SMA0332. NOTAS DE AULA DO CÁLCULO II

2º SEMESTRE DE 2014

1. “Lembranças” do cálculo I e da geometria analítica

Ninguém precisa de definições: a gente as conhece desde ser nascido.
— *Misha Kapranov (com sua namorada na platéia)*, 1990

Para quaisquer $x, y \in \mathbb{R}^n$, vale a *desigualdade triangular* $|x+y| \leq |x| + |y|$. Além disso, $|rx| = |r| \cdot |x|$ para todos $x \in \mathbb{R}^n$ e $r \in \mathbb{R}$.

1.1. Definição. Um subconjunto $U \subset \mathbb{R}^n$ é dito *aberto,* o que denotamos $U \subset_\circ \mathbb{R}^n$, se junto com cada seu ponto $p \in U$ ele contém uma *bola aberta centrada* neste ponto, isto é, existe $r > 0$ tal que $B(p,r) \subset U$, onde $B(p,r) := \big\{x \in \mathbb{R}^n \mid |x-p| < r\big\}$. Um subconjunto $F \subset \mathbb{R}^n$ é dito *fechado,* o que denotamos $F \subset_f \mathbb{R}^n$, se seu complemento $\mathbb{R}^n \setminus F$ é aberto, $(\mathbb{R}^n \setminus F) \subset_\circ \mathbb{R}^n$.

1.2. Observações. 1. *Qualquer bola aberta é aberta.*
2. *Se $U_1, U_2 \subset_\circ \mathbb{R}^n$, então $U_1 \cap U_2 \subset_\circ \mathbb{R}^n$. (A interseção de dois subconjuntos abertos é um subconjunto aberto.)*
3. *Se $U_i \subset_\circ \mathbb{R}^n$ para todo $i \in I$, então $\bigcup\limits_{i \in I} U_i \subset_\circ \mathbb{R}^n$. (A união de qualquer família de subconjuntos abertos é um subconjunto aberto.) Em particular,* $\varnothing$ e $\mathbb{R}^n$ *são abertos.*
4. *Para quaisquer pontos distintos $p_1, p_2 \in \mathbb{R}^n$, $p_1 \neq p_2$, existem subconjuntos abertos $p_i \in U_i \subset_\circ \mathbb{R}^n$, $i = 1, 2$, (chamados* vizinhanças abertas *dos p_i's) que são disjuntos, isto é, $U_1 \cap U_2 = \varnothing$. (Esta propriedade se chama o* axioma de Hausdorff.)

Acabamos de introduzir a topologia padrão *em* $\mathbb{R}^n$.

Demonstração. 1. Seja $q \in B(p,r)$. Então $|q-p| < r$ e $B(q,\varepsilon) \subset B(p,r)$ pela desigualdade triangular, onde $\varepsilon := r - |q-p| > 0$. Com efeito, $x \in B(q,\varepsilon)$ implica $|x-q| < r - |q-p|$. Logo, $|x-p| \leq |x-q| + |q-p| < |q-p| + r - |q-p| = r$.
2. Seja $p \in U_1 \cap U_2$. Então $B(p,r_1) \subset U_1$ e $B(p,r_2) \subset U_2$ para alguns $r_1, r_2 > 0$. Portanto, $B(p,r) \subset U_1 \cap U_2$, onde $r := \min(r_1, r_2) > 0$.
3. É óbvio.
4. Temos $|p_2 - p_1| = 2r > 0$. Basta tomar $U_i := B(p_i, r)$ e aplicar a desigualdade triangular ■

Usando as fórmulas $C \setminus \big(\bigcup\limits_{i\in I} S_i\big) = \bigcap\limits_{i\in I} (C \setminus S_i)$ e $C \setminus \big(\bigcap\limits_{i\in I} S_i\big) = \bigcup\limits_{i\in I} (C \setminus S_i)$, válidas para quaisquer subconjuntos $S_i \subset C$, $i \in I$, de um conjunto C dado, podemos reescrever as propriedades características 2–4 da topologia em termos de subconjuntos fechados.

1.3. Definição. Seja $X \subset \mathbb{R}^n$ um subconjunto.
Para a *topologia induzida* em X, um subconjunto V é *aberto em* X, o que denotamos $V \subset_\circ X$, se $V = X \cap U$ para algum $U \subset_\circ \mathbb{R}^n$. Subconjuntos abertos em X satisfazem as propriedades características semelhantes às 2–4 na Definição 1.1, pois $X \cap \big(\bigcap\limits_{i\in I} U_i\big) = \bigcap\limits_{i\in I} (X \cap U_i)$ e $X \cap \big(\bigcup\limits_{i\in I} U_i\big) = \bigcup\limits_{i\in I} (X \cap U_i)$.

Equivalentemente, um subconjunto G é *fechado em* X, o que denotamos $G \subset_f X$, se $G = X \cap F$ para algum $F \subset_f \mathbb{R}^n$. Como acima, podemos reescrever as propriedades características da topologia em X em termos de subconjuntos fechados.

O sentido intuitivo do conceito de subconjunto aberto é que todos os pontos bastante próximos a um ponto dado deste subconjunto estão no subconjunto. O sentido intuitivo do conceito de subconjunto fechado é dado pelo seguinte lema.

1.4. Lema. *Um subconjunto $F \subset \mathbb{R}^n$ é fechado se e somente se, para toda sequência convergente $x_i \in \mathbb{R}^n$, $i \in \mathbb{N}$, cujos membros estão em F, o limite da sequência também está em F. Em palavras, um subconjunto é fechado se e só se ele é fechado relativamente a passar aos limites.*

Demonstração. Recordamos que uma sequência de pontos $x_i \in \mathbb{R}^n$, $i \in \mathbb{N}$, é *convergente* e tem *limite* $x \in \mathbb{R}^n$, o que denotamos $\lim\limits_{i\to\infty} x_i = x$, se, para todo $\varepsilon > 0$, existe $m \in \mathbb{N}$ tal que $|x_i - x| < \varepsilon$ para todo $i \geq m$. Podemos reformular essa definição: para qualquer vizinhança aberta de x, a sequência fica toda nessa vizinhança a partir de um momento. O axioma de Hausdorff implica que qualquer sequência convergente possui um único limite.

Se F é fechado, $x_i \in F$ para todo $i \in \mathbb{N}$ e $\lim\limits_{i\to\infty} x_i = x$, mas $x \notin F$, então a partir de um momento a sequência fica na vizinhança aberta $\mathbb{R}^n \setminus F$ do ponto x. Uma contradição.

Reciprocamente, se $F \subset \mathbb{R}^n$ é fechado relativamente a fazer limites, então o complemento $\mathbb{R}^n \setminus F$ é aberto. Realmente, se $\mathbb{R}^n \setminus F$ não fosse aberto, poderíamos achar um ponto $x \in \mathbb{R}^n \setminus F$ tal que nenhuma bola aberta centrada em x está contida em $\mathbb{R}^n \setminus F$. Em outras palavras, para qualquer bola aberta $B(x,\varepsilon)$, $\varepsilon > 0$, a interseção $F \cap B(x,\varepsilon)$ não é vazia. Fazendo $\varepsilon := \frac{1}{i}$, $i \in \mathbb{N}$, encontramos um ponto $x_i \in F$ tal que $|x_i - x| < \frac{1}{i}$. É imediato que $\lim\limits_{i\to\infty} x_i = x$. Uma contradição ■

1.5. Definição. Sejam $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$. Uma aplicação $f : X \to Y$ é *contínua* se, para todo subconjunto aberto $V \subset_\circ Y$, a *imagem inversa* $f^{-1}(V) := \big\{x \in X \mid f(x) \in V\big\}$ é aberta em X, $f^{-1}(V) \subset_\circ X$. Uma definição equivalente usa subconjuntos fechados: $G \subset_f Y \Longrightarrow f^{-1}(G) \subset_f X$. Isto segue da propriedade $f^{-1}(Y \setminus S) = X \setminus f^{-1}(S)$, válida para quaisquer subconjunto $S \subset Y$ e aplicação $f : X \to Y$ entre conjuntos.

Note que a aplicação $f : X \to Y$ é contínua se e só se a correspondente aplicação $\hat{f} : X \to \mathbb{R}^n$ é contínua já que a topologia em Y é induzida pela topologia em $\mathbb{R}^n$. Pelo mesmo motivo, a aplicação $f|_{X'} : X' \to Y$ (chamada a *restrição* de f para X') é contínua se a aplicação $f : X \to Y$ é contínua e $X' \subset X$.

Dada uma aplicação $f : X \to Y$ com $X \subset \mathbb{R}^m$ e $Y \subset \mathbb{R}^n$, usando n coordenadas em $\mathbb{R}^n$, podemos interpretar a aplicação $\hat{f} : X \to \mathbb{R}^n$ como sendo n funções $f_i : X \to \mathbb{R}$, $i = 1, \dots, n$, isto é, $\hat{f}(x) = \big(f_1(x), \dots, f_n(x)\big) \in \mathbb{R}^n$ para todo $x \in X$. Chamamos as funções f_i's *componentes* da aplicação $\hat{f}$ (ou de f).

Dadas aplicações $f : X \to Y$ e $g : Y \to Z$ entre conjuntos, lembramos que a *aplicação composta* $g \circ f : X \to Z$ é definida pela regra $(g \circ f)(x) := g\big(f(x)\big)$ para $x \in X$. Uma verificação direta mostra a fórmula $(g \circ f)^{-1}(S) = f^{-1}\big(g^{-1}(S)\big)$ para imagens inversas, onde $S \subset Z$.

A seguinte proposição é uma lista das propriedades básicas de aplicações contínuas. Na maioria dos casos que encontraremos em nossa prática, a proposição junto com a continuidade das funções elementares é uma ferramenta adequada para demonstrar que uma aplicação concreta é contínua.

1.6. Proposição. *Sejam $f_1, f_2 : X \to \mathbb{R}$ funções contínuas, $g_1, g_2 : X \to \mathbb{R}^m$ aplicações contínuas, $f : X \to Y$ e $g : Y \to Z$ aplicações, $\mathbb{R}^m \xrightarrow{L} \mathbb{R}^n$ uma aplicação linear, $X' \subset X$ e $p \in \mathbb{R}^n$, onde $X \subset \mathbb{R}^l$, $Y \subset \mathbb{R}^m$ e $Z \subset \mathbb{R}^n$. Então as seguintes afirmações são válidas.*

1. *A aplicação f é contínua se e só se a correspondente aplicação $\hat{f} : X \to \mathbb{R}^m$ é contínua.*

2. *Se f é contínua, então sua restrição $f|_{X'} : X' \to Y$ é contínua.*

3. *Se f e g são contínuas, então a composta $g \circ f$ é contínua.*

4. *Sejam $U_i \subset_\circ X$, $i \in I$, subconjuntos abertos tais que $X = \bigcup\limits_{i \in I} U_i$. Então a aplicação f é contínua se e só se todas as restrições $f|_{U_i}$, $i \in I$, são contínuas. Em palavras, "ser contínua" é uma* propriedade local *de uma aplicação.*

5. *A aplicação f é contínua se e só se ela* preserva limites. *Isto significa que, toda vez em que temos uma sequência $x_i \in X$, $i \in \mathbb{N}$, convergente a um ponto $x \in X$, $\lim\limits_{i\to\infty} x_i = x$, a sequência $f(x_i)$, $i \in \mathbb{N}$, é convergente e $\lim\limits_{i\to\infty} f(x_i) = f(x)$, ou seja, $\lim\limits_{i\to\infty} f(x_i) = f\big(\lim\limits_{i\to\infty} x_i\big)$.*
6. *A aplicação f é contínua se e só se suas componentes são funções contínuas. Em particular, o "esquecimento" $\mathbb{R}^n \to \mathbb{R}^m$ de algumas coordenadas em $\mathbb{R}^n$ é uma aplicação contínua.*
7. *As aplicações $\mathbb{R}^m \overset{\text{const}}{\longrightarrow} \mathbb{R}^n$ e $\mathbb{R}^n \overset{1_{\mathbb{R}^n}}{\longrightarrow} \mathbb{R}^n$ e as funções $\mathbb{R}^2 \overset{+}{\longrightarrow} \mathbb{R}$, $\mathbb{R}^2 \overset{\times}{\longrightarrow} \mathbb{R}$ e $\mathbb{R}\setminus\{0\} \overset{-1}{\longrightarrow} \mathbb{R}$, dadas pelas regras $x \mapsto p$, $x \mapsto x$, $(x_1,x_2) \mapsto x_1 + x_2$, $(x_1,x_2) \mapsto x_1 \cdot x_2$ e $x \mapsto x^{-1}$, são contínuas.*
8. *As funções $f_1 + f_2$ e $f_1 f_2$ são contínuas. Se $f_2(x) \neq 0$ para todo $x \in X$, então a função f_1/f_2 é contínua.*
9. *A aplicação $f_1 g_1 + f_2 g_2 : X \to \mathbb{R}^m$ é contínua.*
10. *A aplicação linear L é contínua.*[1]

Demonstração. 1. Discutido acima.
2. Discutido acima.
3. Uma consequência direta da fórmula $(g\circ f)^{-1}(S) = f^{-1}\big(g^{-1}(S)\big)$ e da definição de aplicação contínua.
4. Suponhamos que todas as restrições $f|_{U_i} : U_i \to Y$, $i \in I$, são contínuas e seja $V \subset\!\circ\, Y$. Precisamos verificar que $f^{-1}(V) \subset\!\circ\, X$. É fácil ver que $(f|_{U_i})^{-1}(V) = U_i \cap f^{-1}(V)$. Logo, $U_i \cap f^{-1}(V) \subset\!\circ\, U_i$ para todo $i \in I$, ou seja, $U_i \cap f^{-1}(V) = U_i \cap W_i$ para algum $W_i \subset\!\circ\, X$, implicando $U_i \cap f^{-1}(V) \subset\!\circ\, X$. Concluímos de $X = \bigcup\limits_{i\in I} U_i$ que $f^{-1}(V) = \bigcup\limits_{i\in I} \big(U_i \cap f^{-1}(V)\big) \subset\!\circ\, X$.

