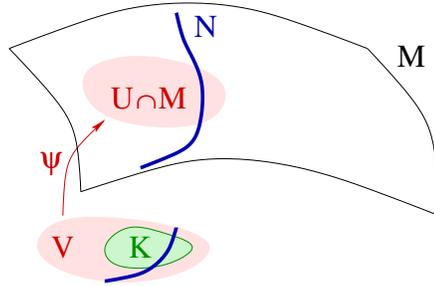
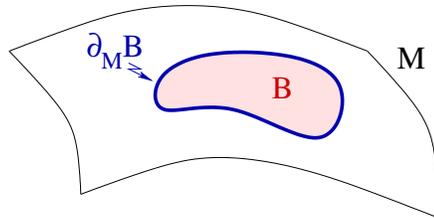
$$K \cap \psi^{-1}(N) = \{x \in K \mid \psi(x) \in N\} \subset \mathbb{R}^d$$

*é um conjunto Jordan-nulo (veja a Figura 6.3 e Definição 5.2.1). Um subconjunto compacto $B \subset M$ é chamado de **Jordan-mensurável com respeito a M** , se o conjunto*

$$\partial_M B := \left\{ x \in M \left| \begin{array}{l} \text{para todos os } \varepsilon > 0 \text{ vale} \\ B_\varepsilon(x) \cap B \neq \emptyset \text{ e} \\ B_\varepsilon(x) \cap (M \setminus B) \neq \emptyset \end{array} \right. \right\} \quad (6.2.1)$$

*é um conjunto Jordan-nulo d -dimensional. Chama-se o conjunto $\partial_M B \subset M$ do **bordo de B com respeito a M** (veja a Figura 6.4).*

Comentário 6.2.2. Sejam $\|\cdot\|$ a norma euclidiana no \mathbb{R}^n e $d_M : M \times M \rightarrow \mathbb{R}$ a função distância definida através de $d_M(x, y) := \|x - y\|$ para $x, y \in M$. Então o bordo $\partial_M B$ em (6.2.1) de um compacto $B \subset M$ é o bordo de B como subconjunto do espaço métrico (M, d_M) . Assim $\partial_M B = \partial B \subset B$. Note que compacidade em M e em \mathbb{R}^n são equivalente; veja o Exercício C.1.3.

Figura 6.3: Um conjunto Jordan-nulo d -dimensional.Figura 6.4: O bordo de B com respeito a M .

Lema 6.2.3. *Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional e $B \subset M$ um conjunto compacto e com respeito a M Jordan-mensurável. Sejam $U \subset \mathbb{R}^n$ e $V \subset \mathbb{R}^d$ conjuntos abertos tal que $U \cap M$ é um domínio de carta, seja $\psi : V \rightarrow U \cap M$ uma C^1 -parametrização deste domínio de carta, e seja $K \subset V$ um conjunto Jordan-mensurável compacto. Então o conjunto*

$$A := K \cap \psi^{-1}(B)$$

é compacto e Jordan-mensurável.

Demonstração. Provaremos primeiro que A é um subconjunto compacto de \mathbb{R}^d . Seja $(x_i)_{i \in \mathbb{N}}$ uma sequência em A . Como $A \subset K$ e K é compacto, existe uma subsequência $(x_{i_\nu})_{\nu \in \mathbb{N}}$ a qual converge para um elemento $x := \lim_{\nu \rightarrow \infty} x_{i_\nu} \in K$. Como B é compacto e $(\psi(x_{i_\nu}))_{\nu \in \mathbb{N}}$ é uma sequência em B , podemos supor através de mais uma transição a uma subsequência (ainda denotado de $(x_{i_\nu})_{\nu \in \mathbb{N}}$) que a sequência $(\psi(x_{i_\nu}))_{\nu \in \mathbb{N}}$ converge para um ponto p , ou seja $p = \lim_{\nu \rightarrow \infty} \psi(x_{i_\nu}) \in B$. Segue daí a equação $p = \psi(\lim_{\nu \rightarrow \infty} x_{i_\nu}) = \psi(x)$, porque ψ é contínuo. Então segue $\psi(x) \in B$ e assim $x \in K \cap \psi^{-1}(B) = A$. Portanto temos mostrado que toda sequência em A possui uma subsequência a qual converge para um elemento de A . Então A é compacto.

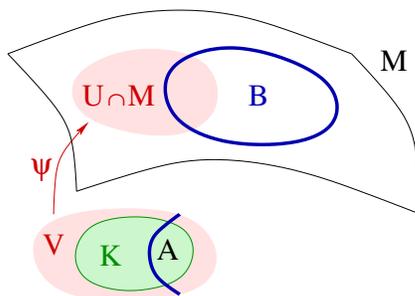


Figura 6.5: Um conjunto Jordan-mensurável com respeito a M .

O que provaremos próximo é a inclusão

$$\partial A \subset \partial K \cup (K \cap \psi^{-1}(\partial_M B)). \tag{6.2.2}$$

(Veja a Figura 6.5.) Como K é Jordan-mensurável, seu bordo ∂K é um conjunto Jordan-nulo e, como B é Jordan-mensurável com respeito a M , seu bordo $\partial_M B$ é um conjunto Jordan-nulo d -dimensional, e assim $K \cap \psi^{-1}(\partial_M B)$ também é um conjunto Jordan-nulo. Portanto a união destes dois conjuntos é um conjunto Jordan-nulo. Segundo (6.2.2), daí ∂A também é um conjunto Jordan-nulo, e por isso A é Jordan-mensurável.

Portanto resta mostrar que vale (6.2.2). Seja $x \in \partial A \setminus \partial K$. Como já temos mostrado que A é um subconjunto compacto de \mathbb{R}^d , também sabemos que A é fechado e segue dali que $\partial A \subset A \subset K$. Por isso vale

$$x \in K \setminus \partial K = \overset{\circ}{K}.$$

Além disso, existe uma sequência $x_i \in \mathbb{R}^d \setminus A$, a qual converge para x . Como seu limite x é um elemento do conjunto aberto $\overset{\circ}{K}$, vale $x_i \in \overset{\circ}{K}$ para todo índice $i \in \mathbb{N}$ suficientemente grande. Portanto podemos supor sem perda de generalidade que $x_i \in \overset{\circ}{K} \setminus A$ para todos os $i \in \mathbb{N}$. Então vale

$$\psi(x_i) \notin B$$

para todos os $i \in \mathbb{N}$ conforme a definição do conjunto A . Depois vale

$$\lim_{i \rightarrow \infty} \psi(x_i) = \psi(x) \in B,$$

porque $x \in \partial A \subset A$. Como $\psi(x_i) \in M \setminus B$, segue disso que $\psi(x) \in \partial_M B$ e daí que $x \in \psi^{-1}(\partial_M B) \cap K$. Portanto temos provado a inclusão

$$\partial A \setminus \partial K \subset K \cap \psi^{-1}(\partial_M B).$$

Disso segue (6.2.2) e com isso o Lema 6.2.3 fica provado. □

Partição da unidade

O **suporte** de uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é o conjunto

$$\text{supp}(f) := \overline{\{x \in \mathbb{R}^n \mid f(x) \neq 0\}}.$$

Ou seja, o suporte é o fecho do conjunto de todos os pontos em \mathbb{R}^n nas quais a função f não se anula.

Teorema 6.2.4 (Partição da unidade). *Seja $K \subset \mathbb{R}^n$ um conjunto compacto e sejam $U_1, U_2, \dots, U_\ell \subset \mathbb{R}^n$ conjuntos cobrindo K , ou seja*

$$K \subset U_1 \cup U_2 \cup \dots \cup U_\ell. \quad (6.2.3)$$

Então existem funções suaves $\rho_1, \rho_2, \dots, \rho_\ell : \mathbb{R}^n \rightarrow [0, 1]$ tal que vale o seguinte:

$$\text{supp}(\rho_i) \subset U_i, \quad \text{supp}(\rho_i) \text{ é compacto} \quad \text{para } i = 1, 2, \dots, \ell \quad (6.2.4)$$

$$\sum_{i=1}^{\ell} \rho_i(x) = 1 \quad \text{para todos os } x \in K. \quad (6.2.5)$$

Demonstração. Para $K = \emptyset$ vale a afirmação trivialmente com $\rho_i \equiv 0$ para todos os i . Seja $K \neq \emptyset$. Então provamos a existência dos ρ_i em quatro passos.

Passo 1. *Existe uma função suave $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, tal que $\phi(x) > 0$ para todos os $x \in \mathbb{R}^n$ com $\|x\| < 1$, e $\phi(x) = 0$ para todos os $x \in \mathbb{R}^n$ com $\|x\| \geq 1$.*

Uma fórmula explícita para uma tal função é

$$\phi(x) := \begin{cases} e^{-\frac{1}{1-\|x\|^2}}, & \text{caso } \|x\| < 1, \\ 0, & \text{caso } \|x\| \geq 1, \end{cases} \quad (6.2.6)$$

para $x \in \mathbb{R}^n$. Esta função pode ser escrito na forma $\phi(x) = g(1 - \|x\|^2)$, onde a função $g : \mathbb{R} \rightarrow \mathbb{R}$ é definida através da fórmula

$$g(s) := \begin{cases} e^{-\frac{1}{s}}, & \text{caso } s > 0, \\ 0, & \text{caso } s \leq 0, \end{cases}$$

para $s \in \mathbb{R}$. Que g é suave foi mostrado no curso Análise I, e com isso segue da regra da cadeia que ϕ também é uma função suave.

Passo 2. Existe um número natural $m \in \mathbb{N}$, elementos $x_1, \dots, x_m \in K$, números reais positivos $\delta_1, \dots, \delta_m$, e índices $i_1, \dots, i_m \in \{1, \dots, \ell\}$, tal que

$$K \subset \bigcup_{j=1}^m B_{\delta_j}(x_j) \tag{6.2.7}$$

$$\overline{B_{3\delta_j}(x_j)} \subset U_{i_j} \quad \text{para } j = 1, \dots, m. \tag{6.2.8}$$

Para todo elemento $x \in K$ seja $i(x) \in \{1, \dots, \ell\}$ o índice menor com $x \in U_{i(x)}$ e seja $\delta(x) > 0$ o número maior no intervalo $(0, 1]$ com $B_{4\delta(x)}(x) \subset U_{i(x)}$. Então vale $\overline{B_{3\delta(x)}(x)} \subset U_{i(x)}$ para todos os $x \in K$ e os conjuntos $B_{\delta(x)}(x)$, para $x \in K$, formam uma cobertura aberta de K . Como K é compacto, existem segundo o Teorema C.1.2 um número finito de elementos $x_1, \dots, x_m \in K$ com

$$K \subset B_{\delta(x_1)}(x_1) \cup B_{\delta(x_2)}(x_2) \cup \dots \cup B_{\delta(x_m)}(x_m).$$

Com $\delta_j := \delta(x_j)$ e $i_j := i(x_j)$ para $j = 1, \dots, m$ seguem daí as afirmações (6.2.7) e (6.2.8) em Passo 2.

Passo 3. Existe uma função suave $\phi_0 : \mathbb{R}^n \rightarrow [0, \infty)$ com

$$\begin{aligned} \phi_0(x) &> 0 \quad \text{para todos os } x \in \mathbb{R}^n \setminus \overline{\bigcup_{j=1}^m B_{2\delta_j}(x_j)} \\ \phi_0(x) &= 0 \quad \text{para todos os } x \in \bigcup_{j=1}^m B_{\delta_j}(x_j). \end{aligned} \tag{6.2.9}$$

Seja $\delta := \min\{\delta_1, \dots, \delta_m\} > 0$ e defina $\sigma_0 : \mathbb{R}^n \rightarrow [0, \infty)$ através de

$$\sigma_0(x) := \inf \left\{ \|x - y\| \mid y \in \bigcup_{j=1}^m B_{2\delta_j}(x_j) \right\} \quad \text{para } x \in \mathbb{R}^n.$$

Esta função é Lipschitz-contínua segundo a desigualdade triangular, anula-se no conjunto $\overline{\bigcup_{j=1}^m B_{2\delta_j}(x_j)}$, e no complemento dele possui valores positivos. Seja $\phi : \mathbb{R}^n \rightarrow [0, 1]$ uma função suave em Passo 1 e defina

$$\phi_0(x) := \int_{\mathbb{R}^n} \sigma_0(\xi) \phi(\delta^{-1}(x - \xi)) \, d\xi_1 \cdots d\xi_n \quad \text{para } x \in \mathbb{R}^n.$$

Então ϕ_0 é suave o que pode-se provar através de diferenciação parcial *dentro da integral*. Demais ϕ_0 satisfaz as condições (6.2.9) em Passo 3.

Passo 4. *Provaremos Teorema 6.2.4.*

Para $j = 1, \dots, m$ definimos a função $\phi_j : \mathbb{R}^n \rightarrow [0, 1]$ através de

$$\phi_j(x) := \phi\left(\frac{x - x_j}{3\delta_j}\right) \quad \text{para } x \in \mathbb{R}^n.$$

Aqui $\phi : \mathbb{R}^n \rightarrow [0, 1]$ é a função suave de Passo 1. Então ϕ_j é uma função suave com suporte compacto

$$\text{supp}(\phi_j) = \overline{B_{3\delta_j}(x_j)} \subset U_{i_j}. \quad (6.2.10)$$

Além disso, vale

$$\sum_{j=0}^m \phi_j(x) > 0 \quad \text{para todos os } x \in \mathbb{R}^n. \quad (6.2.11)$$

Pois para $x \in B_{3\delta_j}(x_j)$, $j > 0$, vale $\phi_j(x) > 0$ e para $x \notin \bigcup_{j=1}^m B_{3\delta_j}(x_j)$ vale

$$x \notin \bigcup_{j=1}^m \overline{B_{2\delta_j}(x_j)} = \overline{\bigcup_{j=1}^m B_{2\delta_j}(x_j)}$$

e neste caso $\phi_0(x) > 0$ segundo o Passo 3. Com isso (6.2.11) fica provado. Segundo o Passo 2 vale $K \subset \bigcup_{j=1}^m B_{\delta_j}(x_j)$ e dali segue segundo o Passo 3 que ϕ_0 anula-se em K , ou seja

$$\phi_0(x) = 0 \quad \text{para todos os } x \in K. \quad (6.2.12)$$

Para $i = 1, \dots, \ell$ seja

$$J_i := \{j \in \{1, \dots, m\} \mid i_j = i\},$$

$$\rho_i(x) := \frac{\sum_{j \in J_i} \phi_j(x)}{\sum_{j=0}^m \phi_j(x)} \quad \text{para } x \in \mathbb{R}^n. \quad (6.2.13)$$

A função $\rho_i : \mathbb{R}^n \rightarrow [0, 1]$ é bem definida segundo (6.2.11). Além disso, ela é suave e tem suporte

$$\text{supp}(\rho_i) \subset \bigcup_{j \in J_i} \text{supp}(\phi_j) = \bigcup_{j \in J_i} \overline{B_{3\delta_j}(x_j)} \subset U_i.$$

Aqui os últimos dois passos seguem de (6.2.10). Para todos os $x \in K$ vale segundo (6.2.12) e (6.2.13) a equação

$$\sum_{i=1}^{\ell} \rho_i(x) = \frac{\sum_{j=1}^m \phi_j(x)}{\sum_{j=0}^m \phi_j(x)} = 1.$$

Com isso o Teorema 6.2.4 fica provado. \square

Definição da integral**Definição 6.2.5 (Integral sobre uma variedade).**

Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional, seja $B \subset M$ um conjunto compacto Jordan-mensurável com respeito a M , seja $N \subset M$ um conjunto Jordan-nulo com respeito a M , e seja $f : M \rightarrow \mathbb{R}$ uma função limitada, contínua em $M \setminus N$. Então a integral de Riemann de f sobre B e o volume d -dimensional de B são definidos como segue.

Escolha subconjuntos abertos

$$U_1, U_2, \dots, U_\ell \subset \mathbb{R}^n$$

e funções contínuas

$$\rho_1, \rho_2, \dots, \rho_\ell : \mathbb{R}^n \rightarrow [0, 1]$$

as quais satisfazem as seguintes condições.

- (a) Para $i = 1, \dots, \ell$ a interseção $U_i \cap M$ é um domínio de carta.
- (b) Para $i = 1, \dots, \ell$ o suporte $\text{supp}(\rho_i)$ é um subconjunto compacto de U_i .
- (c) $B \subset \bigcup_{i=1}^{\ell} U_i$ e $\sum_{i=1}^{\ell} \rho_i(x) = 1$ para todos os $x \in B$.

Então a **integral de Riemann de f sobre B** é o número

$$\int_B f \, dS := \sum_{i=1}^{\ell} \int_{B \cap U_i} \rho_i f \, dS \quad (6.2.14)$$

e o **volume d -dimensional de B** é o número

$$\text{Vol}_d(B) := \int_B 1 \, dS = \sum_{i=1}^{\ell} \int_{B \cap U_i} \rho_i \, dS. \quad (6.2.15)$$

Em Definição 6.2.5 tem-se que dar muito cuidado que as palavras “com respeito a M ” não são omitidas, mais detalhado que trata-se no caso de N de um conjunto Jordan-nulo **com respeito a M** e no caso de B de um conjunto Jordan-mensurável **com respeito a M** . Por exemplo, se M é uma C^1 -subvariedade d -dimensional compacta com $d < n$, então M mesmo, e assim todo subconjunto de M , também é um conjunto Jordan-nulo em \mathbb{R}^n , segundo o Exemplo 5.6.13. Mas geralmente não podemos integrar uma função contínua $f : M \rightarrow \mathbb{R}$ sobre um subconjunto arbitrário de M .

Comentário 6.2.6 (À definição da integral). Tem-se que entender, conforme Definição 6.1.10, os somandos no lado direito de (6.2.14) assim: Como $U_i \cap M$ é um domínio de carta, existe um aberto $V_i \subset \mathbb{R}^d$ e uma C^1 -parametrização $\psi_i : V_i \rightarrow U_i \cap M$. Como $\text{supp}(\rho_i) \cap B$ é um subconjunto compacto de $U_i \cap M$, o subconjunto $\psi_i^{-1}(\text{supp}(\rho_i) \cap B)$ de V_i é compacto. Por isso existe, segundo o Exercício 5.1.5, um conjunto Jordan-mensurável compacto $K_i \subset \mathbb{R}^d$ com $\psi_i^{-1}(\text{supp}(\rho_i) \cap B) \subset K_i \subset V_i$. Segundo a Definição 6.2.1 o conjunto $N_i := K_i \cap \psi_i^{-1}(N)$ é um conjunto Jordan-nulo e segundo o Lema 6.2.3 o conjunto $A_i := K_i \cap \psi_i^{-1}(B)$ é Jordan-mensurável. Por isso o conjunto $B_i := \psi_i(A_i) = \psi_i(K_i) \cap B$ é Jordan-mensurável com respeito a M . Além disso, vale $\text{supp}(\rho_i) \cap B \subset B_i$. A função $((\rho_i f) \circ \psi_i) g^{\psi_i} : A_i \rightarrow \mathbb{R}$ é limitada e contínua em $A_i \setminus N_i$, por isso Riemann-integrável. Tem-se que entender o i -ésimo somando em (6.2.14) como a integral desta função sobre A_i :

$$\int_{B \cap U_i} \rho_i f dS = \int_{B_i} \rho_i f dS = \int_{A_i} \rho_i(\psi_i(x)) f(\psi_i(x)) \sqrt{\det(d\psi_i(x)^T d\psi_i(x))} dx.$$

Que esta integral é independente da escolha da C^1 -parametrização ψ_i e do conjunto K_i , foi provado em Lema 6.1.11.

Comentário 6.2.7 (A integral é bem definida). Provaremos que o lado direito da equação (6.2.14) é independente da escolha dos conjuntos abertos U_i e das funções contínuas ρ_i . Sejam então $V_1, \dots, V_m \subset \mathbb{R}^n$ conjuntos abertos e $\sigma_1, \dots, \sigma_m : \mathbb{R}^n \rightarrow [0, 1]$ funções contínuas, as quais também satisfaçam as condições (a), (b), e (c) em Definição 6.2.5. Ou seja, para todo $j \in \{1, \dots, m\}$ a interseção $V_j \cap M$ é um domínio de carta e $\text{supp}(\rho_j)$ um subconjunto compacto de V_j , os conjuntos V_j cobrem B , e vale $\sum_{j=1}^m \sigma_j(x) = 1$ para todos os $x \in B$. Então recebemos

$$\begin{aligned} \sum_{j=1}^m \int_{B \cap V_j} \sigma_j f dS &= \sum_{j=1}^m \int_{B \cap V_j} \left(\sum_{i=1}^{\ell} \rho_i \right) \sigma_j f dS \\ &= \sum_{i=1}^{\ell} \sum_{j=1}^m \int_{B \cap U_i \cap V_j} \rho_i \sigma_j f dS \\ &= \sum_{i=1}^{\ell} \int_{B \cap U_i} \sum_{j=1}^m \rho_i \sigma_j f dS \\ &= \sum_{i=1}^{\ell} \int_{B \cap U_i} \rho_i f dS. \end{aligned}$$

Comentário 6.2.8. Na situação de Definição 6.2.5 um subconjunto compacto $B \subset M$ é um conjunto Jordan-nulo com respeito a M se e somente se é Jordan-mensurável com respeito a M e de volume nulo $\text{Vol}_d(B) = 0$.

Exercício: Prova esta afirmação; cf. Corolário 5.6.4.

Comentário 6.2.9 (Regras de cálculo). A integral em Definição 6.2.5 satisfaz as regras comuns de cálculo para a integral, como são formulados por exemplo em Teorema 5.3.4. Para ver isso é útil introduzir as notações

$$\begin{aligned} \mathcal{F}(M) &:= \left\{ f : M \rightarrow \mathbb{R} \left| \begin{array}{l} \text{existe um conjunto Jordan-nulo} \\ d\text{-dimensional } N \subset M, \text{ tal que} \\ f \text{ é contínua em } M \setminus N \end{array} \right. \right\} \\ \mathcal{B}(M) &:= \left\{ B \subset M \left| \begin{array}{l} B \text{ é compacto e Jordan-} \\ \text{mensurável com respeito a } M \end{array} \right. \right\}. \end{aligned} \quad (6.2.16)$$

Então $\mathcal{F}(M)$ é um espaço vetorial e a aplicação

$$\mathcal{F}(M) \rightarrow \mathbb{R} : f \mapsto \int_B f \, dS$$

é linear para todos os $B \in \mathcal{B}(M)$. Ou seja, para todos os $f, g \in \mathcal{F}(M)$, todo $\lambda \in \mathbb{R}$, e todo $B \in \mathcal{B}(M)$ vale

$$\int_B (f + g) \, dS = \int_B f \, dS + \int_B g \, dS$$

$$\int_B \lambda f \, dS = \lambda \int_B f \, dS.$$

Além disso, vale para todo $f \in \mathcal{F}(M)$ e todos os $B, C \in \mathcal{B}(M)$ com

$$\text{Vol}_d(B \cap C) = 0$$

a equação

$$\int_{B \cup C} f \, dS = \int_B f \, dS + \int_C f \, dS.$$

Mesmo vale para todo $f \in \mathcal{F}(M)$ e todo $B \in \mathcal{B}(M)$ que

$$f \geq 0 \quad \implies \quad \int_B f \, dS \geq 0$$

e

$$\left| \int_B f \, dS \right| \leq \int_B |f| \, dS.$$

Exercício: Prova estas afirmações com a ajuda de Teorema 5.3.4.

Comentário 6.2.10 (Integrais impróprias). Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade d -dimensional não compacta, seja $\mathcal{B}(M) \subset 2^M$ de (6.2.16), e seja $f \in \mathcal{F}(M)$ uma função, a qual é contínua no complemento de um conjunto Jordan-nulo com respeito a M . A função f chama-se de **(imprópria) Riemann-integrável** se ela satisfaz a condição

$$\sup_{B \in \mathcal{B}(M)} \int_B |f| dS < \infty. \quad (6.2.17)$$

Então existe exatamente um número real c com propriedades como segue.

(i) Para toda sequência $B_i \in \mathcal{B}(M)$ com $B_i \subset \overset{\circ}{B}_{i+1} \forall i$ e $M = \bigcup_{i=1}^{\infty} B_i$ vale

$$\lim_{i \rightarrow \infty} \int_{B_i} f dS = c.$$

(ii) Para todo $\varepsilon > 0$ existe um $B_0 \in \mathcal{B}(M)$, tal que para todo $B \in \mathcal{B}(M)$:

$$B_0 \subset B \quad \implies \quad \left| c - \int_B f dS \right| < \varepsilon.$$

A existência de um tal número $c \in \mathbb{R}$ prova-se exatamente como em Teorema 5.8.4. Este número c é chamado de **integral (imprópria) de f sobre M** e é denotado de $\int_M f dS := c$. Ou seja

$$\int_M f dS := \lim_{i \rightarrow \infty} \int_{B_i} f dS \quad (6.2.18)$$

para toda sequência $B_i \in \mathcal{B}(M)$ com $B_i \subset \overset{\circ}{B}_{i+1} \forall i$ e $M = \bigcup_{i=1}^{\infty} B_i$.

Comentário 6.2.11. Seja $M \subset \mathbb{R}^n$ uma subvariedade d -dimensional compacta, seja $N \subset M$ um conjunto não vazio e Jordan-nulo com respeito a M , e seja $f : M \rightarrow \mathbb{R}$ uma função contínua. Então vale

$$\int_M f dS = \int_{M \setminus N} f dS = \lim_{i \rightarrow \infty} \int_{B_i} f dS$$

para toda sequência $B_i \in \mathcal{B}(M \setminus N)$ com $B_i \subset \overset{\circ}{B}_{i+1}$ para todo índice i e tal que $M \setminus N = \bigcup_{i=1}^{\infty} B_i$. Ou seja, a integral imprópria de f sobre $M \setminus N$ coincide com a integral de f sobre M inteiro.

6.3 Exemplos

Exemplo 6.3.1 (Famílias somáveis). Seja $M \subset \mathbb{R}^n$ uma subvariedade 0-dimensional. Então M é **discreto** (para todo $p \in M$ existe um $\varepsilon > 0$ com $B_\varepsilon(p) \cap M = \{p\}$), e daí – finito ou infinito – contável. A topologia relativa em M é a **topologia discreta** (todo subconjunto de M é aberto), todo subconjunto de M é Jordan-mensurável com respeito a M , e um subconjunto é compacto se e somente se é finito. A integral de uma função $f : M \rightarrow \mathbb{R}$ sobre um subconjunto finito $B \subset M$ é

$$\int_B f \, dS = \sum_{p \in B} f(p).$$

Uma função $f : M \rightarrow \mathbb{R}$ é portanto **absolutamente somável** (ou seja, o conjunto $\{\sum_{p \in B} |f(p)| \mid B \subset M, \#B < \infty\}$ é limitado) se é imprópria Riemann-integrável, e nesta situação a parte (ii) em Comentário 6.2.10 é precisamente a afirmação de [4, Satz 4.3].

Exemplo 6.3.2 (O caso $d = n$). Uma C^1 -subvariedade n -dimensional do \mathbb{R}^n é um subconjunto aberto $M \subset \mathbb{R}^n$. Então $V = U \cap M$ é para todo subconjunto aberto $U \subset \mathbb{R}^n$ um domínio de carta com $\psi = \text{id} : V \rightarrow U \cap M$ como C^1 -parametrização. Nesta situação um subconjunto compacto $B \subset M$ é Jordan-mensurável com respeito a M se e somente se é Jordan-mensurável em \mathbb{R}^n , e a integral de uma função contínua $f : M \rightarrow \mathbb{R}$ sobre B em Definição 6.1.10 coincide com a integral em Definição 5.1.10. A independência da integral da escolha da C^1 -parametrização é exatamente a fórmula da transformação em Teorema 5.7.11.

Exemplo 6.3.3 (Integrais linhas). Seja $M \subset \mathbb{R}^n$ uma C^1 -subvariedade 1-dimensional e seja $U \subset \mathbb{R}^n$ um subconjunto aberto tal que $U \cap M$ é um domínio de carta conexo. Uma C^1 -parametrização deste domínio de carta é uma C^1 -aplicação injetiva $\gamma : I \rightarrow \mathbb{R}^n$ num intervalo aberto $I \subset \mathbb{R}$ com imagem $\gamma(I) = U \cap M$ e velocidade $\dot{\gamma}(t) \neq 0$ para todos os t . Seja $A = [a, b] \subset I$ um subintervalo compacto. Então $B := \gamma(A)$ é Jordan-mensurável com respeito a M , a função g^γ em (6.1.15) é $g^\gamma(t) = \|\dot{\gamma}(t)\|$, e por (6.1.16) a integral de uma função contínua $f : M \rightarrow \mathbb{R}$ sobre $B = \gamma([a, b])$ é a integral linha

$$\int_B f \, dS = \int_a^b f(\gamma(s)) \|\dot{\gamma}(s)\| \, ds. \quad (6.3.1)$$

Para $f = 1$ assim $\text{Vol}_1(B) = \int_B 1 \, dS = L(\gamma|_{[a,b]})$ é o comprimento da curva.

Exemplo 6.3.4 (A circunferência). O círculo unitário

$$M = \mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

é uma subvariedade suave 1-dimensional do \mathbb{R}^2 . Defina a curva $\gamma : \mathbb{R} \rightarrow \mathbb{R}^2$ através de

$$\gamma(t) := (\cos(t), \sin(t)) \quad \text{para } t \in \mathbb{R}.$$

Esta curva é suave, satisfaz a condição $\|\dot{\gamma}(t)\| = 1$ para todos os $t \in \mathbb{R}$, e sua restrição a todo intervalo aberto (a, b) com $b - a \leq 2\pi$ é uma parametrização suave de um domínio de carta em \mathbb{S}^1 . Para $\varepsilon > 0$ o conjunto

$$B_\varepsilon := \gamma([\varepsilon, 2\pi - \varepsilon])$$

é um subconjunto compacto de \mathbb{S}^1 , Jordan-mensurável com respeito a \mathbb{S}^1 , e vale

$$\bigcup_{\varepsilon > 0} B_\varepsilon = \mathbb{S}^1 \setminus \{(1, 0)\}.$$

Se $f : \mathbb{S}^1 \rightarrow \mathbb{R}$ é uma função contínua, então vale

$$\begin{aligned} \int_{\mathbb{S}^1} f \, dS &= \int_{\mathbb{S}^1 \setminus \{(1, 0)\}} f \, dS \\ &= \lim_{\varepsilon \rightarrow 0} \int_{B_\varepsilon} f \, dS \\ &= \lim_{\varepsilon \rightarrow 0} \int_{\varepsilon}^{2\pi - \varepsilon} f(\gamma(t)) \|\dot{\gamma}(t)\| \, dt \\ &= \int_0^{2\pi} f(\gamma(t)) \, dt. \end{aligned}$$

Aqui a primeira igualdade segue de Comentário 6.2.11, a segunda da parte (i) de Comentário 6.2.10, a terceira de Exemplo 6.3.3, e a quarta do fato que $\|\dot{\gamma}(t)\| = 1$ para todos os t . Disso segue a fórmula

$$\text{Vol}_1(\mathbb{S}^1) = \int_{\mathbb{S}^1} 1 \, dS = \int_0^{2\pi} dt = 2\pi$$

para a circunferência do círculo unitário.

Exemplo 6.3.5 (Integração sobre gráficos). Seja $V \subset \mathbb{R}^{n-1}$ um subconjunto aberto e $h : V \rightarrow \mathbb{R}$ uma função suave. Então

$$M := \{(x, h(x)) \mid x = (x_1, \dots, x_{n-1}) \in V\}$$

é uma subvariedade suave $(n - 1)$ -dimensional do \mathbb{R}^n . Neste caso M inteiro é um domínio de carta com $U := V \times \mathbb{R}$ e $U \cap M = M$. A aplicação

$$V \rightarrow M : x \mapsto \psi(x) := (x, h(x))$$

é uma parametrização suave e a matriz jacobiana de ψ no ponto $x \in V$ é

$$d\psi(x) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ \frac{\partial h}{\partial x_1} & \cdots & \frac{\partial h}{\partial x_{n-2}} & \frac{\partial h}{\partial x_{n-1}} \end{pmatrix} \in \mathbb{R}^{n \times (n-1)}.$$

Daí o determinante da matriz $d\psi(x)^T d\psi(x) \in \mathbb{R}^{(n-1) \times (n-1)}$ é dado, segundo o Exemplo D.3.14, através de

$$\det(d\psi(x)^T d\psi(x)) = 1 + \sum_{i=1}^{n-1} \left(\frac{\partial h}{\partial x_i}(x) \right)^2 = 1 + \|\nabla h(x)\|^2.$$

Com isso encontraremos as equações

$$g^\psi(x) = \sqrt{1 + \|\nabla h(x)\|^2}$$

e

$$\int_M f dS = \int_V f(x, h(x)) \sqrt{1 + \|\nabla h(x)\|^2} dx_1 \cdots dx_{n-1} \quad (6.3.2)$$

para toda função $f : M \rightarrow \mathbb{R}$, a qual é contínua fora de um conjunto Jordan-nulo com respeito a M e é imprópria Riemann-integrável (Comentário 6.2.10). Assim a variedade M possui o volume

$$\text{Vol}_{n-1}(M) = \int_V \sqrt{1 + \|\nabla h(x)\|^2} dx_1 \cdots dx_{n-1}. \quad (6.3.3)$$

Isto também possa ser infinito.

Exemplo 6.3.6 (Integração sobre esferas). Para $r > 0$ e $n \in \mathbb{N}$ seja

$$M_r := \mathbb{S}_r^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = r\}$$

a esfera de raio r no \mathbb{R}^n . Então M_r é uma subvariedade suave $(n - 1)$ -dimensional do \mathbb{R}^n e o conjunto

$$U_r := \{x = (x_1, \dots, x_n) \in M_r \mid x_2 = 0 \implies x_1 > 0\}$$

é um domínio de carta, cujo complemento

$$N_r := M_r \setminus U_r = \{x \in M_r \mid x_2 = 0, x_1 \leq 0\}$$

é um conjunto Jordan-nulo com respeito a M_r . Uma parametrização suave de U_r é a aplicação $\psi_r : V \rightarrow U_r$ com

$$V := \{\theta \in \mathbb{R}^{n-1} \mid |\theta_1| < \pi, |\theta_i| < \pi/2 \text{ para } i = 2, \dots, n-1\}$$

e

$$\psi_r(\theta) := r \begin{pmatrix} \cos(\theta_1) \cdots \cos(\theta_{n-1}) \\ \text{sen}(\theta_1) \cos(\theta_2) \cdots \cos(\theta_{n-1}) \\ \text{sen}(\theta_2) \cos(\theta_3) \cdots \cos(\theta_{n-1}) \\ \vdots \\ \text{sen}(\theta_{n-2}) \cos(\theta_{n-1}) \\ \text{sen}(\theta_{n-1}) \end{pmatrix} \quad (6.3.4)$$

para $\theta = (\theta_1, \dots, \theta_{n-1}) \in V$. Nota-se que $\|\psi_r(\theta)\| = r$ e $\cos(\theta_i) > 0$ para $i = 2, \dots, n-1$ e todos os $\theta \in V$, de onde resulta que a aplicação $\psi_r : V \rightarrow U_r$ é bijetiva. Mesmo as derivadas parciais de ψ_r são dois-a-dois ortogonais e tem a norma

$$\left\| \frac{\partial \psi_r}{\partial \theta_i}(\theta) \right\| = \begin{cases} r \cos(\theta_{i+1}) \cdots \cos(\theta_{n-1}), & \text{para } 1 \leq i \leq n-2, \\ r, & \text{para } i = n-1. \end{cases}$$

Daí resulta para a função $g^{\psi_r}(\theta) = \sqrt{d\psi_r(\theta)^T d\psi_r(\theta)}$ a fórmula

$$g^{\psi_r}(\theta) = r^{n-1} g(\theta), \quad g(\theta) := \cos(\theta_2) \cos^2(\theta_3) \cdots \cos^{n-2}(\theta_{n-1}). \quad (6.3.5)$$

Dali, segundo a Definição 6.1.10 e Comentário 6.2.11, a integral de uma função contínua $f : \mathbb{S}_r^{n-1} \rightarrow \mathbb{R}$ é dada através da fórmula

$$\begin{aligned}
 \int_{\mathbb{S}_r^{n-1}} f \, dS &= \int_{U_r} f \, dS \\
 &= \int_V f(\psi_r(\theta)) r^{n-1} g(\theta) \, d\theta \\
 &= r^{n-1} \int_V f(r\psi_1(\theta)) g(\theta) \, d\theta \\
 &= r^{n-1} \int_{\mathbb{S}^{n-1}} f(rx) \, dS(x).
 \end{aligned} \tag{6.3.6}$$

Particularmente resulta com $f \equiv 1$ a equação

$$\text{Vol}_{n-1}(\mathbb{S}_r^{n-1}) = r^{n-1} \text{Vol}_{n-1}(\mathbb{S}^{n-1}). \tag{6.3.7}$$

Exemplo 6.3.7 (O volume da esfera unitária).

O volume da esfera unitária no \mathbb{R}^n é denotado de $\omega_n := \text{Vol}_{n-1}(\mathbb{S}^{n-1})$. De (6.3.5) e (6.3.6) resulta para este volume a fórmula explicita

$$\begin{aligned}
 \omega_n &= \text{Vol}_{n-1}(\mathbb{S}^{n-1}) \\
 &= \int_V g(\theta) \, d\theta \\
 &= \int_{-\pi/2}^{\pi/2} \cdots \int_{-\pi/2}^{\pi/2} \int_{-\pi}^{\pi} \cos(\theta_2) \cos^2(\theta_3) \cdots \cos^{n-2}(\theta_{n-1}) \, d\theta_1 \cdots d\theta_{n-1} \\
 &= 2\pi \prod_{k=1}^{n-2} \int_{-\pi/2}^{\pi/2} \cos^k(\theta) \, d\theta
 \end{aligned}$$

para $n \geq 3$. Portanto vale

$$\omega_{n+1} = \omega_n \int_{-\pi/2}^{\pi/2} \cos^{n-1}(\theta) \, d\theta \tag{6.3.8}$$

para todo inteiro $n \geq 2$. Para \mathbb{S}^0 temos $\omega_1 = 2$ segundo Exemplo 6.3.1 e para \mathbb{S}^1 vale $\omega_2 = 2\pi$ segundo Exemplo 6.3.4. Disso resultam segundo (6.3.8) as fórmulas $\omega_3 = 2\pi \int_{-\pi/2}^{\pi/2} \cos(\theta) \, d\theta = 4\pi$ e $\omega_4 = 4\pi \int_{-\pi/2}^{\pi/2} \cos^2(\theta) \, d\theta = 2\pi^2$. Resumindo temos para $1 \leq n \leq 4$ os valores

$$\omega_1 = 2, \quad \omega_2 = 2\pi, \quad \omega_3 = 4\pi, \quad \omega_4 = 2\pi^2. \tag{6.3.9}$$

Uma fórmula geral para ω_n pode-se deduzir através de indução com ajuda da equação (6.3.8). Para isso definimos

$$c_n := \int_{-\pi/2}^{\pi/2} \cos^{n-1}(\theta) d\theta$$

para $n \in \mathbb{N}$. Então vale $\omega_{n+1} = c_n \omega_n$ para todos os $n \in \mathbb{N}$. Para $n \geq 3$ resulta via integração parcial com $f(\theta) = \cos^{n-2}(\theta)$ e $g(\theta) = \text{sen}(\theta)$ a fórmula

$$\begin{aligned} c_n &= \int_{-\pi/2}^{\pi/2} \cos^{n-1}(\theta) d\theta \\ &= \int_{-\pi/2}^{\pi/2} f(\theta)g'(\theta) d\theta \\ &= - \int_{-\pi/2}^{\pi/2} f'(\theta)g(\theta) d\theta \\ &= (n-2) \int_{-\pi/2}^{\pi/2} \cos^{n-3}(\theta) \text{sen}^2(\theta) d\theta \\ &= (n-2)(c_{n-2} - c_n). \end{aligned}$$

Disso segue

$$c_n = \frac{n-2}{n-1} c_{n-2}$$

para todos os $n \geq 3$ e daí, com $c_1 = \pi$ e $c_2 = 2$,

$$\begin{aligned} c_{2k} &= \frac{2k-2}{2k-1} \frac{2k-4}{2k-3} \cdots \frac{2}{3} 2 = \frac{2^k(k-1)!}{3 \cdot 5 \cdots (2k-1)} \\ c_{2k+1} &= \frac{2k-1}{2k} \frac{2k-3}{2k-2} \cdots \frac{1}{2} \pi = \frac{3 \cdot 5 \cdots (2k-1)}{2^k k!} \pi. \end{aligned}$$

Disso segue $c_{2k+1}c_{2k} = \frac{\pi}{k}$ e $c_{2k}c_{2k-1} = \frac{2\pi}{2k-1}$ e por isso valem $\omega_{2k+2} = \frac{\pi}{k}\omega_{2k}$ e $\omega_{2k+1} = \frac{2\pi}{2k-1}\omega_{2k-1}$. Com $\omega_1 = 2$ e $\omega_2 = 2\pi$ recebemos

$$\omega_{2k} = \frac{2\pi^k}{(k-1)!}, \quad \omega_{2k+1} = \frac{2^{k+1}\pi^k}{3 \cdot 5 \cdots (2k-1)} \quad (6.3.10)$$

para todos os $k \in \mathbb{N}$. A comparação desta fórmula com (5.6.2) em Exemplo 5.6.15 da a relação $\text{Vol}_n(B^n) = \text{Vol}_{n-1}(\mathbb{S}^{n-1})/n$ entre o volume da bola unitária e da esfera unitária no \mathbb{R}^n . Vamos deduzir mais uma vez esta relação em Exemplo 6.3.8 num modo diferente e consideravelmente mais simples.

Exemplo 6.3.8 (Fubini para coordenadas polares).

Seja $\psi_r : V \rightarrow \mathbb{R}^n$ a parametrização de um domínio de carta na esfera do raio r em (6.3.4) e defina a aplicação $\psi : (0, \infty) \times V \rightarrow \mathbb{R}^n$ através da fórmula $\psi(r, x) := \psi_r(x)$ para $r > 0$ e $x \in V$. Esta aplicação é um difeomorfismo ao subconjunto aberto $U := \{x \in \mathbb{R}^n \mid x_2 = 0 \implies x_1 > 0\}$. Seu complemento $N := \mathbb{R}^n \setminus U = \{x \in \mathbb{R}^n \mid x_2 = 0, x_1 \leq 0\}$ é um subconjunto fechado ilimitado. Como ilimitado, o conjunto N não pode ser um conjunto Jordan-nulo, porém o que importa é que N intersecta todo subconjunto compacto do \mathbb{R}^n num conjunto Jordan-nulo. Segundo o Exemplo 6.3.6, a matriz jacobiana de ψ tem o determinante

$$\det(d\psi(r, \theta)) = r^{n-1}g(\theta)$$

para todos os $r > 0$ e todos os $\theta \in V$, onde a função $g : V \rightarrow (0, \infty)$ é dada através de (6.3.5). Sejam ora $0 < a < b$ números reais e

$$I := [a, b], \quad A := \{x \in \mathbb{R}^n \mid a \leq \|x\| \leq b\}.$$

Então resulta da equação (6.3.6) em Exemplo 6.3.6 com ajuda do Teorema de Fubini (Teorema 5.4.1) e a fórmula da transformação (Teorema 5.7.11) para toda função contínua $f : \mathbb{R}^n \rightarrow \mathbb{R}$ a equação

$$\begin{aligned} \int_A f(x) dx &= \int_{I \times V} f(\psi(r, \theta)) \det(d\psi(r, \theta)) dr d\theta \\ &= \int_a^b \left(\int_V f(\psi(r, \theta)) r^{n-1} g(\theta) d\theta \right) dr \\ &= \int_a^b \left(r^{n-1} \int_{\mathbb{S}^{n-1}} f(rx) dS(x) \right) dr \\ &= \int_a^b \left(\int_{S_r^{n-1}} f dS \right) dr. \end{aligned} \tag{6.3.11}$$

Com $f \equiv 1$ e a transição ao limite $a \rightarrow 0$ e $b = 1$ resulta da equação (6.3.11) a fórmula

$$\begin{aligned} \text{Vol}_n(B^n) &= \int_0^1 \text{Vol}_{n-1}(\mathbb{S}_r^{n-1}) dr \\ &= \int_0^1 r^{n-1} dr \text{Vol}_{n-1}(\mathbb{S}^{n-1}) \\ &= \frac{1}{n} \text{Vol}_{n-1}(\mathbb{S}^{n-1}). \end{aligned} \tag{6.3.12}$$

Exemplo 6.3.9 (Volume esférica e Gamma-função).

Se $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função contínua a qual é imprópria Riemann-integrável ao longo de todo o \mathbb{R}^n , então resulta da equação (6.3.11) com a transição ao limite $a \rightarrow 0$ e $b \rightarrow \infty$ a fórmula

$$\int_{\mathbb{R}^n} f(x) dx = \int_0^\infty \left(\int_{S_r^{n-1}} f dS \right) dr. \quad (6.3.13)$$

Esta integral torna notadamente simples, se f é dado através de $f(x) = h(\|x\|)$ para uma função $h : (0, \infty) \rightarrow \mathbb{R}$, e assim é constante em toda esfera S_r^{n-1} . Então vale

$$\int_{\mathbb{R}^n} h(\|x\|) dx = \omega_n \int_0^\infty r^{n-1} h(r) dr. \quad (6.3.14)$$

Para a função $f(x) := e^{-\|x\|^2}$ resulta da equação (6.3.14) junto com (5.8.11) no Exemplo 5.8.9 a fórmula

$$\begin{aligned} \pi^{n/2} &= \int_{\mathbb{R}^n} e^{-\|x\|^2} dx \\ &= \omega_n \int_0^\infty r^{n-1} e^{-r^2} dr \\ &= \frac{\omega_n}{2} \int_0^\infty (r^2)^{\frac{n}{2}-1} e^{-r^2} 2r dr \\ &= \frac{\omega_n}{2} \int_0^\infty s^{\frac{n}{2}-1} e^{-s} ds \\ &= \frac{\omega_n}{2} \Gamma\left(\frac{n}{2}\right). \end{aligned}$$

Disso segue

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}. \quad (6.3.15)$$

Utiliza-se ora a fórmula de recursão

$$\Gamma(x+1) = x\Gamma(x)$$

para a Gamma-função, assim como os valores $\Gamma(1) = 1$ e $\Gamma(1/2) = \sqrt{\pi}$, então recebe-se uma terceira prova para a fórmula (6.3.10) para ω_n . (A primeira prova é a dedução em Exemplo 6.3.7, a segunda prova segue da combinação da equação (5.6.2) em Exemplo 5.6.15 com a equação (6.3.12) em Exemplo 6.3.8.)

Exemplo 6.3.10 (Integração sobre superfícies de rotação).

Sejam $z_0 < z_1$ números reais e seja $f : (z_0, z_1) \rightarrow (0, \infty)$ uma função suave. Então a superfícies de rotação

$$M := \{(x, y, z) \in \mathbb{R}^3 \mid z_0 < z < z_1, x^2 + y^2 = f(z)^2\} \quad (6.3.16)$$

é uma subvariedade suave 2-dimensional do \mathbb{R}^3 . Provaremos que seu conteúdo de área é dado pela fórmula

$$\text{Vol}_2(M) = 2\pi \int_{z_0}^{z_1} f(z) \sqrt{1 + f'(z)^2} dz. \quad (6.3.17)$$

Este conteúdo de área também pode ser infinito.

Para a prova consideramos os domínios de carta

$$M^\pm := \{(x, y, z) \in M \mid \pm x > 0\}.$$

Então M^+ é o gráfico da função $h : V \rightarrow \mathbb{R}$ no domínio

$$V := \{(y, z) \in \mathbb{R}^2 \mid z_0 < z < z_1, -f(z) < y < f(z)\}$$

a qual é definida através de

$$h(y, z) := \sqrt{f(z)^2 - y^2}$$

para $(y, z) \in V$. As derivadas parciais de h são

$$\frac{\partial h}{\partial y} = -\frac{y}{\sqrt{f(z)^2 - y^2}} \quad \frac{\partial h}{\partial z} = \frac{f(z)f'(z)}{\sqrt{f(z)^2 - y^2}},$$

e daí segue

$$1 + \|\nabla h(y, z)\|^2 = 1 + \frac{y^2 + f(z)^2 f'(z)^2}{f(z)^2 - y^2} = \frac{1 + f'(z)^2}{1 - (y/f(z))^2}.$$

Com isso resulta da equação (6.3.3) em Exemplo 6.3.5 a fórmula

$$\begin{aligned} \text{Vol}_2(M^+) &= \int_V \sqrt{\frac{1 + f'(z)^2}{1 - (y/f(z))^2}} dy dz \\ &= \int_{z_0}^{z_1} \left(\int_{-f(z)}^{f(z)} \frac{dy/f(z)}{\sqrt{1 - (y/f(z))^2}} \right) f(z) \sqrt{1 + f'(z)^2} dz \\ &= \int_{-1}^1 \frac{dt}{\sqrt{1 - t^2}} \int_{z_0}^{z_1} f(z) \sqrt{1 + f'(z)^2} dz. \end{aligned}$$

Com $t = \text{sen}(\theta)$ para $-\pi/2 < \theta < \pi/2$ resulta a fórmula

$$\frac{dt}{d\theta} = \cos(\theta) = \sqrt{1 - \text{sen}(\theta)^2} = \sqrt{1 - t^2}$$

e dali segue

$$\int_{-1}^1 \frac{dt}{\sqrt{1 - t^2}} = \int_{-\pi/2}^{\pi/2} d\theta = \pi.$$

Portanto vale

$$\text{Vol}_2(M^+) = \pi \int_{z_0}^{z_1} f(z) \sqrt{1 + f'(z)^2} dz.$$

Como $\text{Vol}_2(M^+) = \text{Vol}_2(M^-)$ e $M \setminus (M^+ \cup M^-)$ é um conjunto Jordan-nulo com respeito a M (união de dois gráficos), segue equação (6.3.17).

Exemplo 6.3.11 (Superfície da bola). Consideramos em Exemplo 6.3.10 o caso especial $f(z) = \sqrt{1 - z^2}$ com $-1 < z_0 < z_1 < 1$, então $M = S(z_0, z_1)$ é aquela parte da superfície da bola unitária no \mathbb{R}^3 a qual é encurtado dos meridianos nas alturas z_0 e z_1 . A derivada de f é $f'(z) = -z/\sqrt{1 - z^2}$. Dali segue $1 + f'(z)^2 = 1/(1 - z^2)$ e daí $f(z)\sqrt{1 + f'(z)^2} = 1$. Com isso resulta da equação (6.3.17) a fórmula

$$\text{Vol}_2(S(z_0, z_1)) = 2\pi(z_1 - z_0). \quad (6.3.18)$$

Ou seja, o conteúdo de área do segmento $S(z_0, z_1)$ da superfície da bola é proporcional à diferença de altura $z_1 - z_0$ (e que o fator de proporcionalidade deve ser então precisamente 2π , segue da consideração de um segmento perto do equador). Esta observação origina de Arquimedes. Com $z_0 = -1$ e $z_1 = 1$ resulta o conteúdo de área conhecido $\omega_3 = 4\pi$ da superfície da bola unitária no \mathbb{R}^3 (equação (6.3.9)).

Exemplo 6.3.12 (Superfície do cone). Mais um caso especial é a superfície $M = K(r, h)$ de um cone da altura h sobre uma área circular do raio r . Este caso corresponde a $z_0 = 0$, $z_1 = h > 0$ e $f(z) = r(1 - z/h)$. A equação (6.3.17) fornece então a fórmula

$$\begin{aligned} \text{Vol}_2(K(r, h)) &= 2\pi \int_0^h r \left(1 - \frac{z}{h}\right) \sqrt{1 + \frac{r^2}{h^2}} dz \\ &= \pi r \sqrt{h^2 + r^2} \int_0^h 2 \left(1 - \frac{z}{h}\right) \frac{dz}{h} \\ &= \pi r \sqrt{h^2 + r^2} \end{aligned} \quad (6.3.19)$$

para a superfície do cone.

6.4 O Teorema de Gauß

O Teorema da divergência de Gauß é o teorema mais importante deste capítulo. É uma generalização do Teorema Fundamental do Cálculo para funções de várias variáveis. O teorema cria uma relação entre a integral da divergência de um campo de vetores sobre uma região no \mathbb{R}^n e uma integral adequada sobre o bordo desta região. Nos provamos o teorema para regiões com *bordo suave*, respectivamente com C^ℓ -bordo para um $\ell \in \mathbb{N}$.

Definição 6.4.1 (Bordo suave). *Seja $\Omega \subset \mathbb{R}^n$ um subconjunto suave limitado, seja $G \subset \mathbb{R}^n$ um subconjunto compacto, e seja $\ell \in \mathbb{N} \cup \{\infty\}$.*

(i) Ω é chamado de **conjunto aberto limitado com C^ℓ -bordo** (respectivamente com **bordo suave** no caso de $\ell = \infty$), se $\partial\Omega$ é uma subvariedade $(n-1)$ -dimensional de \mathbb{R}^n e coincide com $\partial\bar{\Omega}$ (ou seja, se todo subconjunto aberto $U \subset \mathbb{R}^n$ com $U \cap \partial\Omega \neq \emptyset$ contém elementos tanto de Ω como de $\mathbb{R}^n \setminus \bar{\Omega}$).

(ii) G é chamado de **conjunto compacto com C^ℓ -bordo** (respectivamente com **bordo suave** no caso de $\ell = \infty$), se ∂G é uma C^ℓ -subvariedade $(n-1)$ -dimensional de \mathbb{R}^n e coincide com $\partial\mathring{G}$ (ou seja, se todo subconjunto aberto $U \subset \mathbb{R}^n$ com $U \cap \partial G \neq \emptyset$ contém elementos tanto de \mathring{G} como de $\mathbb{R}^n \setminus G$).

Comentário 6.4.2. Para o entendimento da Definição 6.4.1 poderiam ser útil as seguintes anotações.

(i) Seja Ω um subconjunto aberto de um espaço métrico M e seja $G := \bar{\Omega}$. Então vale $\Omega \subset \mathring{G}$ e $\partial G \subset \partial\Omega$, assim como $\Omega = \mathring{G} \iff \partial G = \partial\Omega$.

(ii) Seja G um subconjunto fechado de um espaço métrico M e seja $\Omega := \mathring{G}$. Então vale $\bar{\Omega} \subset G$ e $\partial\Omega \subset \partial G$, assim como $\bar{\Omega} = G \iff \partial\Omega = \partial G$.

Exemplo 6.4.3. (i) Seja $\Omega := \{x \in \mathbb{R}^n \mid 0 < \|x\| < 2, \|x\| \neq 1\}$. Então Ω é aberto e seu bordo é uma subvariedade suave. Embora vale $\partial\bar{\Omega} \subsetneq \partial\Omega$ e por isso Ω não é um “conjunto aberto com bordo suave”. Ainda assim seu fecho $G := \bar{\Omega}$ é um conjunto compacto com bordo suave.

(ii) Seja $B \subset \mathbb{C}$ o disco unitário fechado e defina $G := B \cup [1, 2]$. Então o bordo $\partial G = \mathbb{S}^1 \cup [1, 2] \subset \mathbb{C}$ não é uma subvariedade. Ainda assim seu interior $\Omega := \mathring{G} = \mathring{B}$ é um conjunto aberto com bordo suave.

(iii) Seja $G \subset \mathbb{R}^n$ uma subvariedade suave compacta não-vazia de dimensão $d < n$. Então $\partial G = G$ é uma subvariedade suave, embora vale $\partial\mathring{G} = \emptyset \subsetneq \partial G$ e por isso G não é um “conjunto compacto com bordo suave”.

Vamos caracterizar em Lema 6.4.5 conjuntos abertos com C^ℓ -bordo.

Lema 6.4.4. *Seja $U \subset M$ um subconjunto de um espaço métrico (M, d) , seja $V \subset \mathbb{R}^n$ um subconjunto aberto convexo de \mathbb{R}^n , e seja*

$$\phi = (\phi_1, \dots, \phi_n) : U \rightarrow V$$

um homeomorfismo. Definimos

$$\begin{aligned} U^+ &:= \{p \in U \mid \phi_n(p) > 0\} \\ U^0 &:= \{p \in U \mid \phi_n(p) = 0\} \\ U^- &:= \{p \in U \mid \phi_n(p) < 0\}. \end{aligned} \tag{6.4.1}$$

Seja $\Omega \subset M$ mais um conjunto aberto com

$$U \cap \partial\Omega = U^0. \tag{6.4.2}$$

Então valem as afirmações

$$U^+ \cap \Omega \neq \emptyset \quad \implies \quad U^+ \subset \Omega, \tag{6.4.3}$$

$$U^- \cap \Omega \neq \emptyset \quad \implies \quad U^- \subset \Omega. \tag{6.4.4}$$

Demonstração. Seja $I := [0, 1]$ e sejam

$$p \in U^+ \cap \Omega, \quad x := \phi(p), \quad q \in U^+, \quad y := \phi(q).$$

Então $(1-t)x + ty \in V$ para todos os $t \in I$, porque V é um conjunto convexo, e vale $x_n > 0$ e $y_n > 0$, porque $p, q \in U^+$. Seja $\gamma : I \rightarrow U$ a curva

$$\gamma(t) := \phi^{-1}((1-t)x + ty) \quad \text{para } t \in I.$$

Esta curva é contínua, como ϕ é um homeomorfismo. Demais

$$\phi_n(\gamma(t)) = (1-t)x_n + ty_n > 0 \quad \text{para todos os } t \in I,$$

e disso segue $\gamma(t) \notin \partial\Omega$ para todos os $t \in I$ segundo (6.4.1) e (6.4.2). Daí segue $\gamma(t) \in \Omega \iff \gamma(t) \in \overline{\Omega}$ para todos os $t \in I$, e dali o conjunto

$$A := \{t \in I \mid \gamma(t) \in \Omega\} = \{t \in I \mid \gamma(t) \in \overline{\Omega}\}$$

é tanto I -aberto, como I -fechado. Além disso, este conjunto não é vazio, porque $\gamma(0) = p \in \Omega$ e daí $0 \in A$. Disso segue $A = I$, como I é conexo (Teorema B.2.2). Por isso $1 \in A$, e isso significa $q = \gamma(1) \in \Omega$. Com isso (6.4.3) fica provado. A afirmação (6.4.4) segue ora através de reversão de sinal de ϕ_n . Com isso o Lema 6.4.4 fica provado. \square

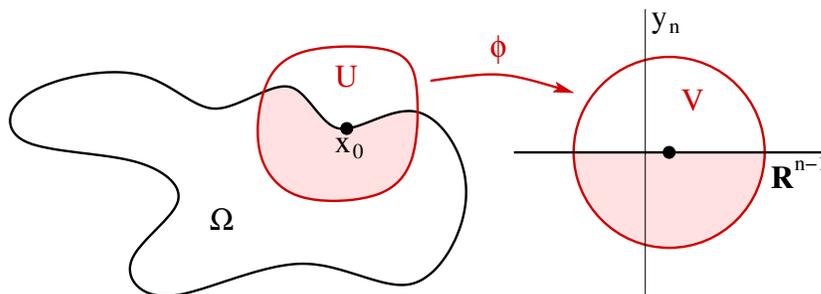


Figura 6.6: Um conjunto aberto com bordo suave.

Lema 6.4.5 (Bordo suave). *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado e seja $\ell \in \mathbb{N} \cup \{\infty\}$. Então as seguintes afirmações são equivalente.*

- (i) Ω é um conjunto aberto limitado com C^ℓ -bordo.
(ii) $\partial\Omega$ é uma C^ℓ -subvariedade $(n-1)$ -dimensional de \mathbb{R}^n e para todo $x_0 \in \partial\Omega$ existe um $\xi_0 \in \mathbb{R}^n$ e um $\delta > 0$ com

$$\xi_0 \notin T_{x_0}\partial\Omega, \quad x_0 + t\xi_0 \notin \Omega \text{ para todos os } t \in [0, \delta]. \quad (6.4.5)$$

- (iii) Para todo $x_0 \in \partial\Omega$ existem conjuntos abertos $U, V \subset \mathbb{R}^n$ com $x_0 \in U$ e um C^ℓ -difeomorfismo $\phi : U \rightarrow V$ (Figura 6.6) com

$$\phi(U \cap \partial\Omega) = \{y \in V \mid y_n = 0\}, \quad \phi(U \cap \Omega) = \{y \in V \mid y_n < 0\}. \quad (6.4.6)$$

- (iv) Para todo $x_0 \in \partial\Omega$ existe um conjunto aberto $U \subset \mathbb{R}^n$ com $x_0 \in U$ e uma C^ℓ -função $g : U \rightarrow \mathbb{R}$ com $\nabla g(x) \neq 0$ para todos os $x \in U \cap \partial\Omega$ e

$$U \cap \partial\Omega = \{x \in U \mid g(x) = 0\}, \quad U \cap \Omega = \{x \in U \mid g(x) < 0\}. \quad (6.4.7)$$

Demonstração. Provaremos (i) \implies (ii). Seja então $\partial\Omega = \partial\bar{\Omega}$ uma C^ℓ -subvariedade $(n-1)$ -dimensional de \mathbb{R}^n e seja $x_0 \in \partial\Omega$. Então existe um C^ℓ -difeomorfismo $\phi : U \rightarrow V$ entre subconjuntos abertos $U, V \subset \mathbb{R}^n$ com $x_0 \in U$, tal que V é convexo e $U \cap \partial\Omega = U^0$ (em notação de Lema 6.4.4). Disso segue que um dos conjuntos $U^+ \cap \Omega$ e $U^- \cap \Omega$ tem que ser vazio, porque senão seja $U \subset \bar{\Omega}$ segundo o Lema 6.4.4, e assim x_0 não seja um ponto do bordo de $\bar{\Omega}$. Se $U^+ \cap \Omega = \emptyset$, então todo vetor $\xi_0 \in \mathbb{R}^n$ com $d\phi_n(x_0)\xi_0 > 0$ satisfaz a condição (6.4.5). Se $U^- \cap \Omega = \emptyset$, então todo vetor $\xi_0 \in \mathbb{R}^n$ com $d\phi_n(x_0)\xi_0 < 0$ satisfaz a condição (6.4.5). Então Ω satisfaz a condição (ii).

Provaremos (ii) \implies (iii). Seja $x_0 \in \partial\Omega$. Como $\partial\Omega$ é uma C^ℓ -subvariedade $(n-1)$ -dimensional, existem conjuntos abertos $U, V \subset \mathbb{R}^n$ com $x_0 \in U$ e um C^1 -difeomorfismo $\phi : U \rightarrow V$ com

$$\phi(U \cap \partial\Omega) = \{y \in V \mid y_n = 0\}$$

ou seja $U \cap \partial\Omega = U^0$ em notação de Lema 6.4.4. Além disso, vale

$$T_{x_0}\partial\Omega = d\phi(x_0)^{-1}(\mathbb{R}^{n-1} \times \{0\}).$$

Através de diminuir U, V se necessário, podemos supor que V é convexo. Ora, sejam $\xi_0 \in \mathbb{R}^n$ e $\delta > 0$ escolhido tal que vale (6.4.5). Como $\xi_0 \notin T_{x_0}\partial\Omega$, vale $d\phi(x_0)\xi_0 \notin \mathbb{R}^{n-1} \times \{0\}$. Dali $d\phi_n(x_0)\xi_0 \neq 0$ e, no que substituímos ϕ_n por $-\phi_n$ se necessário, podemos supor que

$$d\phi_n(x_0)\xi_0 > 0.$$

Disso segue $\phi_n(x_0 + t\xi_0) > 0$ para pequenos $t > 0$. Se $0 < t_0 \leq \delta$ seja tão pequena que $x_0 + t_0\xi_0 \in U$ e $\phi_n(x_0 + t_0\xi_0) > 0$, então $x_0 + t_0\xi_0 \in U^+ \setminus \Omega$ na notação de Lema 6.4.4. Então $U^+ \not\subset \Omega$ e daí $U^+ \cap \Omega = \emptyset$ segundo Lema 6.4.4. Como $(U^+ \cup U^0) \cap \Omega = \emptyset$, a interseção $U^- \cap \Omega \neq \emptyset$ tem que ser vazia, e disso segue $U^- \subset \Omega$ segundo o Lema 6.4.4. Portanto

$$U \cap \Omega = U^-$$

e com isso vale (ii).

A implicação (iii) \implies (iv) segue imediato com $g := \phi_n$.

Provaremos (iv) \implies (i). Seja $x_0 \in \partial\Omega$ e seja $g : U \rightarrow \mathbb{R}$ como em (iv). Então a interseção de $\partial\Omega$ com uma vizinhança suficientemente pequena de x_0 é a pre-imagem de um valor regular sob uma C^ℓ -função real, e daí uma C^ℓ -subvariedade $(n-1)$ -dimensional. Além disso, vale

$$x_0 + t\nabla g(x_0) \in U, \quad g(x_0 + t\nabla g(x_0)) > 0$$

para todo $t > 0$ suficientemente pequeno. Portanto $x_0 + t\nabla g(x_0) \notin \bar{\Omega}$ para todo tal t segundo (6.4.7), e dali segue $x_0 = \lim_{t \searrow 0} (x_0 + t\nabla g(x_0)) \in \partial\bar{\Omega}$. Com isso temos mostrado que $\partial\Omega \subset \partial\bar{\Omega}$. A inclusão reversa vale para todo conjunto aberto Ω segundo a parte (i) de Comentário 6.4.2. Portanto $\partial\Omega = \partial\bar{\Omega}$ e com isso Ω satisfaz a condição (i). Com isso o Lema 6.4.5 fica provado. \square

Lema 6.4.6. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado com C^ℓ -bordo. Então existe um conjunto aberto $U \subset \mathbb{R}^n$ com $\partial\Omega \subset U$ e uma C^ℓ -função $g : U \rightarrow \mathbb{R}$ com as propriedades*

$$\nabla g(x) \neq 0 \quad \text{para todos os } x \in \partial\Omega \quad (6.4.8)$$

$$U \cap \partial\Omega = \{x \in U \mid g(x) = 0\} \quad (6.4.9)$$

$$U \cap \Omega = \{x \in U \mid g(x) < 0\}. \quad (6.4.10)$$

Demonstração. Como Ω é limitado, e dali $\partial\Omega$ compacto, existem segundo a definição (parte (iii) de Lema 6.4.5) e Teorema C.1.2, conjuntos abertos U_1, \dots, U_N com $\partial\Omega \subset \bigcup_{i=1}^N U_i$ e C^ℓ -funções $g_i : U_i \rightarrow \mathbb{R}$ para $i = 1, \dots, N$, as quais tem 0 como valor regular e satisfazem (6.4.7). Segundo o Teorema 6.2.4 existem funções suaves $\rho_i : \mathbb{R}^n \rightarrow [0, 1]$ com suporte compacto, as quais satisfazem as condições $\text{supp}(\rho_i) \subset U_i$ para $i = 1, \dots, N$ e

$$\sum_{i=1}^N \rho_i(x) = 1 \quad \text{para todos os } x \in \partial\Omega.$$

Então

$$U := \left\{ x \in \mathbb{R}^n \mid \sum_{i=1}^N \rho_i(x) > 0 \right\} \subset \bigcup_{i=1}^N U_i$$

é um conjunto aberto, o qual contém $\partial\Omega$, e a função

$$g := \sum_{i=1}^N \rho_i g_i : U \rightarrow \mathbb{R}$$

satisfaz as afirmações do lema. Se $x \in U \cap \partial\Omega$, então todo somando $\rho_i(x)g_i(x)$ com $x \in U_i$ iguale nulo. Se $x \in U \cap \Omega$, então todo somando $\rho_i(x)g_i(x)$ com $x \in U_i$ é menor ou igual nulo e pelo menos um é negativo. Se $x \in U \setminus \bar{\Omega}$, então todo somando $\rho_i(x)g_i(x)$ com $x \in U_i$ é maior ou igual nulo e pelo menos um é positivo. Disso seguem as equações (6.4.9) e (6.4.10). A afirmação (6.4.8) vale, porque $g_i(x) = 0$ para todos os $x \in U_i \cap \partial\Omega$ e os vetores $\nabla g_i(x)$ e $\nabla g_j(x)$ para $x \in U_i \cap U_j \cap \partial\Omega$ são múltiplos positivos um do outro. Com isso temos provado o Lema 6.4.6. \square

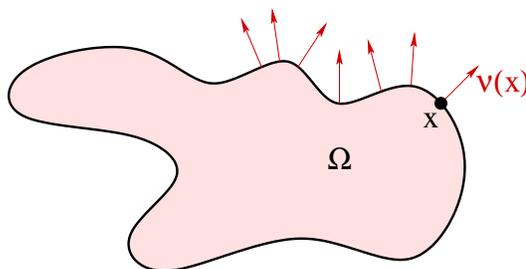


Figura 6.7: O campo de normais unitárias exteriores.

Lema 6.4.7 (O campo de normais unitárias exteriores). *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado com C^1 -bordo. Então vale o seguinte.*

(i) *Para todo $x \in \partial\Omega$ existe exatamente um $\nu(x) \in \mathbb{R}^n$ (Figura 6.7) com*

$$\begin{aligned} \|\nu(x)\| &= 1, & \nu(x) &\perp T_x\partial\Omega, \\ x + t\nu(x) &\notin \Omega \text{ para pequenos } t > 0. \end{aligned} \quad (6.4.11)$$

(ii) *A aplicação $\nu : \partial\Omega \rightarrow \mathbb{S}^{n-1}$ em (i) é contínua.*

(iii) *Se Ω tem um bordo de classe C^ℓ , então existe um conjunto aberto $U \subset \mathbb{R}^n$ com $\partial\Omega \subset U$, tal que pode-se estender a aplicação $\nu : \partial\Omega \rightarrow \mathbb{S}^{n-1}$ a uma $C^{\ell-1}$ -aplicação de U a \mathbb{S}^{n-1} .*

Demonstração. Seja $g : U \rightarrow \mathbb{R}$ uma função que satisfaça as condições de Lema 6.4.6. Então 0 é um valor regular de g com $\partial\Omega = g^{-1}(0)$ segundo (6.4.8) e (6.4.9). Daí vale $T_x\partial\Omega = \nabla g(x)^\perp$ para todos os $x \in \partial\Omega$. A existência em parte (i) segue de Lema 6.4.6 com

$$\nu(x) := \frac{\nabla g(x)}{\|\nabla g(x)\|}. \quad (6.4.12)$$

Este vetor tem a norma um, e é ortogonal a $T_x\partial\Omega$. Demais $\langle \nabla g(x), \nu(x) \rangle = \|\nabla g(x)\| > 0$ e daí segue $g(x + t\nu(x)) > 0$, e daí $x + t\nu(x) \notin \Omega$ por (6.4.10), para todo $t > 0$ suficientemente pequeno. Também segue $g(x_0 - t\nu(x_0)) < 0$ e dali $x_0 - t\nu(x_0) \in \Omega$ para todo $t > 0$ suficientemente pequeno. Como existem só dois vetores unitários em $(T_{x_0}\partial\Omega)^\perp$, com isso parte (i) fica provado. A continuidade em parte (ii) também segue diretamente da fórmula (6.4.12) e da continuidade das derivadas parciais de g . Em parte (iii) a existência de uma $C^{\ell-1}$ -extensão também segue diretamente da fórmula (6.4.12), como a função g pode ser escolhido segundo o Lema 6.4.6, tal que é ℓ vezes continuamente diferenciável. Com isso Lema 6.4.7 fica provado. \square

Definição 6.4.8. *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado com C^1 -bordo. A função $\nu : \partial\Omega \rightarrow \mathbb{S}^{n-1}$ em Lema 6.4.7 é chamado do **campo exterior de normais unitários** de Ω . Também é chamada da **aplicação de Gauß** de Ω .*

Teorema 6.4.9 (Teorema da Divergência de Gauß). *Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado com C^1 -bordo, seja $\nu : \partial\Omega \rightarrow \mathbb{S}^{n-1}$ o campo exterior de normais unitários, seja $W \subset \mathbb{R}^n$ um conjunto aberto com $\bar{\Omega} \subset W$, seja $f : W \rightarrow \mathbb{R}^n$ uma C^1 -aplicação, e seja*

$$\operatorname{div}(f) := \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} : W \rightarrow \mathbb{R}, \quad \langle f, \nu \rangle := \sum_{i=1}^n f_i \nu_i : \partial\Omega \rightarrow \mathbb{R}.$$

Então vale

$$\int_{\Omega} \operatorname{div}(f) = \int_{\partial\Omega} \langle f, \nu \rangle dS. \tag{6.4.13}$$

Demonstração. Primeiro seja mencionado que de Ω trata-se de um conjunto Jordan-mensurável do \mathbb{R}^n , porque seu bordo é uma C^ℓ -subvariedade compacta $(n - 1)$ -dimensional de \mathbb{R}^n e por isso, segundo o Exemplo 5.6.13, é um conjunto Jordan-nulo. Como f é contínuo em uma vizinhança do conjunto $\bar{\Omega}$, portanto a integral no lado esquerdo da equação (6.4.13) é bem definido (Definição 5.5.2).

Provaremos o teorema de divergência para já para funções com suporte em uma vizinhança suficientemente pequena de um ponto do bordo. Seja então dado $x_0 \in \partial\Omega$. Então existe segundo a hipótese (parte (iii) de Lema 6.4.5) um conjunto aberto $U \subset W$ com $x_0 \in U$ e uma C^1 -aplicação $g : U \rightarrow \mathbb{R}$ com $\nabla g(x_0) \neq 0$ e

$$U \cap \partial\Omega = \{x \in U \mid g(x) = 0\}, \quad U \cap \Omega = \{x \in U \mid g(x) < 0\}.$$

Como $\nabla g(x_0)$ é não-nulo, podemos supor sem perda de generalidade que $\partial g / \partial x_n(x_0) > 0$. Mais detalhado, plenamente é possível que $\partial g / \partial x_n(x_0) = 0$; neste caso existe um $i \in \{1, \dots, n - 1\}$ com $\partial g / \partial x_i(x_0) \neq 0$; através de permutar as coordenadas x_i e x_n recebemos então $\partial g / \partial x_n(x_0) \neq 0$; de seguida podemos ainda substituir x_n por $-x_n$, se necessário, para receber a desigualdade $\partial g / \partial x_n(x_0) > 0$. Depois estas preparações g satisfaz as condições

$$g(x_0) = 0, \quad \frac{\partial g}{\partial x_n}(x_0) > 0. \tag{6.4.14}$$

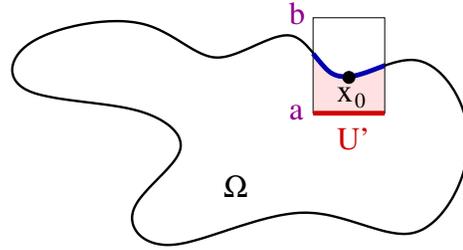


Figura 6.8: Uma vizinhança de um ponto de bordo.

Neste ponto é útil introduzir a notação

$$x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, \quad x = (x_1, \dots, x_n) = (x', x_n) \in \mathbb{R}^n.$$

Particularmente $x'_0 \in \mathbb{R}^{n-1}$ é o vetor cujas coordenadas coincidem com as primeiras $n - 1$ coordenadas de x_0 e $x_{0n} \in \mathbb{R}$ é a última coordenada de x_0 , ou seja $x_0 = (x'_0, x_{0n})$. Então segue das condições em (6.4.14) que existem números reais $a < b$ com

$$a < x_{0n} < b, \quad \{x'_0\} \times [a, b] \subset U, \quad \frac{\partial g}{\partial x_n}(x'_0, x_n) > 0 \text{ para todos os } x_n \in [a, b].$$

Como $g(x'_0, x_{0n}) = g(x_0) = 0$, então segundo o Teorema do Valor Médio vale

$$g(x'_0, a) < 0 < g(x'_0, b).$$

Por isso existe um conjunto aberto $U' \subset \mathbb{R}^{n-1}$ (Figura 6.8) com

$$\begin{aligned} x'_0 \in U', \quad U' \times [a, b] \subset U \\ \frac{\partial g}{\partial x_n}(x) > 0 \quad \text{para todos os } x \in U' \times [a, b] \\ g(x', a) < 0 < g(x', b) \quad \text{para todos os } x' \in U'. \end{aligned} \quad (6.4.15)$$

De (6.4.15) segue segundo o Teorema do Valor Intermediário e o Teorema do Valor Médio que para todo $x' \in U'$ existe exatamente um número real $h(x')$ com

$$a < h(x') < b, \quad g(x', h(x')) = 0. \quad (6.4.16)$$

Segundo o Teorema da Função Implícita em Capítulo 3, a função $h : U' \rightarrow (a, b)$ definida assim é continuamente diferenciável (veja a Figura 6.9).

O vetor unitário exterior no ponto $(x', h(x')) \in \partial\Omega$ nesta situação é o vetor unitário único, o qual é ortogonal ao gráfico da derivada de h no ponto $x' \in U'$, e cujo n -ésima coordenada é positiva. Isto é o vetor $\nu(x', h(x'))$ de \mathbb{R}^n com as coordenadas

$$\begin{aligned} \nu_i(x', h(x')) &= -\frac{\partial_i h(x')}{\sqrt{1 + \|\nabla h(x')\|^2}} && \text{para } i = 1, \dots, n-1 \\ \nu_n(x', h(x')) &= \frac{1}{\sqrt{1 + \|\nabla h(x')\|^2}}. \end{aligned} \tag{6.4.17}$$

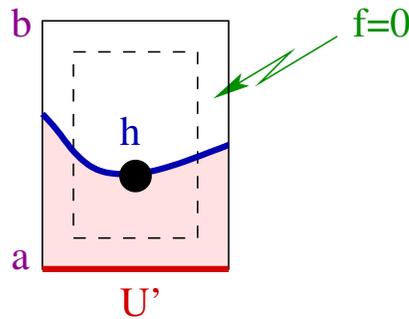


Figura 6.9: O bordo como gráfico de uma função $h : U' \rightarrow (a, b)$.

Ora, seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ um campo de vetores continuamente diferenciáveis, cujo suporte é contido no conjunto aberto $U' \times (a, b)$. Então a integral da função $\langle f, \nu \rangle : \partial\Omega \rightarrow \mathbb{R}$ é dada segundo (6.4.17) e Exemplo 6.3.5 pela fórmula

$$\begin{aligned} &\int_{\partial\Omega} \langle f, \nu \rangle dS \\ &= \int_{U'} \sum_{i=1}^n f_i(x', h(x')) \nu_i(x', h(x')) \sqrt{1 + \|\nabla h(x')\|^2} dx' \\ &= \int_{U'} \left(f_n(x', h(x')) - \sum_{i=1}^{n-1} f_i(x', h(x')) \frac{\partial h}{\partial x_i}(x') \right) dx' \end{aligned} \tag{6.4.18}$$

e a integral de $\text{div}(f)$ sobre Ω é segundo o Teorema de Fubini (Teorema 5.4.1) a integral iterada

$$\int_{\Omega} \text{div}(f) = \sum_{i=1}^n \int_{U'} \int_a^{h(x')} \frac{\partial f_i}{\partial x_i}(x', x_n) dx_n dx'. \tag{6.4.19}$$

A mostrar é ora que as duas integrais em (6.4.18) e (6.4.19) coincidem. Para $i = n$ recebemos segundo o Teorema Fundamental do Cálculo a seguinte fórmula

$$\begin{aligned} \int_a^{h(x')} \frac{\partial f_n}{\partial x_n}(x', x_n) dx_n &= f_n(x', h(x')) - f_n(x', a) \\ &= f_n(x', h(x')). \end{aligned}$$

Através de integração sobre $x' \in U'$ resulta disso

$$\int_{\Omega} \frac{\partial f_n}{\partial x_n} = \int_{\partial\Omega} f_n \nu_n dS.$$

Para $i = 1, \dots, n-1$ e todo número real t o Teorema Fundamental do Cálculo fornece a fórmula

$$\int_{U'} \frac{\partial}{\partial x_i} f_i(x', t + h(x')) dx' = 0.$$

Através de integração sobre $t \in (-\infty, 0]$ encontraremos então

$$\begin{aligned} 0 &= \int_{-\infty}^0 \int_{U'} \frac{\partial}{\partial x_i} f_i(x', t + h(x')) dx' dt \\ &= \int_{-\infty}^0 \int_{U'} \left(\frac{\partial f_i}{\partial x_i}(x', t + h(x')) + \frac{\partial f_i}{\partial x_n}(x', t + h(x')) \frac{\partial h}{\partial x_i}(x') \right) dx' dt \\ &= \int_{U'} \int_{-\infty}^0 \frac{\partial f_i}{\partial x_i}(x', t + h(x')) dt dx' \\ &\quad + \int_{U'} \left(\int_{-\infty}^0 \frac{\partial}{\partial t} f_i(x', t + h(x')) dt \right) \frac{\partial h}{\partial x_i}(x') dx' \\ &= \int_{U'} \int_a^{h(x')} \frac{\partial f_i}{\partial x_i}(x', x_n) dx_n dx' \\ &\quad + \int_{U'} f_i(x', h(x')) \frac{\partial h}{\partial x_i}(x') dx' \\ &= \int_{\Omega} \frac{\partial f_i}{\partial x_i} - \int_{\partial\Omega} f_i \nu_i dS. \end{aligned}$$

Com isso o Teorema de divergência de Gauß fica provado para todo campo de vetores continuamente diferenciável $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ com $\text{supp}(f) \subset U' \times (a, b)$.