A recíproca segue do item 2.
5. Suponhamos que f é contínua e que $\lim\limits_{i\to\infty} x_i = x$ com $x_i, x \in X$. Tomemos uma vizinhança aberta arbitrária $f(x) \in V \subset\!\circ\, Y$. Então $x \in f^{-1}(V) \subset\!\circ\, X$ é uma vizinhança aberta de x. Logo, a partir de um momento k, temos $x_i \in f^{-1}(V)$ para $i \geq k$. Portanto, $f(x_i) \in V$ para $i \geq k$, como desejado.

Reciprocamente, suponhamos que f preserva limites e seja $V \subset\!\circ\, Y$. Tomemos qualquer ponto $x \in f^{-1}(V)$. Precisamos provar que existe uma bola aberta $B(x,\varepsilon)$ centrada em x tal que $X \cap B(x,\varepsilon) \subset f^{-1}(V)$. Senão, para todo $i \in \mathbb{N}$, temos $X \cap B(x,\frac{1}{i}) \not\subset f^{-1}(V)$. Podemos escolher então um ponto $x_i \in X \cap B(x,\frac{1}{i})$ tal que $x_i \notin f^{-1}(V)$. Claro que $\lim\limits_{i\to\infty} x_i = x$. Concluímos daí que $\lim\limits_{i\to\infty} f(x_i) = f(x)$. Mas $f(x) \in V \subset\!\circ\, Y$ é uma vizinhança aberta de $f(x)$ e $x_i \notin f^{-1}(V)$ implica $f(x_i) \notin V$. Uma contradição com a definição de limite.
6. Primeiramente observamos, utilizando o item 4, que a *função de projeção* $\mathbb{R}^m \overset{\pi_j}{\longrightarrow} \mathbb{R}$, definida pela regra $(x_1,\dots,x_m) \mapsto x_j$, é contínua. Com efeito, se $\lim\limits_{i\to\infty} p_i = q$ para $p_i = (p_{1i},\dots,p_{mi}) \in \mathbb{R}^m$ e $q = (q_1,\dots,q_m) \in \mathbb{R}^m$, então $\lim\limits_{i\to\infty} |p_i - q| = 0$ pela definição de limite. Logo, $\lim\limits_{i\to\infty} |p_i - q|^2 = 0$ e $\lim\limits_{i\to\infty} (p_{ji} - q_j)^2 = 0$, pois $0 \leq (p_{ji} - q_j)^2 \leq |p_i - q|^2$. Daí segue $\lim\limits_{i\to\infty} (p_{ji} - q_j) = 0$, como desejado.

De fato, raciocínios parecidos aos acima mostram que o limite de uma sequência de pontos em $\mathbb{R}^m$ se calcula por componentes.

A j-ésima componente de $\hat f$ é nada mais do que $\pi_j \circ \hat f$. Pelo item 3, a função $\pi_j \circ \hat f$ é contínua se $\hat f$ for contínua.

Reciprocamente, suponhamos que todas as componentes $f_1,\dots,f_m$ de $\hat f$ são contínuas e seja $p_i \in X$, $i \in \mathbb{N}$, uma sequência convergente a um ponto $q \in X$, $\lim\limits_{i\to\infty} p_i = q$. Precisamos demonstrar que $\lim\limits_{i\to\infty} \hat f(p_i) = \hat f(q)$. Como foi observado, o limite se calcula por componentes. Deste modo, precisamos verificar que $\lim\limits_{i\to\infty} f_j(p_i) = f_j(q)$. Mas isto segue da continuidade da componente f_j pelo item 5.
7. A aplicação $\mathbb{R}^m \overset{\text{const}}{\longrightarrow} \mathbb{R}^n$ é contínua, pois a imagem inversa de qualquer subconjunto $S \subset \mathbb{R}^n$ é ou vazia (se $p \notin S$) ou todo $\mathbb{R}^m$ (se $p \in S$).

A aplicação $\mathbb{R}^n \overset{1_{\mathbb{R}^n}}{\longrightarrow} \mathbb{R}^n$ é contínua, pois a imagem inversa de qualquer subconjunto $S \subset \mathbb{R}^n$ é ele mesmo.

---
[1]Cabe mencionar que este fato não vale para espaços lineares de dimensão infinita.

Para demonstrar que as funções $\mathbb{R}^2 \xrightarrow{+} \mathbb{R}$ e $\mathbb{R}^2 \xrightarrow{\times} \mathbb{R}$ são contínuas, tomemos uma sequência $(x_i, y_i) \in \mathbb{R}^2$ convergente a um ponto $(x,y) \in \mathbb{R}^2$, $\lim\limits_{i\to\infty}(x_i,y_i) = (x,y)$. Como foi observado na demonstração do item anterior, isto significa que $\lim\limits_{i\to\infty} x_i = x$ e $\lim\limits_{i\to\infty} y_i = y$. Lembramos agora que $\lim\limits_{i\to\infty}(x_i+y_i) = x+y$ e $\lim\limits_{i\to\infty} x_iy_i = xy$. Resta utilizar o item 5.

O mesmo funciona para a função $\mathbb{R}\setminus\{0\} \xrightarrow{-1} \mathbb{R}$. Sejam $x_i \in \mathbb{R}\setminus\{0\}$, $i \in \mathbb{N}$, tais que $\lim\limits_{i\to\infty} x_i = x \in \mathbb{R}\setminus\{0\}$. Sabemos que $\lim\limits_{i\to\infty} x_i^{-1} = x^{-1}$.

8. Consideremos a aplicação $h : X \to \mathbb{R}^2$ cujas componentes são f_1 e f_2. A função $f_1 + f_2$ é a composta da função $+$ (vide item 7) e da aplicação h. Pelo item 3, $f_1 + f_2$ é contínua. Da mesma maneira, a função f_1f_2, sendo a composta de $\times$ e de h, é contínua. Finalmente, $f_1/f_2 = f_1(1/f_2)$. A função $1/f_2$, sendo a composta da função $f_2 : X \to \mathbb{R}\setminus\{0\}$ com a função $^{-1}$ (vide item 7), é contínua pelo item 3.
9. A afirmação segue dos itens 8 e 6.
10. Pelo item 6, podemos supor que $n = 1$, ou seja, $L(x_1,\dots,x_m) = c_1x_1 + \dots + c_mx_m$ para alguns $c_1,\dots,c_m \in \mathbb{R}$. Agora, o fato segue do item 8 e da segunda afirmação do item 6 ■

1.7. Continuidade uniforme e convergência uniforme. Como sabemos, a continuidade de uma função $f : X \to \mathbb{R}$ se expressa assim: $\forall\, \varepsilon > 0\ \forall x_0 \in X\ \exists\, \delta > 0\ \forall x \in X\ |x-x_0| < \delta \implies \big|f(x)-f(x_0)\big| < \varepsilon$. Enfatizamos que δ depende de $x_0 \in X$ e de ε. Se trocarmos os quantificadores, ou seja, se fizermos δ independer de $x_0 \in X$, obtemos a seguinte

1.7.1. Definição. Seja $X \subset \mathbb{R}^n$. Uma função $f : X \to \mathbb{R}$ é dita *uniformemente contínua* sobre X se $\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall x_0, x_1 \in X\ |x_1 - x_0| < \delta \implies \big|f(x_1) - f(x_0)\big| < \varepsilon$.

1.7.2. Definição. Um subconjunto $L \subset \mathbb{R}^n$ é *limitado* se ele está contido em uma bola. Equivalentemente, L é limitado se $|L| \leq r$ para algum $r \in \mathbb{R}$. Dizemos que $\varnothing \neq K \subset \mathbb{R}^n$ é um *compacto* se K é fechado em $\mathbb{R}^n$, $K \subset_f \mathbb{R}^n$, e limitado.

Um exemplo de um compacto é um bloco fechado $[a,b] \times [c,d] \subset \mathbb{R}^2$.

1.7.3. Exercício. Utilizando o fato que as coordenadas em $\mathbb{R}^2$ são contínuas pela Proposição 1.6 (6), prove que o bloco $[a,b] \times [c,d] \subset \mathbb{R}^2$ é fechado.

1.7.4. Definição. Seja $f : X \to \mathbb{R}$ uma função. Definimos a *norma* de f (sobre X) pela fórmula $||f|| := ||f||_X := \sup\big|f(X)\big|$, onde $\sup S$ denota o *supremo* (isto é, a menor cota superior[2]) de $S \subset \mathbb{R}$. Claro que $||f|| = 0$ implica $f = 0$.

1.7.5. Lema. *Sejam $f, g : X \to \mathbb{R}$ funções. Então $||f + g|| \leq ||f|| + ||g||$ e $||fg|| \leq ||f|| \cdot ||g||$.*

Demonstração. Para qualquer $x \in X$, temos $\big|f(x) + g(x)\big| \leq \big|f(x)\big| + \big|g(x)\big| \leq ||f|| + ||g||$ e $\big|f(x) \cdot g(x)\big| = \big|f(x)\big| \cdot \big|g(x)\big| \leq ||f|| \cdot ||g||$ ■

1.7.6. Definição. Dizemos que uma sequência de funções $f_i : X \to \mathbb{R}$, $i \in \mathbb{N}$, converge *uniformemente* a uma função $f : X \to \mathbb{R}$ e escrevemos $\lim\limits_{i\to\infty} f_i = f$ se $\lim\limits_{i\to\infty} ||f_i - f|| = 0$.

O fato que $||f_i - f|| \leq \varepsilon$ significa que o "gráfico" de f_i está situado na "faixa" de largura 2ε em torno do "gráfico" de f. Consequentemente, a convergência uniforme implica que o gráfico de f_i está na faixa, de largura arbitrariamente pequena, em torno do "gráfico" de f para todo i suficientemente grande.

Recordamos agora o

1.7.7. Lema (teorema de Lagrange). *Seja $f : [u,v] \to \mathbb{R}$ uma função contínua derivável em (u,v). Então $f(v) = f(u) + f'(\xi)(v-u)$ para algum $\xi \in (u,v)$.*

[2]A propriedade característica do supremo: $S \leq \sup S$ (isto significa que $s \leq \sup S$ para todo $s \in S$) e, para qualquer $r \in \mathbb{R} \cup \{\infty\}$ tal que $S \leq r$, temos $\sup S \leq r$.

1.7.8. Definição. Uma função $f : [a,b] \to \mathbb{R}$ é dita *integrável* se f é o limite uniforme de uma sequência de funções escada $f_i : [a,b] \to \mathbb{R}$, $i \in \mathbb{N}$, $\lim\limits_{i\to\infty} f_i = f$.

Lembramos também as propriedades básicas da integral de funções de uma variável real.