Resumindo temos mostrado que todo ponto do bordo de Ω possui uma vizinhança aberta $U \subset W$ tal que o Teorema de divergência de Gauß é válido

para todos os campos de vetores com suporte em U . O mesmo vale para pontos x_0 em Ω , como ambas integrais em (6.4.13) anulam-se para C^1 -funções cujos suportes são contidos numa vizinhança aberta $U(x_0) \subset \Omega$ suficientemente pequena. Como $\bar{\Omega}$ é um subconjunto compacto do \mathbb{R}^n , segue disso segundo o Teorema C.1.2 que existem um número finito de subconjuntos abertos $U_1, \dots, U_N \subset W$, tal que

$$\bar{\Omega} \subset U_1 \cup U_2 \cup \dots \cup U_N$$

e tal que vale a equação (6.4.13) para todo $i \in \{1, \dots, N\}$ e todo campo de vetores continuamente diferenciável $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ com $\text{supp}(f) \subset U_i$. Segundo o Teorema 6.2.4 existem funções suaves

$$\rho_1, \dots, \rho_N : \mathbb{R}^n \rightarrow [0, 1]$$

com suporte compacto, as quais satisfazem as condições

$$\text{supp}(\rho_i) \subset U_i \quad \text{para } i = 1, \dots, N$$

e

$$\sum_{i=1}^N \rho_i(x) = 1 \quad \text{para todos os } x \in \bar{\Omega}.$$

Se ora $f : W \rightarrow \mathbb{R}^n$ é um campo de vetores continuamente diferenciável qualquer, então vale (6.4.13) para cada um dos campos de vetores $\rho_i f$ e dali segue

$$\begin{aligned} \int_{\Omega} \text{div}(f) &= \int_{\Omega} \text{div} \left(\sum_{i=1}^N \rho_i f \right) \\ &= \sum_{i=1}^N \int_{\Omega} \text{div}(\rho_i f) \\ &= \sum_{i=1}^N \int_{\partial\Omega} \langle \rho_i f, \nu \rangle dS \\ &= \int_{\partial\Omega} \sum_{i=1}^N \rho_i \langle f, \nu \rangle dS \\ &= \int_{\partial\Omega} \langle f, \nu \rangle dS. \end{aligned}$$

Com isso o Teorema 6.4.9 fica provado. □

6.5 Mais exemplos

Exemplo 6.5.1 (Operador de Laplace – Laplaciano). O operador de Laplace, ou o Laplaciano, no \mathbb{R}^n é o operador diferencial de segunda ordem

$$\Delta := \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}. \quad (6.5.1)$$

Seja ora $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado com bordo suave.

Uma C^2 -função real $u : U \rightarrow \mathbb{R}$ definida num subconjunto aberto $U \subset \mathbb{R}^n$ chama-se de **harmónica** se $\Delta u = 0$. Para toda C^2 -função $u : U \rightarrow \mathbb{R}$ e toda C^1 -função $v : U \rightarrow \mathbb{R}$ vale

$$\operatorname{div}(v\nabla u) = v\Delta u + \langle \nabla v, \nabla u \rangle. \quad (6.5.2)$$

Para $\ell \in \mathbb{N}$ chamamos $u : \bar{U} \rightarrow \mathbb{R}$ uma **C^ℓ -função em \bar{U}** , se pode-se estender-la a uma C^ℓ -função em um subconjunto aberto do \mathbb{R}^n contendo \bar{U} .

Se $u : \bar{\Omega} \rightarrow \mathbb{R}$ é uma C^1 -função, então chama-se de **derivada normal** de u a derivada de u em direção do campo exteriores de normais $\nu : \partial\Omega \rightarrow \mathbb{R}$. Denotamos ela de

$$\frac{\partial u}{\partial \nu} := \langle \nabla u, \nu \rangle : \partial\Omega \rightarrow \mathbb{R}.$$

Com isso recebemos de Teorema 6.4.9 e equação (6.5.2) com $v = 1$ e $v = u$ para toda C^2 -função $u : \bar{\Omega} \rightarrow \mathbb{R}$ a fórmula

$$\int_{\Omega} \Delta u = \int_{\partial\Omega} \frac{\partial u}{\partial \nu} dS \quad (6.5.3)$$

e a **primeira identidade de Green**

$$\int_{\Omega} \|\nabla u\|^2 = \int_{\partial\Omega} u \frac{\partial u}{\partial \nu} dS - \int_{\Omega} u \Delta u. \quad (6.5.4)$$

A equação (6.5.4) mostra que toda C^2 -função harmónica $u : \bar{\Omega} \rightarrow \mathbb{R}$, a qual anula-se no bordo, anula-se em Ω inteiro, porque ela é constante em todo componente conexo de $\bar{\Omega}$ e cada uma tal componente conexo contem um ponto de bordo. A equação (6.5.3) mostra que a derivada normal de uma C^2 -função harmónica $u : \bar{\Omega} \rightarrow \mathbb{R}$ tem média nula sobre o bordo. Mais uma equação importante, a qual segue de (6.5.2) e Teorema 6.4.9 para duas C^2 -funções $u, v : \bar{\Omega} \rightarrow \mathbb{R}$, é a **segunda identidade de Green**

$$\int_{\Omega} (u\Delta v - v\Delta u) = \int_{\partial\Omega} \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) dS. \quad (6.5.5)$$

Exemplo 6.5.2 (Volume). A divergência do campo de vetores $f(x) = x$ é a função $\operatorname{div}(f) \equiv n$ constante e com isso vale $\int_{\Omega} \operatorname{div}(f) = n \operatorname{Vol}_n(\Omega)$. Segundo o Teorema 6.4.9 resulta daí para todo conjunto aberto limitado $\Omega \subset \mathbb{R}^n$ com C^1 -bordo a fórmula

$$\operatorname{Vol}_n(\Omega) = \frac{1}{n} \int_{\partial\Omega} \langle x, \nu(x) \rangle dS(x). \quad (6.5.6)$$

Se $\Omega = B^n \subset \mathbb{R}^n$ é a bola unitária aberta, então $\partial\Omega = \mathbb{S}^{n-1}$ é a esfera unitária e neste caso o campo exterior de normais unitários $\nu : \partial\Omega \rightarrow \mathbb{S}^{n-1}$ é a identidade $\nu(x) = x$. Com isso resulta de (6.5.6) a equação

$$\operatorname{Vol}_n(B^n) = \frac{1}{n} \operatorname{Vol}_{n-1}(\mathbb{S}^{n-1})$$

e com assim temos encontrado mais uma prova da equação (6.3.12).

Exemplo 6.5.3. Seja $\Omega \subset \mathbb{R}^n$ um conjunto aberto limitado com C^1 -bordo, seja $a \in \mathbb{R}^n \setminus \partial\Omega$, und sei $f : \mathbb{R}^n \setminus \{a\} \rightarrow \mathbb{R}^n$ o campo de vetores

$$f(x) := \frac{x - a}{\|x - a\|^n}, \quad x \in \mathbb{R}^n \setminus \{a\}.$$

Então vale

$$\frac{\partial f_i}{\partial x_i}(x) = \frac{\partial}{\partial x_i} \frac{x_i - a_i}{\|x - a\|^n} = \frac{1}{\|x - a\|^n} - \frac{n(x_i - a_i)^2}{\|x - a\|^{n+2}}$$

para $i = 1, \dots, n$ e daí

$$(\operatorname{div}(f))(x) = \frac{n}{\|x - a\|^n} - \sum_{i=1}^n \frac{n(x_i - a_i)^2}{\|x - a\|^{n+2}} = 0.$$

Para $a \in \mathbb{R}^n \setminus \overline{\Omega}$ segue disso segundo o Teorema 6.4.9 a equação

$$\int_{\partial\Omega} \frac{\langle x - a, \nu(x) \rangle}{\|x - a\|^n} dS(x) = \int_{\partial\Omega} \langle f, \nu \rangle dS = \int_{\Omega} \operatorname{div}(f) = 0.$$

Se $a \in \Omega$ e $r > 0$ com $\overline{B}_r(a) = \{x \in \mathbb{R}^n \mid \|x - a\| \leq r\} \subset \Omega$, então $\Omega \setminus \overline{B}_r(a)$ é um subconjunto aberto limitado do \mathbb{R}^n com C^1 -bordo

$$\partial(\Omega \setminus \overline{B}_r(a)) = \partial\Omega \cup \partial B_r(a).$$

Para $x \in \partial B_r(a)$ o vetor $\nu(x) = -\|x - a\|^{-1}(x - a)$ é o vetor unitário normal exterior com respeito ao conjunto aberto $\Omega \setminus \overline{B}_r(a)$ (veja a Figura 6.10).

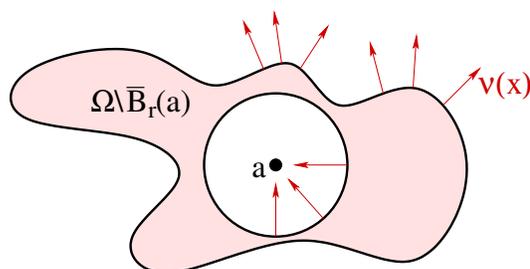


Figura 6.10: O Teorema de Gauß para o complemento de uma bola.

Com isso resulta de Teorema 6.4.9 a equação

$$\begin{aligned}
 0 &= \int_{\Omega \setminus \bar{B}_r(a)} \operatorname{div}(f) \\
 &= \int_{\partial\Omega} \langle f, \nu \rangle dS - \int_{\partial B_r(a)} \frac{\langle f(x), x - a \rangle}{\|x - a\|} dS(x) \\
 &= \int_{\partial\Omega} \langle f, \nu \rangle dS - \omega_n.
 \end{aligned}$$

Resumindo recebemos a fórmula

$$\frac{1}{\omega_n} \int_{\partial\Omega} \frac{\langle x - a, \nu(x) \rangle}{\|x - a\|^n} dS(x) = \begin{cases} 1, & \text{para } a \in \Omega, \\ 0, & \text{para } a \in \mathbb{R}^n \setminus \bar{\Omega}. \end{cases} \quad (6.5.7)$$

Em caso $n = 2$ pode-se ver a integral como *número de circulação* de $\partial\Omega$ em torno de um ponto a no complemento do bordo. (Compare Exemplo 7.6.2.)

Exemplo 6.5.4 (Eletrodinâmica). Sejam dado m pontos $a_1, \dots, a_m \in \mathbb{R}^3$ dois-a-dois diferentes. Suponhamos que encontre-se uma carga elétrica da intensidade $q_i \in \mathbb{R}$ no ponto a_i . Estas cargas geram um campo elétrico o qual opera com a força

$$E(x) = \sum_{i=1}^m q_i \frac{x - a_i}{\|x - a_i\|^3} \quad (6.5.8)$$

em uma carga elétrica unitária positiva no ponto $x \in \mathbb{R}^3 \setminus \{a_1, \dots, a_m\}$. Se $\Omega \subset \mathbb{R}^3$ é um conjunto aberto com C^1 -bordo $\partial\Omega \subset \mathbb{R}^3 \setminus \{a_1, \dots, a_k\}$, então resulta-se de (6.5.7) com $\omega_3 = 4\pi$ a fórmula

$$\int_{\partial\Omega} \langle E, \nu \rangle dS = \sum_{i=1}^m q_i \int_{\partial\Omega} \frac{\langle x - a_i, \nu(x) \rangle}{\|x - a_i\|^3} dS(x) = 4\pi \sum_{a_i \in \Omega} q_i. \quad (6.5.9)$$

O fluxo de carga sobre o bordo de Ω é portanto 4π vezes a carga total na região Ω (Figura 6.11).

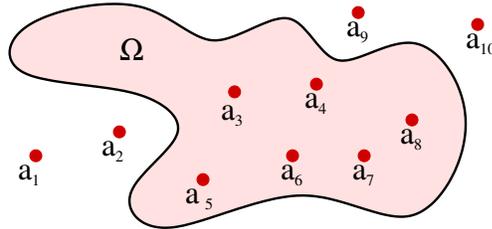


Figura 6.11: Fluxo elétrico de carga.

Exemplo 6.5.5 (Conteúdo de área). Seja $\Omega \subset \mathbb{R}^2$ um conjunto aberto limitado com C^1 -bordo, tal que o bordo $\partial\Omega$ é conexo. Então existe uma C^1 -aplicação $\gamma : \mathbb{R} \rightarrow \partial\Omega$ com

$$\dot{\gamma}(t) \neq 0 \quad \gamma(t+1) = \gamma(t) \quad \text{para todos os } t \in \mathbb{R}$$

tal que a aplicação $\gamma|_{[0,1)} : [0,1) \rightarrow \partial\Omega$ é bijetiva. (Para uma prova veja [3].) Esta parametrização de $\partial\Omega$ pode ser escolhido tal que a curva γ passa pelo bordo de Ω em **sentido anti-horário**. Escrevemos $\gamma(t) =: (x(t), y(t))$, então o vetor unitário normal exterior de Ω no ponto $\gamma(t) \in \partial\Omega$ é dado nestes hipóteses através da fórmula

$$\nu(\gamma(t)) = \frac{(\dot{y}(t), -\dot{x}(t))}{\sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}}$$

(veja a Figura 6.12).

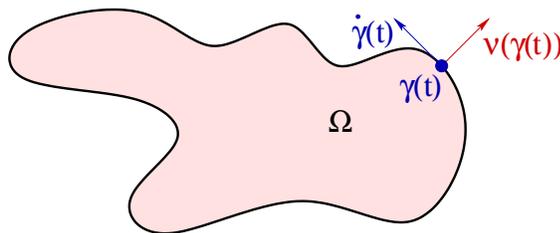


Figura 6.12: O conteúdo de área como integral sobre o bordo.

Como em Exemplo 6.5.2 consideramos o campo de vetores $f(x, y) := (x, y)$ em \mathbb{R}^2 . Então segue de Teorema 6.4.9 com ajuda da integral linha em Exemplo 6.3.3 a equação

$$\begin{aligned}
 \text{Vol}_2(\Omega) &= \frac{1}{2} \int_{\Omega} \text{div}(f) \\
 &= \frac{1}{2} \int_{\partial\Omega} \langle f, \nu \rangle dS \\
 &= \frac{1}{2} \int_0^1 \langle f(\gamma(t)), \nu(\gamma(t)) \rangle \|\dot{\gamma}(t)\| dt \\
 &= \frac{1}{2} \int_0^1 (x(t)\dot{y}(t) - y(t)\dot{x}(t)) dt.
 \end{aligned} \tag{6.5.10}$$

Consideramos o disco unitário aberto $\Omega = B^2 \subset \mathbb{R}^2$ com a parametrização $\gamma(t) := (\cos(2\pi t), \sin(2\pi t))$ do bordo $\partial B^2 = \mathbb{S}^1$, então resulta mais uma vez a fórmula $\text{Vol}_2(B^2) = \pi$ para o conteúdo de área.

Chama-se de “integral da 1-forma $\frac{1}{2}(xdy - ydx)$ ao longo da curva γ ” o último termo em (6.5.10). Esta formulação será tratado com mais detalhe em Capítulo 7. Como vamos ver, é extremamente útil traduzir o Teorema de divergência de Gauß à língua das formas diferenciais. Nem só lida esta tradução a uma generalização natural no Teorema de Stokes (Teorema 7.5.4), mas sim pode-se deduzir certas aplicações muito mais fácil com o formalismo das formas diferenciais, em contraste à formulação com a divergência de campos vetoriais. Um exemplo disso é o teorema de retração (Teorema 7.6.5) no próximo capítulo.

Capítulo 7

Integração de Formas Diferenciais

O objetivo deste capítulo é definir a integral de uma forma diferencial de grau d sobre uma subvariedade $M \subset \mathbb{R}^n$ orientada e d -dimensional. Para isso precisa-se primeiramente uma introdução às noções de formas alternadas (Seção 7.1) e formas diferenciais (Seção 7.2), assim como à noção de orientação de uma subvariedade do \mathbb{R}^n (Seção 7.3). A integral sobre uma subvariedade será introduzido em Seção 7.4. O resultado mais importante neste capítulo é o Teorema de Stokes o qual será provado em Seção 7.5. Neste capítulo consideramos exclusivamente subvariedades suaves só, sem mencionar isso em cada um caso.

7.1 Formas alternadas

Nesta Seção X, Y, Z são espaços vetoriais reais de dimensão finita. Para $k \in \mathbb{N}$ denotamos de S_k o grupo das permutações, ou seja das aplicações bijetivas

$$\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\}.$$

A paridade de uma permutação $\sigma \in S_k$ é o número

$$\varepsilon(\sigma) := (-1)^{\#\nu(\sigma)}, \quad \nu(\sigma) := \{(i, j) \mid 1 \leq i < j \leq k, \sigma(i) > \sigma(j)\}.$$

Um elemento (i, j) de $\nu(\sigma)$ é chamado de **inversão na permutação** σ . A paridade define um homomorfismo de grupos

$$\varepsilon : S_k \rightarrow \{\pm 1\}.$$

Definição 7.1.1. Seja $k \in \mathbb{N}$. Uma aplicação multi-linear $\omega : X^k \rightarrow \mathbb{R}$ chama-se de **k -forma alternada** se ela satisfaz para todos os $i, j \in \{1, \dots, k\}$ com $i < j$ e todos os $\xi_1, \dots, \xi_k \in X$ a condição

$$\omega(\xi_1, \dots, \xi_{i-1}, \xi_j, \xi_{i+1}, \dots, \xi_{j-1}, \xi_i, \xi_{j+1}, \dots, \xi_k) = -\omega(\xi_1, \dots, \xi_k).$$

Uma **0-forma alternada** em X é um número real por definição. Para $k \in \mathbb{N}$ o conjunto das k -formas alternadas em X é denotado de

$$\Lambda^k X^* := \{\omega : X^k \rightarrow \mathbb{R} \mid \omega \text{ é uma } k\text{-forma alternada}\}.$$

Isto é um espaço vetorial real. Para $k = 0$ resulta

$$\Lambda^0 X^* := \mathbb{R}$$

e para $k = 1$ resulta $\Lambda^1 X^* = X^* = \mathcal{L}(X, \mathbb{R})$ – o espaço dual de X .

Exercício 7.1.2. Para todo $k \in \mathbb{N}$ o $\Lambda^k X^*$ é um espaço vetorial real.

Exercício 7.1.3. Uma aplicação multi-linear $\omega : X^k \rightarrow \mathbb{R}$ é uma k -forma alternada se e somente se satisfaz a condição

$$\omega(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) = \varepsilon(\sigma)\omega(\xi_1, \dots, \xi_k)$$

para toda permutação $\sigma \in S_k$ e todos os vetores $\xi_1, \dots, \xi_k \in X$. **Dica 1:** Se $\sigma \in S_k$ é uma transposição, então vale $\varepsilon(\sigma) = -1$. **Dica 2:** Toda permutação pode ser escrito como composição de transposições.

Exercício 7.1.4. Seja $\omega \in \Lambda^k X^*$ e sejam $\xi_1, \dots, \xi_k \in X$ linearmente dependente. Então vale $\omega(\xi_1, \dots, \xi_k) = 0$. Particularmente segue disso que $\Lambda^k X^* = \{0\}$ para $k > \dim X$.

Exemplo 7.1.5. Seja $X = \mathbb{R}^n$ e $k \in \{1, \dots, n\}$. Para uma k -tuplo ordenada

$$I = (i_1, \dots, i_k) \in \mathbb{N}^k, \quad 1 \leq i_1 < i_2 < \dots < i_k \leq n,$$

definimos a aplicação $dx_I : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ através da fórmula

$$dx_I(\xi_1, \xi_2, \dots, \xi_k) := \det \left((\xi_{\mu i_\nu})_{\mu, \nu=1}^k \right)$$

para $\xi_\mu = (\xi_{\mu 1}, \dots, \xi_{\mu n}) \in \mathbb{R}^n$, $\mu = 1, \dots, k$. Segue das propriedades do determinante que dx_I é realmente uma k -forma alternada em \mathbb{R}^n (veja o Apêndice D).

Para $k = 1$ e $I = i \in \{1, \dots, n\}$ a aplicação $dx_i : \mathbb{R}^n \rightarrow \mathbb{R}$ é a projeção na i -ésima coordenada, ou seja

$$dx_i(\xi) = \xi_i, \quad \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n.$$

O conjunto das k -tuplos ordenadas em $\{1, \dots, n\}$ denotamos de

$$\mathcal{I}_k(n) := \{I = (i_1, \dots, i_k) \in \mathbb{N}^k \mid 1 \leq i_1 < i_2 < \dots < i_k \leq n\}.$$

O número de k -tuplos ordenadas em $\{1, \dots, n\}$ é

$$\#\mathcal{I}_k(n) = \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Para $k = 0$ é útil utilizar as convenções $\mathcal{I}_0(n) := \{\emptyset\}$ e $dx_\emptyset := 1 \in \mathbb{R}$.

Lema 7.1.6. *Seja $n \in \mathbb{N}$ e $k \in \{0, 1, \dots, n\}$. Então as k -formas alternadas dx_I para $I \in \mathcal{I}_k(n)$ formam uma base de $\Lambda^k(\mathbb{R}^n)^*$. Particularmente vale*

$$\dim \Lambda^k(\mathbb{R}^n)^* = \binom{n}{k}.$$

Demonstração. Para $k = 0$ a afirmação do lema segue direto das definições. Seja $k \geq 1$ e seja e_1, \dots, e_n a base padrão do \mathbb{R}^n . Então vale para todos os $I, J \in \mathcal{I}_k(n)$ com $J = (j_1, \dots, j_k)$ a equação

$$dx_I(e_{j_1}, \dots, e_{j_k}) = \delta_{IJ} = \begin{cases} 1, & \text{caso } I = J, \\ 0, & \text{caso } I \neq J. \end{cases} \quad (7.1.1)$$

Seja ora dado $\omega \in \Lambda^k(\mathbb{R}^n)^*$ e uma aplicação $\mathcal{I}_k(n) \rightarrow \mathbb{R} : I \mapsto a_I$. Então vale

$$\omega = \sum_{I \in \mathcal{I}_k(n)} a_I dx_I \quad \iff \quad \begin{cases} a_I = \omega(e_{i_1}, \dots, e_{i_k}) \\ \forall I = (i_1, \dots, i_k) \in \mathcal{I}_k(n). \end{cases} \quad (7.1.2)$$

Para mostrar isso escolhemos uma k -tuplo $J = (j_1, \dots, j_k) \in \mathcal{I}_k(n)$ e valorizamos as k -formas ω e $\sum_I a_I dx_I$ nos vetores $(e_{j_1}, \dots, e_{j_k})$. Por causa de (7.1.1) recebemos para ambas k -formas a mesma resposta se e somente se $\omega(e_{j_1}, \dots, e_{j_k}) = a_J$. Por outra parte coincidem dois k -formas em \mathbb{R}^n se e somente se coincidem em toda k -tuplo de vetores de uma base. Com isso temos mostrado (7.1.2).

Segue diretamente de (7.1.2) que as k -formas dx_I formam uma base de $\Lambda^k(\mathbb{R}^n)^*$. Com isso o Lema 7.1.6 fica provado. \square

O produto exterior

Definição 7.1.7. Sejam $k, \ell \in \mathbb{N}$. O **produto exterior** de $\alpha \in \Lambda^k X^*$ e $\beta \in \Lambda^\ell X^*$ é a $(k + \ell)$ -forma alternada $\alpha \wedge \beta \in \Lambda^{k+\ell} X^*$, a qual é definida através de

$$(\alpha \wedge \beta)(\xi_1, \dots, \xi_{k+\ell}) := \sum_{\sigma \in S_{k,\ell}} \varepsilon(\sigma) \alpha(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) \beta(\xi_{\sigma(k+1)}, \dots, \xi_{\sigma(k+\ell)})$$

para $\xi_1, \dots, \xi_{k+\ell} \in X$. Aqui denotamos de

$$S_{k,\ell} := \{\sigma \in S_{k+\ell} \mid \sigma(1) < \dots < \sigma(k), \sigma(k+1) < \dots < \sigma(k+\ell)\}$$

o conjunto dos (k, ℓ) -“shuffles”. Para $c \in \Lambda^0 X^* = \mathbb{R}$ e $\omega \in \Lambda^k X^*$ definimos

$$c \wedge \omega := c\omega.$$

Exemplo 7.1.8. Sejam $\alpha, \beta \in \Lambda^1 X^*$ e $\omega \in \Lambda^2 X^*$. Então vale

$$\begin{aligned} (\alpha \wedge \beta)(\xi, \eta) &= \alpha(\xi)\beta(\eta) - \alpha(\eta)\beta(\xi), \\ (\alpha \wedge \omega)(\xi, \eta, \zeta) &= \alpha(\xi)\omega(\eta, \zeta) - \alpha(\eta)\omega(\xi, \zeta) + \alpha(\zeta)\omega(\xi, \eta) \end{aligned}$$

para todos os $\xi, \eta, \zeta \in X$.

Lema 7.1.9. Sejam $k, \ell, m \in \mathbb{N} \cup \{0\}$.

(i) A aplicação

$$\Lambda^k X^* \times \Lambda^\ell X^* \rightarrow \Lambda^{k+\ell} X^* : (\alpha, \beta) \mapsto \alpha \wedge \beta$$

é bi-linear.

(ii) Para $\alpha \in \Lambda^k X^*$ e $\beta \in \Lambda^\ell X^*$ vale

$$\beta \wedge \alpha = (-1)^{k\ell} \alpha \wedge \beta.$$

(iii) Para $\alpha \in \Lambda^k X^*$, $\beta \in \Lambda^\ell X^*$, e $\gamma \in \Lambda^m X^*$ vale

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma).$$

(iv) Para $\alpha_1, \dots, \alpha_k \in \Lambda^1 X^*$ e $\xi_1, \dots, \xi_k \in X$ vale

$$(\alpha_1 \wedge \dots \wedge \alpha_k)(\xi_1, \dots, \xi_k) = \det \left((\alpha_\nu(\xi_\mu))_{\mu,\nu=1}^k \right). \quad (7.1.3)$$

(v) Para $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$ vale

$$dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \dots \wedge dx_{i_k}.$$

Demonstração. As afirmações (i) e (ii) seguem imediatamente da definição do produto exterior. Para a prova de (iii) definimos

$$S_{k,\ell,m} := \left\{ \sigma \in S_{k+\ell+m} \mid \begin{array}{l} \sigma(1) < \dots < \sigma(k), \\ \sigma(k+1) < \dots < \sigma(k+\ell), \\ \sigma(k+\ell+1) < \dots < \sigma(k+\ell+m) \end{array} \right\}.$$

Então $((\alpha \wedge \beta) \wedge \gamma)(\xi_1, \dots, \xi_{k+\ell+m})$ e $(\alpha \wedge (\beta \wedge \gamma))(\xi_1, \dots, \xi_{k+\ell+m})$ são ambos iguais à expressão

$$\sum_{\sigma} \varepsilon(\sigma) \alpha(\xi_{\sigma(1)}, \dots, \xi_{\sigma(k)}) \beta(\xi_{\sigma(k+1)}, \dots, \xi_{\sigma(k+\ell)}) \gamma(\xi_{\sigma(k+\ell+1)}, \dots, \xi_{\sigma(k+\ell+m)})$$

onde a soma corre sobre todos os $\sigma \in S_{k,\ell,m}$. Disso segue a associatividade do produto exterior.

Se iteramos a fórmula na prova de (iii), então resulta através de indução a equação

$$\begin{aligned} (\alpha_1 \wedge \dots \wedge \alpha_k)(\xi_1, \dots, \xi_k) &= \sum_{\sigma \in S_k} \varepsilon(\sigma) \alpha_1(\xi_{\sigma(1)}) \dots \alpha_k(\xi_{\sigma(k)}) \\ &= \det \left((\alpha_\nu(\xi_\mu))_{\mu,\nu=1}^k \right) \end{aligned}$$

para $\alpha_1, \dots, \alpha_k \in X^*$ e $\xi_1, \dots, \xi_k \in X$. Isto prova (iv), e com $X = \mathbb{R}^n$ e $\alpha_\nu = dx_{i_\nu}$ segue disso a parte (v). Com isso o Lema 7.1.9 fica provado. \square

Exercício 7.1.10. Seja $X = \mathbb{R}^{2n}$ com coordenadas $x_1, \dots, x_n, y_1, \dots, y_n$ e

$$\omega := dx_1 \wedge dy_1 + \dots + dx_n \wedge dy_n \in \Lambda^2(\mathbb{R}^{2n})^*.$$

Então a fórmula

$$\omega^{\wedge n} = (-1)^{n(n-1)/2} n! dx_1 \wedge \dots \wedge dx_n \wedge dy_1 \wedge \dots \wedge dy_n.$$

mostra o produto exterior de ω com si mesmo n -vezes.

“Pull-back” - buscar

Definição 7.1.11. Seja $A : X \rightarrow Y$ uma aplicação linear e $\omega \in \Lambda^k Y^*$. A k -forma alternada buscada através de A ou o pull-back de ω sob A é denotada de $A^* \omega \in \Lambda^k X^*$ e definida através da fórmula

$$(A^* \omega)(\xi_1, \dots, \xi_k) := \omega(A\xi_1, \dots, A\xi_k)$$

para $\xi_1, \dots, \xi_k \in X$.

Lema 7.1.12. *Sejam $A : X \rightarrow Y$ e $B : Y \rightarrow Z$ aplicações lineares.*

(i) *A aplicação $\Lambda^k Y^* \rightarrow \Lambda^k X^* : \omega \mapsto A^* \omega$ é linear.*

(ii) *Para $\omega \in \Lambda^k Z^*$ vale*

$$(B \circ A)^* \omega = A^* B^* \omega.$$

(iii) *Para $\alpha \in \Lambda^k Y^*$ e $\beta \in \Lambda^\ell Y^*$ vale*

$$A^*(\alpha \wedge \beta) = (A^* \alpha) \wedge (A^* \beta).$$

(iv) *Sejam $X = \mathbb{R}^n$, $Y = \mathbb{R}^p$, e*

$$A = (a_{ji})_{i=1, \dots, n}^{j=1, \dots, p} \in \mathbb{R}^{p \times n}$$

(entendido como aplicação linear de \mathbb{R}^n para \mathbb{R}^p). Denotamos as coordenadas em \mathbb{R}^n com $x = (x_1, \dots, x_n)$ e aquelas em \mathbb{R}^p com $y = (y_1, \dots, y_p)$. Então vale para $J = (j_1, \dots, j_k) \in \mathcal{I}_k(p)$ que

$$A^* dy_J = \sum_{I \in \mathcal{I}_k(n)} \det(A_{JI}) dx_I$$

onde

$$A_{JI} := (a_{j_\nu i_\mu})_{\mu, \nu=1}^k \in \mathbb{R}^{k \times k}$$

para $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$.

Demonstração. As afirmações (i), (ii), (iii) seguem imediatamente das definições. Para a prova de (iv) denotamos de e_1, \dots, e_n a base padrão do \mathbb{R}^n e escolhemos uma k -tuplo $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$. Então vale

$$\begin{aligned} (A^* dy_J)(e_{i_1}, \dots, e_{i_k}) &= dy_J(Ae_{i_1}, \dots, Ae_{i_k}) \\ &= \det \left(((Ae_{i_\mu})_{j_\nu})_{\mu, \nu=1}^k \right) \\ &= \det(A_{JI}). \end{aligned}$$

Aqui segue a última equação do fato que a j -ésima coordenada do vetor Ae_i é a entrada $a_{ji} = (Ae_i)_j$ da matriz A (i -ésima coluna e j -ésima linha). Com isso a afirmação segue de (7.1.2). \square

7.2 Formas diferenciais

Nesta seção $n, p, q \in \mathbb{N} \cup \{0\}$ e

$$U \subset \mathbb{R}^n, \quad V \subset \mathbb{R}^p, \quad W \subset \mathbb{R}^q$$

são conjuntos abertos não-vazios. Os elementos de U serão denotado de $x = (x_1, \dots, x_n)$ e aqueles de V de $y = (y_1, \dots, y_p)$.

Definição 7.2.1. *Seja $k \in \mathbb{N} \cup \{0\}$. Uma forma diferencial em U de grau k é uma aplicação suave $\omega : U \times (\mathbb{R}^n)^k \rightarrow \mathbb{R}$, tal que a aplicação*

$$(\mathbb{R}^n)^k \longrightarrow \mathbb{R} : (\xi_1, \dots, \xi_k) \mapsto \omega(x; \xi_1, \dots, \xi_k)$$

é uma k -forma alternada para todo $x \in U$.

Se $\omega : U \times (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ é uma forma diferencial, então denotamos de $\omega_x \in \Lambda^k(\mathbb{R}^n)^*$ a k -forma alternada determinada através de $x \in U$, ou seja

$$\omega_x(\xi_1, \dots, \xi_k) := \omega(x; \xi_1, \dots, \xi_k)$$

para $\xi_1, \dots, \xi_k \in \mathbb{R}^n$. Nesta ortografia podemos entender uma forma diferencial de grau k também como uma aplicação suave $U \rightarrow \Lambda^k(\mathbb{R}^n)^* : x \mapsto \omega_x$. Ora, segundo o Lema 7.1.6, podemos representar a k -forma alternada ω_x na base $\{dx_I\}_{I \in \mathcal{I}_k(n)}$. Ou seja, existem funções suaves $a_I : U \rightarrow \mathbb{R}$, uma para cada uma k -tuplo $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$, tal que para todos os $x \in U$ vale:

$$\omega_x = \sum_{I \in \mathcal{I}_k(n)} a_I(x) dx_I. \quad (7.2.1)$$

A demonstração de Lema 7.1.6 também mostra que pode-se escrever os coeficientes na forma $a_I(x) = \omega_x(e_{i_1}, \dots, e_{i_k})$. O conjunto de formas diferenciais em U de grau k vamos denotar de

$$\Omega^k(U) := C^\infty(U, \Lambda^k(\mathbb{R}^n)^*).$$

Isto é um espaço vetorial real (de dimensão infinita se $0 \leq k \leq n \in \mathbb{N}$). Às vezes e abreviando chamamos os elementos de $\Omega^k(U)$ de k -formas. Porém aqui tem-se que dar atenção para a diferença entre k -formas alternadas (elementos de $\Lambda^k(\mathbb{R}^n)^*$) e k -formas como formas diferenciais (elementos de $\Omega^k(U)$).

“Pull-back” - buscar

Definição 7.2.2. *Seja $f : U \rightarrow V$ uma aplicação suave e $\omega \in \Omega^k(V)$. A forma diferencial buscada através de f ou o pull-back de ω sob f*

$$f^*\omega \in \Omega^k(U)$$

é definida através de

$$(f^*\omega)(x; \xi_1, \dots, \xi_k) := \omega(f(x); df(x)\xi_1, \dots, df(x)\xi_k)$$

para $x \in U$ e $\xi_1, \dots, \xi_k \in \mathbb{R}^n$.

Lema 7.2.3. *Sejam $f : U \rightarrow V$ e $g : V \rightarrow W$ aplicações suaves.*

(i) *A aplicação $\Omega^k(V) \rightarrow \Omega^k(U) : \omega \mapsto f^*\omega$ é linear.*

(ii) *Para $\omega \in \Omega^k(W)$ vale*

$$(g \circ f)^*\omega = f^*g^*\omega.$$

(iii) *Para $\alpha \in \Omega^k(V)$ e $\beta \in \Omega^\ell(V)$ vale*

$$f^*(\alpha \wedge \beta) = (f^*\alpha) \wedge (f^*\beta).$$

(iv) *Se*

$$\omega = \sum_{J \in \mathcal{I}_k(p)} b_J dy_J \in \Omega^k(V)$$

com $b_J \in C^\infty(V)$, então vale

$$f^*\omega = \sum_{J \in \mathcal{I}_k(p)} \sum_{I \in \mathcal{I}_k(n)} (b_J \circ f) \det \left(\frac{\partial f_J}{\partial x_I} \right) dx_I \quad (7.2.2)$$

onde

$$\frac{\partial f_J}{\partial x_I} := \left(\frac{\partial f_{j_\nu}}{\partial x_{i_\mu}} \right)_{\nu, \mu=1}^k \in \mathbb{C}^\infty(U, \mathbb{R}^{k \times k})$$

para $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$ e $J = (j_1, \dots, j_k) \in \mathcal{I}_k(p)$.

Demonstração. As afirmações (i), (ii), (iii) seguem imediatamente através das definições e (iv) segue de Lema 7.1.12 (iv). \square

Exemplo 7.2.4. Seja $V = \mathbb{R}^2$ com as coordenadas (x, y) e $U = \mathbb{R}^2$ com as coordenadas (r, θ) . Seja $f : U \rightarrow V$ a aplicação

$$f(r, \theta) := (r \cos(\theta), r \sin(\theta)).$$

Então temos

$$x = f_1(r, \theta) = r \cos(\theta), \quad y = f_2(r, \theta) = r \sin(\theta).$$

A fórmula (7.2.2) para “pull-back” através de f resulta em

$$f^* dx = \frac{\partial f_1}{\partial r} dr + \frac{\partial f_1}{\partial \theta} d\theta = \cos(\theta) dr - r \sin(\theta) d\theta$$

e

$$f^* dy = \frac{\partial f_2}{\partial r} dr + \frac{\partial f_2}{\partial \theta} d\theta = \sin(\theta) dr + r \cos(\theta) d\theta.$$

Disso segue segundo o Lema 7.2.3 (iii) que

$$f^*(dx \wedge dy) = r dr \wedge d\theta.$$

Temos utilizado nisso que

$$d\theta \wedge dr = -dr \wedge d\theta.$$

Isto coincide com a equação

$$\det(df(r, \theta)) = r.$$

A diferencial

Definição 7.2.5. Seja $k \in \mathbb{N} \cup \{0\}$ e $\omega \in \Omega^k(U)$. A **diferencial de ω** é a forma diferencial $d\omega \in \Omega^{k+1}(U)$, a qual é definida através de

$$d\omega(x; \xi_0, \dots, \xi_k) := \sum_{i=0}^k \sum_{\nu=1}^n (-1)^i \xi_{i\nu} \frac{\partial \omega}{\partial x_\nu}(x; \xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \quad (7.2.3)$$

para $x \in U$ e $\xi_0, \dots, \xi_k \in \mathbb{R}^n$. A aplicação resultante

$$d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$$

é um operador diferencial linear de primeira ordem.

Comentário 7.2.6. Uma 0-forma em U é uma função suave $u : U \rightarrow \mathbb{R}$. A diferencial $du \in \Omega^1(U)$ de uma tal 0-forma $u \in \Omega^0(U) = C^\infty(U)$ é segundo a Definição 7.2.5 a 1-forma definida através de

$$du(x; \xi) := \sum_{\nu=1}^n \xi_{i\nu} \frac{\partial u}{\partial x_\nu}(x)$$

para $x \in U$ e $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$. Então

$$(du)_x = du(x) \in (\mathbb{R}^n)^*$$

com efeito é a derivada de u no ponto $x \in U$ como conhecemos a derivada. Numa ortografia um tique diferente temos que

$$du = \sum_{i=1}^n \frac{\partial u}{\partial x_i} dx_i.$$

Aplicado para a função $U \rightarrow \mathbb{R} : x = (x_1, \dots, x_n) \mapsto x_i$ isto justifica a posteriori a notação dx_i para a projeção na i -ésima coordenada.

Teorema 7.2.7. (i) Se $\omega = \sum_{I \in \mathcal{I}_k(n)} a_I dx_I$ com $a_I \in C^\infty(U)$, então vale

$$d\omega = \sum_{\nu=1}^n \sum_{I \in \mathcal{I}_k(n)} \frac{\partial a_I}{\partial x_\nu} dx_\nu \wedge dx_I = \sum_{I \in \mathcal{I}_k(n)} da_I \wedge dx_I. \quad (7.2.4)$$

(ii) Para $\alpha \in \Omega^k(U)$ e $\beta \in \Omega^\ell(U)$ vale

$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta).$$

(iii) Para todo $\omega \in \Omega^k(U)$ vale

$$d(d\omega) = 0.$$

(iv) Se $f : U \rightarrow V$ é uma aplicação suave, então para todo $\omega \in \Omega^k(V)$ vale

$$d(f^*\omega) = f^*(d\omega).$$

Demonstração. Provaremos que as fórmulas (7.2.3) e (7.2.4) para $d\omega$ coincidem. Se $\omega = \sum_I a_I dx_I$, então encontraremos para $x \in U$ e $\xi_0, \dots, \xi_k \in \mathbb{R}^n$:

$$\begin{aligned} d\omega(x; \xi_0, \dots, \xi_k) &= \sum_{i=0}^k \sum_{\nu=1}^n (-1)^i \xi_{i\nu} \frac{\partial \omega}{\partial x_\nu}(x; \xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \\ &= \sum_{i=0}^k \sum_{\nu=1}^n \sum_I (-1)^i \xi_{i\nu} \frac{\partial a_I}{\partial x_\nu} dx_I(\xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \\ &= \sum_{\nu=1}^n \sum_I \frac{\partial a_I}{\partial x_\nu} \left(\sum_{i=0}^k (-1)^i \xi_{i\nu} dx_I(\xi_0, \dots, \xi_{i-1}, \xi_{i+1}, \dots, \xi_k) \right) \\ &= \sum_{\nu=1}^n \sum_I \frac{\partial a_I}{\partial x_\nu} (dx_\nu \wedge dx_I)(\xi_0, \dots, \xi_k) \end{aligned}$$

onde $\frac{\partial a_I}{\partial x_\nu}$ na verdade abrevia $\frac{\partial a_I}{\partial x_\nu}(x)$. Com isso (i) fica provado.

Provaremos (ii). Como $d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$ é um operador linear, basta mostrar a regra de Leibnitz para $\alpha = a dx_I$ e $\beta = b dx_J$, onde $a, b : U \rightarrow \mathbb{R}$ são funções suaves e $I \in \mathcal{I}_k(n)$, $J \in \mathcal{I}_\ell(n)$. Neste caso recebemos

$$\begin{aligned} d(\alpha \wedge \beta) &= d(ab dx_I \wedge dx_J) = \sum_{\nu=1}^n \frac{\partial(ab)}{\partial x_\nu} dx_\nu \wedge dx_I \wedge dx_J \\ &= \sum_{\nu=1}^n \frac{\partial a}{\partial x_\nu} b dx_\nu \wedge dx_I \wedge dx_J + \sum_{\nu=1}^n a \frac{\partial b}{\partial x_\nu} dx_\nu \wedge dx_I \wedge dx_J \\ &= \left(\sum_{\nu=1}^n \frac{\partial a}{\partial x_\nu} dx_\nu \wedge dx_I \right) \wedge \beta + (-1)^k \alpha \wedge \left(\sum_{\nu=1}^n \frac{\partial b}{\partial x_\nu} dx_\nu \wedge dx_J \right) \\ &= d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta. \end{aligned}$$

Aqui temos utilizado na segunda e terceira equação a fórmula (7.2.4). Com isso (ii) fica provado.

Provaremos (iii). Para uma 0-forma $u \in \Omega^0(U) = C^\infty(U)$ vale

$$\begin{aligned} d(du) &= d\left(\sum_{\nu=1}^n \frac{\partial u}{\partial x_\nu} dx_\nu \right) = \sum_{\mu, \nu=1}^n \frac{\partial}{\partial x_\mu} \frac{\partial u}{\partial x_\nu} dx_\mu \wedge dx_\nu \\ &= \sum_{\mu < \nu} \left(\frac{\partial}{\partial x_\mu} \frac{\partial u}{\partial x_\nu} - \frac{\partial}{\partial x_\nu} \frac{\partial u}{\partial x_\mu} \right) dx_\mu \wedge dx_\nu = 0. \end{aligned}$$

Demais vale, segundo definição, que $d(dx_I) = 0$ para toda k -tuplo $I \in \mathcal{I}_k(n)$. Para $\omega = \sum_I a_I dx_I$ resulta daí de (i) e (ii) a equação

$$d(d\omega) = d\left(\sum_I da_I \wedge dx_I\right) = \sum_I d(da_I) \wedge dx_I = 0.$$

Com isso (iii) fica provado.

Provaremos (iv). Se $u \in \Omega^0(V) = C^\infty(V)$, então $f^*u = u \circ f$ e a fórmula

$$f^*(du) = d(f^*u) \tag{7.2.5}$$

é uma reformulação da regra da cadeia. Com $u(y) = y_j$ vale particularmente

$$f^*dy_j = df_j$$

para $j = 1, \dots, p$. Disso segue para $J = (j_1, \dots, j_k) \in \mathcal{I}_k(p)$

$$f^*dy_J = df_{j_1} \wedge \dots \wedge df_{j_k}.$$

Aplicamos a regra de Leibnitz em (ii) a esta fórmula, então (com indução)

$$d(f^*dy_J) = 0 \tag{7.2.6}$$

Para uma k -forma $\omega = \sum_{J \in \mathcal{I}_k(p)} b_J dy_J$ com $b_J \in C^\infty(V)$ vale dali

$$\begin{aligned} d(f^*\omega) &= d\left(\sum_{J \in \mathcal{I}_k(p)} (f^*b_J)(f^*dy_J)\right) \\ &= \sum_{J \in \mathcal{I}_k(p)} (d(f^*b_J)) \wedge (f^*dy_J) \\ &= \sum_{J \in \mathcal{I}_k(p)} (f^*(db_J)) \wedge (f^*dy_J) \\ &= f^*\left(\sum_{J \in \mathcal{I}_k(p)} (db_J) \wedge dy_J\right) \\ &= f^*(d\omega). \end{aligned}$$

Aqui a segunda equação segue de (7.2.6) e da regra de Leibnitz em (ii), a terceira de (7.2.5), a quarta de Lema 7.2.3 (iii), e a última de (7.2.4). Com isso Teorema 7.2.7 fica provado. \square

Definição 7.2.8. Uma forma diferencial $\omega \in \Omega^k(U)$ é chamada de **fechada** se $d\omega = 0$. Ela é chamada de **exata** se exista uma $(k-1)$ -forma $\alpha \in \Omega^{k-1}(U)$ com $d\alpha = \omega$.

Segundo a parte (iii) de Teorema 7.2.7 toda k -forma exata é fechada. Mas vamos ver que uma k -forma fechada não deve ser necessariamente exata. Porém o Exercício 7.2.10 vai mostrar que para regiões estreladas $U \subset \mathbb{R}^n$ toda k -forma fechada em U com $k \geq 1$ também é exata necessariamente. Em outras palavras, para todo subconjunto $U \subset \mathbb{R}^n$ temos uma sequência finita de operadores lineares

$$\Omega^0(U) \xrightarrow{d} \Omega^1(U) \xrightarrow{d} \Omega^2(U) \xrightarrow{d} \dots \xrightarrow{d} \Omega^{n-1}(U) \xrightarrow{d} \Omega^n(U)$$

com a propriedade que a imagem de cada um operador é contido no núcleo do seguinte. Uma tal sequência de operadores também é chamada de **complexo de cadeia**. Se a imagem do operador $d : \Omega^{k-1}(U) \rightarrow \Omega^k(U)$ é um subespaço próprio do núcleo do operador $d : \Omega^k(U) \rightarrow \Omega^{k+1}(U)$, então o quociente destes espaços é um espaço vetorial de dimensão positiva. Estes espaços quociente são chamados de **grupos de cohomologia de Rham** de U e denotado de $H^k(U; \mathbb{R})$. Eles contem informações interessantes sobre as propriedades do conjunto U . Porém este tema supera longe dos conteúdos deste curso.

Formas diferenciais em dimensão três

Exemplo 7.2.9. Consideramos o espaço euclidiano 3-dimensional \mathbb{R}^3 com as coordenadas x, y, z . Então vale

$$\Lambda^0(\mathbb{R}^3)^* = \mathbb{R}, \quad \Lambda^1(\mathbb{R}^3)^* \cong \mathbb{R}^3, \quad \Lambda^2(\mathbb{R}^3)^* \cong \mathbb{R}^3, \quad \Lambda^3(\mathbb{R}^3)^* \cong \mathbb{R}.$$

O isomorfismo $\Lambda^1(\mathbb{R}^3)^* \cong \mathbb{R}^3$ é dado através da base dx, dy, dz e o isomorfismo $\Lambda^2(\mathbb{R}^3)^* \cong \mathbb{R}^3$ através da base $dy \wedge dz, dz \wedge dx, dx \wedge dy$.

Seja $U \subset \mathbb{R}^3$ um conjunto aberto e $\text{Vet}(U) := C^\infty(U, \mathbb{R}^3)$ o espaço dos campos vetoriais suaves em U . Seja dado também um campo vetorial $f = (a, b, c) : U \rightarrow \mathbb{R}^3$, então a 1-forma e a 2-forma associada são

$$\alpha_f := a dx + b dy + c dz, \quad \omega_f := a dy \wedge dz + b dz \wedge dx + c dx \wedge dy. \quad (7.2.7)$$

Isto fornece isomorfismos $\text{Vet}(U) \rightarrow \Omega^1(U)$ e $\text{Vet}(U) \rightarrow \Omega^2(U)$. A diferencial da 1-forma α_f é

$$d\alpha_f = \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} \right) dy \wedge dz + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} \right) dz \wedge dx + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) dx \wedge dy.$$

Esta 2-forma por sua vez é identificada com o campo vetorial

$$\text{rot}(f) := \begin{pmatrix} \partial c/\partial y - \partial b/\partial z \\ \partial a/\partial z - \partial c/\partial x \\ \partial b/\partial x - \partial a/\partial y \end{pmatrix} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \nabla \times f.$$

Aqui usamos o produto cruz e tratamos a expressão $\nabla := (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ formal como seja um vetor. Igualmente recebemos

$$d\omega_f = \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z} \right) dx \wedge dy \wedge dz = (\nabla \cdot f) dx \wedge dy \wedge dz$$

onde a notação $\nabla \cdot f = \langle \nabla, f \rangle$ indica o produto interno padrão. Resumindo, resulta

$$du = \alpha_{\text{grad}(u)}, \quad d\alpha_f = \omega_{\text{rot}(f)}, \quad d\omega_f = \text{div}(f) dx \wedge dy \wedge dz \quad (7.2.8)$$

para $u \in C^\infty(U)$ e $f \in \text{Vet}(U)$ com

$$\text{grad}(u) = \nabla u, \quad \text{rot}(f) = \nabla \times f, \quad \text{div}(f) = \nabla \cdot f.$$

Ora, segue de Teorema 7.2.7 (ou também através de cálculo direto) que

$$\text{rot}(\text{grad}(u)) = 0, \quad \text{div}(\text{rot}(f)) = 0 \quad (7.2.9)$$

para $u \in C^\infty(U)$ e $f \in \text{Vet}(U)$. Além disso, vale

$$\text{div}(\text{grad}(u)) = \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}. \quad (7.2.10)$$

Isto é o **operador de Laplace**.

Exercícios

Exercício 7.2.10 (Lema de Poincaré). Seja $U \subset \mathbb{R}^n$ aberto e **estrelado** (ou seja, se $x \in U$ então também $tx \in U$ para todo $t \in [0, 1]$). Seja $k \in \mathbb{N}$ e $\omega \in \Omega^k(U)$ uma k -forma fechada. Então ω é exata. **Dica:** Defina $\alpha \in \Omega^{k-1}(U)$ através de

$$\alpha(x; \xi_1, \dots, \xi_{k-1}) := \int_0^1 \omega(tx; x, \xi_1, \dots, \xi_{k-1}) t^{k-1} dt$$

para $x \in U$ e $\xi_1, \dots, \xi_{k-1} \in \mathbb{R}^n$. Então vale $d\alpha = \omega$.

Exercício 7.2.11. Seja $U \subset \mathbb{R}^n$ um conjunto aberto, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ um campo vetorial, e seja $\omega_f \in \Omega^{n-1}(U)$ dado através de

$$\omega_f(x; \xi_1, \dots, \xi_{n-1}) := \det(f(x), \xi_1, \dots, \xi_{n-1})$$

para $x \in U$ e $\xi_1, \dots, \xi_{n-1} \in \mathbb{R}^n$. Então vale

$$\omega_f = \sum_{i=1}^n (-1)^{i-1} f_i(x) dx_1 \wedge \dots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \dots \wedge dx_n \quad (7.2.11)$$

$$d\omega_f = \operatorname{div}(f) dx_1 \wedge \dots \wedge dx_n. \quad (7.2.12)$$

Exercício 7.2.12. O ***-operador** $* : \Omega^k(\mathbb{R}^n) \rightarrow \Omega^{n-k}(\mathbb{R}^n)$ é definido assim

$$*dx_I := \varepsilon(\sigma) dx_{*I}$$

para $I = (i_1, \dots, i_k) \in \mathcal{I}_k(n)$, onde $*I = (j_1, \dots, j_{n-k}) \in \mathcal{I}_{n-k}(n)$ é escolhido, tal que $\{1, \dots, n\} = \{i_1, \dots, i_k, j_1, \dots, j_{n-k}\}$, e tal que a permutação $\sigma \in S_n$ é definida através de $\sigma(\nu) = i_\nu$ para $\nu = 1, \dots, k$ e $\sigma(\nu) = j_{\nu-k}$ para $\nu = k+1, \dots, n$. Como o *-operador é linear, ele é determinado unicamente pelas imagens dos vetores dx_I da base. Ele tem as seguintes propriedades.

(i) Para todo $I \in \mathcal{I}_k(n)$ vale

$$**dx_I = (-1)^{k(n-k)} dx_I, \quad dx_I \wedge *dx_I = dx_1 \wedge \dots \wedge dx_n.$$

Demais $dx_I \wedge *dx_J = 0$ para todos os $I, J \in \mathcal{I}_k(n)$ com $I \neq J$.

(ii) Seja $U \subset \mathbb{R}^n$ um conjunto aberto, $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ um campo vetorial suave, e

$$\alpha_f := \sum_{i=1}^n f_i(x) dx_i \in \Omega^1(U).$$

Então vale que $*\alpha_f = \omega_f$ e $d(*\alpha_f) = \operatorname{div}(f) dx_1 \wedge \dots \wedge dx_n$ (veja o Exercício 7.2.11). Para toda função suave $u : \mathbb{R}^n \rightarrow \mathbb{R}$ vale

$$*d*du = \Delta u := \frac{\partial^2 u}{\partial x_1^2} + \dots + \frac{\partial^2 u}{\partial x_n^2}.$$

(iii) Para $n = 3$ vale

$$*dx = dy \wedge dz, \quad *dy = dz \wedge dx, \quad *dz = dx \wedge dy$$

assim como $*\omega_f = \alpha_f$ e $*d\alpha_f = \alpha_{\operatorname{rot}(f)}$ para todo campo vetorial $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

7.3 Orientação

Nosso objetivo é integrar formas diferenciais de grau d sobre subvariedades de dimensão d . Para isso a noção de orientação faz um papel central. O ponto de partida é a seguinte definição.

Espaços vetoriais orientados

Seja X um espaço vetorial real d -dimensional. Duas bases ordenadas

$$e_1, \dots, e_d \quad \tilde{e}_1, \dots, \tilde{e}_d$$

têm a **mesma orientação** se a $(d \times d)$ matriz $A = (a_{ij})_{i,j=1}^d$ definida pelas fórmulas

$$\tilde{e}_i = \sum_{j=1}^d a_{ij} e_j \quad i = 1, \dots, d$$

tem determinante positivo. A noção “mesma orientação” define uma relação de equivalência no conjunto das bases de X com exatamente duas classes de equivalência. Uma **orientação** de X é a escolha de uma destas classes de equivalência de bases. Se uma orientação de X é dada, então chamamos as bases na classe de equivalência escolhida **positivas** e as outras bases **negativas**. Um **espaço vetorial orientado** é um espaço vetorial real junto com uma orientação.

Uma orientação de X é então determinado através de uma base ordenada. As bases positivas (respectivamente negativas) são então aquelas as quais decorrem desta base através de uma transformação com determinante positivo (respectivamente negativo). Trocamos por exemplo dois elementos de uma base positiva então recebemos uma base negativa. A **orientação padrão** do \mathbb{R}^d é determinada pela base padrão (ou base canônica)

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_d = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Se $\Phi : X \rightarrow Y$ é um isomorfismo de espaços vetoriais entre dois espaços vetoriais orientados, então ele é chamado de **preservando orientação** se

ele transforma bases positivas de X em bases positivas de Y e **reversando orientação** se ele transforma bases positivas de X em bases negativas de Y . Um isomorfismo preservando orientação também transforma evidentemente bases negativas de X em bases negativas de Y e um isomorfismo reversando orientação transforma bases negativas de X em bases positivas de Y . Em outras palavras, se representamos a aplicação linear Φ como matriz com respeito de duas bases positivas de X e Y , então Φ é preservando orientação (respectivamente reversando orientação) se e somente se o determinante desta matriz é positiva (respectivamente negativa).

Subvariedades

Queremos ora transferir esta noção de orientação a C^∞ -subvariedades suaves $M \subset \mathbb{R}^n$ d -dimensionais. Para evocação discutimos mais uma vez as noções fundamentais (Definição 6.1.9). Seja destacado aqui mais uma vez que neste capítulo todos os difeomorfismos são C^∞ -difeomorfismos.

Sejam dados números inteiros d, n com $0 \leq d \leq n$. Lembramos que um subconjunto $M \subset \mathbb{R}^n$ é chamado de *subvariedade suave d -dimensional* de \mathbb{R}^n , se para todo ponto $p_0 \in M$ existe um subconjunto M -aberto $U_0 \subset M$ o qual contém p_0 e é difeomorfo a um subconjunto aberto $V_0 \subset \mathbb{R}^d$. Um tal difeomorfismo $\phi_0 : U_0 \rightarrow V_0$ chamamos de uma *carta* de M , o conjunto M -aberto U_0 chamamos de um *domínio de carta* de M , e a aplicação inversa $\psi_0 := \phi_0^{-1} : V_0 \rightarrow U_0$ chamamos de uma *parametrização suave* do domínio de carta.

Ao pé desta formulação tem-se que considerar várias coisas. Para já o subconjunto $U_0 \subset M$, do qual falamos, **não é** um subconjunto aberto de \mathbb{R}^n (exceto no caso de $d = n$), mas somente é aberto com respeito à topologia relativa de M . Segundo o Lema B.1.2 isso significa que existe um subconjunto aberto $U \subset \mathbb{R}^n$ com $U \cap M = U_0$. Se falamos pois de uma aplicação suave $\phi_0 : U_0 \rightarrow \mathbb{R}^d$, então isso significa que existe uma aplicação suave $\phi : U \rightarrow \mathbb{R}^d$ num conjunto aberto $U \subset \mathbb{R}^n$ tal que $U \cap M = U_0$ e a restrição de ϕ a $U \cap M = U_0$ coincide com ϕ_0 . Portanto, neste sentido uma carta $\phi_0 : U_0 \rightarrow V_0$ é uma aplicação suave, a qual é adicionalmente bijetiva e possui uma aplicação inversa suave, a parametrização $\psi_0 : V_0 \rightarrow U_0$ do domínio de carta referida há pouco. Esta aplicação inversa é evidentemente suave no sentido usual, porque V_0 é um subconjunto aberto de \mathbb{R}^d e podemos considerar ψ_0 como aplicação de V_0 a \mathbb{R}^n (e neste caso ignorar temporariamente que os valores de ψ_0 estão contido realmente em $U_0 \subset M$).

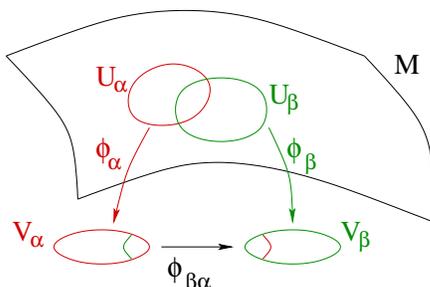


Figura 7.1: Aplicações de transição.

Voltamos ora à nossa subvariedade M , então podemos cobrir M (segundo a definição) com domínios de carta U_α (com índices α num conjunto A) e para isso podemos escolher ainda cartas suaves $\phi_\alpha : U_\alpha \rightarrow V_\alpha$. Um tal sistema de cartas $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$ chama-se de **atlas** de M . Se M é compacto então segue da caracterização de compacidade através da propriedade cobertura em Teorema C.1.2 que a subvariedade M possui um atlas finito. Para quaisquer duas cartas

$$\phi_\alpha : U_\alpha \rightarrow V_\alpha, \quad \phi_\beta : U_\beta \rightarrow V_\beta$$

as **aplicações de transição**

$$\phi_{\beta\alpha} := \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \rightarrow \phi_\beta(U_\alpha \cap U_\beta)$$

fazem um papel importante (veja a Figura 7.1).

Como ϕ_α e ϕ_β particularmente são homeomorfismos (então bijetivo e em ambas direções contínuo), e $U_\alpha \cap U_\beta$ é um subconjunto M -aberto de M , os subconjuntos

$$\phi_\alpha(U_\alpha \cap U_\beta) \subset V_\alpha, \quad \phi_\beta(U_\alpha \cap U_\beta) \subset V_\beta$$

também são conjuntos abertos em \mathbb{R}^d e $\phi_{\beta\alpha}$ é um difeomorfismo suave entre estes conjuntos.

Seja $M \subset \mathbb{R}^n$ uma subvariedade suave d -dimensional. O **espaço vetorial** $T_p M \subset \mathbb{R}^n$ é o conjunto

$$T_p M := \{ \dot{\gamma}(0) \mid \gamma \in C^1(\mathbb{R}, \mathbb{R}^n), \gamma(t) \in M \forall t \in \mathbb{R}, \gamma(0) = p \}.$$

Tínhamos mostrado em Capítulo 3 que $T_p M$ é um subespaço linear d -dimensional de \mathbb{R}^n e que

$$T_p M = \text{im } d\psi_\alpha(x)$$

para toda parametrização suave $\psi_\alpha : V_\alpha \rightarrow U_\alpha$ de um domínio de carta U_α com $p \in U_\alpha$ e $x = \psi_\alpha^{-1}(p) \in V_\alpha$. Particularmente a derivada

$$d\psi_\alpha(x) : \mathbb{R}^d \rightarrow T_p M, \quad p := \psi_\alpha(x)$$

é um isomorfismo de espaços vetoriais para todos os $\alpha \in A$ e $x \in V_\alpha$.

Exercício 7.3.1. Seja $M \subset \mathbb{R}^n$ uma subvariedade suave d -dimensional. Então o **fibrado tangente**

$$TM := \{(p, v) \in \mathbb{R}^n \times \mathbb{R}^n \mid p \in M, v \in T_p M\}$$

é uma subvariedade suave $2d$ -dimensional de \mathbb{R}^{2n} .

Exemplo 7.3.2. Para a lembrança mais uma vez este exemplo. O conjunto

$$M := \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$$

não é uma subvariedade de \mathbb{R}^2 . Ainda que o subconjunto

$$U_0 := \{(x, y) \in \mathbb{R}^2 \mid -1 < x < 1, y = 0\} \subset M$$

é difeomorfo a um subconjunto de \mathbb{R} , o conjunto U_0 não é M -aberto. Em Capítulo 3 tínhamos mostrado que um subconjunto $M \subset \mathbb{R}^n$ é uma subvariedade suave d -dimensional se e somente se para todo $p_0 \in M$ existe um subconjunto aberto $U \subset \mathbb{R}^n$ e uma aplicação suave

$$f : U \rightarrow \mathbb{R}^{n-d}$$

tal que $p_0 \in U$, o valor 0 de f é regular, e

$$U \cap M = \{x \in U \mid f(x) = 0\}.$$

No nosso exemplo seria $n = 2$ e $d = 1$, porém o ponto

$$p_0 = (0, 0)$$

não possui nenhuma tal vizinhança: Caso $f : U \rightarrow \mathbb{R}$ anule-se em $U \cap M$, então $df(0, 0) = 0$ também e por isso 0 não é um valor regular de f .

Subvariedades orientadas

Definição 7.3.3. *Seja $M \subset \mathbb{R}^n$ uma subvariedade suave d -dimensional. Uma **orientação de M** é a escolha de uma orientação para todo espaço tangente $T_p M$, tal que estas orientações são vinculadas da seguinte maneira. Para todo ponto $p_0 \in M$ existe um domínio de carta $U_0 \subset M$ com $p_0 \in U_0$ e uma parametrização suave*

$$\psi_0 : V_0 \rightarrow U_0$$

(num subconjunto aberto $V_0 \subset \mathbb{R}^d$), tal que a derivada

$$d\psi_0(x) : \mathbb{R}^d \rightarrow T_{\psi_0(x)} M$$

para todo $x \in V_0$ é um isomorfismo preservando orientação (ou seja, o isomorfismo $d\psi_0(x)$ traslada a base padrão do \mathbb{R}^d numa base positiva de $T_{\psi_0(x)} M$). Um tal ψ_0 chamamos de **parametrização orientada** de U_0 e a aplicação inversa

$$\phi_0 := \psi_0^{-1} : U_0 \rightarrow V_0$$

carta orientada. Chama-se M de **orientável** se existe uma tal orientação. Uma **subvariedade orientada** de \mathbb{R}^n é uma subvariedade junto com uma orientação.

Se $M \subset \mathbb{R}^n$ é uma subvariedade orientada d -dimensional, então existe obviamente um atlas $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$, o qual é composto exclusivamente de cartas orientadas. Neste caso todas as aplicações de transição $\phi_{\beta\alpha}$ **preservam orientação**, ou seja, para $\alpha, \beta \in A$ e $x \in \phi_\alpha(U_\alpha \cap U_\beta)$ vale

$$\det(d\phi_{\beta\alpha}(x)) > 0. \quad (7.3.1)$$

Falamos então de um **atlas orientado**. Vice versa, todo atlas orientado induz uma orientação de M ; definimos a orientação de $T_p M$ como a orientação induzida pelo isomorfismo

$$d\psi_\alpha(x) : \mathbb{R}^d \rightarrow T_p M$$

onde $\alpha \in A$ é escolhido, tal que $p \in U_\alpha$ e

$$x := \phi_\alpha(p), \quad \psi_\alpha := \phi_\alpha^{-1}.$$

Segundo (7.3.1) esta definição de orientação de $T_p M$ é independente de α .

Exemplo 7.3.4. A esfera unitária

$$\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = 1\}$$

é uma subvariedade orientada de \mathbb{R}^n . Uma base v_1, \dots, v_{n-1} de $T_x\mathbb{S}^{n-1} = x^\perp$ é positiva se e somente se x, v_1, \dots, v_{n-1} é uma base positiva do \mathbb{R}^n .

Exemplo 7.3.5. O n -toro padrão

$$\mathbb{T}^n := \{z = (z_1, \dots, z_n) \in \mathbb{C}^n \mid |z_i| = 1 \text{ para } i = 1, \dots, n\}$$

é orientado. O espaço tangente $T_z\mathbb{T}^n = \{(\mathbf{i}t_1z_1, \dots, \mathbf{i}t_nz_n) \mid t_\nu \in \mathbb{R}\}$ possui um isomorfismo canônico a \mathbb{R}^n (isto é $t \mapsto (\mathbf{i}t_1z_1, \dots, \mathbf{i}t_nz_n)$) e aquele induz a orientação padrão do espaço tangente.

Exemplo 7.3.6. Seja $M \subset \mathbb{R}^n$ uma subvariedade suave $(n-1)$ -dimensional e $\nu : M \rightarrow \mathbb{S}^{n-1}$ uma aplicação suave com $T_xM = \nu(x)^\perp$ para todos os $x \in M$. Então ν determina uma orientação de M : Uma base v_1, \dots, v_{n-1} de T_xM é positiva se e somente se $\det(\nu(x), v_1, \dots, v_{n-1}) > 0$. Exemplo 7.3.4 é o caso especial $M = \mathbb{S}^{n-1}$ com $\nu(x) = x$.

Exemplo 7.3.7. Seja $f = (f_1, \dots, f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma aplicação suave e seja 0 um valor regular de f . Então $M := f^{-1}(0)$ é uma subvariedade suave orientável de \mathbb{R}^n de dimensão $d := n - m$. Para $x \in M$ pode-se definir uma orientação de T_xM da seguinte maneira. Uma base v_1, \dots, v_d de $T_xM = \ker df(x)$ é positiva se e somente se

$$\det(\nabla f_1(x), \dots, \nabla f_m(x), v_1, \dots, v_d) > 0.$$

Exemplo 7.3.6 é o caso especial $m = 1$ com $\nu(x) = \|\nabla f(x)\|^{-1} \nabla f(x)$.

Exemplo 7.3.8. A fita de Möbius $M \subset \mathbb{R}^3$ em Figura 7.2 não é orientável.

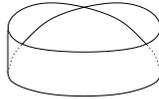


Figura 7.2: A fita de Möbius.

7.4 Integração

Seja $V \subset \mathbb{R}^d$ um subconjunto aberto e $\tau \in \Omega^d(V)$ uma forma diferencial de grau d . As coordenadas em V denotamos de $y = (y_1, \dots, y_d)$. Então existe uma função suave $h : V \rightarrow \mathbb{R}$ (i.e. C^∞), tal que

$$\tau = h(y) dy_1 \wedge \dots \wedge dy_d.$$

Ora, suponhamos que τ possua **suporte compacto** em V . Isto é, o conjunto fechado

$$\text{supp}(\tau) := \overline{\{y \in V \mid h(y) \neq 0\}}$$

(aqui o fecho deve ser entendido em \mathbb{R}^d) é compacto e contido em V . Equivalente a isso é a condição que o fecho do conjunto $\{y \in V \mid h(y) \neq 0\}$ com respeito à topologia relativa de V é um conjunto compacto. Reformulado mais uma vez, existe um conjunto compacto $K \subset V$, tal que h anula-se no complemento $V \setminus K$. É uma consequência disso que a integral Riemann de h é definida em V : podemos construir um prédio de blocos B com $K \subset B \subset V$ (Figura 7.3 e Exercício 5.1.5) e integrar a função h sobre B . Escrevemos a integral sobre V e definimos

$$\int_V \tau := \int_V h(y) dy_1 \cdots dy_d. \quad (7.4.1)$$

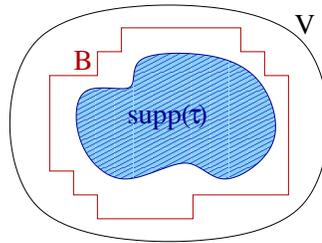


Figura 7.3: Forma diferencial com suporte compacto.

Ora, dado mais um subconjunto aberto $V' \subset \mathbb{R}^d$ e um difeomorfismo suave

$$\phi : V' \rightarrow V.$$

Denotando as coordenadas em V' com $x = (x_1, \dots, x_d)$, então vale

$$\phi^* \tau = h(\phi(x)) \det(d\phi(x)) dx_1 \wedge \dots \wedge dx_d$$

(veja a equação (7.2.2) em Lema 7.2.3). Se ω possui suporte compacto em V , assim a d -forma buscada $\phi^*\tau \in \Omega^d(V')$ também possui suporte compacto em V' . Se o determinante de $d\phi(x)$ é positivo em todo o lugar, então segue da fórmula da transformação em Teorema 5.7.11 que

$$\int_{\phi^{-1}(V)} \phi^*\tau = \int_V \tau. \quad (7.4.2)$$

Esta equação vale só para difeomorfismos ϕ preservando orientação. No caso de um difeomorfismo revertendo orientação a sinal da integral muda. Estas observações nós permitem transferir a integral a domínios de carta em subvariedades orientadas.

Integração sobre domínios de carta

Seja $M \subset \mathbb{R}^n$ uma subvariedade orientada d -dimensional de \mathbb{R}^n , seja $W \subset \mathbb{R}^n$ um subconjunto aberto contendo M , e seja $\omega \in \Omega^d(W)$. Seja

$$U \subset M$$

um domínio de carta e seja $\psi : V \rightarrow U$ uma parametrização orientada suave de U . Suponhamos que o suporte

$$\text{supp}(\omega) := \overline{\{x \in \mathbb{R}^n \mid \omega_x \neq 0\}}$$

intersecta o domínio de carta U num conjunto compacto. Então $\psi^*\omega$ possui suporte compacto em V e definimos a **integral de ω sobre U** através de

$$\int_U \omega := \int_V \psi^*\omega \quad (7.4.3)$$

onde o lado direito foi definido em (7.4.1). A integral não depende da escolha da parametrização: Pois se $\psi' : V' \rightarrow U$ é mais uma parametrização de U , então o difeomorfismo $\phi := \psi^{-1} \circ \psi' : V' \rightarrow V$ preserva orientação e vale

$$\int_{V'} (\psi')^*\omega = \int_{V'} (\psi \circ \phi)^*\omega = \int_{V'} \phi^*\psi^*\omega = \int_V \psi^*\omega.$$

A última equação segue de (7.4.2) com $\tau = \psi^*\omega$.

Integral sobre subvariedades compactas

Seja $M \subset \mathbb{R}^n$ uma subvariedade compacta orientada d -dimensional, contido num conjunto aberto $W \subset \mathbb{R}^n$, e $\omega \in \Omega^d(W)$. Então existe um atlas orientado finito $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$. Escolhemos uma partição da unidade $\rho_\alpha : \mathbb{R}^n \rightarrow [0, 1]$, tal que

$$\text{supp}(\rho_\alpha) \cap M \subset U_\alpha, \quad \sum_{\alpha} \rho_\alpha(x) = 1 \quad \forall x \in M. \quad (7.4.4)$$

A existência de uma tal partição da unidade segue de Teorema 6.2.4. (Para já escolhemos subconjuntos abertos $U'_\alpha \subset \mathbb{R}^n$ com $U'_\alpha \cap M = U_\alpha$. Depois aplicamos Teorema 6.2.4 à cobertura de M pelos conjuntos U'_α .)

Definição 7.4.1. *Sejam M , ω , U_α , ρ_α como acima. A integral de ω sobre M é definida como uma soma de integrais de tipo (7.4.3), mais detalhado*

$$\int_M \omega := \sum_{\alpha} \int_{U_\alpha} \rho_\alpha \omega. \quad (7.4.5)$$

Lema 7.4.2. *A integral de ω sobre M é independente da escolha do atlas orientado e da partição da unidade.*

Demonstração. Se $\{\tilde{U}_\beta, \tilde{\phi}_\beta\}_{\beta \in \tilde{A}}$ é mais um atlas com partição da unidade correspondente $\tilde{\rho}_\beta : \mathbb{R}^n \rightarrow [0, 1]$, então vale

$$\sum_{\alpha} \int_{U_\alpha} \rho_\alpha \omega = \sum_{\alpha} \sum_{\beta} \int_{U_\alpha \cap \tilde{U}_\beta} \rho_\alpha \tilde{\rho}_\beta \omega = \sum_{\beta} \int_{\tilde{U}_\beta} \tilde{\rho}_\beta \omega.$$

Com isso o lema fica provado. \square

Comentário 7.4.3. Denotando as coordenadas em \mathbb{R}^n de x_1, \dots, x_n , e as coordenadas em \mathbb{R}^d de y_1, \dots, y_d , e supondo que ω é dado na forma $\omega = \sum_{I \in \mathcal{I}_d(n)} a_I dx_I$ com $a_I \in C^\infty(W)$, então a integral em Definição 7.4.1 tem a forma explícita

$$\int_M \omega = \sum_{\alpha} \sum_{I \in \mathcal{I}_d(n)} \int_{V_\alpha} \rho_\alpha(\psi_\alpha(y)) a_I(\psi_\alpha(y)) \det(d\psi_{\alpha I}(y)) dy_1 \cdots dy_d. \quad (7.4.6)$$

Aqui denotamos de $\psi_\alpha := \phi_\alpha^{-1} : V_\alpha \rightarrow U_\alpha$ a parametrização orientada de U_α e de $\psi_{\alpha I}$ a composição de ψ_α com a projeção às coordenadas denotadas de $I \in \mathcal{I}_d(n)$.

Integral sobre subconjuntos compactos

Como anteriormente seja $M \subset \mathbb{R}^n$ uma subvariedade orientada d -dimensional, contido num conjunto aberto $W \subset \mathbb{R}^n$, e $\omega \in \Omega^d(W)$. Porém não solicitamos mais que M é compacto e integramos ω em vez disso sobre um subconjunto compacto adequado $G \subset M$.

Para $G \subset M$ denotamos de $\partial G := \partial_M G$ o bordo de G com respeito à topologia relativa em M , também chamado de **M -bordo de G** (veja (6.2.1)). Isto é, um ponto $p \in M$ pertence a ∂G se existe tanto uma sequência em G como uma sequência em $M \setminus G$, a qual converge para p . Suponhamos que ∂G é um conjunto Jordan-nulo d -dimensional (Definição 6.2.1). Ou seja, o conjunto $K \cap \psi^{-1}(\partial G) = \{y \in K \mid \psi(y) \in \partial G\}$ é um conjunto Jordan-nulo para toda parametrização $\psi : V \rightarrow U$ de um domínio de carta e todo subconjunto compacto $K \subset V$. O seguinte lema nós lembra mais uma vez a inclusão (6.2.2) na prova de Lema 6.2.3.

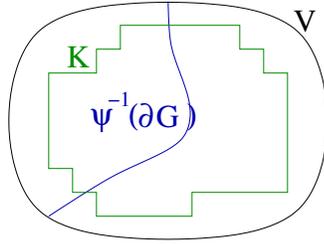


Figura 7.4: O conjunto $\psi^{-1}(\partial G)$.

Lema 7.4.4. *Seja $G \subset M$ um subconjunto compacto cujo M -bordo ∂G é um conjunto nulo d -dimensional, $\psi : V \rightarrow U$ uma parametrização orientada de um domínio de carta, e $K \subset V$ um conjunto compacto (veja a Figura 7.4). Então vale*

$$\partial(K \cap \psi^{-1}(G)) \subset \partial K \cup (K \cap \psi^{-1}(\partial G)). \quad (7.4.7)$$

Demonstração. Seja $y \in \partial(K \cap \psi^{-1}(G))$ e $y \notin \partial K$. Então existem duas sequências $y_i \in K \cap \psi^{-1}(G)$ e $y'_i \in V \setminus (K \cap \psi^{-1}(G))$ as quais convergem para y . Como K é compacto e assim fechado, vale $y \in K \subset V$, e como $\psi(y_i) \in G$, também vale $\psi(y) = \lim_{i \rightarrow \infty} \psi(y_i) \in G$. Particularmente y é um elemento do conjunto aberto $K \setminus \partial K$ e daí vale $y'_i \in K$ para i suficientemente grande. Disso segue $\psi(y'_i) \notin G$ para i grande e por isso $\psi(y)$ é tanto o limite de uma sequência $\psi(y_i) \in G$, como de uma sequência $\psi(y'_i) \in M \setminus G$. Então vale $\psi(y) \in \partial G$ e com isso $y \in K \cap \psi^{-1}(\partial G)$. Com isso o lema fica provado. \square

Podemos proceder como acima. Se o suporte de $\omega \in \Omega^d(W)$ intersecta o domínio de carta U num conjunto compacto e se $\psi : V \rightarrow U$ é uma parametrização orientada suave, então a d -forma buscada $\psi^*\omega \in \Omega^d(V)$ possui suporte compacto em V , e escolhemos um prédio de blocos $B \subset V$ o qual contem o suporte K de $\psi^*\omega$. Segundo o Lema 7.4.4 o bordo do conjunto $K \cap \psi^{-1}(G)$ é um conjunto Jordan-nulo. Com isso $K \cap \psi^{-1}(G)$ é Jordan-mensurável e $\psi^*\omega$ anula-se fora deste conjunto. Consequentemente definimos a integral de ω sobre $G \cap U$ como a integral da d -forma buscada

$$\psi^*\omega =: h(y)dy_1 \wedge \cdots \wedge dy_d$$

sobre a região buscada $\psi^{-1}(G)$:

$$\int_{(G \cap U) \subset M} \omega := \int_{\psi^{-1}(G) \subset V} \psi^*\omega := \int_{(K \cap \psi^{-1}(G)) \subset B \subset V} h(y)dy_1 \cdots dy_d. \quad (7.4.8)$$

Como anteriormente segue da fórmula da transformação em Teorema 5.7.11 que a integral é independente da escolha da parametrização orientada ψ .

Se $\omega \in \Omega^d(W)$ é uma d -forma escolhemos um atlas orientado $\{U_\alpha, \phi_\alpha\}_{\alpha \in A}$ com a propriedade adicional que U_α é limitado e $\bar{U}_\alpha \subset M$ para todo $\alpha \in A$ (onde entendemos \bar{U}_α como o fecho do conjunto U_α no espaço ambiente \mathbb{R}^n). Então escolhemos uma partição da unidade $\rho_\alpha : \mathbb{R}^n \rightarrow [0, 1]$ tal que $\text{supp}(\rho_\alpha)$ é compacto para todo $\alpha \in A$, ρ_α é não-nulo só para um número finito de α , e

$$\text{supp}(\rho_\alpha) \cap M \subset U_\alpha, \quad \sum_{\alpha} \rho_\alpha(x) = 1 \quad \forall x \in G.$$

Nas nossas hipóteses $\text{supp}(\rho_\alpha) \cap M$ é compacto: Se p_i é uma sequência em $\text{supp}(\rho_\alpha) \cap M$, então uma subsequência converge para um $p \in \text{supp}(\rho_\alpha)$; como $p_i \in U_\alpha$ e $\bar{U}_\alpha \subset M$ segue $p \in M$ e com isso $p \in \text{supp}(\rho_\alpha) \cap M$. Para a construção do ρ_α procedemos como acima. Cobrimos G para já utilizando um número finito de U_α , escolhemos depois para este α conjuntos abertos limitados U'_α com $U'_\alpha \cap M = U_\alpha$, e aplicamos afinal Teorema 6.2.4 sobre a existência de uma partição da unidade à cobertura finita do conjunto compacto G pelos U'_α . Ora definimos a **integral de ω sobre G** através de

$$\int_G \omega := \sum_{\alpha} \int_{G \cap U_\alpha} \rho_\alpha \omega. \quad (7.4.9)$$

Como anteriormente esta integral é independente da escolha dos U_α e ρ_α .