1.7.9. Proposição. *Sejam $f, f_1, f_2, h_i : [a,c] \to \mathbb{R}$, $i \in \mathbb{N}$, funções integráveis, $h : [a,c] \to \mathbb{R}$ uma função, $b \in [a,c]$ e $r_1, r_2 \in \mathbb{R}$. Então os seguintes fatos são válidos.*

- *As funções $f|_{[a,b]}$ e $f|_{[b,c]}$ são integráveis e $\int_a^c f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t + \int_b^c f(t)\,\mathrm{d}t$.*
- *A função $r_1 f_1 + r_2 f_2$ é integrável e $\int_a^c \big(r_1 f_1(t) + r_2 f_2(t)\big)\,\mathrm{d}t = r_1 \int_a^c f_1(t)\,\mathrm{d}t + r_2 \int_a^c f_2(t)\,\mathrm{d}t$.*
- *Se $f_1(t) \le f_2(t)$ para todo $t \in [a,c]$, então $\int_a^c f_1(t)\,\mathrm{d}t \le \int_a^c f_2(t)\,\mathrm{d}t$.*
- *A função $|f|$ é integrável e $\big|\int_a^c f(t)\,\mathrm{d}t\big| \le \int_a^c \big|f(t)\big|\,\mathrm{d}t$.*
- $\big|\int_a^c f(t)\,\mathrm{d}t\big| \le ||f||(c-a)$.
- *A função $f_1 f_2$ é integrável.*
- *Se $\big|f(t)\big| \ge m > 0$ para todo $t \in [a,c]$ e algum $m \in \mathbb{R}$, a função $\frac{1}{f}$ é integrável.*
- *Se $\lim\limits_{n\to\infty} h_i = h$ é limite uniforme, então h é uma função integrável e $\int_a^c h(t)\,\mathrm{d}t = \lim\limits_{i\to\infty} \int_a^c h_i(t)\,\mathrm{d}t$.*

Em outras palavras, limite uniforme comuta com integral. Em particular, toda função contínua é integrável.

Nossas "reminiscências" já estão ultrapassando todos os limites. Vamos por isto adiar a demonstração do seguinte lema até um momento adequado.

1.7.10. Lema. *Sejam $K \subset \mathbb{R}^m$ um compacto e $g : K \to \mathbb{R}$ uma função contínua. Então a função g é uniformemente contínua. Além disso, existem $k_0, k_1 \in K$ tais que $g(k_0) \le g(k) \le g(k_1)$ para todo $k \in K$.*

Os conceitos e fatos introduzidos acima possibilitam em certas circunstâncias comutar a derivação com a integração:

1.7.11. Lema. *Seja $f : [a,b] \times (u,v) \to \mathbb{R}$ uma função contínua tal que $f(t,s)$ é derivável em $s \in (u,v)$ para cada $t \in [a,b]$. Suponhamos que a função $\frac{\mathrm{d}f}{\mathrm{d}s}(t,s)$ é contínua em $[a,b] \times (u,v)$. Então a função $h(s) := \int_a^b f(t,s)\,\mathrm{d}t$ é derivável e $h'(s) := \int_a^b \frac{\mathrm{d}f}{\mathrm{d}s}(t,s)\,\mathrm{d}t$. Em outras palavras, $\frac{\mathrm{d}}{\mathrm{d}s}\int_a^b f(t,s)\,\mathrm{d}t = \int_a^b \frac{\mathrm{d}}{\mathrm{d}s} f(t,s)\,\mathrm{d}t$ se a função $\frac{\mathrm{d}}{\mathrm{d}s} f(t,s)$ é contínua em $[a,b] \times (u,v)$.*

Demonstração. Seja $s \in (u,v)$. Precisamos mostrar que, para qualquer sequência $s \ne s_i \in (u,v)$, $i \in \mathbb{N}$, convergente a s, $\lim\limits_{i\to\infty} s_i = s$, vale $\lim\limits_{i\to\infty} \frac{h(s_i)-h(s)}{s_i-s} = \int_a^b \frac{\mathrm{d}}{\mathrm{d}s} f(t,s)\,\mathrm{d}t$, isto é, $\lim\limits_{i\to\infty} \int_a^b \frac{f(t,s_i)-f(t,s)}{s_i-s}\,\mathrm{d}t = \int_a^b \frac{\mathrm{d}}{\mathrm{d}s} f(t,s)\,\mathrm{d}t$. Pela Proposição 1.7.9, é suficiente provar que a função $g(t) := \frac{\mathrm{d}}{\mathrm{d}s} f(t,s)$ é o limite uniforme da sequência de funções $g_i(t) := \frac{f(t,s_i)-f(t,s)}{s_i-s}$, $i \in \mathbb{N}$. É claro que existe um intervalo fechado $[c,d] \subset (u,v)$ tal que $s, s_i \in [c,d]$ para todo $i \in \mathbb{N}$. Sendo $g(t,x) := \frac{\mathrm{d}}{\mathrm{d}s} f(t,x)$ contínua em $[a,b] \times (u,v)$, ela é uniformemente contínua em $[a,b] \times [c,d]$ pelos Exercício 1.7.3 e Lema 1.7.10.

Tomemos $\varepsilon > 0$ qualquer. Sendo $g(t,x)$ uniformemente contínua em $[a,b] \times [c,d]$, existe $\delta > 0$ tal que $|x-s| < \delta \Longrightarrow \big|g(t,x) - g(t,s)\big| < \varepsilon$. Devido a $\lim\limits_{i\to\infty} s_i = s$, existe $m \in \mathbb{N}$ tal que $|s_i - s| < \delta$ para todo $i \ge m$. Pelo teorema de Lagrange (o Lema 1.7.7), para cada $i \ge m$ e cada $t \in [a,b]$, podemos encontrar ξ_i entre s_i e s tal que $\frac{f(t,s_i)-f(t,s)}{s_i-s} = g(t,\xi_i)$. Já que $|s_i - s| < \delta$ implica $|\xi_i - s| < \delta$, obtemos $\big|g(t,\xi_i) - g(t,s)\big| < \varepsilon$ para todo $i \ge m$, ou seja, $\big|g_i(t) - g(t)\big| < \varepsilon$ para todo $i \ge m$ e $t \in [a,b]$. Isto significa que $||g_i - g|| < \varepsilon$ para todo $i \ge m$ ∎

1.8. Definição. Uma aplicação contínua $c : [a,b] \to \mathbb{R}^n$ se chama *caminho de classe* C^0 em $\mathbb{R}^n$ ou simplesmente *caminho* em $\mathbb{R}^n$ ou *curva parametrizada* em $\mathbb{R}^n$. Um caminho $c : [a,b] \to \mathbb{R}^n$ é dito *diferenciável* se existe a derivada $\dot{c}(t_0) := c'(t_0) := \lim\limits_{\substack{[a,b]\ni t\to t_0\\ t\ne t_0}} \frac{c(t)-c(t_0)}{t-t_0}$ para todo $t_0 \in [a,b]$. Se, por sua vez, $\dot{c} : [a,b] \to \mathbb{R}^n$ é um caminho de classe C^0, ou seja, é contínuo, dizemos que c é um caminho *de classe* C^1. Por indução, um caminho *de classe* C^k é um caminho diferenciável cuja derivada é de classe C^{k-1}. Se um caminho é de classe C^k para todo $k \in \mathbb{N}$, dizemos que o caminho é *de classe* C^∞.

Às vezes, vale a pena pensar no parâmetro t em $c(t)$ como se t fosse o tempo. Com esta visão, o valor $c(t)$ indica no espaço a posição do ponto no momento t. Fazendo t variar, $c(t)$ descreve o movimento do ponto no espaço e, em particular, a trajetória do ponto. A derivada $\dot{c}(t_0)$ (quando existir) pode ser considerada como *vetor velocidade* no momento $t_0 \in [a, b]$. A segunda derivada $\ddot{c}(t_0)$ pode ser interpretada como *vetor aceleração.*

Pela Proposição 1.6 (6), um caminho $c : [a, b] \to \mathbb{R}^n$ é nada mais do que n funções contínuas $c_1, \ldots, c_n : [a, b] \to \mathbb{R}$, as componentes de c. Deste modo, um caminho diferenciável $c : [a, b] \to \mathbb{R}^n$ é simplesmente n funções deriváveis.

1.9. Integração de caminho. A gente pode integrar caminhos do mesmo jeito como acabamos de derivá-los. Por exemplo, os itens 1, 2, 3 (interpretando a desigualdade $(p_1, \ldots, p_n) \leq (q_1, \ldots, q_n)$ como $p_i \leq q_i$ para todo $i = 1, \ldots, n$) e 8 da Proposição 1.7.9 são obviamente válidos para caminhos. Os itens 4 e, consequentemente, 5 também permanecem válidos para caminhos, mas o 4 exige uma demonstração. "Lembraremos" tudo isto quando necessitarmos (veja a Proposição 2.11.5).

1.10. Topologia e normas em $\mathbb{R}^n$. Na Definição 1.1, introduzimos a topologia padrão em $\mathbb{R}^n$ usando de fato uma norma "padrão" em $\mathbb{R}^n$.

1.10.1. Definição. Seja V um espaço $\mathbb{R}$-linear. Uma aplicação $|\ | : V \to \mathbb{R}$ se chama *norma* em V se são válidos os seguintes axiomas.

- $|v| \geq 0$ para todo $v \in V$ com $|v| = 0$ apenas no caso $v = 0$.
- $|rv| = |r| \cdot |v|$ para todos $v \in V$ e $r \in \mathbb{R}$.
- $|v_1 + v_2| \leq |v_1| + |v_2|$ para todos $v_1, v_2 \in V$.

Duas normas $|\ |$ e $|\ |'$ em V são *equivalentes* se existem constantes c, c' tais que $|v|' \leq c|v|$ e $|v| \leq c'|v|'$ para todo $v \in V$.

1.10.2. Lema. *Se V é um espaço $\mathbb{R}$-linear de dimensão finita,*[3] *então todas as normas em V são equivalentes.*

Demonstração. Basta considerar a norma "padrão" $|\ |$ e uma norma arbitrária $|\ |'$ em V.

Note que $\big||p|' - |q|'\big| \leq |p-q|'$. Realmente, $|p|' = |p-q+q|' \leq |p-q|' + |q|'$ implica $|p|' - |q|' \leq |p-q|'$. Pela simetria, $|q|' - |p|' \leq |q-p|' = \big|(-1)(p-q)\big|' = |p-q|'$.

Primeiramente observamos que a função $|\ |'$ é contínua em V. Com efeito, pela Proposição 1.6 (5), basta ver que $\lim\limits_{i\to\infty} p_i = q$ implica $\lim\limits_{i\to\infty} |p_i|' = |q|'$, ou seja, $\lim\limits_{i\to\infty} \big||p_i|' - |q|'\big| = 0$. Logo, é suficiente provar que $\lim\limits_{i\to\infty} |p_i - q|' = 0$. Escrevendo q e p_i em termos da base "padrão" $e_1, \ldots, e_n \in V = \mathbb{R}^n$, obtemos $p_i = \sum\limits_{j=1}^{n} p_{ij} e_j$ e $q = \sum\limits_{j=1}^{n} q_j e_j$ com $p_{ij}, q_j \in \mathbb{R}$. Como foi observado na demonstração da Proposição 1.6 (6), temos $\lim\limits_{i\to\infty} p_{ij} = q_j$, ou seja, $\lim\limits_{i\to\infty} |p_{ij} - q_j| = 0$ para todo $j = 1, \ldots, n$. Agora concluímos que $|p_i - q|' = \Big|\sum\limits_{j=1}^{n} (p_{ij} - q_j) e_j\Big|' \leq \sum\limits_{j=1}^{n} |p_{ij} - q_j| \cdot |e_j|'$ tende a 0 quando $i \to \infty$.

Sendo a imagem inversa do conjunto fechado $\{1\} \subset \mathbb{R}$ em relação à função contínua $|\ |$, a esfera $\mathbb{S}^{n-1} := \big\{p \in \mathbb{R}^n \mid |p| = 1\big\}$ é compacta, pois é obviamente limitada. Pelo Lema 1.7.10, a função contínua $|\ |'$ assume em $\mathbb{S}^{n-1}$ seu mínimo $\frac{1}{c'} > 0$ e seu máximo $c > 0$. Para qualquer $0 \neq v \in V$, temos $\frac{v}{|v|} \in \mathbb{S}^{n-1}$. Consequentemente, $\frac{1}{c'} \leq |\frac{v}{|v|}|' \leq c$, isto é, $\frac{1}{c'}|v| \leq |v|' \leq c|v|$ ■

Concluímos do Lema 1.10.2 que a topologia padrão em $\mathbb{R}^n$ independe da escolha de norma em $\mathbb{R}^n$.

[3]Cabe mencionar que este fato não vale para espaços lineares de dimensão infinita.

2. Funções e aplicações diferenciáveis

Na rodovia, todo mundo anda ao longo de uma reta.
Só que a reta também está andando!
— *Um motorista bêbado*

Lidando no cálculo I com derivadas de funções de uma variável real, apesar de saber relacionar a derivada com a reta tangente ao gráfico da função, talvez não chegamos ao entender pleno que a derivada $f'(p)$ da função $f(x)$ em p é de fato a melhor aproximação de $f(x)$ por uma função linear na vizinhança de p.