7.5 O Teorema de Stokes

Nesta seção formulamos e provamos o Teorema de Stokes para subconjuntos compactos G de uma subvariedade suave $M \subset \mathbb{R}^n$, a qual possui um bordo suave (compare Definição 6.4.1).

Definição 7.5.1 (Bordo suave). *Seja $M \subset \mathbb{R}^n$ uma subvariedade suave d -dimensional e seja G um subconjunto compacto de M . Chamamos G de um conjunto compacto com bordo suave, se seu M -bordo ∂G é uma subvariedade suave $(d-1)$ -dimensional do \mathbb{R}^n , e se este M -bordo coincide com o M -bordo do M -interior de G (isto é, se todo conjunto aberto $U \subset \mathbb{R}^n$ com $U \cap \partial G \neq \emptyset$ contém tanto elementos de $G \setminus \partial G$, como elementos de $M \setminus G$).*

Definição 7.5.2 (Vetor normal exterior unitário). *Seja $M \subset \mathbb{R}^n$ uma subvariedade suave d -dimensional, seja $G \subset M$ um subconjunto suave com bordo suave, e seja $p \in \partial G = \partial_M G$. Um vetor $v \in T_p M$ com $\|v\| = 1$ é chamado **vetor normal exterior unitário de G no ponto p** se $v \perp T_p \partial G$ e existe uma curva suave $\gamma : (-\delta, \delta) \rightarrow M$ com $\gamma(0) = p$, $\dot{\gamma}(0) = v$, e $\gamma(t) \notin G$ para $0 < t < \delta$.*

Lema 7.5.3. *Seja $M \subset \mathbb{R}^n$ uma subvariedade suave orientada d -dimensional e seja $G \subset M$ um subconjunto compacto com bordo suave. Então vale o seguinte.*

(i) *Se $\phi : U \rightarrow V \subset \mathbb{R}^d$ é uma carta de M , então $\phi(\partial G \cap U) \subset V$ é uma subvariedade suave $(d-1)$ -dimensional de \mathbb{R}^d .*

(ii) *∂G é um conjunto Jordan-nulo com respeito a M .*

(iii) *Para todo ponto $p \in \partial G$ existe um domínio de carta $U \subset M$ com $p \in U$, um subconjunto aberto $V \subset \mathbb{R}^d$, e uma carta $\phi : U \rightarrow V$ com*

$$\phi(U \cap G) = \{x \in V \mid x_d \leq 0\}. \quad (7.5.1)$$

(iv) *Para todo elemento $p \in \partial G$ existe exatamente um vetor normal exterior unitário $\nu(p) \in \mathbb{S}^{n-1}$. A aplicação $\nu : \partial G \rightarrow \mathbb{S}^{n-1}$ assim definida chama-se **campo normal exterior de G** .*

(v) *A aplicação $\nu : \partial G \rightarrow \mathbb{S}^{n-1}$ em (iv) pode ser estendida a uma aplicação suave numa vizinhança aberta de ∂G com valores em \mathbb{S}^{n-1} .*

(vi) *Existe uma orientação de ∂G , tal que uma base v_2, \dots, v_d de $T_p \partial G$ é positiva se e somente se $\nu(p), v_2, \dots, v_d$ é uma base positiva de $T_p M$. Ela é chamada de **orientação canônica do bordo de ∂G** .*

Demonstração. Prova de (i). Seja $x_0 \in \phi(U \cap \partial G)$ e $p_0 := \phi^{-1}(x_0) \in U \cap \partial G$. Então existe um domínio de carta $U_0 \subset U \cap \partial G$ de ∂G com $x_0 \in U_0$ e um difeomorfismo $\phi_0 : U_0 \rightarrow V_0$ a um conjunto aberto $V_0 \subset \mathbb{R}^{d-1}$. Particularmente $U_0 \subset U \cap \partial G$ é aberto com respeito à topologia relativa de $U \cap \partial G$. Dali o conjunto $U'_0 := \phi(U_0) \subset \phi(U \cap \partial G)$ é aberto com respeito à topologia relativa de $\phi(U \cap \partial G) \subset \mathbb{R}^d$. O conjunto U'_0 contém o ponto $x_0 = \phi(p_0)$. Além disso, a aplicação $\phi'_0 := \phi_0 \circ \phi^{-1} : U'_0 \rightarrow V_0 \subset \mathbb{R}^{d-1}$ é um difeomorfismo. Isto mostra que o conjunto $\phi(U \cap \partial G)$ é uma subvariedade suave $(d-1)$ -dimensional de \mathbb{R}^d . Com isso a parte (i) fica provada.

Provaremos parte (ii). Seja $U \subset M$ um domínio de carta, seja $V \subset \mathbb{R}^d$ um conjunto aberto, e seja $\phi : U \rightarrow V$ uma carta de M (Definição 6.1.9). Como ∂G é um subconjunto M -fechado de M , então $U \cap \partial G$ é um subconjunto fechado de U com respeito à topologia relativa de U (Lema B.1.2). Como ϕ é um homeomorfismo segue disso que $\phi(U \cap \partial G)$ é um subconjunto fechado de V com respeito à topologia relativa de V . Demais $\phi(U \cap \partial G)$ é, segundo (i), uma subvariedade suave $(d-1)$ -dimensional de \mathbb{R}^d . Disso segue que o conjunto $\phi(U \cap \partial G) \cap K$, para todo subconjunto compacto $K \subset V$, é um subconjunto compacto de uma subvariedade $(d-1)$ -dimensional do \mathbb{R}^d , e com isso segundo o Exemplo 5.6.13 um conjunto Jordan-nulo. Segundo a Definição 6.2.1 o bordo ∂G é dali um conjunto Jordan-nulo com respeito M . Com isso a parte (ii) fica provada.

Provaremos parte (iii). Seja dado $p_0 \in \partial G$. Vamos mostrar em dois passos que existe uma carta $\phi : U \rightarrow V$ com $p_0 \in U$, a qual satisfaz (7.5.1).

Passo 1. *Existe uma carta $\phi : U \rightarrow V$ de M com $p_0 \in U$ e com $V \subset \mathbb{R}^d$ aberto e convexo, e tal que $\phi(U \cap \partial G) = \{x \in V \mid x_d = 0\}$.*

Seja $U \subset M$ um domínio de carta com $p_0 \in U$, seja $V \subset \mathbb{R}^d$ um conjunto aberto, e seja $\phi : U \rightarrow V$ uma carta de M . Segundo (i) o conjunto $\phi(U \cap \partial G)$ é uma subvariedade do \mathbb{R}^d e contém o ponto $x_0 := \phi(p_0)$. Dali existe um subconjunto aberto $V_0 \subset V$ com $y_0 \in V_0$, um conjunto aberto $V' \subset \mathbb{R}^d$ e um difeomorfismo $\psi_0 : V_0 \rightarrow V'$ com

$$\psi_0(V_0 \cap \phi(U \cap \partial G)) = \{x' \in V' \mid x'_d = 0\}.$$

Como $x_0 = \phi(p_0) \in V_0$, o conjunto $U' := \phi^{-1}(V_0) \subset U$ é um domínio de carta em M com $p_0 \in U'$ e a derivada $\phi' := \psi_0 \circ \phi : U' \rightarrow V'$ é uma carta de M com $\phi'(U' \cap \partial G) = \psi_0(V_0 \cap \phi(U \cap \partial G)) = \{x' \in V' \mid x'_d = 0\}$. Através de diminuição de U' e V' , se necessário, podemos supor que V' é um subconjunto convexo de \mathbb{R}^d . Com isso o Passo 1 (i) fica provado.

Passo 2. Existe um domínio de carta $U \subset M$ com $p \in U$, um subconjunto aberto $V \subset \mathbb{R}^n$, e uma carta $\phi : U \rightarrow V$, a qual satisfaz (7.5.1).

Seja $\phi : U \rightarrow V$ como em Passo 1. Além disso, seja $\Omega := \overset{\circ}{G}$ o M -interior de G . Então vale $\partial\Omega = \partial G$ segundo a hipótese, conseqüentemente o homeomorfismo $\phi : U \rightarrow V$ satisfaz as condições de Lema 6.4.4. Particularmente vale que $U \cap \partial\Omega = U \cap \partial G = U^0$ segundo o Passo 1 (na notação de Lema 6.4.4). Se $U^- \cap \Omega$ e $U^+ \cap \Omega$ são ambos vazios, então vale $U \cap \Omega = \emptyset$ e dali $p_0 \notin \partial\Omega$, em contradição a $\partial\Omega = \partial G$. Se $U^- \cap \Omega$ e $U^+ \cap \Omega$ são ambos não-vazios, então $U \subset \bar{\Omega} = G$ segundo o Lema 6.4.4 e dali $p_0 \in \overset{\circ}{G}$, em contradição a $p_0 \in \partial G$. Portanto exatamente um dos conjuntos $U^\pm \cap \Omega$ é não-vazio. Através de troca de sinal de ϕ_d , se necessário, podemos supor que $U^- \cap \Omega \neq \emptyset$. Então é vazio $U^+ \cap G = U^+ \cap (\Omega \cup \partial\Omega) = \emptyset$. E vale $U^- \cup U^0 \subset \Omega \cup \partial\Omega = G$ segundo o Lema 6.4.4. Com isso Passo 2 e parte (ii) ficam provados.

Provaremos parte (iv). Seja $p \in \partial G$ e seja $\phi : U \rightarrow V$ como em parte (iii). Então com $\psi := \phi^{-1} : V \rightarrow U$ e $x := \phi(p)$ vale

$$T_p M = \text{im} d\psi(x), \quad T_p \partial G = \left\{ \sum_{i=1}^{d-1} \frac{d\psi}{dx_i}(x) \xi_i \mid \xi_1, \dots, \xi_{d-1} \in \mathbb{R} \right\}.$$

Se $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$ com $\xi_d > 0$, então $\gamma(t) := \psi(x + t\xi) \in M$ satisfaz a condição $\gamma(t) \notin G$ para $t > 0$ pequeno, como $U \cap G = \{\psi(x) \mid x \in V, x_d \leq 0\}$ segundo (7.5.1). Um tal vetor ξ com $d\psi(x)\xi \perp T_p \partial G$ é dado através da fórmula $\xi = (d\psi(x)^T d\psi(x))^{-1} e_d$, onde $e_d := (0, \dots, 0, 1) \in \mathbb{R}^d$ é o vetor da base padrão. Disso resulta a fórmula

$$\nu(p) = \frac{d\psi(x)(d\psi(x)^T d\psi(x))^{-1} e_d}{\|d\psi(x)(d\psi(x)^T d\psi(x))^{-1} e_d\|} \quad (7.5.2)$$

para o vetor normal exterior unitário de G o ponto p . Este vetor é único, porque o subespaço linear $T_p M \cap (T_p \partial G)^\perp$ é de dimensão um, e toda curva $\gamma : (-\delta, \delta) \rightarrow U$ com $\gamma(0) = p$ e $\dot{\gamma}(0) = -\nu(p)$ satisfaz a condição $\gamma(t) \in G$ para pequenos $t > 0$. Com isso a parte (iv) fica provada.

Em parte (v) a extensão local do campo vetorial normal exterior unitário $\nu : \partial G \rightarrow \mathbb{S}^{n-1}$ a uma função suave da fórmula (7.5.2) segue no que se põe $x = \phi(p)$ e utiliza o fato que pode-se estender ϕ a uma função suave de um subconjunto aberto do \mathbb{R}^n para V . Uma extensão global de ν pode-se então construir mediante uma partição da unidade por meio de Teorema 6.2.4.

Parte (vi) segue direto de parte (v) e com isso Lema 7.5.3 fica provado. \square

A parte (ii) de Lema 7.5.3 mostra que podemos integrar uma d -forma sobre G , e a parte (vi) de Lema 7.5.3 mostra que podemos integrar uma $(d-1)$ -forma com a ajuda da orientação-bordo canônica sobre ∂G . Com isso estamos pronto para a formulação do **Teorema de Stokes**.

Teorema 7.5.4 (Stokes). *Seja $M \subset \mathbb{R}^n$ uma subvariedade suave orientada d -dimensional, seja $G \subset M$ um subconjunto compacto com bordo suave, e seja $\omega \in \Omega^{d-1}(W)$ uma $(d-1)$ -forma num conjunto aberto $W \subset \mathbb{R}^n$, o qual contem M . Então vale*

$$\int_G d\omega = \int_{\partial G} \omega. \quad (7.5.3)$$

Demonstração. Para já consideramos o caso $d = n$. Então $M = W \subset \mathbb{R}^d$ é um subconjunto aberto e $G \subset W$ um conjunto compacto com bordo suave (Definição 6.4.1). A $(d-1)$ -forma em W podemos então escrever na forma

$$\omega = \sum_{j=1}^d (-1)^{j-1} f_j(x) dx_1 \wedge \cdots \wedge dx_{j-1} \wedge dx_{j+1} \wedge \cdots \wedge dx_d$$

onde

$$f = (f_1, \dots, f_d) : W \rightarrow \mathbb{R}^d$$

é um campo suave. Seja $\nu : \partial G \rightarrow \mathbb{R}^d$ o campo normal exterior. Vamos provar a equação

$$\int_{\partial G} \omega = \int_{\partial G} \langle f, \nu \rangle dS. \quad (7.5.4)$$

Com

$$d\omega = \operatorname{div}(f) dx_1 \wedge \cdots \wedge dx_d$$

(Exercício 7.2.11) segue então

$$\begin{aligned} \int_G d\omega &= \int_G \operatorname{div}(f) dx_1 \cdots dx_d \\ &= \int_{\partial G} \langle f, \nu \rangle dS \\ &= \int_{\partial G} \omega. \end{aligned}$$

Aqui a segunda equação segue de Teorema 6.4.9 de Gauß.

Provaremos ora a equação (7.5.4). Seja $V \subset \mathbb{R}^{d-1}$ aberto e seja

$$\psi = (\psi_1, \dots, \psi_d) : V \rightarrow U \subset \partial G$$

uma parametrização suave orientada de um domínio de carta U de ∂G . Provaremos para já que o vetor normal exterior unitário $\nu(\psi(y)) \in \mathbb{R}^d$ para todos os $y \in V$ é dado através da fórmula

$$\nu_j(\psi(y)) \sqrt{\det(d\psi(y)^T d\psi(y))} = (-1)^{j-1} \det(d\psi_{I_j}(y)) \quad (7.5.5)$$

para $j = 1, \dots, d$, onde $\psi_{I_j} : V \rightarrow \mathbb{R}^{d-1}$ denota a aplicação, a qual resulta de ψ através de abandonar a j -ésima coordenada, isto é

$$\psi_{I_j}(y) := (\psi_1(y), \dots, \psi_{j-1}(y), \psi_{j+1}(y), \dots, \psi_d(y)).$$

Para a prova de (7.5.5) fixamos um elemento $y \in V$ e denotamos de $\eta = (\eta_1, \dots, \eta_d) \in \mathbb{R}^d$ o vetor definido pelo lado direito da equação (7.5.5), isto é

$$\eta_j := (-1)^{j-1} \det(d\psi_{I_j}(y)), \quad j = 1, \dots, d.$$

Para $\xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}^d$ consideramos a matriz

$$A_\xi := \left(\xi, \frac{\partial \psi}{\partial y_1}(y), \dots, \frac{\partial \psi}{\partial y_{d-1}}(y) \right) \in \mathbb{R}^{d \times d}.$$

Desenvolvemos o determinante de A_ξ com respeito à primeira coluna, então resulta segundo a parte (iv) de Teorema D.3.5 a equação

$$\det(A_\xi) = \sum_{j=1}^d (-1)^{j-1} \xi_j \det(d\psi_{I_j}(y)) = \langle \xi, \eta \rangle. \quad (7.5.6)$$

Com $\xi \in \text{im } d\psi(y)$ anula-se o lado esquerdo de (7.5.6). Portanto η está ortogonal a $\text{im } d\psi(y) = T_{\psi(y)}\partial G$. Para $\xi = \eta$ recebemos $\det(A_\eta) = \|\eta\|^2 > 0$. Então os vetores

$$\eta, \frac{\partial \psi}{\partial y_1}, \dots, \frac{\partial \psi}{\partial y_{d-1}}$$

formam uma base positiva de \mathbb{R}^d e, como os vetores $\partial\psi/\partial y_1, \dots, \partial\psi/\partial y_{d-1}$ formam segundo a escolha de ψ uma base positiva de $T_{\psi(y)}\partial G$, segue que o vetor η mostra para fora. Além disso, vale

$$A_\eta^T A_\eta = \begin{pmatrix} \|\eta\|^2 & 0 \\ 0 & d\psi(y)^T d\psi(y) \end{pmatrix},$$

e dali $\|\eta\|^2 \det(d\psi(y)^T d\psi(y)) = \det(A_\eta^T A_\eta) = \det(A_\eta)^2 = \|\eta\|^4$. Portanto vale $\|\eta\|^2 = \det(d\psi(y)^T d\psi(y))$ e disso segue (7.5.5).

Segundo (7.5.5) recebemos no caso de $\text{supp}(f) \cap \partial G \subset U = \psi(V)$ a equação

$$\begin{aligned} \int_{\partial G} \langle f, \nu \rangle dS &= \int_{U=\psi(V)} \langle f, \nu \rangle dS \\ &:= \int_V \langle f(\psi(y)), \nu(\psi(y)) \rangle \sqrt{\det(d\psi(y)^T d\psi(y))} dy_1 \cdots dy_{d-1} \\ &= \int_V \sum_{j=1}^d (-1)^{j-1} f_j(\psi(y)) \det(d\psi_{I_j}(y)) dy_1 \cdots dy_{d-1} \\ &= \int_{\psi^{-1}(U)} \psi^* \omega =: \int_U \omega = \int_{\partial G} \omega. \end{aligned}$$

Aqui segue o passo dois das equações (6.1.15) e (6.1.16) em Definição 6.1.10. O terceiro passo segue então de (7.5.5) e o quarto passo de (7.2.2). Com isso temos provado (7.5.4) para $(d-1)$ -formas com suporte pequeno. O caso de uma $(d-1)$ -forma arbitrária ω em W pode-se reduzir ao caso local com ajuda de partições da unidade. Com isso o Teorema de Stokes fica provado no caso de $d = n$.

Consideramos ora o caso $d < n$. Escolhemos uma cobertura finita de G por domínios de carta $U_\alpha \subset M$ tal que $\bar{U}_\alpha \subset M$ para todos os α . Além disso, escolhemos parametrizações orientadas suaves $\psi_\alpha : V_\alpha \rightarrow U_\alpha$ em subconjuntos abertos $V_\alpha \subset \mathbb{R}^d$ e uma partição da unidade $\rho_\alpha : \mathbb{R}^n \rightarrow [0, 1]$, tal que $\text{supp}(\rho_\alpha)$ é compacto para todo $\alpha \in A$ e

$$\text{supp}(\rho_\alpha) \cap M \subset U_\alpha, \quad \sum_{\alpha} \rho_\alpha(x) = 1 \quad \forall x \in G.$$

Introduzimos as seguintes abreviações:

$$\omega_\alpha := \psi_\alpha^*(\rho_\alpha \omega) \in \Omega^{d-1}(V_\alpha), \quad G_\alpha := \psi_\alpha^{-1}(G), \quad \partial G_\alpha := \psi_\alpha^{-1}(\partial G).$$

Então ∂G_α é o bordo de G_α na topologia relativa de V_α . Como ω_α tem suporte compacto em V_α , deduzimos do caso $d = n$, que

$$\int_{G_\alpha} d\omega_\alpha = \int_{\partial G_\alpha} \omega_\alpha.$$

Isto é segundo a definição equivalente à equação

$$\int_{G \cap U_\alpha} d(\rho_\alpha \omega) = \int_{\partial G \cap U_\alpha} \rho_\alpha \omega.$$

Forme a soma sobre todos os α para concluir a prova de Teorema 7.5.4. \square

7.6 Exemplos e aplicações

Exemplo 7.6.1 (Conteúdo de área). Consideramos a 1-forma

$$\alpha := \frac{1}{2}(xdy - ydx), \quad d\alpha = dx \wedge dy.$$

Para uma região Jordan-mensurável $G \subset \mathbb{R}^2$ a integral de $d\alpha$ sobre G é igual á medida Jordan, respectivamente ao conteúdo de área, ou ao “volume 2-dimensional”. Se o bordo de G é uma 1-variedade, então vale segundo o Teorema de Stokes

$$\text{Vol}_2(G) = \int_G dx \wedge dy = \frac{1}{2} \int_{\partial G} (xdy - ydx).$$

Exemplo 7.6.2 (Número de circulação). Consideramos a 1-forma

$$\alpha := \frac{1}{2\pi} \frac{xdy - ydx}{x^2 + y^2}$$

em $\mathbb{R}^2 \setminus \{(0, 0)\}$. Esta forma é fechada (Exercício). Porém não é exata. No caso contrário sua integral ao longo de toda curva fechada

$$\Gamma \subset \mathbb{R}^2 \setminus \{0\}$$

deve anular-se segundo o Teorema de Stokes. Isto seja o caso $d = 1$ com $\partial\Gamma = \emptyset$ e uma 1-forma df . Porém pode-se mostra facilmente que a integral de α sobre \mathbb{S}^1 (com a orientação padrão sentido anti-horário) iguale 1. Mais geral, se $\gamma : [0, 1] \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\}$ é uma curva suave com

$$\gamma(0) = \gamma(1)$$

e escrevemos $\gamma(t) = (x(t), y(t))$, então

$$\iota(\gamma) := \int_{[0,1]} \gamma^* \alpha = \frac{1}{2\pi} \int_0^1 \frac{x(t)\dot{y}(t) - y(t)\dot{x}(t)}{x(t)^2 + y(t)^2} dt$$

é um número inteiro (Exercício). Pode-se interpretar $\iota(\gamma)$ como número de circulação em torno da origem.

Exemplo 7.6.3 (O Teorema clássico de Stokes). Seja $M \subset \mathbb{R}^3$ uma subvariedade suave 2-dimensional orientada e $G \subset M$ um subconjunto compacto com bordo suave. Então o bordo ∂G (como anteriormente entendido com respeito à topologia relativa de M) é uma subvariedade compacta 1-dimensional de \mathbb{R}^3 . Seja demais $U \subset \mathbb{R}^3$ um subconjunto aberto, o qual contem M , e

$$f = (a, b, c) : U \rightarrow \mathbb{R}^3$$

um campo vetorial suave em U . Consideramos a 1-forma

$$\alpha_f := a dx + b dy + c dz \in \Omega^1(U).$$

Como temos visto em equação (7.2.8) em Exemplo 7.2.9

$$d\alpha_f = \omega_{\text{rot}(f)} = u dy \wedge dz + v dz \wedge dx + w dx \wedge dy, \quad (7.6.1)$$

onde $u := \partial c / \partial y - \partial b / \partial z$, $v := \partial a / \partial z - \partial c / \partial x$, $w := \partial b / \partial x - \partial a / \partial y$, são as coordenadas do campo vetorial $\text{rot}(f) = (u, v, w) : U \rightarrow \mathbb{R}^3$. Segundo o Teorema 7.5.4 und (7.6.1) vale

$$\int_G \omega_{\text{rot}(f)} = \int_{\partial G} \alpha_f. \quad (7.6.2)$$

Pode-se reformular esta equação da seguinte maneira. Seja $\nu : M \rightarrow \mathbb{S}^2$ o **campo normal unitário positivo de M** , ou seja $\nu(p) \perp T_p M$ tem norma um e dois vetores $v, w \in T_p M$ formam uma base positiva do espaço tangente exatamente se $\det(\nu(p), v, w) > 0$. Demais seja $\tau : \partial G \rightarrow \mathbb{S}^2$ o **campo tangente unitário positivo**, ou seja $\tau(p) \in T_p \partial G$ é o vetor tangente unitário positivo para todo $p \in \partial G$. Como a letra ν foi utilizada de outra forma, neste exemplo o **campo normal exterior de G em M** (veja Lema 7.5.3) será denotado de

$$\eta : \partial G \rightarrow \mathbb{S}^2.$$

Para $p \in T_p \partial G$ este $\eta(p) \in T_p M \cap (T_p \partial G)^\perp$ é então o único vetor unitário ortogonal a ambos $\nu(p)$ e $\tau(p)$, o qual segundo as convenções introduzidas em Lema 7.5.3 para a orientação-bordo canônica de ∂G satisfaz a equação

$$\det(\nu(p), \eta(p), \tau(p)) = 1. \quad (7.6.3)$$

Vice versa, a equação (7.6.3) estabelece através do já dado campo normal exterior η de G em M a relação entre a orientação de G via ν e aquela de ∂G via τ . (Veja a Figura 7.5.)

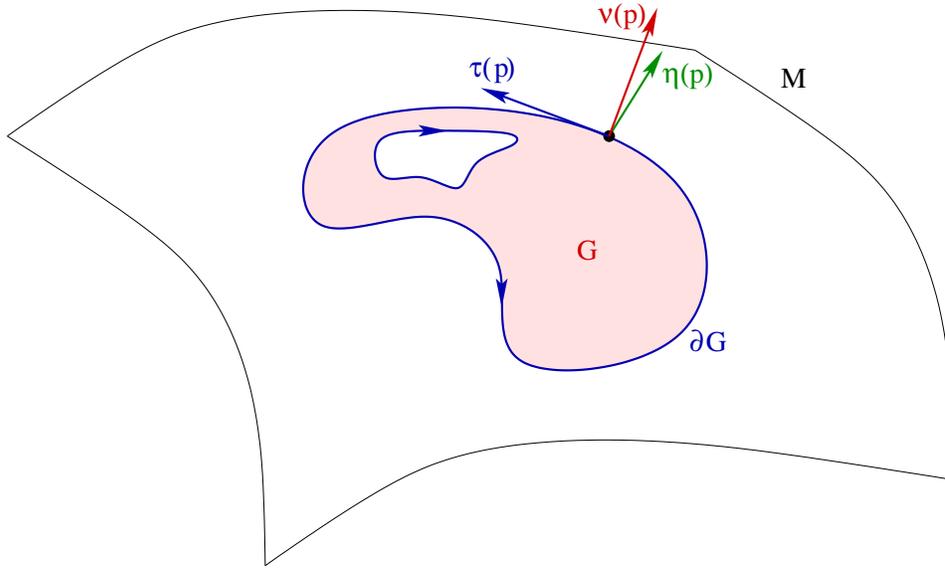


Figura 7.5: O Teorema clássico de Stokes.

Ora, segue de equação (7.6.1) junto com a equação (7.5.4) na prova de Teorema 7.5.4 (com $d = 3$ e G em vez de ∂G e $\text{rot}(f)$ em vez de f) a fórmula

$$\int_G \omega_{\text{rot}(f)} = \int_G \langle \text{rot}(f), \nu \rangle dS. \quad (7.6.4)$$

Escolhemos um componente conexo $\Gamma \subset \partial G$ e uma parametrização positiva $\gamma : \mathbb{R} \rightarrow \Gamma$ pelo comprimento de arco, então vale $\gamma(t + T) = \gamma(t)$ para todos os $t \in \mathbb{R}$ e um $T > 0$ adequado (o comprimento de Γ), e $\dot{\gamma}(t) = \tau(\gamma(t))$ para todos os $t \in \mathbb{R}$. Particularmente $\|\dot{\gamma}(t)\| = 1$ para todos os $t \in \mathbb{R}$ e disso segue

$$\int_\Gamma \alpha_f = \int_0^T \alpha_f(\gamma(t); \dot{\gamma}(t)) dt = \int_0^T \langle f(\gamma(t)), \tau(\gamma(t)) \rangle dt = \int_\Gamma \langle f, \tau \rangle dS.$$

Formamos ora a soma sobre todos os componentes conexos de ∂G , então recebemos a fórmula

$$\int_{\partial G} \alpha_f = \int_{\partial G} \langle f, \tau \rangle dS. \quad (7.6.5)$$

Com as equações (7.6.4) e (7.6.5) a fórmula (7.6.2) transforma-se em

$$\int_G \langle \text{rot}(f), \nu \rangle dS = \int_{\partial G} \langle f, \tau \rangle dS. \quad (7.6.6)$$

Isto é o Teorema de Stokes na forma clássica.

Exemplo 7.6.4. Consideramos a 2-forma

$$\omega := \frac{xdy \wedge dz + ydz \wedge dx + zdx \wedge dy}{(x^2 + y^2 + z^2)^{3/2}}$$

em $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$. Esta 2-forma é fechada mas não exata. Sua integral sobre a esfera unitária $\mathbb{S}^2 \subset \mathbb{R}^3$ é

$$\int_{\mathbb{S}^2} \omega = \text{Vol}_2(\mathbb{S}^2) = 4\pi.$$

(Exercício: Prove esta afirmação com o Teorema de Stokes.)

O seguinte Teorema de Retração é uma aplicação interessante do Teorema de Stokes. Como esta aplicação afeta o caso $d = n$, também pode-se provar o Teorema de Retração evidentemente direto com ajuda do teorema de divergência de Gauß (Teorema 6.4.9). Porém a prova com o formalismo de formas diferenciais é consideravelmente mais simples. Se quer-se prova-lo direto com ajuda da divergência de um campo vetorial adequado, recebe-se fórmulas bem complicadas, as quais saem em refazer de novo, para o caso $d = n$, os argumentos na demonstração do Teorema de Stokes (Teorema 7.5.4).

Teorema 7.6.5 (Teorema de Retração). *Seja $\Omega \subset \mathbb{R}^n$ um subconjunto aberto limitado com bordo suave, e seja $U \subset \mathbb{R}^n$ um conjunto aberto, tal que $\overline{\Omega} \subset U$. Então não existe nenhuma aplicação suave $\phi : U \rightarrow \mathbb{R}^n$ com*

$$\phi(\overline{\Omega}) \subset \partial\Omega$$

e $\phi(x) = x$ para todos os $x \in \partial\Omega$.

Demonstração. Suponhamos que exista uma tal aplicação suave $\phi : U \rightarrow \mathbb{R}^n$. Então vale $\phi(x + t\xi) \in \partial\Omega$ para todo $x \in \Omega$, todo vetor $\xi \in \mathbb{R}^n$, e todo $t \in \mathbb{R}$ suficientemente pequeno. Disso segue

$$d\phi(x)\xi \in T_{\phi(x)}\partial\Omega$$

para todos os $x \in \Omega$ e todos os $\xi \in \mathbb{R}^n$. Portanto as colunas da matriz jacobiana $d\phi(x)$ são linearmente dependente para todos os $x \in \Omega$, e dali por causa da continuidade das primeiras derivadas de ϕ , também são para todos os $x \in \overline{\Omega}$. Disso segue a equação

$$\det(d\phi(x)) = 0 \quad \text{para todos os } x \in \overline{\Omega}. \quad (7.6.7)$$

Consideramos ora a $(n - 1)$ -forma

$$\omega := x_1 dx_2 \wedge \cdots \wedge dx_n.$$

Esta tem como diferencial

$$d\omega = dx_1 \wedge dx_2 \wedge \cdots \wedge dx_n$$

a forma volume padrão no \mathbb{R}^n . Dali segue de (7.6.7) e parte (iv) de Lema 7.1.12 com $k = n$ a fórmula

$$\phi^* d\omega = \det(d\phi) d\omega = 0.$$

Daí vale, segundo o Teorema 7.2.7 e Teorema 7.5.4,

$$\begin{aligned} 0 &= \int_G \phi^* d\omega \\ &= \int_G d\phi^* \omega \\ &= \int_{\partial G} \phi^* \omega \\ &= \int_{\partial G} \omega \\ &= \int_G d\omega \\ &= \text{Vol}_n(G) \\ &> 0. \end{aligned}$$

Aqui a quarta igualdade segue da hipótese que a aplicação

$$\phi|_{\partial G} : \partial G \rightarrow \partial G$$

é a identidade. Temos então gerado uma contradição a esta hipótese, a qual mostra que uma aplicação suave $\phi : U \rightarrow \partial\Omega$ com $\phi(x) = x$ para todos os $x \in \partial\Omega$ não pode existir. Com isso o Teorema 7.6.5 fica provado. \square

O Teorema de Retração tem várias consequências interessantes. A primeira afeta pontos fixos de uma aplicação suave $f : B \rightarrow B$, onde

$$B := \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$$

denota a bola unitária em \mathbb{R}^n com a norma euclidiana.

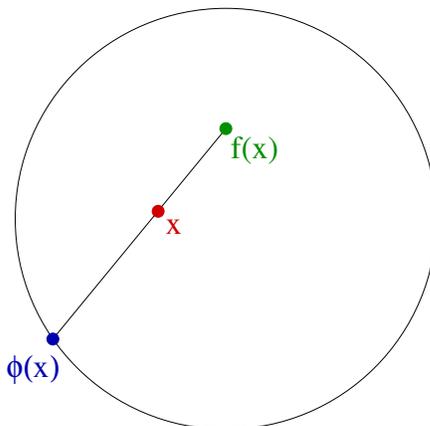


Figura 7.6: O Teorema de Ponto Fixo de Brouwer.

Corolário 7.6.6. *Toda aplicação suave $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ com $f(B) \subset B$ possui um ponto fixo $x = f(x) \in B$.*

Demonstração. Denotamos o bordo de B de

$$S := \partial B = \{x \in \mathbb{R}^n \mid \|x\| = 1\} = \mathbb{S}^{n-1}.$$

Ora suponhamos que exista uma aplicação suave $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, a qual leva a bola B em si mesma mas não possui nenhum ponto fixo em B . Então

$$U := \left\{ x \in \mathbb{R}^n \mid f(x) \neq x, \|x\|^2 < 1 + \frac{\langle x, x - f(x) \rangle^2}{\|x - f(x)\|^2} \right\}$$

é um subconjunto aberto de \mathbb{R}^n , o qual contém B , e existe uma aplicação suave $\phi : U \rightarrow S$, a qual satisfaz a condição $\phi(x) = x$ para todos os $x \in S$. Uma fórmula explícita para uma tal aplicação é, para $x \in U$, a seguinte

$$\begin{aligned} \phi(x) &:= x - t(x)(x - f(x)) \\ t(x) &:= \frac{1}{\|x - f(x)\|} \left(\sqrt{1 - \|x\|^2 + \frac{\langle x, x - f(x) \rangle^2}{\|x - f(x)\|^2}} - \frac{\langle x, x - f(x) \rangle}{\|x - f(x)\|} \right). \end{aligned}$$

Geometricamente recebe-se esta fórmula no que vai-se no rasto da reta de $f(x)$ na direção x , enquanto até encontra-se a esfera S . Este ponto na esfera é o ponto imagem $\phi(x)$ (veja a Figura 7.6). Mas como uma tal aplicação $\phi : U \rightarrow S$ não existe segundo o Teorema 7.6.5, então f deve ter sim, em contrário à nossa hipótese, um ponto fixo em B , e com isso o Corolário 7.6.6 fica provado. \square

Corolário 7.6.7 (Teorema de Ponto Fixo de Brouwer).

Toda aplicação contínua $f : B \rightarrow B$ possui um ponto fixo $x = f(x) \in B$.

Demonstração. A prova requer o teorema de aproximação de Weierstraß, o qual diz que para toda função contínua $f : B \rightarrow \mathbb{R}$ existe uma sequência de polinômios $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ (em n variáveis), a qual converge em B uniformemente para f . Este resultado também vale evidentemente para aplicações com valores em \mathbb{R}^n .

Suponhamos ora que f não possua um ponto fixo. Então vale

$$\varepsilon := \inf_{x \in B} \|x - f(x)\| > 0.$$

Segundo o teorema de aproximação de Weierstraß existe um polinômio

$$p : \mathbb{R}^n \rightarrow \mathbb{R}^n,$$

o qual satisfaz a condição

$$\|p(x) - f(x)\| \leq \frac{\varepsilon}{2} \quad \text{para todos os } x \in B.$$

Ora definimos a aplicação $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ através de

$$g(x) := \frac{p(x)}{1 + \varepsilon/2} \quad \text{para } x \in \mathbb{R}^n.$$

Esta aplicação g é suave e leva a bola unitária B em si mesma, porque para todo ponto $x \in B$ a desigualdade

$$\|p(x)\| \leq \|f(x)\| + \|p(x) - f(x)\| \leq 1 + \frac{\varepsilon}{2}$$

é satisfeita. Para todo ponto $x \in B$ vale igualmente a desigualdade

$$\begin{aligned} \|g(x) - x\| &= \left\| f(x) - x + \frac{p(x) - f(x)}{1 + \varepsilon/2} - \frac{\varepsilon/2}{1 + \varepsilon/2} f(x) \right\| \\ &\geq \|f(x) - x\| - \frac{\|p(x) - f(x)\| + (\varepsilon/2) \|f(x)\|}{1 + \varepsilon/2} \\ &\geq \varepsilon - \frac{\varepsilon/2 + \varepsilon/2}{1 + \varepsilon/2} \\ &> 0. \end{aligned}$$

Portanto g não possui nenhum ponto fixo em B , em contradição a Corolário 7.6.6. Com isso o Corolário 7.6.7 fica provado. \square

Corolário 7.6.7 é um caso especial do **Teorema de Ponto Fixo de Brouwer**. Aquele foi provado por Luitzen Egbertus Jan Brouwer no ano 1910 e diz que toda aplicação contínua de um subconjunto compacto convexo do \mathbb{R}^n em si mesmo possui um ponto fixo. O Teorema de Ponto Fixo de Brouwer foi estendido 1930 por Juliusz Schauder a subconjuntos compactos convexos de um espaço de Banach. Um teorema, o qual foi provado independente por Oskar Perron em 1907 e por Georg Frobenius em 1912, diz que toda matriz quadrada com entradas positivas possui um autovetor com entradas não-negativas. Este teorema pode ser deduzido de Corolário 7.6.7.

Corolário 7.6.8 (Perron–Frobenius).

Seja $A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$ uma matriz com $a_{ij} > 0$ para todo $i, j = 1, \dots, n$. Então existe um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ e um número real λ com

$$x_i \geq 0 \text{ para } i = 1, \dots, n, \quad \sum_{i=1}^n x_i = 1, \quad \lambda > 0, \quad Ax = \lambda x. \quad (7.6.8)$$

Demonstração. Para $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ defina a norma $\|\xi\|_1 := \sum_i |\xi_i|$ e a norma $\|\xi\|_2 := \sqrt{\sum_i \xi_i^2}$. O símlice padrão é o conjunto

$$\Delta := \left\{ x \in \mathbb{R}^n \mid x_i \geq 0 \text{ para } i = 1, \dots, n, \sum_{i=1}^n x_i = 1 \right\}$$

e definimos a aplicação contínua $f : \Delta \rightarrow \Delta$ através de

$$f(x) := \frac{Ax}{\|Ax\|_1} \quad \text{para } x \in \Delta.$$

O símlice Δ é homeomorfo à bola $B := \{y \in \mathbb{R}^{n-1} \mid \|y\|_2 \leq 1\}$. Um homeomorfismo explícito $\phi = (\phi_1, \dots, \phi_n) : B \rightarrow \Delta$ é dado através de

$$\begin{aligned} \phi_i(y) &:= \frac{1}{n} \left(1 - \frac{y_i \|y\|_2}{\max\{y_1, \dots, y_{n-1}, -\sum_{j=1}^{n-1} y_j\}} \right), \quad i = 1, \dots, n-1 \\ \phi_n(y) &:= \frac{1}{n} \left(1 + \frac{\sum_{j=1}^{n-1} y_j \|y\|_2}{\max\{y_1, \dots, y_{n-1}, -\sum_{j=1}^{n-1} y_j\}} \right) \end{aligned} \quad (7.6.9)$$

para $y \in B$. Ora a aplicação $g := \phi^{-1} \circ f \circ \phi : B \rightarrow B$ é contínua e possui daí, segundo o Corolário 7.6.7, um ponto fixo $y = g(y) \in B$. Disso segue que $x := \phi(y) \in \Delta$ é um ponto fixo de f . Isto é equivalente à equação $Ax = \|Ax\|_1 x$, e daí x satisfaz a afirmação de Corolário 7.6.8 com $\lambda := \|Ax\|_1 > 0$. \square

Apêndice A

Fecho e Interior

A.1 Conceitos básicos da topologia

Seja (X, d) um **espaço métrico**, ou seja X é um conjunto e $d : X \times X \rightarrow \mathbb{R}$ é uma função com as seguintes propriedades.