2.1. Definição. Sejam V, V' espaços $\mathbb{R}$-lineares de dimensão finita, seja $U \subset\circ V$ um subconjunto aberto e seja $f : U \to V'$ uma aplicação. Dizemos que f é *derivável* em $p \in U$ se existe uma aplicação linear $L : V \to V'$ tal que a aplicação $g(h) := f(p+h) - f(p) - Lh$, definida em uma vizinhança de $0 \in V$, é *pequena em comparação* com $|h|$. Isto significa que $\lim\limits_{0\neq h\to 0} \frac{|g(h)|}{|h|} = 0$ (ou, equivalentemente, que $\lim\limits_{0\neq h\to 0} \frac{g(h)}{|h|} = 0$).

A aplicação linear L é única se existir. Realmente, se a aplicação $g'(h) := f(p+h) - f(p) - L'h$ também é pequena em comparação com $|h|$, de $\big|g'(h) - g(h)\big| \le \big|g'(h)\big| + \big|g(h)\big|$ segue $\lim\limits_{0\neq h\to 0} \frac{|(L'-L)h|}{|h|} = 0$. Supondo que $L' - L \neq 0$, encontramos $v \in V$ tal que $(L' - L)v \neq 0$. Agora, $0 = \lim\limits_{0<t\to 0} \frac{|(L'-L)(tv)|}{|tv|} = \frac{|(L'-L)v|}{|v|}$, uma contradição.

Chamamos L a *derivada* de f em p e denotamos $D_p f := L$. Note que, pelo Lema 1.10.2, estes conceitos independem da escolha de normas em V e V'.

O conceito de derivada de uma função de uma variável real e, mais ainda, o de derivada de um caminho $c : (a,b) \to \mathbb{R}^n$ em $t_0 \in (a,b)$, são casos particulares da Definição 2.1, pois $\dot c(t_0) = \lim\limits_{0\neq h\to 0} \frac{c(t_0+h)-c(t_0)}{h}$ implica que a função $g(h) := c(t_0+h) - c(t_0) - \dot c(t_0)h$, definida em uma vizinhança aberta de $0 \in \mathbb{R}$, é pequena em comparação com $|h|$. A única diferença é que agora a derivada $\dot c(t_0)$ não é um vetor, mas é uma aplicação linear $\dot c(t_0) : \mathbb{R} \to \mathbb{R}^n$, dada pela regra $h \mapsto \dot c(t_0)h$, sendo $\dot c(t_0)$ como na Definição 1.8.

2.2. Lema (regra da cadeia[4]). *Sejam V_1, V_2, V_3 espaços $\mathbb{R}$-lineares de dimensão finita, $U_1 \subset\circ V_1$ e $U_2 \subset\circ V_2$ subconjuntos abertos e $U_1 \xrightarrow{f_1} U_2 \xrightarrow{f_2} V_3$ aplicações. Se f_1 é derivável em $p \in U_1$ e f_2 é derivável em $f_1(p) \in U_2$, então a composta $f_2 \circ f_1 : U_1 \to V_3$ é derivável em p e $D_p(f_2 \circ f_1) = D_{f_1(p)} f_2 \circ D_p f_1$.*

Demonstração. As aplicações $g_1 : U \to V_2$ e $g_2 : U' \to V_3$, dadas pelas regras

$$g_1(h) := f_1(p+h) - f_1(p) - (D_p f_1)h, \qquad g_2(h') := f_2\big(f_1(p) + h'\big) - f_2\big(f_1(p)\big) - (D_{f_1(p)} f_2)h'$$

e definidas respectivamente em vizinhanças abertas $0 \in U \subset\circ V_1$ e $0 \in U' \subset\circ V_2$, são pequenas em comparação com $|h|$ e $|h'|$. Já que $D_p f_1$ é contínua pela Proposição 1.6 (10) e $\lim\limits_{h\to 0} g_1(h) = 0$, podemos reescolher a vizinhança $0 \in U \subset\circ V_1$ fazendo-a tão pequena que $h' := (D_p f_1)h + g_1(h) \in U'$ para todo $h \in U$. Agora, pela linearidade de $D_{f_1(p)} f_2$, temos

$$f_2\big(f_1(p+h)\big) = f_2\big(f_1(p) + (D_p f_1)h + g_1(h)\big) = f_2\big(f_1(p) + h'\big) = f_2\big(f_1(p)\big) + (D_{f_1(p)} f_2)h' + g_2(h') =$$

$$= f_2\big(f_1(p)\big) + (D_{f_1(p)} f_2)(D_p f_1)h + (D_{f_1(p)} f_2)\big(g_1(h)\big) + g_2\big((D_p f_1)h + g_1(h)\big)$$

[4]Tenho certeza que o inventor desta “terminologia” está no inferno, na cadeia, pois regra é regra.

para todo $h \in U$. Portanto,

$$f_2\big(f_1(p+h)\big) - f_2\big(f_1(p)\big) - (D_{f_1(p)}f_2)(D_pf_1)h = (D_{f_1(p)}f_2)\big(g_1(h)\big) + g_2\big((D_pf_1)h + g_1(h)\big).$$

Sendo $D_{f_1(p)}f_2$ linear e contínua pela Proposição 1.6 (10), a função $(D_{f_1(p)}f_2)\big(g_1(h)\big)$ é pequena em comparação com $|h|$, pois $\lim\limits_{0\neq h\to 0} \frac{(D_{f_1(p)}f_2)(g_1(h))}{|h|} = D_{f_1(p)}f_2 \lim\limits_{0\neq h\to 0}\frac{g_1(h)}{|h|} = 0$ pela Proposição 1.6 (5).

Resta provar que $\lim\limits_{0\neq h\to 0} \frac{\left|g_2\left((D_pf_1)h+g_1(h)\right)\right|}{|h|} = 0$. Sabemos que $|g_2(h')| = |h'|g(h')$, onde $g : U' \to \mathbb{R}^{\geq 0}$ é uma função que satisfaz $\lim\limits_{h'\to 0} g(h') = 0$. Logo,

$$\frac{\left|g_2\left((D_pf_1)h+g_1(h)\right)\right|}{|h|} = \frac{\left|(D_pf_1)h+g_1(h)\right|}{|h|}g\big((D_pf_1)h+g_1(h)\big) \leq \Big(\big|(D_pf_1)\tfrac{h}{|h|}\big| + \tfrac{|g_1(h)|}{|h|}\Big)g\big((D_pf_1)h+g_1(h)\big).$$

A aplicação D_pf_1 é linear e, pela Proposição 1.6 (10), é contínua. De $\lim\limits_{h\to 0} g_1(h) = 0$, concluímos que $(D_pf_1)h + g_1(h)$ tende a 0 quando h tende a 0. Levando em conta que $\lim\limits_{0\neq h\to 0}\frac{|g_1(h)|}{|h|} = 0$, resta observar que os valores da função $0 \neq h \mapsto \big|(D_pf_1)\frac{h}{|h|}\big|$ são limitados. Note que os pontos $\frac{h}{|h|}$, $h \neq 0$, estão na esfera unitária $\mathbb{S}^{n-1}$ (onde n é a dimensão de V_1). Na demonstração do Lema 1.10.2 foi observado que $\mathbb{S}^{n-1}$ é um compacto. Pelo Lema 1.7.10, os valores da função $\big|(D_pf_1)x\big|$, $x \in \mathbb{S}^{n-1}$, são limitados, pois essa função é contínua pela Proposição 1.6 ∎

2.3. Matriz jacobiana. Em termos de coordenadas $(x_1,\dots,x_m)$ em $\mathbb{R}^m$ e $(y_1,\dots,y_n)$ em $\mathbb{R}^n$, a derivada em $p \in U \subset\!\circ\, \mathbb{R}^m$ da aplicação $f : U \to \mathbb{R}^n$, pode ser descrita pela *matriz jacobiana*: Denotemos por $f_i(x_1,\dots,x_m)$ e $D_i(h_1,\dots,h_m)$, $i = 1,\dots,n$, as componentes de $f(x)$ e de $(D_pf)h$ e seja $e_1,\dots,e_m \in \mathbb{R}^m$ a base que corresponde às coordenadas $(x_1,\dots,x_m)$. Substituindo $h := te_j$, $0 \neq t \in \mathbb{R}$, em $\frac{g(h)}{|h|} = \frac{f(p+h)-f(p)-(D_pf)h}{|h|}$, obtemos

$$0 = \lim_{0\neq t\to 0} \frac{f_i(p+te_j) - f_i(p) - D_i(te_j)}{|te_j|} =$$

$$= \lim_{0\neq t\to 0}\Big(\frac{f_i(p_1,\dots,p_{j-1},p_j+t,p_{j+1},\dots,p_m) - f_i(p_1,\dots,p_{j-1},p_j,p_{j+1},\dots,p_m)}{t} - D_ie_j\Big)\frac{t}{|t|}.$$

Isto implica que a matriz $[D_ie_j]_{ij}$ de D_pf é montada pelas *derivadas parciais* em p das componentes f_i's de f, ou seja, $[D_ie_j]_{ij} = \big[\frac{\partial f_i}{\partial x_j}(p)\big]_{ij}$. Assim, $f_i(p+h) = f_i(p) + \sum\limits_{j=1}^{m}\frac{\partial f_i}{\partial x_j}(p)h_j + g_i(h)$ para todos $i = 1,\dots,n$, onde g_i denota a i-ésima componente de g.

2.4. Definição. Sejam V_1, V_2 espaços $\mathbb{R}$-lineares e $U \subset\!\circ\, V_1$. Uma aplicação contínua $f : U \to V_2$ se chama *aplicação de classe* C^0. Uma aplicação $f : U \to V_2$ é *derivável sobre* U se ela é derivável em cada $p \in U$. Neste caso, f é de classe C^0 pela Proposição 1.6 (5). Mais ainda, a regra $p \mapsto D_pf$ define uma aplicação $Df : U \to \mathrm{Lin}_{\mathbb{R}}(V_1,V_2)$. Sabemos que $\mathrm{Lin}_{\mathbb{R}}(V_1,V_2)$ é um espaço $\mathbb{R}$-linear de dimensão finita. Caso a aplicação Df seja contínua, chamamos f uma *aplicação de classe* C^1. Por indução, se Df é de classe C^{k-1}, dizemos que f é uma *aplicação de classe* C^k. Se f é de classe C^k para todo $k \in \mathbb{N}$, a aplicação f é *de classe* C^∞.

Caso $f : U \to \mathbb{R}$ seja uma função de classe C^k, escrevemos $f \in C^k(U)$, onde $k \in \mathbb{N}\cup\{\infty\}$.

Por 2.3, f é de classe C^k se e só se todas as derivadas parciais de ordem k das componentes de f são contínuas sobre U, isto é, $\frac{\partial^k f_i}{\partial x_{j_1}\dots\partial x_{j_k}} \in C^0(U)$ para todos $i = 1,\dots,n$ e $j_1,\dots,j_k \in \{1,\dots,m\}$.

2.5. Exemplo. Seja $\mathbb{R}^n \supset\!\circ\, U \xrightarrow{f} \mathbb{R}$ uma função. Dizemos que um ponto $p \in U$ é um *mínimo* (*máximo*) *local* de f se existe uma vizinhança aberta $p \in W \subset\!\circ\, U$ tal que $f(w) \geq f(p)$ ($f(w) \leq f(p)$) para todo $w \in W$. Lidando no cálculo I com funções de uma variável real, estudamos condições necessárias (a primeira derivada é nula no ponto) e suficientes (a primeira derivada é nula no ponto e a segunda é positiva/negativa) para um ponto ser um mínimo/máximo local de uma função de classe C^2. Para funções de várias variáveis, a situação é parecida, mas um pouco mais complicada.

Consideremos uma função $f : \mathbb{R}^n \to \mathbb{R}$ dada pelo polinômio de grau ≤ 2 nas coordenadas $(x_1, \dots, x_n)$ em $\mathbb{R}^n$, $f(x_1, \dots, x_n) = c_0 + \sum\limits_{j=1}^{n} c_j x_j + \sum\limits_{j,k=1}^{n} c_{jk} x_j x_k$, onde $c_0, c_j, c_{jk} \in \mathbb{R}$ e $c_{jk} = c_{kj}$. As derivadas parciais de ordem 1 de f são $\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) = c_i + \sum\limits_{k=1}^{n} c_{ik} x_k + \sum\limits_{j=1}^{n} c_{ji} x_j = c_i + 2 \sum\limits_{k=1}^{n} c_{ik} x_k$, $i = 1, \dots, n$. As derivadas parciais de ordem 2 de f são constantes $\frac{\partial^2 f}{\partial x_j \partial x_i}(x_1, \dots, x_n) = 2c_{ij}$, $i, j = 1, \dots, n$. As derivadas parciais de ordem ≥ 3 de f são portanto nulas. Assim, $f \in C^\infty(\mathbb{R}^n)$. A decomposição $f(x_1, \dots, x_n) = c_0 + \sum\limits_{j=1}^{n} c_j x_j + \sum\limits_{j,k=1}^{n} c_{jk} x_j x_k$ pode ser vista como a decomposição $f(x) = f(0) + (D_0 f)x + g(x)$, onde $g(x)$ é uma função pequena em comparação com $|x|$. Isto implica $(D_0 f)(x_1, \dots, x_n) = \sum\limits_{j=1}^{n} c_j x_j$.

Quando a origem 0 é um mínimo local de f ? Suponhamos que 0 é um mínimo local de f. Se $D_0 f \neq 0$, podemos encontrar $p \in \mathbb{R}^n$ tal que $(D_0 f)p < 0$. Agora, para todos suficientemente pequenos $t > 0$, temos $t(D_0 f)p + g(tp) < 0$. Portanto, $f(tp) = f(0) + t(D_0 f)p + g(tp) < f(0)$ para tais t, contradizendo 0 ser um mínimo local de f. Logo, $D_0 f = 0$ e, sem perda de generalidade, podemos supor que $c_0 = 0$, ou seja, que $f(x_1, \dots, x_n) = xCx^t$, onde $x := [x_1 \dots x_n]$, x^t denota a matriz x transposta e C é a matriz simétrica $C := [c_{ij}]$. Dizemos que uma matriz simétrica C é *positiva* se $xCx^t \geq 0$ para todo x e se $xCx^t = 0$ apenas para $x = 0$. É imediato que 0 é um mínimo local da função xCx^t se a matriz simétrica C é positiva. (Claro que este mínimo é de fato global.) O critério de Sylvester possibilita decidir se uma matriz simétrica é positiva. Por exemplo, no caso $n = 2$, o critério diz que a matriz $C = [c_{ij}]$ é positiva se e só se $c_{11} > 0$ e $\det C > 0$.

Podemos resumir o resultado obtido: um polinômio f de grau ≤ 2 tem um mínimo local em p se $D_p f = 0$ e a matriz *hessiana* $[\frac{\partial^2 f}{\partial x_i \partial x_j}(p)]_{ij}$ de f em p é positiva. Posteriormente perceberemos que o mesmo fato vale para qualquer função de classe C^3.

2.6. Lema (de Hadamard)**.** *Seja $f : U \to \mathbb{R}$ uma função de classe C^∞ definida em uma vizinhança aberta $p \in U \subset\!\circ\, \mathbb{R}^n$ do ponto $p = (p_1, \dots, p_n) \in \mathbb{R}^n$. Então existem uma vizinhança aberta $p \in W \subset\!\circ\, U$ e funções $g_i : W \to \mathbb{R}$ de classe C^∞ tais que $f(x) = f(p) + \sum\limits_{i=1}^{n} (x_i - p_i) g_i(x)$ para todo $x \in W$.*

Demonstração. Seja $W := B(p, r)$, $r > 0$, uma bola aberta centrada em p tal que $W \subset U$ e seja $x \in W$. Definimos uma função $u : [0, 1] \to \mathbb{R}$ pela regra $u(t) := f\big(p + t(x - p)\big)$. Pelo Lema 2.2 e por 2.3, a função $u(t)$ é de classe[5] C^∞ e $u'(t) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}\big(p + t(x - p)\big)(x_i - p_i)$.

Pelo teorema fundamental do cálculo,

$$f(x) - f(p) = u(1) - u(0) = \int_0^1 u'(t)\,\mathrm{d}t = \int_0^1 \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\big(p + t(x - p)\big)(x_i - p_i)\,\mathrm{d}t =$$

$$= \sum_{i=1}^{n} (x_i - p_i) \int_0^1 \frac{\partial f}{\partial x_i}\big(p + t(x - p)\big)\,\mathrm{d}t.$$

[5]De fato, para cada $x \in W$ fixo, a função u pode ser definida em um intervalo aberto que contém $[0, 1]$; a função u é de classe C^∞ neste intervalo aberto.

Resta fazer $g_i(x) := \int_0^1 \frac{\partial f}{\partial x_i}\big(p + t(x-p)\big)\,\mathrm{d}t$ e observar que g_i é de classe C^∞ aplicando várias vezes o Lema 1.7.11 ■

2.7. Lema. A função $f : \mathbb{R} \to \mathbb{R}^{\geq 0}$, definida pela regra $f(x) := \begin{cases} 0 & \text{se} \quad x \leq 0 \\ e^{-\frac{1}{x}} & \text{se} \quad x > 0 \end{cases}$, é de classe C^∞, $f \in C^\infty(\mathbb{R})$, e $f(x) = 0$ se e só se $x \leq 0$.

Demonstração. Basta mostrar, por indução sobre $n \in \mathbb{N}$, que

$$f^{(n)}(x) = \begin{cases} 0 & \text{se} \quad x \leq 0 \\ e^{-\frac{1}{x}} x^{-2n} p_n(x) & \text{se} \quad x > 0 \end{cases}, \tag{2.8}$$

onde $p_n(x)$ é um polinômio, pois sabemos que $\lim\limits_{0<x\to 0} e^{-\frac{1}{x}} x^k = \lim\limits_{y\to\infty} \frac{y^{-k}}{e^y} = 0$ para qualquer $k \in \mathbb{Z}$. (Note que (2.8) define uma função contínua.) Para $n = 0$, tomemos $p_0(x) := 1$. Finalmente,

$$\tfrac{\mathrm{d}}{\mathrm{d}x}\big(e^{-\frac{1}{x}} x^{-2n} p_n(x)\big) = e^{-\frac{1}{x}} \tfrac{1}{x^2} x^{-2n} p_n(x) - 2n e^{-\frac{1}{x}} x^{-2n-1} p_n(x) + e^{-\frac{1}{x}} x^{-2n} p_n'(x) =$$

$$= e^{-\frac{1}{x}} x^{-2(n+1)} \big(p_n(x) - 2nx p_n(x) + x^2 p_n'(x)\big),$$

ou seja, $p_{n+1}(x) := p_n(x) - 2nxp_n(x) + x^2 p_n'(x)$ é um polinômio ■

2.9. Corolário. *Sejam* $0 < r_0 < r_1$ *e* $p \in \mathbb{R}^n$. *Então existe uma função* $g : \mathbb{R}^n \to [0,1]$ *de classe* C^∞, $g \in C^\infty(\mathbb{R}^n)$, *tal que* $g(x) = 0$ *para* $|x-p| \geq r_1$ *e* $g(x) = 1$ *para* $|x-p| \leq r_0$.

Demonstração. Fazendo $g_0(x) := f\big(|x-p|^2 - r_0^2\big)$ e $g_1(x) := f\big(r_1^2 - |x-p|^2\big)$, onde f é a função do Lema 2.7, vemos que $g_0(x) > 0$ para $|x-p| > r_0$, que $g_0(x) = 0$ para $|x-p| \leq r_0$, que $g_1(x) > 0$ para $|x-p| < r_1$ e que $g_1(x) = 0$ para $|x-p| \geq r_1$. Logo, $g_0(x) + g_1(x) > 0$ para todo $x \in \mathbb{R}^n$ e $g(x) := \frac{g_0(x)}{g_0(x)+g_1(x)}$ é a função desejada, pois $g \in C^\infty(\mathbb{R}^n)$ pelo Lema 2.2 ■

2.10. Vetores tangentes, derivadas e campos vetoriais. A derivada $\dot c(t_0)$ de $c : (a,b) \to V$ (onde V é um espaço $\mathbb{R}$-linear de dimensão finita e $t_0 \in (a,b)$), introduzida na Definição 1.8, pode ser interpretada como um vetor tangente ao caminho c no ponto $p := c(t_0)$, pelo menos intuitivamente. Para chegar a uma definição adequada e rigorosa de vetor tangente, é bem útil entender o que este vetor "faz" com funções definidas numa vizinhança aberta de p.

Seja $V \mathrel{\circ\!\!\supset} U \xrightarrow{f} \mathbb{R}$, $p \in U$, uma função de classe C^∞. A composta $f \circ c$, definida numa vizinhança aberta de $t_0 \in (a,b)$, é uma função de classe C^1, supondo que c é de classe C^1. Assim, podemos fazer a derivada

$$\dot c(t_0) f := \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=t_0} (f \circ c) = \lim_{t_0 \neq t \to t_0} \frac{f\big(c(t)\big) - f\big(c(t_0)\big)}{t - t_0}.$$

Deste modo, o vetor tangente $\dot c(t_0)$ associa a cada função f de classe C^∞, definida numa vizinhança aberta de $p = c(t_0)$, um número $\dot c(t_0) f \in \mathbb{R}$. Claro que este número depende apenas do comportamento de f na proximidade de p. Em outras palavras, se diminuirmos a vizinhança U para W, $p \in W \mathrel{\subset\!\!\circ} U$, temos $\dot c(t_0)(f|_W) = \dot c(t_0) f$.

Sejam f_1, f_2 funções de classe C^∞, definidas em vizinhanças abertas $p \in U_1, U_2 \mathrel{\subset\!\!\circ} V$, e sejam $r_1, r_2 \in \mathbb{R}$. As funções $r_1 f_1 + r_2 f_2$ e $f_1 f_2$ são definidas numa vizinhança aberta de p (por exemplo, em $U_1 \cap U_2$) e são de classe C^∞. Portanto, podemos calcular os números $\dot c(t_0)(r_1 f_1 + r_2 f_2)$ e $\dot c(t_0)(f_1 f_2)$. Lembrando como derivar funções de uma variável real, obtemos $\dot c(t_0)(r_1 f_1 + r_2 f_2) = r_1\big(\dot c(t_0) f_1\big) + r_2\big(\dot c(t_0) f_2\big)$ (o vetor tangente $\dot c(t_0)$ é linear) e $\dot c(t_0)(f_1 f_2) = f_1(p)\big(\dot c(t_0) f_2\big) + f_2(p)\big(\dot c(t_0) f_1\big)$ (o vetor tangente $\dot c(t_0)$ satisfaz a regra de Leibniz).

Chegamos à seguinte

2.10.1. Definição. Seja $p \in U \mathrel{\subset\!\!\circ} V$, onde V é um espaço $\mathbb{R}$-linear de dimensão finita. Um *vetor tangente* v a U em p é uma regra que associa um número $vf \in \mathbb{R}$ a cada função f de classe C^∞ definida numa vizinhança aberta de p (ou seja, $U \circ\!\!\supset W \xrightarrow{f} \mathbb{R}$ com $p \in W$) de modo que, para quaisquer funções f_1, f_2 de classe C^∞ definidas em vizinhanças abertas de p, valem as seguintes propriedades:

- $v(r_1 f_1 + r_2 f_2) = r_1(vf_1) + r_2(vf_2)$ para todos $r_1, r_2 \in \mathbb{R}$ (vf é $\mathbb{R}$-linear em f);
- $v(f_1 f_2) = f_1(p)(vf_2) + f_2(p)(vf_1)$ (vf satisfaz a regra de Leibniz);
- vf depende apenas do comportamento de f numa (pequena) vizinhança de p; mais formalmente, isto significa que $v(f|_{W'}) = vf$ se $p \in W' \subset\!\!\circ\, W$.

Grosso modo, um vetor tangente a U em p é simplesmente uma derivação local de funções definidas na proximidade de p.

Denotamos por T_pU o conjunto de todos os vetores tangentes a U em p. De fato, T_pU é um espaço $\mathbb{R}$-linear, pois podemos definir $(rv)f := r(vf)$ e $(v_1 + v_2)f := v_1 f + v_2 f$ para quaisquer $v, v_1, v_2 \in \mathrm{T}_pU$, $r \in \mathbb{R}$ e função f de classe C^∞ definida em uma vizinhança aberta de p. (A verificação do fato que $rv, v_1 + v_2 \in \mathrm{T}_pU$ é imediata.) Chamamos T_pU *espaço tangente* a U em p.

2.10.2. Exemplo. Seja $p \in U \subset\!\!\circ\, V$, onde V é um espaço $\mathbb{R}$-linear de dimensão finita e seja $v \in V$. Para algum $\varepsilon > 0$ suficientemente pequeno, temos um caminho linear $(-\varepsilon, \varepsilon) \xrightarrow{c_{p,v}} U$ definido pela regra $c_{p,v}(t) := p + tv$. Seja $U \circ\!\!\supset W \xrightarrow{f} \mathbb{R}$ uma função de classe C^∞ definida numa vizinhança aberta de p, $p \in W$. A *derivada direcional* $v_p f$ de f em p (na direção de v) é nada mais do que

$$v_p f := \dot{c}_{p,v}(0) f = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=t_0} f(p + tv) = \lim_{0 \neq t \to 0} \frac{f(p + tv) - f(p)}{t}.$$

Considerando uma base linear "padrão" $b_1, \dots, b_n \in \mathbb{R}^n$ e as correspondentes coordenadas $(x_1, \dots, x_n)$, é fácil verificar que $(b_i)_p = \frac{\partial}{\partial x_i}\big|_p$ para todos $i = 1, \dots, n$.