(d1) Para todos os $x, y \in X$ vale $d(x, y) \geq 0$ e $d(x, y) = 0 \iff x = y$.

(d2) Para todos os $x, y \in X$ vale $d(x, y) = d(y, x)$.

(d3) Para todos os $x, y, z \in X$ vale $d(x, z) \leq d(x, y) + d(y, z)$.

A desigualdade em (d3) é chamada de **desigualdade triangular**. A função d é chamada de **função distância**. Um subconjunto $U \subset X$ é chamado de **aberto** se para todo elemento $x \in U$ existe um $\varepsilon > 0$ tal que

$$B_\varepsilon(x) := \{y \in X \mid d(x, y) < \varepsilon\} \subset U.$$

Particularmente o conjunto $B_r(x)$ é aberto para todo $x \in X$ e todo $r > 0$; chamado de **bola aberta em X com centro x e raio r** . Conjuntos abertos têm as seguintes propriedades.

(A1) X e \emptyset são subconjuntos abertos de X .

(A2) Se U_1, \dots, U_n são subconjuntos abertos de X , então $\bigcap_{i=1}^n U_i$ também é.

(A3) Se I é um conjunto e $U_i \subset X$ é um subconjunto aberto para todo $i \in I$, então a união $\bigcup_{i \in I} U_i$ também é aberto.

Diz-se que uma sequência $(x_n)_{n \in \mathbb{N}}$ em X **converge** para $x \in X$ se para todo $\varepsilon > 0$ existe um $n_0 \in \mathbb{N}$, tal que para todo $n \in \mathbb{N}$ vale o seguinte

$$n \geq n_0 \quad \implies \quad d(x_n, x) < \varepsilon.$$

Um subconjunto $A \subset X$ chama-se de **fechado** se para toda sequência $(x_n)_{n \in \mathbb{N}}$ em A a qual converge para um elemento $x \in X$ seu limite x também é um elemento de A .

Um subconjunto $A \subset X$ é fechado se e somente se seu complemento $U := X \setminus A$ é aberto (veja [6]). Conjuntos fechados têm as seguintes propriedades.

- (F1) O conjunto vazio e o espaço X por inteiro são subconjuntos fechados de X .
- (F2) Se A_1, \dots, A_n são subconjuntos fechados de X , então $\bigcup_{i=1}^n A_i$ também é fechado.
- (F3) Se I é um conjunto e $A_i \subset X$ é um subconjunto fechado para todo $i \in I$, então a interseção $\bigcap_{i \in I} A_i$ também é fechado.

Sejam (X, d_X) e (Y, d_Y) dois espaços métricos. Uma aplicação $f : X \rightarrow Y$ chama-se de **contínua** se para todo $x_0 \in X$ e todo $\varepsilon > 0$ existe um $\delta > 0$, tal que para todo $x \in X$ vale o seguinte

$$d_X(x, x_0) < \delta \quad \implies \quad d_Y(f(x), f(x_0)) < \varepsilon.$$

Para uma aplicação $f : X \rightarrow Y$ qualquer as seguintes afirmações são equivalentes.

- (C1) f é contínua.
- (C2) Se $(x_n)_{n \in \mathbb{N}}$ é uma sequência em X a qual converge para $x \in X$, então $f(x_n)$ converge para $f(x) \in Y$.
- (C3) Pre-imagens de abertos são abertos: Se $V \subset Y$ é aberto, então sua pre-imagem $f^{-1}(V) := \{x \in X \mid f(x) \in V\}$ é aberto em X .

(Veja Análise I [6].) Uma aplicação $f : X \rightarrow Y$ chama-se de **homeomorfismo** se é bijetivo e f e f^{-1} são contínuas.

A.2 Os conceitos fecho e interior

Esta seção trata os conceitos do *fecho* e do *interior* de um subconjunto de um espaço métrico. Seja então (X, d) um espaço métrico. O exemplo padrão é um espaço vetorial normado $(X, \|\cdot\|)$ munido da função distância definida assim $d(x, y) := \|x - y\|$ para $x, y \in X$. Porém nesta seção (X, d) pode ser um espaço métrico arbitrário. Para $x \in X$ e $r > 0$ denotamos, como é nossa praxe, com o símbolo $B_r(x) := \{y \in X \mid d(x, y) < r\}$ a bola aberta de raio r e centro x .

Definição A.2.1. *Seja dado um subconjunto $A \subset X$. O fecho de A é o conjunto*

$$\bar{A} := \{x \in X \mid \forall \varepsilon > 0 \text{ vale } B_\varepsilon(x) \cap A \neq \emptyset\}.$$

O interior de A é o conjunto

$$\overset{\circ}{A} := \{x \in X \mid \exists \varepsilon > 0, \text{ tal que } B_\varepsilon(x) \subset A\}.$$

O bordo de A é o conjunto

$$\partial A := \bar{A} \setminus \overset{\circ}{A} = \left\{ x \in X \mid \begin{array}{l} \forall \varepsilon > 0 \text{ vale } B_\varepsilon(x) \cap A \neq \emptyset \\ \text{e } B_\varepsilon(x) \cap (X \setminus A) \neq \emptyset \end{array} \right\}.$$

Exemplo A.2.2. *Seja $X = \mathbb{R}$ o conjunto dos números reais com a função distância padrão $d(x, y) = |x - y|$. Sejam $a < b$ dois números reais e seja A o intervalo semi-aberto*

$$A := [a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}.$$

Então o fecho, o interior, e o bordo de A são dados através de

$$\bar{A} = [a, b], \quad \overset{\circ}{A} = (a, b), \quad \partial A = \{a, b\}$$

respectivamente.

Em todo espaço métrico (X, d) o fecho de um subconjunto A sempre é um conjunto fechado, e o interior de A sempre um conjunto aberto. Com isso o bordo também é um conjunto fechado. A seguinte proposição resume as propriedades mais importantes do fecho e do interior.

Proposição A.2.3. *Seja (X, d) um espaço métrico e $A \subset X$.*

(i) *Seja $x \in X$. Então $x \in \bar{A}$ se e somente se existe uma sequência $x_n \in A$ a qual converge para x .*

(ii) *$X \setminus \overset{\circ}{A} = \overline{X \setminus A}$, e $\partial A = \bar{A} \cap \overline{X \setminus A}$.*

(iii) *$\bar{A} = A \cup \partial A$ e $\overset{\circ}{A} = A \setminus \partial A$.*

(iv) *\bar{A} é o menor subconjunto fechado de X contendo A .*

(v) *$\overset{\circ}{A}$ é o maior subconjunto aberto de X contido em A .*

(vi) *A é fechado se e somente se $\bar{A} = A$ (ou seja $\partial A \subset A$).*

(vii) *A é aberto se e somente se $\overset{\circ}{A} = A$ (ou seja $\partial A \cap A = \emptyset$).*

Demonstração. Provaremos (i). Seja $x_n \in A$ uma sequência a qual converge para um elemento $x \in X$. Seja $\varepsilon > 0$. Segundo a definição de convergência existe um número natural $n_0 \in \mathbb{N}$, tal que para todo $n \in \mathbb{N}$ vale

$$n \geq n_0 \quad \implies \quad d(x, x_n) < \varepsilon.$$

Com isto vale $x_n \in B_\varepsilon(x)$ para todo $n \geq n_0$ e daí segue $B_\varepsilon(x) \cap A \neq \emptyset$. Como isto vale para todo $\varepsilon > 0$, segue dali $x \in \bar{A}$. Vice-versa, seja dado $x \in A$. Então $B_{1/n}(x) \cap A \neq \emptyset$ para todo número natural $n \in \mathbb{N}$. Segundo o axioma da escolha, versão contável, existe então uma sequência $x_n \in A$, tal que para todo $n \in \mathbb{N}$ vale a desigualdade $d(x, x_n) < 1/n$. Segue dali $\lim_{n \rightarrow \infty} d(x, x_n) = 0$ e daí $x = \lim_{n \rightarrow \infty} x_n$. Com isso (i) fica provado.

Provaremos (ii). Segundo a definição de fecho e interior vale

$$\begin{aligned} \overline{X \setminus A} &= \{x \in X \mid \forall \varepsilon > 0 \text{ vale } B_\varepsilon(x) \cap (X \setminus A) \neq \emptyset\} \\ &= \{x \in X \mid \forall \varepsilon > 0 \text{ vale } B_\varepsilon(x) \not\subset A\} \\ &= X \setminus \overset{\circ}{A}. \end{aligned}$$

Segue daí

$$\partial A = \bar{A} \setminus \overset{\circ}{A} = \bar{A} \cap (X \setminus \overset{\circ}{A}) = \bar{A} \cap \overline{X \setminus A}.$$

Com isso (ii) fica provado.

Provaremos (iii). Vale $\overset{\circ}{A} \subset A \subset \bar{A}$. Segundo a definição do bordo segue daí

$$A \cup \partial A = A \cup (\bar{A} \setminus \overset{\circ}{A}) = A \cup \bar{A} = \bar{A}$$

e

$$A \setminus \partial A = A \setminus (\bar{A} \setminus \overset{\circ}{A}) = A \setminus (A \setminus \overset{\circ}{A}) = \overset{\circ}{A}.$$

Com isso (iii) fica provado.

Provaremos (iv). Para já provaremos que \bar{A} é fechado. Ora, seja $x_n \in \bar{A}$ uma sequência a qual converge para um elemento $x \in X$. Seja dado um número $\varepsilon > 0$. Segundo a definição da convergência existe um $n \in \mathbb{N}$ com

$$d(x, x_n) < \frac{\varepsilon}{2}.$$

Como $x_n \in \bar{A}$, vale $B_{\varepsilon/2}(x_n) \cap A \neq \emptyset$. Escolha um elemento $a \in B_{\varepsilon/2}(x_n) \cap A$. Então vale a desigualdade triangular

$$d(x, a) \leq d(x, x_n) + d(x_n, a) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Segue daí

$$B_\varepsilon(x) \cap A \neq \emptyset.$$

Como $\varepsilon > 0$ foi escolhido arbitrariamente, segue dali $x \in \bar{A}$. Com isto temos mostrado que \bar{A} é um subconjunto fechado de X . Seja ora $B \subset X$ um conjunto fechado qualquer contendo A . Resta mostrar que $\bar{A} \subset B$. Seja ora $x \in \bar{A}$. Então existe segundo (i) uma sequência $x_n \in A$ a qual converge para x . Portanto x_n é uma sequência em B e, como B é fechado, segue disso que seu limite x também é um elemento de B . Com isso (iv) fica provado.

A afirmação (v) segue de (ii) e (iv). A afirmação (vi) segue imediatamente de (iii) e (iv). A afirmação (vii) segue imediatamente de (iii) e (v). Com isso a Proposição A.2.3 fica provada. \square

Exercício A.2.4. Num espaço métrico (X, d) denotamos frequentemente a bola fechada com centro $x \in X$ e raio $r > 0$ com

$$\bar{B}_r(x) := \{y \in X \mid d(x, y) \leq r\}.$$

Este conjunto sempre é fechado.

(i) Dado um espaço vetorial normado $(X, \|\cdot\|)$ munido da função distância induzida

$$d(x, y) := \|x - y\|$$

para $x, y \in X$, mostre que $\bar{B}_r(x)$ é o fecho de $B_r(x)$.

(ii) Encontre um exemplo de um espaço métrico (X, d) , de um ponto $x \in X$, e um número $r > 0$, tal que $\bar{B}_r(x)$ não é o fecho de $B_r(x)$, ou seja, a inclusão $\bar{B}_r(x) \subsetneq \overline{B_r(x)}$ é estrita.

Exercício A.2.5. Seja (X, d) um espaço métrico e sejam $A, B \subset X$. Então vale

$$\overline{A \cup B} = \overline{A} \cup \overline{B}, \quad \overline{A \cap B} \subset \overline{A} \cap \overline{B}, \quad (\text{A.2.1})$$

e

$$\overset{\circ}{A} \cup \overset{\circ}{B} \subset A \overset{\circ}{\cup} B, \quad \overset{\circ}{A} \cap \overset{\circ}{B} = A \overset{\circ}{\cap} B. \quad (\text{A.2.2})$$

Para as escolhas $X := \mathbb{R}$ e $A := \mathbb{Q}$ e $B := \mathbb{R} \setminus \mathbb{Q}$ recebemos $\overset{\circ}{A} = \overset{\circ}{B} = A \cap B = \emptyset$ e $\overline{A} = \overline{B} = A \cup B = X$.

Exercício A.2.6. Seja (X, d) um espaço métrico e seja $A \subset X$. Então vale

$$\overset{\circ}{A} \subset \overline{\overset{\circ}{A}}, \quad \partial \overline{A} \subset \partial A, \quad (\text{A.2.3})$$

e

$$\overline{\overset{\circ}{A}} \subset \overline{A}, \quad \partial \overset{\circ}{A} \subset \partial A. \quad (\text{A.2.4})$$

Para $A := \mathbb{Q}$ e $X := \mathbb{R}$ estas são inclusões estritas com $\overset{\circ}{A} = \partial \overset{\circ}{A} = \partial \overline{A} = \emptyset$ e $\overline{A} = \partial A = X$. Para intervalos em \mathbb{R} vale igualdade em (A.2.3) e (A.2.4).

Exercício A.2.7. Sejam (X, d_X) e (Y, d_Y) espaços métricos, e $f : X \rightarrow Y$ uma aplicação contínua. Então vale

$$f(\overline{A}) \subset \overline{f(A)} \quad (\text{A.2.5})$$

para $A \subset X$ e

$$f^{-1}(\overline{B}) \supset \overline{f^{-1}(B)}, \quad f^{-1}(\overset{\circ}{B}) \subset f^{-1}(\overset{\circ}{B}) \quad (\text{A.2.6})$$

para $B \subset Y$. Para $A = X = Y = \mathbb{R}$ e $f = \tan^{-1}$ tem-se uma inclusão estrita em (A.2.5). Para $X = Y = \mathbb{R}$, $f(x) := x^2$, e $A = \{-1\} \cup (-1/2, 1) \cup (1, 2)$ recebemos $f(A) = [0, 4)$, $f(\overset{\circ}{A}) = (0, 4)$, e $f(\overset{\circ}{A}) = [0, 1) \cup (1, 4)$; portanto neste exemplo o interior do conjunto $f(A)$ nem é contido em $f(\overset{\circ}{A})$, nem o contem o conjunto $f(\overset{\circ}{A})$. Em outras palavras, para o interior não existe uma relação análoga da inclusão (A.2.5).

As inclusões em (A.2.6) geralmente são estritas: Seja $f \equiv c : \mathbb{R} \rightarrow \mathbb{R}$ uma função constante nos números reais. Escolha um subconjunto aberto $B \subset \mathbb{R}$ cujo bordo contem c . Então $f^{-1}(\overline{B}) = \mathbb{R} \not\subset \emptyset = \overline{\emptyset} = \overline{f^{-1}(B)}$. Escolha um subconjunto fechado $B \subset \mathbb{R}$ cujo bordo contem c . Então c não é elemento do interior de B , por isso $f^{-1}(\overset{\circ}{B}) = \emptyset \not\supset \mathbb{R} = \overset{\circ}{\mathbb{R}} = f^{-1}(\overset{\circ}{B})$.

Apêndice B

Espaços Conexos

B.1 A topologia relativa

Seja (X, d) um espaço métrico e $Y \subset X$ um subconjunto qualquer. Então a restrição da função distância $d : X \times X \rightarrow \mathbb{R}$ a $Y \times Y$ é uma função distância de novo, denotada

$$d_Y := d|_{Y \times Y} : Y \times Y \rightarrow \mathbb{R}.$$

Assim (Y, d_Y) torna-se um espaço métrico.

Definição B.1.1. *A topologia induzida em Y através da função distância restrita d_Y também chama-se de **topologia relativa** em Y . Um subconjunto $V \subset Y$ é chamado de **Y -aberto** se é aberto com respeito à topologia relativa. Um subconjunto $B \subset Y$ é chamado de **Y -fechado** se é fechado com respeito à topologia relativa.*

O seguinte lema caracteriza os subconjuntos Y -abertos e Y -fechados de Y com a ajuda dos subconjuntos abertos e fechados de X .

Lema B.1.2. *Seja (X, d) um espaço métrico e seja $Y \subset X$ um subconjunto qualquer.*

- (i) *Um subconjunto $V \subset Y$ é Y -aberto se e somente se existe um subconjunto aberto $U \subset X$, tal que $V = U \cap Y$.*
- (ii) *Um subconjunto $B \subset Y$ é Y -fechado se e somente se existe um subconjunto fechado $A \subset X$, tal que $B = A \cap Y$.*

Demonstração. Provaremos (i). Se $V \subset Y$ é aberto com respeito d_Y , então existe para todo $y \in V$ um $\varepsilon > 0$, tal que

$$B_\varepsilon(y; Y) := \{x \in Y \mid d(x, y) < \varepsilon\} \subset V.$$

Para $y \in V$ seja $\varepsilon(y) := \sup\{\varepsilon \in (0, 1] \mid B_\varepsilon(y; Y) \subset V\}$. Então $B_{\varepsilon(y)}(y; Y) \subset V$ para todo $y \in V$. Por isso o conjunto

$$U := \bigcup_{y \in V} B_{\varepsilon(y)}(y; X)$$

é aberto em X e vale

$$U \cap Y = \bigcup_{y \in V} (B_{\varepsilon(y)}(y; X) \cap Y) = \bigcup_{y \in V} B_{\varepsilon(y)}(y; Y) = V.$$

Se existe, do outro lado, um subconjunto aberto $U \subset X$ com $U \cap Y = V$, então $y \in V$ implica $y \in U$, daí existe um $\varepsilon > 0$ com $B_\varepsilon(y; X) \subset U$, e por isso vale $B_\varepsilon(y; Y) = B_\varepsilon(y; X) \cap Y \subset U \cap Y = V$. Consequentemente V é aberto com respeito d_Y . Com isso a parte (i) fica provada.

Para mostrar (ii) suponhamos para já que $B \subset Y$ seja fechado com respeito d_Y . Então

$$V := Y \setminus B$$

é aberto com respeito d_Y . Segundo (i) existe então um conjunto aberto $U \subset X$, tal que $U \cap Y = V$. Assim

$$A := X \setminus U$$

é um subconjunto fechado de X e vale

$$A \cap Y = (X \setminus U) \cap Y = Y \setminus (U \cap Y) = Y \setminus V = B.$$

Do outro lado, se $A \subset X$ é fechado e $B = A \cap Y$, então

$$U := X \setminus A$$

é um subconjunto aberto de X e vale

$$U \cap Y = (X \setminus A) \cap Y = Y \setminus (A \cap Y) = Y \setminus B.$$

Daí segundo (i) o conjunto $Y \setminus B$ é aberto com respeito d_Y , e consequentemente B é fechado com respeito d_Y . Com isso o Lema B.1.2 fica provado. \square

Exemplo B.1.3. Seja $X = \mathbb{R}$ munido da métrica padrão

$$d(x, y) := |x - y|$$

para $x, y \in \mathbb{R}$.

(i) Seja

$$Y := [0, 1]$$

então $[0, b)$, para $0 < b \leq 1$, é um subconjunto Y -aberto de Y . Porém um subconjunto $B \subset Y$ é Y -fechado se e somente se é fechado (porque o próprio Y é um subconjunto fechado de X).

(ii) Seja

$$Y := (0, 1)$$

então $(0, b]$, para $0 < b < 1$, é um subconjunto Y -fechado de Y . Porém um subconjunto $V \subset Y$ é Y -aberto se e somente se é aberto (porque o próprio Y é um subconjunto aberto de X).

B.2 O conceito conexo

Definição B.2.1. Um espaço métrico (X, d) chama-se de **conexo** se não pode-se-lê representar como uma união disjunta de dois subconjuntos abertos não-vazios, ou seja, se para quaisquer dois subconjuntos abertos $U, V \subset X$ vale o seguinte:

$$U \cup V = X, \quad U \cap V = \emptyset \quad \implies \quad U = \emptyset \quad \text{ou} \quad V = \emptyset$$

ou seja, o conjunto vazio e o espaço completo são os únicos subconjuntos de X , as quais são tanto aberto como fechado.

Um subconjunto $Y \subset X$ de um espaço métrico (X, d) chama-se de **conexo**, se é conexo com respeito à métrica induzida d_Y . Segundo o Lema B.1.2 isso significa que para quaisquer dois subconjuntos abertos $U, V \subset X$ vale a seguinte afirmação:

$$A \subset U \cup V, \quad A \cap U \cap V = \emptyset \quad \implies \quad A \cap U = \emptyset \quad \text{ou} \quad A \cap V = \emptyset.$$

O seguinte teorema mostra que os subconjuntos conexos de \mathbb{R} são precisamente os intervalos.

Teorema B.2.2. *Seja $X = \mathbb{R}$ munido da métrica padrão*

$$d(x, y) := |x - y|.$$

Um subconjunto $I \subset \mathbb{R}$ é conexo se e somente se é um intervalo.

Demonstração. Se $I \subset \mathbb{R}$ não é um intervalo, então existem três números reais $a, b, c \in \mathbb{R}$ com

$$a, b \in I, \quad c \notin I, \quad a < c < b.$$

Definimos

$$U := \{x \in \mathbb{R} \mid x < c\}, \quad V := \{x \in \mathbb{R} \mid x > c\}$$

assim são conjuntos abertos com

$$I \subset U \cup V, \quad I \cap U \neq \emptyset, \quad I \cap V \neq \emptyset.$$

Então I não é conexo.

Suponhamos por outro lado que I seja um intervalo e $A \subset I$ seja um subconjunto o qual é tanto aberto, como também fechado (com respeito à métrica d_I), e qual nem é vazio, nem o espaço inteiro. Então

$$B := I \setminus A$$

também é não-vazio e escolhemos $a \in A$ e $b \in B$. Podemos supor sem perda da generalidade que $a < b$. Seja

$$c := \sup(A \cap [a, b]).$$

Então vale $c \in [a, b] \subset I$ e existe uma sequência $a_k \in A$ com $a_k \leq c$ a qual converge para c . Como A é fechado com respeito d_I , segue daí que $c \in A$ e por isso $c < b$. Com isto o intervalo semi-aberto $(c, b]$ é contido em B . Portanto vale $c + 1/n \in B$ para todo número natural n suficientemente grande, e dali

$$c = \lim_{n \rightarrow \infty} (c + 1/n) \in B$$

porque B também é fechado com respeito d_I . Também temos mostrado que c é um elemento tanto de A , como de B , em contradição a $A \cap B = \emptyset$. Com isso o Teorema B.2.2 fica provado. \square

Teorema B.2.3. *Sejam (X, d_X) e (Y, d_Y) espaços métricos e $f : X \rightarrow Y$ uma aplicação contínua; se $A \subset X$ é conexo então também é $f(A) \subset Y$.*

Demonstração. Sejam $U, V \subset Y$ dois conjuntos abertos, tal que

$$f(A) \subset U \cup V, \quad f(A) \cap U \cap V = \emptyset.$$

A mostrar é que um dos conjuntos $f(A) \cap U$ ou $f(A) \cap V$ é vazio.

Como f é contínuo, os conjuntos $f^{-1}(U)$ e $f^{-1}(V)$ são subconjuntos abertos de X . Ora provaremos que estes subconjuntos tem as seguintes propriedades:

$$A \subset f^{-1}(U) \cup f^{-1}(V), \quad A \cap f^{-1}(U) \cap f^{-1}(V) = \emptyset. \quad (\text{B.2.1})$$

A saber se $a \in A$, então $f(a) \in f(A) \subset U \cup V$, e daí $f(a) \in U$ ou $f(a) \in V$; no primeiro caso vale $a \in f^{-1}(U)$ e no segundo caso vale $a \in f^{-1}(V)$. Isto prova a primeira afirmação em (B.2.1). A segunda afirmação em (B.2.1) mostra-se muito melhor indiretamente. Se existe um elemento $a \in A \cap f^{-1}(U) \cap f^{-1}(V)$, então vale $f(a) \in f(A) \cap U \cap V$, na contradição à nossa hipótese sobre U e V . Com isso ambas as afirmações em (B.2.1) ficam provadas.

Como A é conexo, segue de (B.2.1), que pelo menos um dos conjuntos $A \cap f^{-1}(U)$ ou $A \cap f^{-1}(V)$ é vazio. Mas se $A \cap f^{-1}(U) = \emptyset$, então isto significa que sob f nenhum elemento de A é levado a U , e conseqüentemente $f(A) \cap U = \emptyset$ também é vazio; do mesmo modo com V em vez de U . Com isso pelo menos um dos conjuntos $f(A) \cap U$ ou $f(A) \cap V$ é o conjunto vazio, como afirmado. Com isso o Teorema B.2.3 fica provado. \square

Teorema B.2.4. (Teorema do valor intermediário)

Seja (X, d) um espaço métrico conexo, seja $f : X \rightarrow \mathbb{R}$ uma função contínua, e sejam dados dois pontos $x, y \in X$ e um número real $c \in \mathbb{R}$, as quais satisfazam as desigualdades

$$f(x) < c < f(y).$$

Então existe um elemento $z \in X$ com $f(z) = c$.

Demonstração. Como X é conexo segue do Teorema B.2.3 que $f(X)$ também é conexo. Portanto $f(X)$ é um intervalo segundo o Teorema B.2.2. Como $f(x)$ e $f(y)$ são elementos de $f(X)$, segue da definição de um intervalo que $c \in f(X)$. Com isso o Teorema B.2.4 fica provado. \square

B.3 Conjuntos conexo por caminho

Seja (X, d) um espaço métrico.

Um subconjunto $A \subset X$ chama-se de **conexo por caminho** se para quaisquer dois elementos $x_0, x_1 \in A$ existe uma aplicação contínua

$$\gamma : [0, 1] \rightarrow A$$

a qual conecta os pontos

$$\gamma(0) = x_0, \quad \gamma(1) = x_1$$

um com o outro.

Teorema B.3.1. (i) *Todo subconjunto conexo por caminho de um espaço métrico (X, d) é conexo.*

(ii) *Todo subconjunto aberto e conexo de um espaço vetorial normado $(X, \|\cdot\|)$ é conexo por caminho.*

Demonstração. Provaremos (i). Seja (X, d) um espaço métrico e $A \subset X$ um subconjunto conexo por caminho. Fica a mostrar que A é conexo. Suponhamos por absurdo que não seja. Então existem conjuntos abertos $U_0, U_1 \subset X$, tal que

$$A \subset U_0 \cup U_1, \quad A \cap U_0 \cap U_1 = \emptyset, \quad A \cap U_0 \neq \emptyset, \quad A \cap U_1 \neq \emptyset.$$

Seja $x_0 \in A \cap U_0$ e $x_1 \in A \cap U_1$. Como A é conexo por caminho existe uma aplicação contínua $\gamma : [0, 1] \rightarrow A$ com $\gamma(0) = x_0$ e $\gamma(1) = x_1$. O intervalo $I := [0, 1]$ é a união disjunta dos dois conjuntos I -abertos

$$I_0 := \gamma^{-1}(U_0) = \{t \in I \mid \gamma(t) \in U_0\}$$

e

$$I_1 := \gamma^{-1}(U_1) = \{t \in I \mid \gamma(t) \in U_1\}.$$

Ambos estes conjuntos são não-vazios porque $0 \in I_0$ e $1 \in I_1$. Mas isso contradiz o fato que o intervalo $I = [0, 1]$ é conexo segundo o Teorema B.2.2. Esta contradição mostra que o conjunto A deve ter sido conexo, em contradição à nossa hipótese original, e com isso a parte (i) fica provada.

Provaremos (ii). Seja então $(X, \|\cdot\|)$ um espaço vetorial real e normado e seja $U \subset X$ um subconjunto aberto e conexo. Seja $x_0 \in U$ e seja $U_0 \subset U$ o conjunto de todos os pontos $x \in U$, as quais podem ser conectados em U através de um caminho contínuo, ou seja

$$U_0 := \left\{ x \in U \mid \begin{array}{l} \text{existe uma aplicação contínua} \\ \gamma : [0, 1] \rightarrow U \text{ com } \gamma(0) = x_0 \text{ e } \gamma(1) = x \end{array} \right\}.$$

Este conjunto é não-vazio porque $x_0 \in U_0$.

Provaremos que U_0 é um subconjunto aberto de X . Seja $x \in U_0$. Como U é aberto, existe um $\varepsilon > 0$, tal que $B_\varepsilon(x) = \{y \in X \mid \|x - y\| < \varepsilon\} \subset U$. Como $x \in U_0$, existe uma aplicação contínua $\beta : [0, 1] \rightarrow U$ com

$$\beta(0) = x_0, \quad \beta(1) = x.$$

Seja $y \in B_\varepsilon(x)$ e defina a aplicação $\gamma : [0, 1] \rightarrow X$ através de

$$\gamma(t) := \begin{cases} \beta(2t), & \text{para } 0 \leq t \leq 1/2, \\ (2 - 2t)x + (2t - 1)y, & \text{para } 1/2 < t \leq 1. \end{cases}$$

Esta aplicação é contínua, porque $\beta(1) = x$; assume valores em U , porque $B_\varepsilon(x) \subset U$; e satisfaz $\gamma(0) = x_0$ e $\gamma(1) = y$. Consequentemente $y \in U_0$ para todo $y \in B_\varepsilon(x)$. Portanto temos mostrado que para todo $x \in U_0$ existe um $\varepsilon > 0$ com $B_\varepsilon(x) \subset U_0$. Então U_0 é um subconjunto aberto de X .

Provaremos que U_0 é um subconjunto fechado de U com respeito à topologia relativa. Seja então $(x_k)_{k \in \mathbb{N}}$ uma sequência em U_0 a qual converge para um elemento $x \in U$. Escolha $\varepsilon > 0$, tal que $B_\varepsilon(x) \subset U$. Então existe um $k \in \mathbb{N}$ com $\|x_k - x\| < \varepsilon$. Como $x_k \in U_0$ existe uma aplicação contínua $\beta : [0, 1] \rightarrow U$ com

$$\beta(0) = x_0, \quad \beta(1) = x_k.$$

Defina $\gamma : [0, 1] \rightarrow X$ através de

$$\gamma(t) := \begin{cases} \beta(2t), & \text{para } 0 \leq t \leq 1/2, \\ (2 - 2t)x_k + (2t - 1)x, & \text{para } 1/2 < t \leq 1. \end{cases}$$

Esta aplicação é contínua, porque $\beta(1) = x_k$; assume valores em U , porque $B_\varepsilon(x) \subset U$; e satisfaz $\gamma(0) = x_0$ e $\gamma(1) = x$. Consequentemente $x \in U_0$. Portanto é mostrado que U_0 é um subconjunto de U não-vazio, U -aberto, e U -fechado. Como U é um subconjunto conexo de X segue daí que $U_0 = U$. Como $x_0 \in U$ foi escolhido arbitrariamente, isso significa que U é conexo por caminho. Com isso o Teorema B.3.1 fica provado. \square

B.4 Componentes conexos

Lema B.4.1. *Seja (X, d) um espaço métrico, seja I um conjunto arbitrário de índices, e seja $A_i \subset X$ um subconjunto conexo para todo $i \in I$. Se*

$$\bigcap_{i \in I} A_i \neq \emptyset$$

então o conjunto

$$A := \bigcup_{i \in I} A_i$$

é conexo.

Demonstração. Sejam $U, V \subset X$ dois subconjuntos abertos, tal que

$$A \subset U \cup V, \quad A \cap U \cap V = \emptyset.$$

Seja escolhido e fixado um elemento

$$x_0 \in \bigcap_{i \in I} A_i.$$

Então vale, ou $x_0 \in U$, ou $x_0 \in V$. Suponhamos $x_0 \in U$ e seja $i \in I$. Como $A_i \subset U \cup V$ e $A_i \cap U \cap V = \emptyset$ e A_i é conexo, vale ou $A_i \cap U = \emptyset$, ou $A_i \cap V = \emptyset$. Como $x_0 \in A_i \cap U$, segue daí $A_i \cap V = \emptyset$ e com isto $A_i \subset U$. Consequentemente $A_i \subset U$ para todo $i \in I$ e por isto $A \subset U$ e $A \cap V = \emptyset$. Por isso A é conexo. \square

Seja (X, d) um espaço métrico. Definimos em X a seguinte relação de equivalência: chama-se dois elementos $x, y \in X$ de **equivalente** (notação $x \sim y$) se existe um subconjunto conexo $A \subset X$ contendo x e y , ou seja $x, y \in A$. Segundo o Lema B.4.1 isto é realmente uma relação de equivalência. Mais além segue do Lema B.4.1 que as classes de equivalência desta relação de equivalência são conexas; são chamadas de **componentes conexos** X . Para $x_0 \in X$ chama-se o subconjunto

$$A_0 := \{x \in X \mid x \sim x_0\}$$

o **componente conexo** de x_0 ; isto é o maior subconjunto conexo de X o qual contem x_0 .

B.5 Exemplos

Exemplo B.5.1. Seja $X := \mathbb{R}^n$ munido da métrica euclidiana e seja $K \subset \mathbb{R}^n$ um subconjunto convexo. Este subconjunto é conexo por caminho, porque para $x_0, x_1 \in K$ a fórmula $\gamma(t) := (1-t)x_0 + tx_1$ define uma aplicação contínua $\gamma : [0, 1] \rightarrow K$ com $\gamma(0) = x_0$ e $\gamma(1) = x_1$. Então K é conexo por caminho segundo o Teorema B.3.1.

Exemplo B.5.2. Seja $X := \mathbb{Q}$ o conjunto dos número racionais munido da métrica padrão $d(x, y) := |x - y|$. Então todo componente conexo é composto de um único elemento só. Um tal espaço também chama-se **totalmente desconexo**.

Exemplo B.5.3. Seja $X = \mathbb{R}^n$ munido da métrica euclidiana. Este espaço é conexo. O complemento $Y := \mathbb{R}^n \setminus \{0\}$ do vetor nulo é conexo para $n \geq 2$, mas não para $n = 1$. **Exercício:** Prove estas afirmações. Conclua daí que para $n > 1$ não pode existir nenhum homeomorfismo de \mathbb{R} para \mathbb{R}^n .

Exemplo B.5.4. Seja $X = \mathbb{R}^n$ munido da métrica euclidiana. Então a esfera unitária $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$ é conexo. Seu complemento $Y := \mathbb{R}^n \setminus \mathbb{S}^{n-1}$ tem dois componentes conexos para $n > 1$, e três para $n = 1$. **Exercício:** Prove estas afirmações. **Dica:** Use o conceito “conexo por caminho”.

Exemplo B.5.5. O conjunto

$$A := \{(x, y) \mid x > 0, y = \sin(1/x)\} \cup \{(0, y) \mid -1 \leq y \leq 1\}$$

é um subconjunto fechado e conexo de \mathbb{R}^2 , mas não é conexo por caminho.

Exercício: Prove estas afirmações.

Exemplo B.5.6. O grupo

$$\mathrm{GL}^+(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \det(A) > 0\}$$

é conexo para todo número natural n . (Uma demonstração elegante desta afirmação encontra-se em [2, página 36/37]; veja também o Teorema D.4.1)). Segue daí que o conjunto das matrizes $n \times n$ com determinante negativo também é conexo. Consequentemente o grupo $\mathrm{GL}(n, \mathbb{R})$ é composto de exatamente dois componentes conexos.

Apêndice C

Espaços Métricos Compactos

A primeira seção deste apêndice caracteriza os conjuntos compactos sequenciais de um espaço métrico. Na segunda seção vamos provar o Teorema de Arzelà–Ascoli.

C.1 O conceito de compacidade

Seja X um conjunto e seja $K \subset X$. Uma **cobertura** de K é uma coleção $\mathcal{U} = \{U_i\}_{i \in I}$, indexada por um conjunto I , de subconjuntos $U_i \subset X$ tal que a união destes subconjuntos contem o conjunto K , ou seja $K \subset \bigcup_{i \in I} U_i$. Também pode-se descrever uma cobertura como um conjunto $\mathcal{U} \subset 2^X$ de subconjuntos de X com a propriedade que a união deles contem K , ou seja $K \subset \bigcup_{U \in \mathcal{U}} U$. A relação entre as duas formulações do mesmo fenômeno é que um subconjunto $U \subset X$ é um elemento de \mathcal{U} se e somente se existe um $i \in I$ tal que $U_i = U$. Se os conjuntos indexados U_i são dois-a-dois diferente (ou seja $U_i \neq U_j$ para $i \neq j$), então a aplicação $I \rightarrow \mathcal{U} : i \mapsto U_i$ é bijetiva.

Uma cobertura \mathcal{U} (resp. $\{U_i\}_{i \in I}$) é chamada de **finita**, se o conjunto \mathcal{U} (resp. o conjunto I de índices) é finito.

Uma **subcobertura** de \mathcal{U} (resp. $\{U_i\}_{i \in I}$) é um tal subconjunto de \mathcal{U} que ainda é uma cobertura (resp. um subconjunto $J \subset I$ tal que a coleção $\{U_i\}_{i \in J}$ ainda é uma cobertura).

Se (X, d) é um espaço métrico, então chama-se de **cobertura aberta** de um subconjunto $K \subset X$ uma cobertura composto de conjuntos abertos só; no caso de $\{U_i\}_{i \in I}$ isto significa que o conjunto U_i é aberto para todo $i \in I$.

Definição C.1.1. *Seja (X, d) um espaço métrico e seja $K \subset X$.*

(i) *O subconjunto K chama-se de **totalmente limitado**, se ou é vazio, ou para todo $\varepsilon > 0$ existe um número finito de elementos $\xi_1, \dots, \xi_m \in K$, tal que*

$$K \subset \bigcup_{i=1}^m B_\varepsilon(\xi_i).$$

*Aqui denotamos de $B_\varepsilon(\xi) := \{x \in X \mid d(x, \xi) < \varepsilon\}$ a **bola aberta em (X, d) de raio ε com centro ξ** .*

(ii) *O subconjunto K chama-se de **completo**, se toda sequência de Cauchy em K converge para um elemento de K .*

(iii) *O subconjunto K chama-se de **compacto**, se toda sequência em K possui uma subsequência convergente a qual converge para um elemento de K .*

O seguinte teorema mostra que os subconjuntos compactos de um espaço métrico são determinados só pela topologia do espaço (logo pela coletividade dos conjuntos abertos). Nos espaços topológicos gerais usa-se a condição (ii) no Teorema C.1.2 como definição do conceito da compacidade.

Teorema C.1.2. *Seja (X, d) um espaço métrico e seja $K \subset X$. Então são equivalentes as seguintes afirmações.*

(i) *K é compacto.*

(ii) *Toda cobertura aberta de K possui uma subcobertura finita.*

(iii) *K é completo e totalmente limitado.*

Demonstração. O conjunto vazio $K = \emptyset$ satisfaz todas as três condições. Suponhamos então que K não seja vazio. Para evitar confusão denotamos nesta demonstração o conjunto de todos os subconjuntos abertos de X de $\mathcal{T}(X, d) \subset 2^X$. (\mathcal{T} como em “topologia”.)

“(i) \implies (ii)”. Suponhamos que toda sequência em K possua uma subsequência a qual converge para um elemento de K . Seja

$$\mathcal{U} \subset \mathcal{T}(X, d)$$

uma cobertura aberta de X . Provaremos em dois passos que \mathcal{U} possui uma subcobertura finita.

Passo 1. *Existe um $\varepsilon > 0$, tal que para todo $x \in K$ existe um $U \in \mathcal{U}$ com $B_\varepsilon(x) \subset U$.*

Como fórmula lógica pode-se escrever esta afirmação da seguinte forma:

$$(\exists \varepsilon > 0)(\forall x \in K)(\exists U \in \mathcal{U})(B_\varepsilon(x) \subset U).$$

Provaremos isto indiretamente e suponhamos que o contrário seja o caso:

$$(\forall \varepsilon > 0)(\exists x \in K)(\forall U \in \mathcal{U})(B_\varepsilon(x) \not\subset U).$$

Escolha $\varepsilon := 1/n$ com $n \in \mathbb{N}$. Então existe segundo o axioma da escolha, versão contável, uma sequência $(x_n)_{n \in \mathbb{N}}$ em K , tal que para todo $n \in \mathbb{N}$ vale o seguinte:

$$B_{1/n}(x_n) \not\subset U \quad \forall U \in \mathcal{U}.$$

Como K é compacto, a sequência $(x_n)_{n \in \mathbb{N}}$ possui uma subsequência $(x_{n_i})_{i \in \mathbb{N}}$ a qual converge para um elemento $x_0 := \lim_{i \rightarrow \infty} x_{n_i} \in K$. Seja $U \in \mathcal{U}$, tal que $x_0 \in U$. Como U é aberto, existe um número $\varepsilon > 0$ com $B_\varepsilon(x_0) \subset U$. Como x_{n_i} converge para x_0 , existe um número natural $N \in \mathbb{N}$ tal que para todos os $i \in \mathbb{N}$ vale o seguinte:

$$i \geq N \quad \implies \quad d(x_{n_i}, x_0) < \frac{\varepsilon}{2}.$$

Seja $i > N$ com $1/n_i < \varepsilon/2$. Então $B_{1/n_i}(x_{n_i}) \subset B_{\varepsilon/2}(x_{n_i}) \subset B_\varepsilon(x_0) \subset U$, na contradição à construção da sequência (x_n) . Isso prova o Passo 1.

Passo 2. *\mathcal{U} possui uma subcobertura finita.*

Suponhamos que \mathcal{U} tenha uma subcobertura finita. Seja $\varepsilon > 0$ um número satisfazendo a afirmação do Passo 1. Vamos construir indutivamente uma sequência $(x_n)_{n \in \mathbb{N}}$ em K e uma sequência $(U_n)_{n \in \mathbb{N}}$ em \mathcal{U} , tal que

$$B_\varepsilon(x_1) \subset U_1$$

e

$$B_\varepsilon(x_n) \subset U_n, \quad x_n \notin U_1 \cup \dots \cup U_{n-1}$$

para todos os $n \in \mathbb{N}$ com $n \geq 2$. Para já escolhamos um elemento $x_1 \in K$. Então existe segundo o Passo 1 um elemento $U_1 \in \mathcal{U}$ com $B_\varepsilon(x_1) \subset U_1$. Suponhamos ora que tenhamos construído x_1, \dots, x_k e U_1, \dots, U_k , tal que nossas condições são satisfeitos para $n = 1, \dots, k$. Como \mathcal{U} não possui nenhuma subcobertura finita, deve existir um elemento $x_{k+1} \in K$ o qual não é contido em nenhum dos subconjuntos U_1, \dots, U_k . Segundo o Passo 1 existe um $U_{k+1} \in \mathcal{U}$ com $B_\varepsilon(x_{k+1}) \subset U_{k+1}$. Com isso o argumento da indução é concluído e temos construído as sequências desejadas.

Pela construção das seqüências vale $x_n \notin U_m$ para todo $m, n \in \mathbb{N}$, $n > m$. Como $B_\varepsilon(x_m) \subset U_m$, segue daí que $x_n \notin B_\varepsilon(x_m)$ e assim $d(x_n, x_m) \geq \varepsilon$, para todos os $m, n \in \mathbb{N}$ com $n > m$. Trocamos n e m , então recebemos a desigualdade $d(x_n, x_m) \geq \varepsilon$ para todo $m, n \in \mathbb{N}$ com $n \neq m$. Consequentemente a seqüência $(x_n)_{n \in \mathbb{N}}$ não tem nenhuma subsequência convergente, na contradição a (i). Com isso é mostrado que (ii) segue de (i).

“(ii) \implies (iii)”. Suponhamos agora que toda cobertura aberta possui uma subcobertura finita. Que K é totalmente limitado segue através da escolha da cobertura $\mathcal{U} := \{B_\varepsilon(x) \mid x \in K\}$. Segundo (ii) existem elementos $\xi_1, \dots, \xi_N \in K$ com $K \subset \bigcup_{i=1}^N B_\varepsilon(\xi_i)$. Com isto K é totalmente limitado.

Provaremos ora que K é completo. Seja então $(x_n)_{n \in \mathbb{N}}$ uma seqüência de Cauchy em K . Suponhamos que esta seqüência não convirja para um elemento de K . Então, segundo um resultado do curso Análise I, nenhuma subsequência de $(x_n)_{n \in \mathbb{N}}$ converge para um elemento de K . Com isto a seqüência $(x_n)_{n \in \mathbb{N}}$ não possui nenhumas pontos limites em K (veja Análise I). Isso significa o seguinte: para todo $\xi \in K$ existe um $\varepsilon(\xi) > 0$, tal que a bola $B_{\varepsilon(\xi)}(\xi)$ contém só um número finito de membros da nossa seqüência $(x_n)_{n \in \mathbb{N}}$. Segue daí que a cobertura aberta

$$\mathcal{U} := \{B_{\varepsilon(\xi)}(\xi) \mid \xi \in K\}$$

de K não pode conter nenhuma subcobertura finita. Isto é uma contradição a (ii), e com isso é mostrado, que (iii) segue de (ii).

“(iii) \implies (i)”. Suponhamos ora que K seja completo e totalmente limitado. Seja $(x_n)_{n \in \mathbb{N}}$ uma seqüência em K . Temos que encontrar uma subsequência convergente de (x_n) a qual converge para um elemento de K . Para isso vamos construir indutivamente uma seqüência de subconjuntos infinitos $T_k \subset \mathbb{N}$, $k = 0, 1, 2, 3, \dots$, com $\mathbb{N} \supset T_0 \supset T_1 \supset T_2 \supset \dots$ e

$$n, m \in T_k \quad \implies \quad d(x_n, x_m) \leq 2^{-k}.$$

Como (X, d) é totalmente limitado, pode-se cobrir K com um número finito de bolas $B_{1/2}(\xi_i)$ com $i = 1, \dots, m$. Uma destas bolas, dizemos $B_{1/2}(\xi_i)$, dever conter um número infinito de membros da nossa seqüência. Defina

$$T_0 := \{n \in \mathbb{N} \mid x_n \in B_{1/2}(\xi_i)\}.$$

Pela construção isto é um subconjunto infinito de \mathbb{N} e vale

$$d(x_n, x_m) \leq d(x_n, \xi_i) + d(\xi_i, x_m) < 1 \quad \text{para todos os } n, m \in T_0.$$

Seja ora $k \geq 1$ e sejam $T_0 \supset T_1 \supset \dots \supset T_{k-1}$ já construído. Como (X, d) é totalmente limitado, podemos cobrir X com um número finito de bolas $B_{2^{-k-1}}(\xi_i)$ com $i = 1, \dots, m$. Uma destas bolas, dizemos $B_{2^{-k-1}}(\xi_i)$, deve conter um número infinito de membros x_n com $n \in T_{k-1}$. Defina o conjunto $T_k := \{n \in T_{k-1} \mid x_n \in B_{2^{-k-1}}(\xi_i)\}$. Então T_k é um subconjunto infinito de T_{k-1} e vale $d(x_n, x_m) \leq d(x_n, \xi_i) + d(\xi_i, x_m) < 2^{-k}$ para todos os $n, m \in T_k$. Com isso todos os conjuntos T_k para todos os $k \in \mathbb{N}$ são construídos.

Como cada um subconjunto T_k é infinito, existe uma sequência $n_k \in T_k$, tal que $n_k < n_{k+1}$ para todos os $k \in \mathbb{N}$. Então vale $n_\ell \in T_\ell \subset T_k$ para $\ell \geq k$ e daí todos os $k, \ell \in \mathbb{N}$ com $\ell \geq k$ satisfazem a desigualdade $d(x_k, x_\ell) \leq 2^{-k}$. Com isso a sequência $(x_{n_k})_{k \in \mathbb{N}}$ é Cauchy e, como K é completo, ela converge para um elemento de K . Então temos mostrado que toda sequência em X possui uma subsequência convergente. Portanto (i) segue de (iii). Com isso o Teorema C.1.2 fica provado. \square

Exercício C.1.3. Seja (X, d) um espaço métrico e $Y \subset X$ um subconjunto munido da função distância induzida d_Y . Um subconjunto $K \subset Y$ é compacto em (X, d) se e somente se é compacto em (Y, d_Y) . **Dica:** “ \Leftarrow ” Um subconjunto $V_i \subset Y$ é Y -aberto se existe um X -aberto $U_i \subset X$ com $U_i \cap Y = V_i$.

C.2 O Teorema de Arzelà–Ascoli

Segundo o Teorema de Heine–Borel um subconjunto do \mathbb{R}^n (munido da métrica padrão, ou seja, a euclidiana) é compacto se e somente se é fechado e limitado. Particularmente a bola unitária fechada em \mathbb{R}^n é compacta. Isto também vale para qualquer espaço vetorial normado de dimensão finita, ou equivalentemente para qualquer norma em \mathbb{R}^n . Em contrapartida o seguinte exercício mostra que esta propriedade precisamente caracteriza os espaços vetoriais normados de dimensão finita.

Exercício C.2.1. Seja V espaço vetorial normado com bola unitária compacta $B := \{x \in V \mid \|x\| \leq 1\}$. Então V é de dimensão finita.

Dicas: 1. Todo subespaço linear $W \subset V$ de dimensão finita é fechado.

2. Se $W \subset V$ é um subespaço linear fechado, então todo vetor no complemento $v \in V \setminus W$ fica na distância positiva $d(v, W) := \inf_{w \in W} \|v - w\| > 0$.

- 3.** Se $W \subsetneq V$ é um subespaço linear fechado, então existe um vetor $v \in V$ com $\|v\| = 1$ e $d(v, W) \geq 1/2$. (Seja $v_0 \in V \setminus W$ e escolha $w_0 \in W$ tal que $\|v_0 - w_0\| \leq 2d(v_0, W)$; defina $v := \|v_0 - w_0\|^{-1}(v_0 - w_0)$).
- 4.** Se V é de dimensão infinita, então existe uma sequência $(v_n)_{n \in \mathbb{N}}$ em V com $\|v_n\| = 1$ e $\|v_n - v_m\| \geq 1/2$ para todos os $n, m \in \mathbb{N}$ com $n \neq m$. Uma tal sequência não possui uma subsequência convergente.

Exercício C.2.1 nós dá um critério importante para a dimensionalidade finita de um espaço vetorial normado. Também mostra que o princípio de Heine-Borel não vale em espaços vetoriais normados de dimensão infinita: La existem subconjuntos fechados e limitados as quais não são compactos. Com isso leva-se a questão as quais subconjuntos de um tal espaço de dimensão infinita são compacto então. Uma resposta a esta questão é da significância fundamental em muitas aplicações. Um critério especialmente importante para compacidade de subconjuntos do espaço das funções contínuas nós dá o Teorema de Arzelà-Ascoli.

Seja (X, d) um espaço métrico compacto e $\mathcal{C}(X)$ o espaço das funções contínuas $f : X \rightarrow \mathbb{R}$. Este espaço munido da norma do supremo

$$\|f\| := \sup_{x \in X} |f(x)|$$

é um espaço de Banach, ou seja, um espaço vetorial normado o qual é completo (veja [5, Satz 7.1]). Também vale que toda função contínua $f : X \rightarrow \mathbb{R}$ num espaço métrico compacto (X, d) é **uniformemente contínua**, ou seja, para todo $\varepsilon > 0$ existe um $\delta > 0$, tal que todos os $x, y \in X$ com $d(x, y) < \delta$ satisfazem a desigualdade $|f(x) - f(y)| < \varepsilon$. Vamos chamar uma família $\mathcal{F} \subset \mathcal{C}(X)$ de **equicontínua** se o número $\delta > 0$ nesta definição pode ser escolhido independente da função $f \in \mathcal{F}$.

Definição C.2.2. *Uma família $\mathcal{F} \subset \mathcal{C}(X)$ chama-se de equicontínua, se para todo $\varepsilon > 0$ existe um $\delta > 0$, tal que para todo $f \in \mathcal{C}(X)$ e todos os $x, y \in X$ vale:*

$$f \in \mathcal{F}, \quad d(x, y) < \delta \quad \implies \quad |f(x) - f(y)| < \varepsilon.$$

Exercício C.2.3. Todo subconjunto finito de $\mathcal{C}(X)$ é equicontínua. Sejam c e μ dois números reais positivos. Então o subconjunto

$$\mathcal{F} := \left\{ f \in \mathcal{C}(X) \mid \|f\| \leq c, \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)^\mu} \leq c \right\}$$

é fechado, limitado, e equicontínuo. Encontre uma sequência em $\mathcal{C}([0, 1])$ a qual é limitada, mas não equicontínua.

Teorema C.2.4 (Arzelà–Ascoli). *Seja (X, d) um espaço métrico compacto. Um subconjunto $\mathcal{F} \subset \mathcal{C}(X)$ do espaço das funções reais contínuas em X é compacto com respeito da norma do supremo se e somente se tem as seguintes propriedades:*

- (i) \mathcal{F} é fechado.
- (ii) \mathcal{F} é limitado.
- (iii) \mathcal{F} é equicontínua.

Demonstração. Suponhamos para já que \mathcal{F} seja um subconjunto compacto de $\mathcal{C}(X)$. Então \mathcal{F} é fechado segundo um resultado do curso Análise I [6]. Demais \mathcal{F} é limitado porque uma sequência $f_n \in \mathcal{F}$ com $\|f_n\| \rightarrow \infty$ não possui uma subsequência convergente. Então resta mostrar que o conjunto \mathcal{F} é equicontínuo. Seja $\varepsilon > 0$. Como, segundo o Teorema C.1.2, o conjunto $\mathcal{F} \subset \mathcal{C}(X)$ é totalmente limitado, existem um número finito de funções $f_1, \dots, f_m \in \mathcal{F}$ com

$$\mathcal{F} \subset \bigcup_{i=1}^m B_{\varepsilon/3}(f_i; \mathcal{C}(X)).$$

Como (X, d) é compacto, cada uma das funções $f_i : X \rightarrow \mathbb{R}$ é uniformemente contínua segundo um teorema da Análise I [6]. Consequentemente existe para todo $i \in \{1, \dots, m\}$ uma constante $\delta_i > 0$, tal que para todos os $x, y \in X$ vale:

$$d(x, y) < \delta_i \quad \implies \quad |f_i(x) - f_i(y)| < \varepsilon/3.$$

Escolha

$$\delta := \min\{\delta_1, \dots, \delta_m\} > 0.$$

Seja $f \in \mathcal{F}$. Então existe um índice $i \in \{1, \dots, m\}$ com $\|f - f_i\| < \varepsilon/3$. Dali satisfazem todos os $x, y \in X$ com $d(x, y) < \delta \leq \delta_i$ a desigualdade

$$|f(x) - f(y)| \leq |f(x) - f_i(x)| + |(f_i(x) - f_i(y))| + |f_i(y) - f(y)| < \varepsilon.$$

Portanto o conjunto \mathcal{F} é uniformemente contínuo.

Vice-versa, suponhamos ora que o conjunto $\mathcal{F} \subset \mathcal{C}(X)$ seja fechado, limitado e equicontínuo. Vamos provar em quatro passos que o conjunto \mathcal{F} é compacto.

Passo 1. *Existe uma seqüência $(x_k)_{k \in \mathbb{N}}$ em X com a seguinte propriedade. Para todo $\delta > 0$ existe um $m = m(\delta) \in \mathbb{N}$, tal que*

$$X = \bigcup_{k=1}^{m(\delta)} B_\delta(x_k).$$

Construímos a seqüência indutivamente. Primeiramente sabemos, segundo o Teorema C.1.2, que o espaço métrico (X, d) é totalmente limitado. Portanto existem, para $\delta = 1$, um número finito de pontos $x_1, \dots, x_{m_1} \in X$, tal que

$$X = \bigcup_{k=1}^{m_1} B_1(x_k).$$

Do mesmo modo existem, para $\delta = 1/2$, um número finito de pontos de X , dizemos $x_{m_1+1}, \dots, x_{m_2}$, tal que

$$X = \bigcup_{k=m_1+1}^{m_2} B_{1/2}(x_k) = \bigcup_{k=1}^{m_2} B_{1/2}(x_k).$$

Depois temos construído os $x_1, \dots, x_{m_{n-1}}$, escolhemos $x_{m_{n-1}+1}, \dots, x_{m_n} \in X$, tal que

$$X = \bigcup_{k=m_{n-1}+1}^{m_n} B_{1/n}(x_k) = \bigcup_{k=1}^{m_n} B_{1/n}(x_k).$$

Com isso vale a afirmação do Passo 1 com $m(\delta) = m_n$, onde n é escolhido, tal que $1/n < \delta$.

Passo 2. *Seja $(f_n)_{n \in \mathbb{N}}$ uma seqüência em \mathcal{F} . Então existe uma subseqüência $(f_{n_i})_{i \in \mathbb{N}}$, tal que existe o limite*

$$y_k := \lim_{i \rightarrow \infty} f_{n_i}(x_k)$$

para todo $k \in \mathbb{N}$.

Isto é um argumento típico de seqüências diagonais. Seja para já $k = 1$, então a seqüência de números reais $(f_n(x_1))_{n \in \mathbb{N}}$ é limitada e daí possui, segundo o

Teorema de Bolzano–Weierstraß, uma subsequência convergente. Em outras palavras, existe uma função monótona $g_1 : \mathbb{N} \rightarrow \mathbb{N}$ estritamente crescente tal que o limite $y_1 := \lim_{i \rightarrow \infty} f_{g_1(i)}(x_1)$ existe. Ora a sequência $(f_{g_1(i)}(x_2))_{n \in \mathbb{N}}$ também é limitada e por isso também possui uma subsequência convergente. Portanto existe uma função monótona $g_2 : \mathbb{N} \rightarrow \mathbb{N}$ estritamente crescente, tal que o limite $y_2 := \lim_{i \rightarrow \infty} f_{g_1 \circ g_2(i)}(x_2)$ existe. Ora encontramos através de indução uma sequência de funções monótonas $g_k : \mathbb{N} \rightarrow \mathbb{N}$ estritamente crescentes tal que o limite

$$y_k := \lim_{i \rightarrow \infty} f_{g_1 \circ \dots \circ g_k(i)}(x_k)$$

existe para todo $k \in \mathbb{N}$. Seja ora

$$n_i := g_1 \circ \dots \circ g_i(i)$$

para $i \in \mathbb{N}$. Então $(f_{n_i(i)}(x_k))_{i \geq k}$ é uma subsequência de $(f_{g_1 \circ \dots \circ g_k(i)}(x_k))_{k \in \mathbb{N}}$ e converge assim para y_k . Porque isto vale para todo $k \in \mathbb{N}$, o Passo 2 fica provado.

Passo 3. $(f_{n_i})_{i \in \mathbb{N}}$ é uma sequência de Cauchy em $\mathcal{C}(X)$.

Seja $\varepsilon > 0$. Como \mathcal{F} é equicontínuo, existe um $\delta > 0$ tal que para todo $f \in \mathcal{C}(X)$ e todos os $x, y \in X$ vale o seguinte:

$$f \in \mathcal{F}, \quad d_X(x, y) < \delta \quad \implies \quad |f(x) - f(y)| < \varepsilon/3. \quad (\text{C.2.1})$$

Segundo o Passo 1 existe um número natural $m = m(\delta) \in \mathbb{N}$, tal que

$$X = \bigcup_{k=1}^m B_\delta(x_k). \quad (\text{C.2.2})$$

Segundo o Passo 2 existe um $N \in \mathbb{N}$, tal que para todos os $i, j, k \in \mathbb{N}$ vale o seguinte:

$$k \leq m, \quad i, j \geq N \quad \implies \quad |f_{n_i}(x_k) - f_{n_j}(x_k)| < \varepsilon/3. \quad (\text{C.2.3})$$

Afirmamos que com esta escolha de N vale a desigualdade $\|f_{n_i} - f_{n_j}\| \leq \varepsilon$ para todos os $i, j \geq N$. Seja então $x \in X$ e escolha um número $k \in \{1, \dots, m\}$ tal que $d(x, x_k) < \delta$ (veja (C.2.2)). Então segue de (C.2.1) a desigualdade

$$|f_{n_i}(x_k) - f_{n_i}(x)| < \varepsilon/3 \quad (\text{C.2.4})$$

para todos os $i \in \mathbb{N}$. De (C.2.3) e (C.2.4) segue a desigualdade

$$\begin{aligned} |f_{n_i}(x) - f_{n_j}(x)| &\leq |f_{n_i}(x) - f_{n_i}(x_k)| + |f_{n_i}(x_k) - f_{n_j}(x_k)| \\ &\quad + |f_{n_j}(x_k) - f_{n_j}(x)| \\ &< \varepsilon \end{aligned}$$

para todos os $i, j \geq N$. Com isso temos mostrado que a subsequência $(f_{n_i})_{i \in \mathbb{N}}$ construída no Passo 2 é uma sequência de Cauchy em $\mathcal{C}(X)$.

Passo 4. A sequência $(f_{n_i})_{i \in \mathbb{N}}$ converge para uma função $f \in \mathcal{F}$.

Como $\mathcal{C}(X)$ é um espaço de Banach, a sequência $(f_{n_i})_{i \in \mathbb{N}}$ no Passo 3 converge para um $f \in \mathcal{C}(X)$. Como \mathcal{F} é fechado, o limite pertence a \mathcal{F} . Assim temos mostrado que toda sequência em \mathcal{F} possui uma subsequência convergente com limite em \mathcal{F} . Portanto o conjunto \mathcal{F} é compacto e com isso Teorema C.2.4 fica provado. \square

O Teorema da Arzelà-Ascoli e sua prova podem ser transferido palavra por palavra ao espaço $\mathcal{C}(X, V)$ das funções contínuas em X com valores num espaço de Banach V de dimensão finita. No caso de dimensão infinita precisa-se a condição adicional que o conjunto

$$\mathcal{F}(x) := \{f(x) \mid f \in \mathcal{F}\}$$

a chamada assim **\mathcal{F} -órbita do ponto x** , é um subconjunto compacto de V para todos os $x \in X$ (veja [1, Cor 1.1.12]).

Exercício C.2.5. Encontre um subconjunto do espaço $\mathcal{BC}(\mathbb{R})$ (das funções contínuas limitadas $f : \mathbb{R} \rightarrow \mathbb{R}$) o qual com efeito é fechado, limitado, e equicontínuo, mas não compacto.

Apêndice D

O Determinante

D.1 Grupos

Definição D.1.1. Um **grupo** é um triplo $(G, \cdot, \mathbb{1})$, composto de um conjunto G , uma aplicação

$$G \times G \rightarrow G : (g, h) \mapsto g \cdot h$$

(a **operação do grupo**), e um elemento $\mathbb{1} \in G$ (o **elemento neutro**), as quais satisfazem as seguintes condições.

- (i) A operação do grupo é associativo, ou seja, vale $g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3$ para todos os $g_1, g_2, g_3 \in G$.
- (ii) Vale $\mathbb{1} \cdot g = g \cdot \mathbb{1} = g$ para todo $g \in G$.
- (iii) Para todo $g \in G$ existe um elemento $h \in G$, tal que $g \cdot h = h \cdot g = \mathbb{1}$.

O elemento h em (iii) é determinado unicamente por g . É chamado do **inverso de g** e denotado g^{-1} . Um grupo $(G, \cdot, \mathbb{1})$ é chamado de **comutativo** (ou **abeliano**) se

$$g \cdot h = h \cdot g \quad \forall g, h \in G.$$

Exemplo D.1.2. Os números inteiros formam um grupo \mathbb{Z} com a adição como operação do grupo e o número 0 como o elemento neutro. Igualmente para os números racionais \mathbb{Q} , reais \mathbb{R} , e complexos \mathbb{C} .

Exemplo D.1.3. Os números racionais não-nulos formam um grupo com a multiplicação como operação do grupo e o número 1 como elemento neutro. Igualmente para os números reais e complexos.

Exemplo D.1.4. O conjunto de dois elementos $\mathbb{Z}_2 := \{0, 1\}$ é um grupo com a operação do grupo $0 + 0 := 0$, $0 + 1 := 1$, $1 + 0 := 1$ e $1 + 1 := 0$. O zero 0 é o elemento neutro.

Exemplo D.1.5. O conjunto de dois elementos $\{\pm 1\}$ é um grupo com a operação do grupo $1 \cdot 1 := 1$, $(-1) \cdot (-1) := 1$, $1 \cdot (-1) := -1$ e $(-1) \cdot 1 := -1$. O número 1 é o elemento neutro.

Exemplo D.1.6. O conjunto $\mathbb{S}^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ de todos os números complexos de módulo 1 é um grupo com a multiplicação como operação do grupo e o número 1 como elemento neutro.

Exemplo D.1.7. Seja X um conjunto. As aplicações bijetivas $f : X \rightarrow X$ formam um grupo com a composição como operação do grupo e a aplicação identidade como elemento neutro.

Exemplo D.1.8. Seja $X := \{1, \dots, n\}$. Uma **permutação (de X)** é uma aplicação bijetiva $\sigma : X \rightarrow X$. O conjunto de todas as permutações de X denotamos de S_n o qual é um grupo segundo o Exemplo D.1.6. Para $n \geq 3$ este grupo não é comutativo.

Definição D.1.9. *Sejam G e H dois grupos. (As duas operações de grupo denotamos de \cdot e os dois elementos neutrais de $\mathbb{1}_G$ e $\mathbb{1}_H$.) Uma aplicação*

$$\rho : G \rightarrow H$$

*chama-se de **homomorfismo de grupos** (ou **homomorfismo**) se vale para todos os $a, b \in G$ a equação*

$$\rho(a \cdot b) = \rho(a) \cdot \rho(b), \quad \rho(\mathbb{1}_G) = \mathbb{1}_H.$$

*Um **isomorfismo de grupos** (ou **isomorfismo**) é um homomorfismo de grupos o qual é bijetivo.*

Exemplo D.1.10. A aplicação

$$\{0, 1\} \rightarrow \{\pm 1\} : a \mapsto (-1)^a$$

é um isomorfismo.

Exemplo D.1.11. A aplicação

$$\mathbb{R} \rightarrow \mathbb{S}^1 : t \mapsto \exp(it)$$

é um homomorfismo de grupos.

Exemplo D.1.12. A função exponencial

$$\exp : \mathbb{R} \rightarrow (0, \infty)$$

é um isomorfismo do grupo aditivo dos números reais para o grupo multiplicativo dos números reais positivos. A aplicação inversa

$$\log : (0, \infty) \rightarrow \mathbb{R}$$

também é um isomorfismo de grupos.

Exemplo D.1.13. Seja $(G, \cdot, \mathbb{1})$ um grupo, X um conjunto, e $\mathcal{G}(X)$ o grupo das aplicações bijetivas $f : X \rightarrow X$. Uma **ação de grupo de G em X** é uma aplicação

$$\phi : G \times X \rightarrow X$$

tal que para todo $x \in X$ e todos os $a, b \in G$ valem as equações

$$\phi(a \cdot b, x) = \phi(a, \phi(b, x)), \quad \phi(\mathbb{1}, x) = x.$$

Se ϕ é uma tal ação de grupo e definimos as aplicações $\phi_a : X \rightarrow X$ para $a \in G$ através de

$$\phi_a(x) := \phi(a, x) \quad \text{para } x \in X$$

então ϕ_a é bijetiva para todo $a \in G$ (com inversa $\phi_{a^{-1}}$) e a aplicação

$$G \rightarrow \mathcal{G}(X) : a \mapsto \phi_a$$

é um homomorfismo de grupos, porque vale

$$\phi_{a \cdot b} = \phi_a \circ \phi_b, \quad \phi_{\mathbb{1}} = \text{id}$$

para todos os $a, b \in G$. Vice-versa, todo homomorfismo de grupos de G para $\mathcal{G}(X)$ define uma ação de grupo de G em X .

Exemplo D.1.14. Associamos a uma permutação $\sigma \in S_n$ o número natural

$$N(\sigma) := \# \{(i, j) \in \mathbb{N} \times \mathbb{N} \mid 1 \leq i < j \leq n, \sigma(i) > \sigma(j)\}. \quad (\text{D.1.1})$$

A **paridade** de σ é definida através de

$$\varepsilon(\sigma) := (-1)^{N(\sigma)}. \quad (\text{D.1.2})$$

O seguinte lema mostra que a aplicação $\varepsilon : S_n \rightarrow \{\pm 1\}$ é um homomorfismo de grupos.

Lema D.1.15. Vale $\varepsilon(\text{id}) = 1$ e, para todos os $\sigma, \tau \in S_n$,

$$\varepsilon(\tau \circ \sigma) = \varepsilon(\tau)\varepsilon(\sigma).$$

Demonstração. Definimos

$$N(\sigma, \tau) := \# \{(i, j) \mid i < j, \sigma(i) > \sigma(j), \tau(\sigma(i)) < \tau(\sigma(j))\}.$$

Então vale

$$\begin{aligned} N(\tau \circ \sigma) &= \# \{(i, j) \mid i < j, \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &= \# \{(i, j) \mid i < j, \sigma(i) > \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &\quad + \# \{(i, j) \mid i < j, \sigma(i) < \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &= \# \{(i, j) \mid i < j, \sigma(i) > \sigma(j)\} \\ &\quad - \# \{(i, j) \mid i < j, \sigma(i) > \sigma(j), \tau(\sigma(i)) < \tau(\sigma(j))\} \\ &\quad + \# \{(i, j) \mid \sigma(i) < \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &\quad - \# \{(i, j) \mid i > j, \sigma(i) < \sigma(j), \tau(\sigma(i)) > \tau(\sigma(j))\} \\ &= N(\sigma) + N(\tau) - 2N(\sigma, \tau). \end{aligned}$$

Dali segue imediatamente a afirmação. \square

Exercício D.1.16. Para $i, j \in \{1, \dots, n\}$ com $i \neq j$ definimos a permutação $\sigma_{ij} \in S_n$ como a bijeção a qual troca i e j :

$$\sigma_{ij}(\nu) := \begin{cases} \nu, & \text{se } \nu \neq i \text{ e } \nu \neq j, \\ i, & \text{se } \nu = j, \\ j, & \text{se } \nu = i. \end{cases}$$

Uma tal permutação chama-se de **transposição**. Mostre que toda transposição tem paridade -1 e que toda permutação pode ser escrito como composição de transposições.

D.2 Espaços vetoriais

Definição D.2.1. *Um espaço vetorial real é composto de um grupo abeliano $(V, +, 0)$ e uma aplicação*

$$\mathbb{R} \times V \rightarrow V : (\lambda, v) \mapsto \lambda \cdot v$$

(a **multiplicação escalar**) com as seguintes propriedades.

(i) Para todos os $\lambda, \mu \in \mathbb{R}$ e $v \in V$ vale

$$\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v, \quad 1 \cdot v = v.$$

(ii) Para todos os $\lambda, \mu \in \mathbb{R}$ e $v, w \in V$ vale

$$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v, \quad \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w.$$

Os elementos de V chama-se de **vetores** e aqueles de \mathbb{R} **escalares**.

Comentário D.2.2. Para evitar confusões às vezes pode ser útil usar um outro símbolo para o vetor nulo em V , por exemplo 0_V . Com esta notação valem as regras

$$\lambda \cdot 0_V = 0_V, \quad 0 \cdot v = 0_V$$

para todo $\lambda \in \mathbb{R}$ e $v \in V$. (Prove estas equações!) Geralmente entretanto denotamos o vetor nulo em V com o mesmo símbolo 0 como o número real zero. Além disso, no seguinte vamos evitar frequentemente o ponto “.” e vamos escrever λv em vez de $\lambda \cdot v$.

Definição D.2.3. *Sejam V e W dois espaços vetoriais. Uma aplicação $T : V \rightarrow W$ chama-se de **linear** se satisfaz as equações*

$$T(v_1 + v_2) = Tv_1 + Tv_2, \quad T(\lambda \cdot v) = \lambda \cdot Tv$$

para todos os $\lambda \in \mathbb{R}$ e $v, v_1, v_2 \in V$. Uma aplicação linear também é chamada de **homomorfismo de espaços vetoriais**. Um **isomorfismo de espaços vetoriais** (ou simplesmente um **isomorfismo**) é um homomorfismo de espaços vetoriais o qual é bijetivo.

Exercício D.2.4. Se $T : V \rightarrow W$ é uma aplicação linear bijetiva, então a aplicação inversa $T^{-1} : W \rightarrow V$ também é linear.

Definição D.2.5. *Seja V um espaço vetorial real. Uma sequência finita*

$$v_1, \dots, v_n$$

*de vetores em V chama-se de **linearmente independente** se para todos os $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ vale o seguinte:*

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0 \quad \implies \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

*Chamamos a expressão $\lambda_1 v_1 + \dots + \lambda_n v_n$ uma **combinação linear** dos vetores v_1, \dots, v_n . Chamamos os vetores v_1, \dots, v_n um **sistema (finito) de geradores** de V se todo vetor em V pode ser representado como combinação linear destes vetores, ou seja*

$$\forall v \in V \exists \lambda_1, \dots, \lambda_n \in \mathbb{R} : v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

*Uma **base finita** de V é um sistema finito e linearmente independente de geradores v_1, \dots, v_n . Um espaço vetorial V chama-se de **dimensão finita** se possui uma base finita e de **dimensão infinita** se não possui uma base finita.*

Comentário D.2.6. Também pode-se estender o conceito de base a sistemas infinitos $\{v_i\}_{i \in I}$ de vetores de V . Aqui I é um conjunto (possivelmente infinito) e $I \rightarrow V : i \mapsto v_i$ uma aplicação. Um tal sistema chama-se de **linearmente independente** se toda escolha finita de vetores v_{i_1}, \dots, v_{i_N} (com $i_1, \dots, i_N \in I$ dois-a-dois diferente) é linearmente independente. É chamado de **sistema de geradores** se todo vetor $v \in V$ pode ser representado como combinação linear de um número finito de v_i 's. Ora, uma **base** é um sistema de geradores linearmente independente. Com esta noção generalizada pode-se mostrar que todo espaço vetorial possui uma base. (A prova resta no chamado assim *Lema de Zorn* o qual por sua vez pode ser deduzido do *axioma da escolha*. Embora nós vamos-nos interessar aqui só para bases finitas. (Espaços vetoriais de dimensão infinita fazem um papel importante na matemática, mas suas bases não.)

Exemplo D.2.7. O espaço \mathbb{R}^n de todas as n -tuplos $x = (x_1, \dots, x_n)$ de números reais é um espaço vetorial. Os vetores $e_i := (0, \dots, 0, 1, 0, \dots, 0)$ (com 1 no i -ésimo lugar) formam uma base de \mathbb{R}^n , a qual é chamada de **base padrão** ou **base canônica**. Um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ é da forma

$$x = \sum_{i=1}^n x_i e_i.$$

Comentário D.2.8. Seja V um espaço vetorial e $v_1, \dots, v_n \in V$. Então a aplicação $T : \mathbb{R}^n \rightarrow V$ definida através de

$$Tx := \sum_{i=1}^n x_i v_i, \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

é linear. A aplicação T é injetiva se os vetores v_1, \dots, v_n são linearmente independente; ela é surjetiva se e somente se os vetores v_1, \dots, v_n formam um sistema de geradores de V . Portanto T é um isomorfismo de espaços vetoriais se os vetores v_1, \dots, v_n formam uma base de V . Todo isomorfismo de espaços vetoriais $T : \mathbb{R}^n \rightarrow V$ é desta forma com $v_i := T e_i$.

Teorema D.2.9. *Seja V um espaço vetorial real com uma base f_1, \dots, f_n . Sejam v_1, \dots, v_m vetores em V . Então vale o seguinte.*

- (i) *Se v_1, \dots, v_m é uma base de V , então vale $m = n$.*
- (ii) *Se os vetores v_1, \dots, v_m formam um sistema de geradores, então $m \geq n$.*
- (iii) *Se os vetores v_1, \dots, v_m são linearmente independente, então $m \leq n$.*

O número n chama-se da **dimensão** de V e denota-se de $\dim V$.

Demonstração. Segundo o comentário prévio, a base e_1, \dots, e_n de V determina um isomorfismo $T : \mathbb{R}^n \rightarrow V$. Por isso é suficiente provar as afirmações para caso $V = \mathbb{R}^n$ munido da base base padrão e_1, \dots, e_n . Sejam então $v_1, \dots, v_m \in \mathbb{R}^m$.

Passo 1. *Se os vetores v_1, \dots, v_m formam uma base do \mathbb{R}^n , então $m = n$.*

Provaremos a afirmação através de indução. Para $n = 1$ a afirmação segue dos seguintes fatos elementares.

- (a) Todo número real não-nulo forma uma base de \mathbb{R} .
- (b) Um sistema com mais como um elemento de \mathbb{R} é linearmente dependente.

Ora, seja $n \geq 2$. Suponhamos que o resultado seja válido para o \mathbb{R}^{n-1} e que os vetores v_1, \dots, v_m formam uma base do \mathbb{R}^n . Escrevemos $v_\nu =: (v_{\nu 1}, \dots, v_{\nu n})$ para $\nu = 1, \dots, m$. Como o vetor e_n pode ser escrito como combinação linear dos vetores v_1, \dots, v_m , existe pelo menos um $\nu \in \{1, \dots, m\}$, tal que $v_{\nu n} \neq 0$. Podemos supor sem perda de generalidade que isto seja $\nu = m$. Seja

$$w_\nu := v_\nu - \frac{v_{\nu n}}{v_{m n}} v_m, \quad \nu = 1, \dots, m-1.$$

Então os vetores w_1, \dots, w_{m-1}, v_m formam uma base do \mathbb{R}^n e vale $w_{\nu n} = 0$ para $\nu = 1, \dots, m-1$. Segue que os vetores w_1, \dots, w_{m-1} formam (depois deixar fora a última coordenada) uma base do \mathbb{R}^{n-1} . Pela hipótese da indução vale então $m-1 = n-1$ e daí $m = n$.

Passo 2. *Se os vetores v_1, \dots, v_m são linearmente independente vale $m \leq n$.*

Podemos formar dos vetores v_1, \dots, v_m através de adicionar de no máximo n outros vetores e_{i_1}, \dots, e_{i_k} (as quais escolhemos dos elementos da base canônica) uma base $v_1, \dots, v_m, e_{i_1}, \dots, e_{i_k}$. (Isto segue através de indução: Se os vetores v_1, \dots, v_m não formam uma base, então existe um índice $i_1 \in \{1, \dots, n\}$ tal que os vetores v_1, \dots, v_m, e_{i_1} são linearmente independente. Se estes vetores ainda não formam uma base, então podemos adicionar mais elementos da base canônica sem cercear a independência linear.) A base nova é composta de $m+k$ elementos com $k \geq 0$. Vale daí $m+k = n$, segundo o Passo 1, e por isso $m \leq n$.

Passo 3. *Se os v_1, \dots, v_m formam um sistema de geradores, então $m \geq n$.*

Recebemos uma base do \mathbb{R}^n através de deixar fora vetores adequados. Então a base resultante contem no máximo m elementos. Portanto a afirmação do Passo 3 segue do Passo 1. Com isso o Teorema D.2.9 fica provado. \square

D.3 O Determinante

Seja V um espaço vetorial real. Uma aplicação

$$\phi : V^n = V \times V \times \dots \times V \rightarrow \mathbb{R}$$

chama-se de **multilinear** se ϕ é linear em cada uma variável, ou seja, valem as equações

$$\begin{aligned} & \phi(v_1, \dots, v_{i-1}, v_i + w_i, v_{i+1}, \dots, v_n) \\ &= \phi(v_1, \dots, v_n) + \phi(v_1, \dots, v_{i-1}, w_i, v_{i+1}, \dots, v_n) \end{aligned}$$

e

$$\phi(v_1, \dots, v_{i-1}, \lambda v_i, v_{i+1}, \dots, v_n) = \lambda \phi(v_1, \dots, v_n)$$

para todos os $i \in \{1, \dots, n\}$, $v_1, \dots, v_n \in V$, $w_i \in V$, e $\lambda \in \mathbb{R}$. No caso de $n = 2$ chamamos uma tal aplicação de **bilinear**. Geralmente, chamamos uma aplicação multilinear $\phi : V^n \rightarrow \mathbb{R}$ também de **n -linear**.

Comentário D.3.1. O espaço V^n mesmo é um espaço vetorial. Então faz sentido falar de aplicações lineares $\phi : V^n \rightarrow \mathbb{R}$. Mas isto é completamente diferente de uma aplicação multilinear. Consideramos por exemplo a caso $V = \mathbb{R}$ e $n = 2$. Então a aplicação

$$\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} : (x, y) \mapsto xy$$

é bilinear, mas não é linear. Pelo contrário a aplicação

$$\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} : (x, y) \mapsto x + y$$

é linear, mas não é bilinear. (Se $\phi : V^n \rightarrow \mathbb{R}$ é uma aplicação multilinear, então vale $\phi(v_1, \dots, v_n) = 0$ logo que um dos vetores v_1, \dots, v_n seja nulo.)

Exercício D.3.2. Seja V um espaço vetorial real e seja $\phi : V^n \rightarrow \mathbb{R}$ tanto linear, como multilinear. Se $n \geq 2$, então $\phi \equiv 0$.

Teorema D.3.3. Para todo $n \in \mathbb{N}$ existe exatamente uma aplicação n -linear

$$\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}$$

chamada do **determinante**, com as seguintes propriedades.