Na verdade, todos os vetores tangentes são derivadas direcionais:

2.10.3. Proposição. *Seja $p \in U \subset\!\!\circ\, V$, onde V é um espaço $\mathbb{R}$-linear de dimensão finita. Então a aplicação $i_p : V \to \mathrm{T}_pU$ dada pela regra $i_p : v \mapsto v_p$ é um isomorfismo de espaços $\mathbb{R}$-lineares.*

Demonstração. O fato que i_p é linear segue diretamente do Lema 2.2. Com efeito, lembrando como relacionam-se a derivada considerada como um vetor e a derivada interpretada como uma aplicação linear, obtemos $v_p f = \big((D_p f) \circ (D_0 c_{p,v})\big)1$ pelo Lema 2.2. Claro que $D_0 c_{p,v} : r \mapsto rv$. Logo, $v_p f = (D_p f)v$ é linear em v.

Seja $l : V \to \mathbb{R}$ uma função linear. É imediato que $v_p l = lv$. Em particular, o núcleo de i_p é nulo e, assim, a aplicação i_p é injetora.

Seja $w \in \mathrm{T}_pU$. Para todas as funções f_1, f_2 de classe C^∞ definidas em vizinhanças abertas de p, de $f_1(p) = f_2(p) = 0$ deduzimos $w(f_1 f_2) = 0$ pela regra de Leibniz. Além disso, $w1 = 0$, pois $w1 = w(1 \cdot 1) = w1 + w1$ pela regra de Leibniz. Isto implica que $wc = 0$ para qualquer função constante c.

Escolhendo uma base linear $b_1, \dots, b_n \in V$, identificamos V com $\mathbb{R}^n$ e denotamos por $(x_1, \dots, x_n)$ as correspondentes coordenadas. Sendo cada coordenada uma função de classe C^∞, definimos $c_i := wx_i$ para todo $i = 1, \dots, n$, $v := \sum\limits_{j=1}^n c_j b_j$ e $w' := w - v_p$. Agora, $w' x_i = (w - v_p)x_i = c_i - x_i\big(\sum\limits_{j=1}^n c_j b_j\big) = c_i - c_i = 0$ para todo i. Resta mostrar que $w' = 0$.

Seja f uma função arbitrária de classe C^∞ definida numa vizinhança aberta de p. Pelo Lema 2.6, diminuindo a vizinhança, obtemos $f(x) = f(p) + \sum\limits_{i=1}^n (x_i - p_i) g_i(x)$ para algumas funções g_i's de classe C^∞ definidas numa vizinhança aberta de p. Aplicando o mesmo lema às funções g_i's, obtemos $f(x) = f(p) + \sum\limits_{i=1}^n (x_i - p_i) g_i(p) + \sum\limits_{i,j=1}^n (x_i - p_i)(x_j - p_j) h_{ij}(x)$. Daí segue $w' f = 0$ ■

2.10.4. Definição. Seja V um espaço $\mathbb{R}$-linear de dimensão finita e seja $U \subset\!\!\circ\, V$ um subconjunto aberto. Um *campo vetorial suave* F sobre U ou um *operador diferencial de ordem* 1 sobre U é uma aplicação $F : C^\infty(U) \to C^\infty(U)$ tal que $F(r_1 f_1 + r_2 f_2) = r_1(Ff_1) + r_2(Ff_2)$ (linearidade de F) e $F(f_1 f_2) = (Ff_1)f_2 + f_1(Ff_2)$ (a regra de Leibniz para F) para todos $f_1, f_2 \in C^\infty(U)$ e $r_1, r_2 \in \mathbb{R}$.

Sejam F, F_1, F_2 campos vetoriais suaves sobre U e seja $g \in C^\infty(U)$. As regras $(gF)f := g(Ff)$, $(F_1+F_2)f := F_1 f + F_2 f$ e $[F_1,F_2]f := F_1(F_2 f) - F_2(F_1 f)$, onde $f \in C^\infty(U)$, definem campos vetoriais suaves gF, F_1+F_2 e $[F_1,F_2]$ sobre U. Com efeito, para F_1+F_2 (*soma* de campos) e gF, a verificação é imediata. Vamos verificar que o *comutador* de campos $[F_1,F_2]$ é um campo. Para quaisquer $f_1, f_2 \in C^\infty(U)$ e $r_1, r_2 \in \mathbb{R}$, temos

$$[F_1,F_2](r_1f_1+r_2f_2) = F_1\big(F_2(r_1f_1+r_2f_2)\big) - F_2\big(F_1(r_1f_1+r_2f_2)\big) =$$

$$= r_1\big(F_1(F_2f_1)\big) + r_2\big(F_1(F_2f_2)\big) - r_1\big(F_2(F_1f_1)\big) - r_2\big(F_2(F_1f_2)\big) = r_1\big([F_1,F_2]f_1\big) + r_2\big([F_1,F_2]f_2\big),$$

$$[F_1,F_2](f_1f_2) = F_1\big(F_2(f_1f_2)\big) - F_2\big(F_1(f_1f_2)\big) =$$

$$= F_1\big((F_2f_1)f_2\big) + F_1\big(f_1(F_2f_2)\big) - F_2\big((F_1f_1)f_2\big) - F_2\big(f_1(F_1f_2)\big) =$$

$$= \big(F_1(F_2f_1)\big)f_2 + (F_2f_1)(F_1f_2) + (F_1f_1)(F_2f_2) + f_1\big(F_1(F_2f_2)\big) -$$

$$-\big(F_2(F_1f_1)\big)f_2 - (F_1f_1)(F_2f_2) - (F_2f_1)(F_1f_2) - f_1\big(F_2(F_1f_2)\big) =$$

$$= \big([F_1,F_2]f_1\big)f_2 + f_1\big([F_1,F_2]f_2\big).$$

2.10.5. Exemplo. Para qualquer subconjunto aberto $U \subset\!\circ\, \mathbb{R}^n$, o operador $\frac{\partial}{\partial x_i}$, $i = 1, \dots, n$, providencia um campo suave sobre U. A gente se acostumou a imaginar este campo como sendo formado pelos vetores unitários, todos na direção do i-ésimo eixo coordenado.

2.10.6. Proposição. *Sejam $U \subset\!\circ\, \mathbb{R}^n$ um subconjunto aberto e F um campo suave sobre U. Então $F = \sum\limits_{i=1}^{n} f_i \frac{\partial}{\partial x_i}$, onde $f_1, \dots, f_n \in C^\infty(U)$. O campo F pode ser também visto como uma família suave $F_p \in \mathrm{T}_pU$, $p \in U$, de vetores tangentes; de fato, $F_p = \sum\limits_{i=1}^{n} f_i(p) \frac{\partial}{\partial x_i}\big|_p$. Para qualquer subconjunto aberto $W \subset\!\circ\, U$, podemos restringir F para W obtendo um campo suave $F|_W$ sobre W.*

Demonstração. Utilizando as funções construídas no Corolário 2.9, observemos que, para toda função $f \in C^\infty(U)$ e qualquer subconjunto aberto $W \subset\!\circ\, U$, de $f|_W = 0$ segue $(Ff)|_W = 0$. Realmente, suponhamos que $f|_W = 0$, mas $(Ff)(p) \neq 0$ para algum ponto $p \in W$. Sem perda de generalidade, podemos supor que W é uma bola aberta centrada em p, ou seja, que $W = B(p,R)$ com $R > 0$. Pelo Corolário 2.9, para $0 < r_0 < r_1 < R$, construímos uma função $g \in C^\infty(\mathbb{R}^n)$ e definimos $h := g|_U$. Obviamente, $fh = 0$. Logo, $0 = F(fh) = (Ff)h + f(Fh)$. Calculando os valores em p, temos $0 = (Ff)(p) \cdot h(p) + f(p) \cdot (Fh)(p) = (Ff)(p)$, pois $h(p) = 1$ e $f(p) = 0$. Uma contradição.

Seja $p \in U$. Definamos o vetor tangente $F_p \in \mathrm{T}_pU$.

Para isto, tomemos uma função arbitrária $\mathbb{R}^n \circ\!\supset W \xrightarrow{f} \mathbb{R}$ de classe C^∞ definida numa vizinhança aberta de p e, mantendo f a mesma numa vizinhança aberta de p, a estendemos a uma função $h \in C^\infty(U)$ novamente utilizando o Corolário 2.9. Para um apropriado $R > 0$, temos $B(p,R) \subset W \cap U$. Como acima, pelo Corolário 2.9, para $0 < r_0 < r_1 < R$, construímos uma função $g \in C^\infty(\mathbb{R}^n)$ e definimos $h(x) := \begin{cases} f(x)g(x) & \text{se } \ x \in B(p,R) \\ 0 & \text{se } \ x \in U \setminus B(p,R) \end{cases}$. A verificação de $h \in C^\infty(U)$ pode ser realizada localmente. Sobre $B(p,R)$ a função h é de classe C^∞, pois $h = fg$ com f e g de classe C^∞. Denotando por $\overline{B}(p,r_1) := \big\{x \in \mathbb{R}^n \mid |x-p| \leq r_1\big\}$ a bola fechada, vemos que a função g é nula no subconjunto aberto $U \setminus \overline{B}(p,r_1)$ e que $U \setminus \overline{B}(p,r_1) \supset U \setminus B(p,R)$. Resta observar que $U = B(p,R) \cup \big(U \setminus \overline{B}(p,r_1)\big)$. A função h coincide com a função f numa vizinhança do ponto p (de fato, na bola aberta $B(p,r_0)$).

A definição $F_pf := (Fh)(p)$ independe da escolha de h, pois, para uma outra escolha $h' \in C^\infty(U)$, a função $h - h'$ é nula numa vizinhança aberta de p, implicando $\big(F(h-h')\big)(p) = 0$ pela observação acima. Pela mesma razão, F_pf depende apenas do comportamento de f numa (pequena) vizinhança de p. A linearidade de F e a regra de Leibniz para F implicam facilmente as correspondentes propriedades de F_p.

Pelos Exemplo 2.10.2 e Proposição 2.10.3, $F_p = \sum\limits_{i=1}^{n} f_i(p)\frac{\partial}{\partial x_i}\big|_p$ para funções apropriadas $f_i : U \to \mathbb{R}$, $i = 1, \dots, n$. Seja $f \in C^\infty(U)$. Então, pelas considerações acima, $F_p f = (Ff)(p)$. Aplicando isso para a coordenada $f := x_j|_U \in C^\infty(U)$, obtemos $C^\infty(U) \ni F(x_j|_U) = f_j$ ■

2.10.7. Lema. $\big[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}\big] = 0$.

Demonstração. Provaremos um fato mais forte: Seja $\mathbb{R}^2 \supset\!\circ\, U \xrightarrow{f} \mathbb{R}$ uma função tal que as derivadas parciais $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ e $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ existem em todos os pontos de U, sendo $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ contínua em $p \in U$. Então a derivada $\frac{\partial^2 f}{\partial x_2 \partial x_1}(p)$ existe e $\frac{\partial^2 f}{\partial x_2 \partial x_1}(p) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(p)$.

Seja $p = (p_1, p_2)$. Primeiramente provamos um análogo bidimensional do teorema de Lagrange (o Lema 1.7.7) do valor médio. Suponhamos que, para alguns $r_1, r_2 \in \mathbb{R}$, os pontos $(p_1 + t_1 r_1, p_2 + t_2 r_2)$, onde (t_1, t_2) percorre o quadrado unitário $[0,1]\times[0,1]$, estão todos em U, ou seja, $(p_1+t_1r_1, p_2+t_2r_2) \in U$ para qualquer $(t_1, t_2) \in [0,1] \times [0,1]$. Então

$$f(p_1 + r_1, p_2 + r_2) - f(p_1, p_2 + r_2) - f(p_1 + r_1, p_2) + f(p_1, p_2) = r_1 r_2 \frac{\partial^2 f}{\partial x_1 \partial x_2}(p_1 + t_1 r_1, p_2 + t_2 r_2)$$

para um apropriado $(t_1, t_2) \in (0,1) \times (0,1)$.

Com efeito, a função $g(x_2) := f(p_1 + r_1, x_2) - f(p_1, x_2)$ é contínua e derivável para x_2 variando entre p_2 e $p_2 + r_2$. Pelo Lema 1.7.7, $g(p_2 + r_2) - g(p_2) = r_2 g'(p_2 + t_2 r_2)$ para um certo $t_2 \in (0,1)$. Note que $g'(p_2 + t_2 r_2) = \frac{\partial f}{\partial x_2}(p_1 + r_1, p_2 + t_2 r_2) - \frac{\partial f}{\partial x_2}(p_1, p_2 + t_2 r_2)$. Aplicando novamente o Lema 1.7.7, nessa vez, para a função $h(x_1) := \frac{\partial f}{\partial x_2}(x_1, p_2 + t_2 r_2)$ que é contínua e derivável para x_1 variando entre p_1 e $p_1 + r_1$, obtemos $g'(p_2 + t_2 r_2) = r_1 h'(p_1 + t_1 r_1)$ para um certo $t_1 \in (0,1)$, ou seja, $g(p_2 + r_2) - g(p_2) = r_2 r_1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(p_1 + t_1 r_1, p_2 + t_2 r_2)$. Resta observar que

$$g(p_2 + r_2) - g(p_2) = f(p_1 + r_1, p_2 + r_2) - f(p_1, p_2 + r_2) - f(p_1 + r_1, p_2) + f(p_1, p_2).$$

Precisamos provar que $\lim\limits_{p_2 \neq x_2 \to p_2} \dfrac{\frac{\partial f}{\partial x_1}(p_1, x_2) - \frac{\partial f}{\partial x_1}(p_1, p_2)}{x_2 - p_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(p_1, p_2)$, ou seja, que, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $0 \neq |x_2 - p_2| < \delta$ implica $\left|\dfrac{\frac{\partial f}{\partial x_1}(p_1, x_2) - \frac{\partial f}{\partial x_1}(p_1, p_2)}{x_2 - p_2} - d\right| \leq \varepsilon$, onde $d := \frac{\partial^2 f}{\partial x_1 \partial x_2}(p_1, p_2)$. Fixamos $\varepsilon > 0$.

Sendo $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ contínua em (p_1, p_2) e sendo U aberto, encontramos $\delta > 0$ tal que as desigualdades $|r_1|, |r_2| < \delta$ implicam $(p_1 + t_1 r_1, p_2 + t_2 r_2) \in U$ e $\big|\frac{\partial^2 f}{\partial x_1 \partial x_2}(p_1 + t_1 r_1, p_2 + t_2 r_2) - d\big| < \varepsilon$ para todo $(t_1, t_2) \in (0,1) \times (0,1)$. Pelo análogo bidimensional do teorema do valor médio demonstrado acima, as desigualdades $0 < |r_1|, |r_2| < \delta$ implicam

$$\textbf{(2.10.8)} \qquad \Big|\frac{f(p_1 + r_1, p_2 + r_2) - f(p_1, p_2 + r_2) - f(p_1 + r_1, p_2) + f(p_1, p_2)}{r_1 r_2} - d\Big| < \varepsilon.$$

Passando ao limite em (2.10.8) com $r_1 \to 0$, obtemos

$$\left|\frac{\frac{\partial f}{\partial x_1}(p_1, p_2 + r_2) - \frac{\partial f}{\partial x_1}(p_1, p_2)}{r_2} - d\right| \leq \varepsilon$$

para qualquer $r_2 \in \mathbb{R}$ satisfazendo $0 < |r_2| < \delta$ ■

2.10.9. Campos vetoriais e equações diferenciais ordinárias. Os campos vetorias são ferramentas indispensáveis quando lidamos com equações diferenciais ordinárias. Seja F um campo vetorial sobre $U \subset\!\circ\, V$, onde V é um espaço $\mathbb{R}$-linear de dimensão finita, e seja $p_0 \in U$. Queremos achar um caminho derivável $c : (a,b) \to U$ com $t_0 \in (a,b)$ tal que $c(t_0) = p_0$ e $\dot{c}(t) = F_{c(t)}$ para todo $t \in (a,b)$. (Tal caminho se chama *curva integral* de F.) Em outras palavras, buscamos um caminho que, no momento dado t_0, passa por um ponto dado $p_0 \in U$ e cujos vetores tangentes estão no campo dado F. Este problema pode ser formulado também na língua de equações diferenciais ordinárias (o que não fazemos aqui); o caminho $c(t)$ em questão será uma solução de um sistema apropriado de equações diferenciais ordinárias. Podemos também fazer o campo F depender da variável do tempo t e considerar o problema mais geral: achar um caminho $c : (a,b) \to U$ tal que $c(t_0) = p_0$ e $\dot{c}(t) = F_{c(t)}(t)$ para todo $t \in (a,b)$. Não planejamos estudar essa matéria nas presentes notas de aula. Um aluno/leitor interessado no assunto pode consultar o apéndice das notas.

2.11. Polinômios de Taylor. A fórmula $f(p+h) = f(p) + (D_pf)h + g(h)$ providencia a melhor aproximação de uma função f de classe C^1 por uma constante mais uma função linear. Uma aproximação será esperadamente mais exata se tentamos aproximar f por um polinômio.

2.11.1. Teorema. *Sejam $U \subset\circ \mathbb{R}^n$, $f \in C^{k+1}(U)$ e $[p, p+h] \subset U$, onde $[p, p+h]$ denota o segmento de reta que liga os pontos p e $p+h$. Então*

$$f(p+h) = f(p) + \sum_{s=1}^{k} \frac{1}{s!} \sum_{i_1,\dots,i_s=1}^{n} \frac{\partial^s f}{\partial x_{i_1} \dots \partial x_{i_s}}(p) h_{i_1} \dots h_{i_s} +$$

$$+ \int_0^1 \frac{(1-t)^k}{k!} \sum_{i_1,\dots,i_{k+1}=1}^{n} \frac{\partial^{k+1} f}{\partial x_{i_1} \dots \partial x_{i_{k+1}}}(p+th) h_{i_1} \dots h_{i_{k+1}} \, \mathrm{d}t,$$

onde $h = (h_1, \dots, h_n)$.

Demonstração. Em seguida, usamos repetidamente a fórmula $\frac{\mathrm{d}}{\mathrm{d}t} g(p+th) = \sum\limits_{i=1}^{n} \frac{\partial g}{\partial x_i}(p+th) h_i$, válida para qualquer $g \in C^1(U)$ pelo Lema 2.2. Pelo teorema fundamental do cálculo,

$$f(p+h) - f(p) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} f(p+th) \, \mathrm{d}t = \int_0^1 \frac{\mathrm{d}(t-1)}{\mathrm{d}t} \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p+th) h_i \, \mathrm{d}t.$$

Integrando por partes, obtemos

$$f(p+h) - f(p) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) h_i - \int_0^1 (t-1) \sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial f}{\partial x_i}(p+th) h_i \, \mathrm{d}t =$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(p) h_i + \int_0^1 (1-t) \sum_{i_1,i_2=1}^{n} \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(p+th) h_{i_1} h_{i_2} \, \mathrm{d}t$$

pela fórmula acima. Acabamos de demostrar o teorema no caso $k = 1$.

Utilizando a indução sobre k, resta observar que $\frac{\mathrm{d}}{\mathrm{d}t} \frac{(1-t)^{k+1}}{(k+1)!} = -\frac{(1-t)^k}{k!}$ e integrar por partes o último termo na fórmula do teorema ■

2.11.2. Corolário. *Sejam $f \in C^{k+1}\big(B(p,R)\big)$ e $0 < r < R$. Então existe uma função $g : B(0,r) \to \mathbb{R}$ tal que $\lim\limits_{0 \neq h \to 0} \frac{g(h)}{|h|^k} = 0$ e*

$$f(p+h) = f(p) + \sum_{s=1}^{k} \frac{1}{s!} \sum_{i_1,\dots,i_s=1}^{n} \frac{\partial^s f}{\partial x_{i_1} \dots \partial x_{i_s}}(p) h_{i_1} \dots h_{i_s} + g(h)$$

para todo $h \in B(0,r)$.

Demonstração. Pelo Lema 1.7.10, as derivadas parciais $\dfrac{\partial^{k+1} f}{\partial x_{i_1} \dots \partial x_{i_{k+1}}}$ sendo contínuas são limitadas sobre o compacto $\overline{B}(p,r)$. A função $\frac{(1-t)^k}{k!}$ é limitada sobre $[0,1]$ e $\left|\frac{h_i}{|h|}\right| \leq 1$ para todos i e $0 \neq h \in B(0,r)$. Para concluir que $\lim\limits_{0\neq h\to 0} \frac{g(h)}{|h|^k} = 0$, resta aplicar o item 5 da Proposição 1.7.9 (de fato, demonstramos que a função $\frac{g(h)}{|h|^{k+1}}$ é limitada para $0 \neq h \in B(0,r)$) ∎

2.11.3. Critério. *Seja* $\mathbb{R}^n \supset\!\!\circ\, U \xrightarrow{f} \mathbb{R}$ *uma função derivável em* $p \in U$.
Se p *é um mínimo/máximo local de* f, *então* $D_p f = 0$.
Suponhamos que $f \in C^3(U)$, *que* $D_p f = 0$ *e que a matriz hessiana* $H := \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right]_{ij}$ *de* f *em* p *é positiva/negativa. Então* p *é um mínimo/máximo local de* f.

Demonstração. Aqui, repetimos basicamente os argumentos apresentados no item 2.5. Suponhamos que $D_p f \neq 0$. Então $(D_p f)h > 0$ para algum $h \in \mathbb{R}^n$. Pela Definição 2.1, para pequenos $t \in \mathbb{R}$, temos $f(p+th) = f(p) + t(D_p f)h + g(th)$. De $\lim\limits_{0\neq t\to 0} \frac{|g(th)|}{|th|} = 0$, concluímos que $|g(th)| < |t| \cdot (D_p f)h$ para suficientemente pequenos t's. Logo, para tais t's, o sinal de $t(D_p f)h + g(th)$ é igual ao de t. Isto contradiz a hipótese que p é um mínimo/máximo local de f.

Para a segunda parte do critério, podemos aplicar o Corolário 2.11.2, pois $B(p,R) \subset U$ para algum $R > 0$. Assim, $f(p+h) = f(p) + \frac{1}{2} hHh^t + g(h)$, onde $h := [h_1 \dots h_n] \in B(0,r)$, $0 < r < R$ e $\lim\limits_{0\neq h\to 0} \frac{g(h)}{|h|^2} = 0$. Por simplicidade, suponhamos que H é positiva. A função $h \mapsto hHh^t$ é contínua. Pelo Lema 1.7.10, essa função atinge seu mínimo $m > 0$ (positivo!) no compacto $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \mid |x|^2 = 1\}$, implicando que $hHh^t \geq m|h|^2$ para todo $h \in \mathbb{R}^n$.

Já que $\lim\limits_{0\neq h\to 0} \frac{|g(h)|}{|h|^2} = 0$, existe $\varepsilon > 0$ tal que $|g(h)| < \frac{m}{3}|h|^2$ para todo $h \in B(0,\varepsilon)$. Para tais h's, $\frac{1}{2} hHh^t + g(h) \geq 0$. Além disso, $\frac{1}{2} hHh^t + g(h) = 0$ para $h \in B(0,\varepsilon)$ apenas quando $h = 0$ ∎

2.11.4. Pontos críticos de uma função. Seja $\mathbb{R}^n \supset\!\!\circ\, U \xrightarrow{f} \mathbb{R}$ uma função derivável em todos os pontos de U. Dizemos que $p \in U$ é um *ponto crítico* de f se $D_p f = 0$. Pelo Critério 2.11.3, os pontos críticos de f são candidatos naturais a extremos local de f e, como mostra a segunda parte do critério, são mínimos/máximos locais caso f seja de classe C^3 e a matriz hessiana de f em p seja positiva/negativa. Mas, como no caso da função $f(x_1,x_2) := x_1^2 - x_2^2$ e $p = 0$, pode ocorrer que a matriz hessiana no ponto crítico não é positiva, nem negativa. Isto pode acontecer quando a matriz hessiana é degenerada (isto significa que seu determinante é nulo). Por exemplo, este é o caso quando a função f tem extremo local em todos os pontos de uma curva passando por p. Supondo que a matriz hessiana não é degenerada, o ponto crítico em questão, apesar de ser isolado (isto significa que numa vizinhança aberta deste ponto não há outros pontos críticos), não é necessariamente um ponto de extremo local. O ponto pode apresentar um ponto de sela no gráfico de f, de modo que em algumas direções temos um mínimo local de f, em outras, um máximo local. Usando o teorema de inércia de Sylvester, podemos estudar e classificar tais pontos críticos, mas paramos aqui, no início da teoria de Morse, que trata deste assunto.