(i) Para toda permutação $\sigma \in S_n$ e quaisquer n vetores $a_1, \dots, a_n \in \mathbb{R}^n$ vale

$$\det(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \varepsilon(\sigma) \det(a_1, \dots, a_n).$$

(ii) Se e_1, \dots, e_n é a base canónica do \mathbb{R}^n , então vale

$$\det(e_1, \dots, e_n) = 1.$$

Se escrevemos

$$a_i = (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n, \quad i = 1, \dots, n$$

com $a_{ij} \in \mathbb{R}$, então o determinante dos vetores a_1, \dots, a_n é dado pela fórmula

$$\det(a_1, \dots, a_n) := \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{i\sigma(i)}. \quad (\text{D.3.1})$$

Aqui $\varepsilon(\sigma) \in \{\pm 1\}$ denota a paridade da permutação $\sigma \in S_n$ a qual é definida através de (D.1.1) e (D.1.2).

Demonstração. Suponhamos para já que a aplicação $\det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ seja definido através de (D.3.1) e provaremos que ela satisfaz as condições (i) e (ii). Se $a_i = e_i$, então vale

$$a_{ij} = \delta_{ij} = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases}$$

Com isto o único somando não-nulo no lado direito de (D.3.1) é aquele que corresponde a $\sigma = \text{id}$ e por isso vale (ii). Além disso, vale

$$\begin{aligned} \det(a_{\tau(1)}, \dots, a_{\tau(n)}) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{\tau(i)\sigma(i)} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n a_{j\sigma(\tau^{-1}(j))} \\ &= \varepsilon(\tau) \sum_{\sigma \in S_n} \varepsilon(\sigma \circ \tau^{-1}) \prod_{j=1}^n a_{j\sigma\tau^{-1}(j)} \\ &= \varepsilon(\tau) \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{j=1}^n a_{j\sigma(j)} \\ &= \varepsilon(\tau) \det(a_1, \dots, a_n) \end{aligned}$$

para $\tau \in S_n$ e $a_i = (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n$. Portanto \det também satisfaz a condição (i).

Provaremos a unicidade. Seja então $\det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ dado através de (D.3.1) e seja $\phi : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$ uma aplicação n -linear qualquer, a qual satisfaz as condições (i) e (ii). Então a diferença

$$\psi := \phi - \det : (\mathbb{R}^n)^n \rightarrow \mathbb{R}$$

é uma aplicação n -linear, a qual satisfaz as condições (i) e $\psi(e_1, \dots, e_n) = 0$. Sejam dados $a_i = (a_{i1}, \dots, a_{in}) = \sum_{j=1}^n a_{ij} e_j \in \mathbb{R}^n$ para $i = 1, \dots, n$. Então

$$\begin{aligned} \psi(a_1, \dots, a_n) &= \psi \left(\sum_{j_1=1}^n a_{1j_1} e_{j_1}, \sum_{j_2=1}^n a_{2j_2} e_{j_2}, \dots, \sum_{j_n=1}^n a_{nj_n} e_{j_n} \right) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} a_{2j_2} \cdots a_{nj_n} \psi(e_{j_1}, \dots, e_{j_n}) = 0. \end{aligned}$$

Aqui usamos a equação $\psi(e_{j_1}, \dots, e_{j_n}) = 0$ para todos os j_1, \dots, j_n . Portanto $\psi \equiv 0$ e com isto $\phi \equiv \det$. Com isso o Teorema D.3.3 fica provado. \square

Uma aplicação n -linear $\phi : V \times \cdots \times V \rightarrow \mathbb{R}$ é chamada de **n -forma alternada** (em V) se ela satisfaz a condição (i) no Teorema D.3.3. Então no \mathbb{R}^n existe exatamente uma n -forma alternada

$$\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R}$$

a qual satisfaz a condição $\det(e_1, \dots, e_n) = 1$, e esta chama-se do **determinante**. Consideramos os n vetores

$$a_i =: (a_{i1}, \dots, a_{in}) \in \mathbb{R}^n$$

para $i = 1, \dots, n$ como linhas de uma matriz

$$A = (a_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$$

então também podemos ver o determinante como aplicação

$$\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$$

no espaço das matrizes quadradas.

Exemplo D.3.4. Se $n = 1$, então $\mathbb{R}^{1 \times 1} = \mathbb{R}$ e $\det : \mathbb{R}^{1 \times 1} \rightarrow \mathbb{R}$ é a aplicação identidade. Para $n = 2$ temos

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Para $n = 3$ e

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

vale

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned}$$

O seguinte teorema resume umas propriedades importantes do determinante.

Teorema D.3.5. (i) Para toda matriz $A \in \mathbb{R}^{n \times n}$ vale

$$\det(A^T) = \det(A). \quad (\text{D.3.2})$$

(ii) Para todos os $A, B \in \mathbb{R}^{n \times n}$ vale

$$\det(AB) = \det(A) \det(B), \quad \det(\mathbf{1}) = 1. \quad (\text{D.3.3})$$

(iii) Para todos os $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times m}$ vale

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det(A) \det(C). \quad (\text{D.3.4})$$

(iv) Seja $A = (a_{ij})_{i,j=1,\dots,n} \in \mathbb{R}^{n \times n}$ e para $k = 1, \dots, n$ seja

$$A_k := (a_{ij})_{\substack{i \in \{1, \dots, n\} \setminus \{k\} \\ j=2, \dots, n}} \in \mathbb{R}^{(n-1) \times (n-1)}$$

a matriz a qual resulta de A através de deixar de fora a primeira coluna e a k -ésima linha. Então vale

$$\det(A) = \sum_{k=1}^n (-1)^{k-1} a_{k1} \det(A_k). \quad (\text{D.3.5})$$

Demonstração. Provaremos a parte (i). Seja $A = (a_{ij})_{i,j=1}^n$ e seja

$$A^T := (a_{ji})_{i,j=1}^n$$

a matriz transposta. Então vale

$$\begin{aligned} \det(A^T) &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} \\ &= \sum_{\tau \in S_n} \varepsilon(\tau) a_{1\tau(1)} \cdots a_{n\tau(n)} = \det(A). \end{aligned}$$

Aqui a penúltima equação segue da mudança da notação $\tau := \sigma^{-1}$ e o fato que $\varepsilon(\sigma) = \varepsilon(\sigma^{-1})$. Com isso a parte (i) fica provada.

Provaremos a parte (ii). Sejam

$$A = (a_{ij})_{i,j=1}^n, \quad B = (b_{jk})_{j,k=1}^n$$

duas matrizes em $\mathbb{R}^{n \times n}$. As entradas da j -ésima linha da matriz B definem um vetor o qual denotamos

$$b_j := (b_{j1}, \dots, b_{jn}) \in \mathbb{R}^n.$$

Então segundo (D.3.1) vale

$$\det(b_{j_1}, \dots, b_{j_n}) = \sum_{\tau \in S_n} \varepsilon(\tau) b_{j_1 \tau(1)} \cdots b_{j_n \tau(n)}$$

para $j_1, \dots, j_n \in \{1, \dots, n\}$. Segundo a definição do determinante anula-se esta expressão logo que dois dos índices j_i coincidem. (Veja a condição (i) no Teorema D.3.3.) Com isso A e B satisfazem a equação

$$\begin{aligned} \det(AB) &= \sum_{\tau \in S_n} \varepsilon(\tau) \left(\sum_{j_1=1}^n a_{1j_1} b_{j_1 \tau(1)} \right) \cdots \left(\sum_{j_n=1}^n a_{nj_n} b_{j_n \tau(n)} \right) \\ &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n \sum_{\tau \in S_n} \varepsilon(\tau) a_{1j_1} b_{j_1 \tau(1)} \cdots a_{nj_n} b_{j_n \tau(n)} \\ &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} \cdots a_{nj_n} \sum_{\tau \in S_n} \varepsilon(\tau) b_{j_1 \tau(1)} \cdots b_{j_n \tau(n)} \\ &= \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} \cdots a_{nj_n} \det(b_{j_1}, \dots, b_{j_n}) \\ &= \sum_{\sigma \in S_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det(b_{\sigma(1)}, \dots, b_{\sigma(n)}) \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} \det(b_1, \dots, b_n) \\ &= \det(A) \det(B). \end{aligned}$$

Aqui usamos o fato que na quarta linha só aqueles somandos podem ser não-nulo para as quais j_1, \dots, j_n são dois-a-dois diferente. Neste caso os índices definem uma permutação $\sigma \in S_n$ pela fórmula $\sigma(i) := j_i$ para $i = 1, \dots, n$. Com isso a parte (ii) fica provada.

Provaremos a parte (iii). Todo somando em (D.3.1) contem exatamente um fator da cada uma linha e da cada uma coluna de A . Para a matriz

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

isso significa que todo somando o qual contem um fator do bloco B (acima do lado direito) também deve conter um fator do bloco embaixo do lado esquerda e por isso anula-se. Em outras palavras, toda permutação $\rho \in S_{n+m}$ a qual associa com um elemento $i \in \{1, \dots, n\}$ um $\rho(i) \in \{n+1, \dots, n+m\}$, também deve associar com um elemento $j \in \{n+1, \dots, n+m\}$ um $\rho(j) \in \{1, \dots, n\}$. Por isso só precisamos considerar permutações $\rho \in S_{n+m}$ as quais levam os dois subconjuntos $\{1, \dots, n\}$ e $\{n+1, \dots, n+m\}$ a se mesmo. Toda tal permutação tem a forma $\rho = \sigma \# \tau$ para $\sigma \in S_n$ e $\tau \in S_m$, com $\sigma \# \tau(i) := \sigma(i)$ para $1 \leq i \leq n$ e $\sigma \# \tau(n+j) = n + \tau(j)$ para $1 \leq j \leq m$. Esta permutação tem paridade $\varepsilon(\sigma \# \tau) = \varepsilon(\sigma)\varepsilon(\tau)$. Daí segue imediatamente a parte (iii).

Prova da parte (iv). Suponhamos para já que exista um $k \in \{1, \dots, n\}$ tal que $a_{i1} = 0$ para todos os $i \in \{1, \dots, n\} \setminus \{k\}$. Sob esta hipótese definimos as permutações $\sigma_k \in S_n$ através de

$$\sigma_k(i) := \begin{cases} k, & \text{para } i = 1, \\ i - 1, & \text{para } i = 1, \dots, k - 1, \\ i, & \text{para } i = k, \dots, n. \end{cases}$$

Esta permutação tem paridade

$$\varepsilon(\sigma_k) = (-1)^{k-1} \tag{D.3.6}$$

e vale

$$B := (a_{\sigma_k(i)j})_{i,j=1,\dots,n} = \begin{pmatrix} a_{k1} & * \\ 0 & A_k \end{pmatrix}. \tag{D.3.7}$$

Aqui o símbolo $*$ denota as entradas restantes da k -ésima linha da matriz A . Segue de (D.3.6) e (D.3.7) a equação

$$\det(A) = \varepsilon(\sigma_k) \det(B) = (-1)^{k-1} \det(B) = (-1)^{k-1} a_{k1} \det(A_k).$$

Aqui a primeira igualdade segue da parte (i), a segunda igualdade de (D.3.6), a terceira igualdade de (D.3.7) e parte (iii). Ora segue a parte (iv) da multilinearidade do determinante. Com isso o Teorema D.3.5 fica provado. \square

Exercício D.3.6. O determinante é determinado unicamente através das propriedades (ii) e (iii) em Teorema D.3.5.

Teorema D.3.7. Para toda matriz $A \in \mathbb{R}^{n \times n}$ são equivalentes:

- (i) $\det(A) \neq 0$.
- (ii) Para todos os $x \in \mathbb{R}^n$ vale: $Ax = 0 \implies x = 0$.
- (iii) Para todo $y \in \mathbb{R}^n$ existe um $x \in \mathbb{R}^n$, tal que $Ax = y$.

Demonstração. Segundo o Teorema D.2.9, um sistema de n vetores reais $a_1, \dots, a_n \in \mathbb{R}^n$ é linearmente independente se e somente se é um sistema de geradores do \mathbb{R}^n . Aplicado às n colunas da matriz A isso significa que (ii) é equivalente a (iii). Pois fica a mostrar que (i) é equivalente a (ii).

Provaremos que (i) implica (ii). Argumentamos indiretamente e supomos que existe um vetor $x = (x_1, \dots, x_n) \in \mathbb{R}^n \setminus \{0\}$ tal que $Ax = 0$. Denotamos as colunas da matriz A de $a_1, \dots, a_n \in \mathbb{R}^n$, isso significa que

$$x_1 a_1 + \dots + x_n a_n = 0.$$

Como $x \neq 0$, existe um $i \in \{1, \dots, n\}$ com $x_i \neq 0$. Sem perda de generalidade supomos que $x_1 \neq 0$. Segundo o Teorema D.3.5 (i) vale então

$$\begin{aligned} \det(A) &= x_1^{-1} \det(x_1 a_1, a_2, \dots, a_n) \\ &= x_1^{-1} \sum_{i=1}^n \det(x_i a_i, a_2, \dots, a_n) \\ &= x_1^{-1} \det\left(\sum_{i=1}^n x_i a_i, a_2, \dots, a_n\right) \\ &= x_1^{-1} \det(0, a_2, \dots, a_n) \\ &= 0. \end{aligned}$$

Prova que (ii) implica (i). Se (ii) é satisfeito, então as colunas a_1, \dots, a_n de A são linearmente independente e assim formam uma base segundo o Teorema D.2.9. Dali segue que a aplicação $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida através de

$$T_A(x) := Ax = \sum_{i=1}^n x_i a_i \quad \text{para } x = (x_1, \dots, x_n) \in \mathbb{R}^n$$

é bijetiva. Seja $S : \mathbb{R}^n \rightarrow \mathbb{R}^n$ a inversa, ou seja $T_A \circ S = S \circ T_A = \text{id}$. Então S é linear e daí existe uma matriz $B \in \mathbb{R}^{n \times n}$ tal que $S = T_B$. Dali segue

$$AB = BA = \mathbb{1} \tag{D.3.8}$$

e daí, segundo a parte (ii) do Teorema D.3.5, vale $\det(A) \det(B) = 1$. Portanto $\det(A) \neq 0$. Com isso Teorema D.3.7 fica provado. \square

Definição D.3.8. A demonstração do Teorema D.3.7 mostra que para toda matriz A do tipo $(n \times n)$ com $\det(A) \neq 0$ existe uma matriz B do tipo $(n \times n)$ a qual satisfaz a equação (D.3.8). Esta matriz é determinada unicamente por A ; ele chama-se da **inversa de A** , respectivamente **matriz inversa de A** , e denota-se de A^{-1} . Segundo o Teorema D.3.5 vale $\det(A^{-1}) = \det(A)^{-1}$.

Comentário D.3.9. Todas as definições e resultados deste capítulo valem do mesmo jeito e com as mesmas demonstrações para matrizes complexas. É suficiente substituir em todo lugar \mathbb{R} por \mathbb{C} .

Definição D.3.10. Seja $A \in \mathbb{R}^{n \times n}$ e seja $\lambda \in \mathbb{C}$. O número complexo λ chama-se de **autovalor** da matriz real A se existe um vetor complexo $x \in \mathbb{C}^n$ o qual satisfaz as condições

$$Ax = \lambda x, \quad x \neq 0.$$

Segundo o Teorema D.3.7 o número λ é um autovalor de A se e somente se satisfaz a equação $\det(\lambda \mathbb{1} - A) = 0$. O polinômio $p_A : \mathbb{C} \rightarrow \mathbb{C}$ definido através da fórmula

$$p_A(z) := \det(z\mathbb{1} - A)$$

para $z \in \mathbb{C}$ chama-se do **polinômio característico da matriz A** .

Exercício D.3.11. O polinômio característico de uma matriz $A \in \mathbb{R}^{n \times n}$ tem grau n . **Dica:** Equação (D.3.1) para o determinante.

Exercício D.3.12. Se $\lambda \in \mathbb{C}$ é um autovalor de uma matriz $A \in \mathbb{R}^{n \times n}$, então $\bar{\lambda}$ também é um autovalor de A .

Corolário D.3.13. O determinante de uma matriz $A \in \mathbb{R}^{n \times n}$ é o produto dos seus autovalores com multiplicidade, ou seja, se o polinômio característico é escrito na forma

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

com $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$, então vale

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Demonstração. Segundo a definição vale $p_A(0) = (-1)^n \det(A)$ e daí segue imediatamente a afirmação. \square

Exemplo D.3.14. Sejam $a_1, \dots, a_n \in \mathbb{R}$ e seja $A \in \mathbb{R}^{(n+1) \times n}$ a matriz

$$A := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \\ a_1 & \cdots & \cdots & a_n \end{pmatrix}. \quad (\text{D.3.9})$$

Então o determinante da matriz quadrada $A^T A \in \mathbb{R}^{n \times n}$ é dado através de

$$\det(A^T A) = 1 + a_1^2 + a_2^2 + \cdots + a_n^2. \quad (\text{D.3.10})$$

Prova: Sejam $P, Q \in \mathbb{R}^{(n+1) \times (n+1)}$ as matrizes definidas assim

$$P := \begin{pmatrix} 1 & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & a_n \\ -a_1 & \cdots & \cdots & -a_n & 1 \end{pmatrix}$$

$$Q := \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ a_1 & \cdots & \cdots & a_n & 1 \end{pmatrix}.$$

Então $\det(Q) = 1$ e $PQ \in \mathbb{R}^{(n+1) \times (n+1)}$ é a matriz bloco

$$PQ = \begin{pmatrix} A^T A & a \\ 0 & 1 \end{pmatrix}, \quad a := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Segundo a parte (ii) e (iii) do Teorema D.3.5 vale daí

$$\det(A^T A) = \det(PQ) = \det(P) = 1 + a_1^2 + \cdots + a_n^2.$$

Aqui mostra-se a última igualdade através de operações elementares nas linhas, no que adiciona-se para $i = 1, \dots, n$ o produto da i -ésima linha de P com a_i à última linha, e depois aplica-se mais uma vez parte (iii) do Teorema D.3.5.

D.4 O grupo linear geral

O grupo das matrizes $(n \times n)$ reais e invertíveis é chamado de **grupo linear geral** e é denotado de

$$\mathrm{GL}(n, \mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\}.$$

Que isto é um grupo segue de Teorema D.3.5 e Teorema D.3.7. A operação do grupo é multiplicação de matrizes e o elemento neutro é a matriz unidade $\mathbb{1}$. O grupo $\mathrm{GL}(n, \mathbb{R})$ pode ser identificado com o grupo das aplicações lineares bijetivas $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$; tais aplicações também são chamadas de **automorfismos** do \mathbb{R}^n . A correspondência associa com toda matriz $A \in \mathrm{GL}(n, \mathbb{R})$ o automorfismo do \mathbb{R}^n definido através de $T_A(x) := Ax$. Teorema D.3.7 mostra que a aplicação linear $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é bijetiva se e somente se $\det(A) \neq 0$. Segundo o Teorema D.3.5 o conjunto

$$\mathrm{GL}^+(n, \mathbb{R}) := \{A \in \mathrm{GL}(n, \mathbb{R}) \mid \det(A) > 0\}$$

é um subgrupo de $\mathrm{GL}(n, \mathbb{R})$. A aplicação

$$\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$$

é contínua. Portanto $\mathrm{GL}(n, \mathbb{R})$ e $\mathrm{GL}^+(n, \mathbb{R})$ são subconjuntos abertos de $\mathbb{R}^{n \times n}$ (são as pre-imagens dos subconjuntos abertos $\mathbb{R} \setminus \{0\}$ e $(0, \infty)$ de \mathbb{R} sob a aplicação contínua $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$). A restrição da aplicação $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ em $\mathrm{GL}(n, \mathbb{R})$ define um homomorfismo de grupos de $\mathrm{GL}(n, \mathbb{R})$ para os números reais não-nulos (veja o Exemplo D.1.5 na Seção D.1).

Para o seguinte teorema precisamos o conceito de um espaço métrico conexo por caminho, discutido no Apêndice B com mais detalhe. Aqui só seja mencionado que um espaço métrico (X, d) é chamado de **conexo por caminho**, se para todos os $x_0, x_1 \in X$ existe uma aplicação contínua $\gamma : [0, 1] \rightarrow X$ a qual conecta os pontos $\gamma(0) = x_0$ e $\gamma(1) = x_1$ um com o outro.

Teorema D.4.1. *O grupo $\mathrm{GL}^+(n, \mathbb{R})$ é conexo por caminho.*

Demonstração. Veja a página 365. □

A seguinte demonstração resta nas seguintes três lemas. Origina-se de Thomas Honold e é adotado diretamente de [2, página 36].

Lema D.4.2. Para toda matriz $A \in \text{GL}(n, \mathbb{R})$ existe um $\ell \in \{0, 1, \dots, n\}$, uma matriz $U \in \text{GL}^+(n - \ell, \mathbb{R})$ sem autovalores negativos, uma matriz

$$\Lambda = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_\ell \end{pmatrix} \in \text{GL}(\ell, \mathbb{R}), \quad \lambda_i < 0$$

e uma matriz $P \in \text{GL}(n, \mathbb{R})$, tal que a equação

$$P^{-1}AP = \begin{pmatrix} \Lambda & * \\ 0 & U \end{pmatrix}$$

é satisfeito por estas matrizes.

Demonstração. Vamos provar o lema através de indução sobre n . Para $n = 1$ a afirmação obviamente é verdadeiro com $P = 1$ e ou $\ell = 0$, $U = A$ (no caso de $A > 0$) ou $\ell = 1$, $\Lambda = A$ (no caso de $A < 0$).

Ora, seja $n > 1$ e suponhamos que a afirmação seja provada para todas as matrizes em $\text{GL}(n - 1, \mathbb{R})$. Seja dado $A \in \text{GL}(n, \mathbb{R})$. Se A não possua autovalores negativos, então a afirmação vale com $P := \mathbb{1}$, $\ell := 0$, e $U := A$. Se A possua um autovalor negativo $\lambda < 0$, então existe um vetor $v \in \mathbb{R}^n$ com

$$Av = \lambda v, \quad v \neq 0.$$

Seja $Q \in \mathbb{R}^{n \times n}$ uma matriz invertível qualquer cuja primeira coluna iguale v . Então existe uma matriz $A_0 \in \text{GL}(n - 1, \mathbb{R})$, tal que

$$Q^{-1}AQ = \begin{pmatrix} \lambda & * \\ 0 & A_0 \end{pmatrix}.$$

Segundo a hipótese da indução existem matrizes $P_0 \in \text{GL}(n - 1, \mathbb{R})$ e Λ_0 , U_0 como na afirmação, tal que

$$P_0^{-1}A_0P_0 = \begin{pmatrix} \Lambda_0 & * \\ 0 & U_0 \end{pmatrix}.$$

Portanto vale a afirmação para A com

$$P := Q \begin{pmatrix} 1 & 0 \\ 0 & P_0 \end{pmatrix}, \quad \Lambda := \begin{pmatrix} \lambda & * \\ 0 & \Lambda_0 \end{pmatrix}, \quad U := U_0.$$

Com isso o Lema fica provado. \square

Lema D.4.3. *Seja dado $A \in GL^+(n, \mathbb{R})$. Então existem duas matrizes B, C em $GL^+(n, \mathbb{R})$ as quais satisfazem a equação*

$$A = B^2C$$

e não tem autovalores negativos.

Demonstração. Sejam $P \in GL(n, \mathbb{R})$, $\Lambda \in GL(\ell, \mathbb{R})$, e $U \in GL(n - \ell, \mathbb{R})$ como no Lema D.4.2. Segundo o Teorema D.3.5 vale

$$\det(A) = \det(\Lambda) \det(U) = \lambda_1 \cdots \lambda_\ell \det(U).$$

Como U não tem autovalores negativos, vale $\det(U) > 0$. Como $\det(A) > 0$, o número $\ell = 2k$ é par. Definimos

$$S := \begin{pmatrix} H & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & H & 0 \\ 0 & \cdots & \cdots & 0 & \mathbb{1}_{n-2k} \end{pmatrix}, \quad H := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Esta matriz não tem autovalores negativos (os autovalores são $\pm i$ e 1). Além disso, vale

$$S^{-2} = S^2 = \begin{pmatrix} -\mathbb{1}_{2k} & 0 \\ 0 & \mathbb{1}_{n-2k} \end{pmatrix}$$

e com isso

$$T := S^{-2}P^{-1}AP = \begin{pmatrix} -\mathbb{1}_{2k} & 0 \\ 0 & \mathbb{1}_{n-2k} \end{pmatrix} \begin{pmatrix} \Lambda & * \\ 0 & U \end{pmatrix} = \begin{pmatrix} -\Lambda & * \\ 0 & U \end{pmatrix}.$$

Esta matriz não tem autovalores negativos tampouco. Então vale

$$A = P(S^2T)P^{-1} = B^2C$$

onde as matrizes

$$B := PSP^{-1}, \quad C := PTP^{-1}$$

não têm autovalores negativos. Com isso o Lema D.4.3 fica provado. \square

Lema D.4.4. *Seja $A \in \text{GL}^+(n, \mathbb{R})$ uma matriz sem autovalores negativos. Então vale $\det((1-t)\mathbb{1} + tA) > 0$ para todos os $t \in [0, 1]$.*

Demonstração. Definimos a função $f : [0, 1] \rightarrow \mathbb{R}$ através de

$$f(t) := \det((1-t)\mathbb{1} + tA).$$

Esta função é contínua e vale $f(0) > 0$ e $f(1) > 0$. A afirmação do lema diz que $f(t) > 0$ para todos os $t \in [0, 1]$. Se não vale isso, então existe um número $t \in [0, 1]$ com $f(t) \leq 0$. Então, segundo o Teorema do Valor Médio, também existe um número $t_0 \in [0, 1]$ com $f(t_0) = 0$. Este número não pode ser 0 nem 1. Mas neste caso o número

$$\lambda := -\frac{1-t_0}{t_0}$$

é um autovalor negativo de A , em contradição a nossa hipótese. Esta contradição prova Lema D.4.4. \square

Demonstração do Teorema D.4.1. Seja $A \in \text{GL}^+(n, \mathbb{R})$ e escolha $B, C \in \text{GL}^+(n, \mathbb{R})$ como no Lema D.4.3. Para $t \in [0, 1]$ defina

$$A_t := ((1-t)\mathbb{1} + tB)^2((1-t)\mathbb{1} + tC).$$

Então vale $A_0 = \mathbb{1}$ e $A_1 = A$. Como as matrizes B e C não têm autovalores negativos, segue do Lema D.4.4 que

$$\det(A_t) = \det((1-t)\mathbb{1} + tB)^2 \det((1-t)\mathbb{1} + tC) > 0$$

para todos os $t \in [0, 1]$. Portanto a aplicação $[0, 1] \rightarrow \text{GL}^+(n, \mathbb{R}) : t \mapsto A_t$ é um caminho contínuo o qual conecta as matrizes $\mathbb{1}$ e A . Com isso $\text{GL}^+(n, \mathbb{R})$ é conexo por caminho. \square

Exercício D.4.5. O grupo $\text{GL}(n, \mathbb{R})$ é a união disjunta dos subconjuntos conexo por caminho $\text{GL}^+(n, \mathbb{R})$ e

$$\text{GL}^-(n, \mathbb{R}) := \{A \in \text{GL}(n, \mathbb{R}) \mid \det(A) < 0\}.$$

O grupo $\text{GL}(n, \mathbb{R})$ mesmo não é conexo por caminho.

Exercício D.4.6. Toda matriz $A \in \mathbb{R}^{n \times n}$ satisfaz a equação

$$\det(e^A) = e^{\text{tr}(A)} \tag{D.4.1}$$

onde $\text{tr}(A)$ é o **traço** da matriz A , ou seja, a soma das entradas na diagonal.

Apêndice E

O Teorema de Banach do Ponto Fixo

Seja (X, d) um espaço métrico e seja $f : X \rightarrow X$ uma aplicação de X em si mesmo. Chama-se a aplicação f de **contração** se existe uma constante

$$0 \leq \alpha < 1$$

tal que vale a desigualdade

$$d(f(x), f(y)) \leq \alpha d(x, y) \tag{E.0.1}$$

para todos os $x, y \in X$. Um **ponto fixo de f** é um elemento $x \in X$ o qual satisfaz a equação $f(x) = x$.

Teorema E.0.7 (Teorema de Banach do Ponto Fixo). *Seja (X, d) um espaço métrico completo e não vazio e seja $f : X \rightarrow X$ uma contração. Então f possui exatamente um ponto fixo.*

Demonstração. Provaremos unicidade primeiro. Sejam $x, y \in X$ pontos fixos de f . Então vale $f(x) = x$ e $f(y) = y$ e dali

$$d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y)$$

segundo (E.0.1). Daí segue a desigualdade

$$(1 - \alpha)d(x, y) \leq 0.$$

Como $\alpha < 1$, segue daqui $d(x, y) = 0$ e daí $x = y$.

Para a prova da existência escolhemos para já um elemento $x_0 \in X$. Um tal elemento existe, porque X não é vazio. Ora definimos indutivamente uma sequência $(x_n)_{n \in \mathbb{N}}$ in X através de

$$x_1 := f(x_0), \quad x_{n+1} := f(x_n) \quad (\text{E.0.2})$$

para $n \in \mathbb{N}$. Então vale

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq \alpha d(x_{n-1}, x_n)$$

para todos os $n \in \mathbb{N}$ segundo (E.0.1). Daqui recebe-se pela indução a desigualdade

$$d(x_n, x_{n+1}) \leq \alpha^n d(x_0, x_1)$$

para todos os $n \in \mathbb{N}$ e daí, para quaisquer dois números inteiros $0 < n < m$,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &\leq (\alpha^n + \alpha^{n+1} + \cdots + \alpha^{m-1})d(x_0, x_1) \\ &\leq \frac{\alpha^n}{1 - \alpha} d(x_0, x_1). \end{aligned}$$

Isto mostra que a sequência $(x_n)_{n \in \mathbb{N}}$ é Cauchy. Como (X, d) é completo, converge portanto a sequência $(x_n)_{n \in \mathbb{N}}$ para um elemento $x \in X$, e vale

$$\begin{aligned} x &= \lim_{n \rightarrow \infty} x_n \\ &= \lim_{n \rightarrow \infty} x_{n+1} \\ &= \lim_{n \rightarrow \infty} f(x_n) \\ &= f\left(\lim_{n \rightarrow \infty} x_n\right) \\ &= f(x). \end{aligned}$$

Portanto x é um ponto fixo de f e assim o Teorema E.0.7 fica provado. \square

Referências Bibliográficas

- [1] T. Bühler, D.A. Salamon, *Functional Analysis*. ETHZ, 2016.
<https://people.math.ethz.ch/~salamon/PREPRINTS/funcana.pdf>
- [2] K. Königsberger, *Analysis 2*. 5. Auflage, Springer Verlag, 2003.
- [3] John Milnor, *Topology from the Differentiable Viewpoint*. Based on notes by D. Weaver. Revised reprint of the 1965 original. *Princeton Landmarks in Mathematics*, Princeton University Press.
- [4] D.A. Salamon, *Summierbare Familien*, ETH Zürich, 2016.
<http://www.math.ethz.ch/~salamon/PREPRINTS/ana1-sum.pdf>
- [5] D.A. Salamon, *Das Riemannsche Integral*, Analysis I, ETH Zürich, 2016.
<http://www.math.ethz.ch/~salamon/PREPRINTS/ana1-int.pdf>
- [6] D.A. Salamon, *Analysis I*, ETHZ, Herbstsemester 2014. Videolink:
<http://www.video.ethz.ch/lectures/d-math/2014/autumn/401-1261-07L.html>
- [7] D.A. Salamon, *Analysis II*, ETHZ, Frühjahrssemester 2015. Videolink:
<http://www.video.ethz.ch/lectures/d-math/2015/spring/401-1262-07L.html>

Índice Remissivo

- $[x - y, x + y] \subset X$ “intervalo”, 2
- $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, 75
- $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ esfera unitária, 5
- $\partial_M B$ bordo de B com respeito M , 242
- $e_i \in \mathbb{R}^n$ vetores unitários canônicos, 5, 294, 350
- C^ℓ -função em \bar{U} , 273
- $d^X f(x, y) := df_y(x)$, 112
- $d^Y f(x, y) := df_x(y)$, 112
- $f_y(x) := f(x, y) =: f_x(y)$, 112
- $\mathcal{L}(X, Y)$ espaço de operadores lineares contínuos, 8
 - $X^* = \mathcal{L}(X, \mathbb{R})$ espaço dual, 13
- $\overline{S(f, P)}$ soma superior, 164
- $\underline{S(f, P)}$ soma inferior, 164
- ação de grupo, 347
- aberto/a
 - bola $-$, 319
 - conjunto $-$, 319
- absolutamente somável, 252
- álgebra
 - Banach $-$ real, 14
 - real, 14
- aplicação
 - contínua, 320
 - continuamente diferenciável, 49
 - ℓ -vezes, 74
 - duas vezes, 74
 - diferenciável, 39
 - exponencial, 23
 - parcialmente diferenciável, 43
 - duas vezes, 69
 - suave, 74, 76, 102
 - aplicação de
 - Gauß, 268
 - transição, 296
 - argumento
 - de sequência diagonal, 342
- atlas, 118, 296
- Banach
 - álgebra de $-$ real, 14
 - espaço de $-$, 2
- base, 350
 - canônica de \mathbb{R}^n , 5, 43, 294, 350
 - padrão de \mathbb{R}^n , 294, 350
- bola aberta, 319, 336
- bordo, 321
 - $M-$ $-$, 303
 - com respeito a M (M -bordo), 242
 - orientação canônica, 305
 - suave, 262, 305
- campo
 - de normais unitários, 268
 - de vetores, 135
 - divergência nula, 229
 - singularidade, 230
 - normal exterior, 305
- Cantor

- conjunto de $-$, 232
- carta, 118
 - domínio de $-$, 238, 295
- Cauchy
 - sequência de $-$, 2
- circulação
 - número de $-$, 275, 311
- compacto
 - conjunto $-$, 336
- completo, 2
- componente conexo, 332
- comprimento, 232
- conexo, 327
 - componente $-$, 332
- conjunto
 - aberto, 319
 - compacto, 336
 - convexo, 96
 - de Cantor, 232
 - fechado, 320
 - Jordan-nulo, 167
 - com respeito a M , 242
- conjunto nulo
 - Jordan, 167
- contínuo, 320
 - equi-, 340
- continuamente diferenciável, 49
 - ℓ -vezes, 74
 - duas vezes, 74
- contração, 367
- convexo/a
 - combinação $-$, 96
 - conjunto $-$, 96
 - função $-$, 96
 - função estritamente $-$, 96
 - subconjunto $-$, 7
- cubo, 197
- derivada, 40
 - direcional, 43
 - normal, 273
 - parcial, 43
 - segunda, 69
 - superior, 74
- derivada parcial
 - da ordem ℓ , 74
 - segunda, 69
- desigualdade
 - de Hölder, 2
 - de Minkowski, 2
- determinante, 353
- difeomorfismo, 102
 - preservando volume, 226
- diferenciável, 39
 - continuamente
 - ℓ -vezes, 74
 - duas vezes, 74
 - continuamente $-$, 49
 - parcial
 - duas vezes, 69
 - parcialmente, 43
- diferencial, 40
 - de uma função, 49
- discreto/a
 - subconjunto, 252
 - subconjunto $-$, 119
 - topologia, 252
- divergência, 229
 - nula, 229
- domínio da carta, 238
- domínio de carta, 118, 295
- EDO equação diferencial ordinária, 25
- endomorfismo, 106
- equação de

- Euler–Lagrange, 117
- Hamilton, 117
- Newton, 61, 117
- equicontínuo, 340
- escalar, 349
- espaço
 - Banach, 2
 - dual, 13
 - Hilbert, 2
 - vetorial
 - normado, 1
- espaço métrico, 319
 - completo, 336
 - compacto, 336
 - conexo, 327
 - conexo por caminho, 330
 - totalmente limitado, 336
- espaço tangente, 127
- espaço vetorial, 349
 - orientado, 294
- Euler–Lagrange
 - equação de –, 117
- existência
 - intervalo máximo, 144
- exponencial
 - aplicação –, 23, 158
 - função –, 347
- exterior
 - campo normal –, 305
 - vetor normal – unitário, 305
- fórmula
 - transformação
 - linear, 208
 - transformação de variáveis, 213
- fechado
 - conjunto –, 320
- fecho, 321
- fluxo de um campo de vetores
 - continuidade, 147
 - diferenciabilidade, 151
- fluxo de um campo vetorial, 144
- forma diferencial, 285
 - exata, 291
 - fechada, 291
- função
 - analítica, 87
 - convexa, 96
 - distância, 319
 - estritamente convexa, 96
 - harmónica, 273
 - própria, 199
- Gauß
 - aplicação de –, 268
- gradiente, 59
- Green
 - primeira identidade de –, 273
 - segunda identidade de –, 273
- Gronwall
 - Lema de –, 147
- grupo, 345
 - abeliano, 345
 - linear, 362
- Hamilton
 - equação de –, 117
- hessiana, 92
- Hilbert
 - espaço –, 2
- homomorfismo, 346, 349
- identidade de
 - Green
 - primeira, 273
 - segunda, 273
- integral

- ao longo de uma variedade, 248
- Riemann, 165
- sobre um conjunto
 - Jordan-mensurável, 187
 - sobre um domínio carta, 240
 - sobre um prédio de blocos, 165
- interior, 321
- intervalo máximo
 - da existência, 144
- inversão na permutação, 279
- isomorfismo, 346, 349

- Jordan-mensurável, 187
 - com respeito a M , 242
- Jordan-nulo
 - conjunto – com respeito a M , 242
 - conjunto –, 167

- Kepler
 - problema de –, 61

- Laplaciano, 273, 292
- Legendre
 - transformação de –, 115
- limitado
 - operador linear –, 8
- logaritmo, 30, 103

- mínimo local, 89
- máximo local, 89
- matriz
 - elementar, 211
 - hessiana, 92
 - jacobiana, 48
- multi-índice, 75
 - ordem, 75
- multiplicador de Lagrange, 130
- número de circulação, 275, 311

- Newton
 - equação de –, 117
 - equações de –, 61
- norma, 1
 - matriz, 10, 16
 - operador, 8

- operador
 - de Laplace, 273, 292
- operador linear
 - limitado, 8

- parametrização
 - de um domínio de carta, 238
- parcialmente diferenciável, 43
 - duas vezes, 69
- partição, 160
 - da unidade, 245
- permutação, 346
- Perron–Frobenius, 318
- polinômio, 78
 - de Taylor, 79
- ponto crítico, 90
 - não-degenerado, 96, 110, 115
- ponto fixo, 367
- prédio de blocos, 161
- preservar
 - volume, 226
- problema
 - de Kepler, 61
 - de valor inicial, 135
- produto interno, 1
 - padrão no \mathbb{R}^n , 42
- pull-back
 - forma alternada, 283
 - forma diferencial, 286

- raio de convergência, 18
- rectificável, 232

- regra
 - da cadeia, 53
 - de Leibniz, 53
- regular
 - valor $-$, 122
- Riemann-integrável, 165
 - impróprio, 220, 251
 - localmente, 220
- série de Taylor, 84
- séries de potências
 - composição, 27
 - de matrizes, 63
 - numa álgebra de Banach, 18
- segmento, 96
- sequência de Cauchy, 2
- singular
 - valor $-$, 122
- singularidade, 230
- soma
 - de Riemann, 173
 - generalizada, 203
 - inferior, 164
 - superior, 164
- suave
 - aplicação $-$, 76
 - bordo $-$, 262
- subvariedade, 118, 295
 - orientada, 298
- superfície
 - de rotação, 120
 - dos romanos, 126
- suporte, 245
 - compacto, 300
- Taylor
 - polinômio de $-$, 79
 - série de $-$, 84
- Teorema da
 - Função Implícita, 113
 - Barreira, 65
 - Função Inversa, 105
 - Inversa Limitada, 105
 - retração, 314
- teorema da
 - extensão limitada, 33
- Teorema de
 - Arzelà–Ascoli, 341
 - Banach do Ponto Fixo, 367
 - Brouwer de Ponto Fixo, 317
 - Fubini, 182, 189
 - para coordenadas polares, 258
 - Gauß, 268
 - Perron–Frobenius, 318
 - Sard, 123, 195, 213
 - Schwarz, 71
 - ser aberto, 110
 - ser difeomorfo, 110
 - Stokes, 308
 - clássico, 312
- teorema de
 - Hahn-Banach, 34
- topologia
 - canônica em dimensão finita, 9
 - relativa, 325
- toro, 92
- traço, 365
- transformação de Legendre, 115
- transição
 - aplicação de $-$, 296
- valor
 - regular, 122
 - singular, 122
- vetor, 349
 - tangente, 127

vizinhança tubular, 134

volume

de uma variedade, 248

da bola unitária B^n , 199

da esfera unitária S^{n-1} , 256, 259

de um bloco, 159

de um cilindro, 190

de um cone, 190

de um domínio carta, 240

de um prédio de blocos, 166

do símlice unitário Δ^n , 186

preservar $-$, 226