Na próxima subseção, precisaremos de alguns fatos técnicos, inclusive da generalização do item 4 da Proposição 1.7.9 para caminhos integráveis.

2.11.5. Proposição. *Seja* $f : [a,c] \to \mathbb{R}^n$ *um caminho integrável* (*isto significa que as componentes* $f_j : [a,c] \to \mathbb{R}$, $j = 1,\dots,n$, *de* f *são funções integráveis*). *Então a função* $|f| : [a,c] \to \mathbb{R}$ *é integrável e* $\left|\int_a^c f(t)\,\mathrm{d}t\right| \leq \int_a^c |f(t)|\,\mathrm{d}t$, *onde* $|\cdot|$ *denota uma norma arbitrária, fixa em* $\mathbb{R}^n$.

Demonstração. A desigualdade do lema é válida para qualquer caminho constante $f(t) := v \in \mathbb{R}^n$, pois $\left|\int_a^c f(t)\,\mathrm{d}t\right| = |(c-a)v|$ e $\int_a^c |f(t)|\,\mathrm{d}t = (c-a)|v|$ neste caso. Se a desigualdade semelhante vale para caminhos integráveis $f|_{[a,b]}:[a,b]\to\mathbb{R}^n$ e $f|_{[b,c]}:[b,c]\to\mathbb{R}^n$, então a mesma vale para f, pois

$$\big|\int_a^c f(t)\,\mathrm{d}t\big| = \big|\int_a^b f(t)\,\mathrm{d}t+\int_b^c f(t)\,\mathrm{d}t\big| \le \big|\int_a^b f(t)\,\mathrm{d}t\big|+\big|\int_b^c f(t)\,\mathrm{d}t\big| \le$$

$$\le \int_a^b \big|f(t)\big|\,\mathrm{d}t+\int_b^c \big|f(t)\big|\,\mathrm{d}t=\int_a^c \big|f(t)\big|\,\mathrm{d}t.$$

Daí concluímos que a desigualdade vale para caminhos cujas componentes são funções escada.

Pela Definição 1.7.8, cada componente f_j de f é o limite uniforme de uma sequência de funções escada $f_{ji}:[a,c]\to\mathbb{R}$, $i\in\mathbb{N}$, $\lim\limits_{i\to\infty} f_{ji}=f_j$. Denotemos por g_i o caminho (integrável) com componentes $f_{1i},\dots,f_{ni}$. Temos a convergência $\lim\limits_{i\to\infty}\int_a^c g_i(t)\,\mathrm{d}t=\int_a^c f(t)\,\mathrm{d}t$ que se verifica por componentes utilizando o último item da Proposição 1.7.9.

Por outro lado, a sequência de funções $|g_i|$, $i\in\mathbb{N}$, converge uniformemente a função $|f|$. Realmente, usando a norma padrão $|\cdot|_0$ em $\mathbb{R}^n$, pelo Lema 1.10.2, encontramos uma constante $a>0$ tal que $|v|\le a\cdot|v|_0$ para todo $v\in\mathbb{R}^n$. Logo, para todo $t\in[a,c]$,

$$\big|g_i(t)-f(t)\big|\le a\big|g_i(t)-f(t)\big|_0=a\sqrt{\sum_{j=1}^n\big(g_{ij}(t)-f_j(t)\big)^2}\le a\sqrt{\sum_{j=1}^n||g_{ij}-f_j||^2}.$$

Já que $\big|g_i(t)\big|-\big|f(t)\big|\le\big|g_i(t)-f(t)\big|$ e $\big|f(t)\big|-\big|g_i(t)\big|\le\big|f(t)-g_i(t)\big|=\big|g_i(t)-f(t)\big|$ pelas propriedades da norma $|\cdot|$, obtemos $\Big|\big|g_i(t)\big|-\big|f(t)\big|\Big|\le\big|g_i(t)-f(t)\big|$. Portanto, $\Big|\big|g_i(t)\big|-\big|f(t)\big|\Big|\le a\sqrt{\sum_{j=1}^n||g_{ij}-f_j||^2}$, onde a parte direita tende a 0 quando $i\to\infty$. Isto implica que $\lim\limits_{i\to\infty}\big|\big||g_i|-|f|\big|\big|=0$.

Sendo funções escada as componentes das g_i's, sabemos que $\big|\int_a^c g_i(t)\,\mathrm{d}t\big|\le\int_a^c\big|g_i(t)\big|\,\mathrm{d}t$ para todo $i\in\mathbb{N}$. Passando ao limite em ambas as partes dessa desigualdade, pelo último item da Proposição 1.7.9, chegamos ao desejado ■

2.11.6. Definição. Sejam V e V' espaços $\mathbb{R}$-lineares de dimensão finita munidos respectivamente de normas $|\cdot|$ e $|\cdot|'$ e seja $L:V\to V'$ uma aplicação linear. Definimos a *norma* de L pela fórmula $|L|:=\sup\limits_{0\ne v\in V}\frac{|Lv|'}{|v|}=\sup\limits_{|v|=1}|Lv|'$. Sendo $\big\{v\in V\mid |v|=1\big\}$ um compacto e sendo L contínua, concluímos pelo Lema 1.7.10 que $|L|<\infty$. Claro que a norma $|L|$ pressupõe (e depende das) escolhas de normas em V e V'.

2.11.7. Corolário (teorema do valor médio)**.** *Sejam V e V' espaços $\mathbb{R}$-lineares de dimensão finita munidos respectivamente de normas $|\cdot|$ e $|\cdot|'$, seja $V\circ\!\!\supset U\xrightarrow{f}V'$ uma aplicação de classe C^1 e seja $[p_1,p_2]\subset U$, onde $[p_1,p_2]$ denota o segmento de reta que liga os pontos p_1,p_2. Então $\big|f(p_2)-f(p_1)\big|'\le\sup\limits_{x\in[p_1,p_2]}|D_xf|\cdot|p_2-p_1|$.*

Demonstração. Pelo teorema fundamental do cálculo, $f(p_2)-f(p_1)=\int_0^1\frac{\mathrm{d}}{\mathrm{d}t}f\big(p_1+t(p_2-p_1)\big)\,\mathrm{d}t$. Pelo Lema 2.2, $\frac{\mathrm{d}}{\mathrm{d}t}f\big(p_1+t(p_2-p_1)\big)=(D_{p_1+t(p_2-p_1)}f)(p_2-p_1)$. Pelas Proposições 2.11.5 e 1.7.9 e pela Definição 2.11.6, $\big|f(p_2)-f(p_1)\big|'\le\int_0^1\big|(D_{p_1+t(p_2-p_1)}f)(p_2-p_1)\big|'\,\mathrm{d}t\le\sup\limits_{x\in[p_1,p_2]}|D_xf|\cdot|p_2-p_1|$ ■

2.12. Teoremas da função inversa e da função implícita. O teorema da função inversa reduz o problema de existência da inversa local de uma aplicação ao problema de existência da inversa de sua derivada num ponto, isto é, a um problema de álgebra linear. Grosso modo, o teorema afirma que uma aplicação possui inversa local na vizinhança de um ponto se sua derivada neste ponto é um isomorfismo linear:

2.12.1. Teorema (da função inversa)**.** *Sejam V,V' espaços $\mathbb{R}$-lineares de dimensão finita, $p\in U\subset\!\!\circ\, V$ e $f:U\to V'$ uma aplicação de classe C^k, $k\ge1$, tais que $D_pf:V\to V'$ é um isomorfismo linear. Então* *existem vizinhanças abertas* $p \in W \subset\circ U$ *e* $f(p) \in W' \subset\circ V'$ *e uma aplicação* $g : W' \to W$ *de classe* C^k *tais que* $f(W) = W'$, $g \circ \big(f|_W\big) = 1_W$ *e* $\big(f|_W\big) \circ g = 1_{W'}$.

O teorema da função implícita serve, por exemplo, para resolver (pelo menos teoricamente) alguns sistemas de equações não-lineares. Novamente, o problema de existência local de soluções se reduz a um semelhante problema de álgebra linear.

Seja $U_1 \times U_2 \xrightarrow{E} V_3$ uma aplicação de classe C^1, onde $U_1 \subset\circ V_1$, $U_2 \subset\circ V_2$ e V_1, V_2, V_3 são espaços $\mathbb{R}$-lineares de dimensão finita. Fixando um ponto $p_1 \in U_1$, obtemos uma aplicação $E(p_1, -) : U_2 \to V_3$ de classe C^1 dada por $U_2 \ni y \mapsto E(p_1, y) \in V_3$. Denotaremos por $D''_{(p_1,p_2)}E : V_2 \to V_3$ a derivada de $E(p_1, -)$ em $p_2 \in U_2$. Podemos pensar em tal derivada como se fosse a parcial, mas por um grupo y de variáveis. Para algum $p_3 \in V_3$, temos $E(p_1, p_2) = p_3$.

Queremos definir implicitamente uma aplicação s, $s(p_1) = p_2$, por meio da "equação" $E\big(x, s(x)\big) = p_3$. Caso a aplicação linear $D''_{(p_1,p_2)}E : V_2 \to V_3$ seja um isomorfismo, a "equação" pode ser localmente resolvida:

2.12.2. Teorema (da função implícita). *Seja* $U_1 \times U_2 \xrightarrow{E} V_3$ *uma aplicação de classe* C^k, $k \geq 1$, *tal que* $E(p_1, p_2) = p_3$ e $D''_{(p_1,p_2)}E : V_2 \to V_3$ *é um isomorfismo linear, onde* $p_1 \in U_1 \subset\circ V_1$, $p_2 \in U_2 \subset\circ V_2$ *e* V_1, V_2, V_3 *são espaços* $\mathbb{R}$*-lineares de dimensão finita. Então, para uma bola aberta* $p_1 \in B \subset\circ U_1$ *(suficientemente pequena), existe uma única aplicação contínua* $s : B \to U_2$ *tal que* $s(p_1) = p_2$ *e* $E\big(x, s(x)\big) = p_3$ *para todo* $x \in B$*. Essa aplicação* s *é de classe* C^k.

Na demonstração deste teorema precisaremos do seguinte lema cuja demonstração será adiada.

2.12.3. Lema. *Se* $C \subset\circ [0, 1]$ *e* $C \subset_f [0, 1]$, *então* $C = \varnothing$ *ou* $C = [0, 1]$.

Demonstração do Teorema 2.12.2. Definamos a aplicação $f : U_1 \times U_2 \to V_1 \times V_3$ pela regra $f(x, y) := \big(x, E(x, y)\big)$. É facil ver que f é de classe C^k. Levando em conta que

$$D_p f = \begin{bmatrix} 1 & 0 \\ D'_{(p_1,p_2)}E & D''_{(p_1,p_2)}E \end{bmatrix}$$

na forma matricial em blocos, podemos portanto aplicar o Teorema 2.12.1 a f e $p := (p_1, p_2)$.

Escrevemos a inversa local $(V_1 \times V_3) \circ\supset W' \xrightarrow{g} W \subset\circ (U_1 \times U_2) \subset\circ (V_1 \times V_2)$ de f na forma $g(x, z) = \big(g_1(x, z), g_2(x, z)\big)$, onde $(p_1, p_2) \in W$, $(p_1, p_3) \in W'$ e g_1, g_2 são de classe C^k. Para uma bola aberta B tal que $p_1 \in B \subset\circ U_1$, temos $(B, p_3) \subset W'$.

A aplicação $s : B \to U_2$ dada por $s(x) := g_2(x, p_3)$ é claramente de classe C^k. De $f(p_1, p_2) = (p_1, p_3)$ segue $g(p_1, p_3) = (p_1, p_2)$. Daí, $s(p_1) = p_2$. A igualdade $\big(f|_W\big) \circ g = 1_{W'}$ significa que $\Big(g_1(x, z), E\big(g_1(x, z), g_2(x, z)\big)\Big) = (x, z)$ para todo $(x, z) \in W'$. Logo, $g_1(x, z) = x$ e $E\big(x, g_2(x, z)\big) = z$. Consequentemente, $E\big(x, s(x)\big) = p_3$ para todo $x \in B$.

Resta mostrar a unicidade de s. Seja $s_0 : B \to U_2$ uma aplicação contínua satisfazendo $s_0(p_1) = p_2$ e $E\big(x, s_0(x)\big) = p_3$ para todo $x \in B$.

Suponhamos que $s(b) = s_0(b)$ para algum $b \in B$. Então s e s_0 coincidem numa vizinhança aberta de b. Com efeito, $\big(b, s_0(b)\big) = \big(b, s(b)\big) = \big(b, g_2(b, p_3)\big) = g(b, p_3) \in W$. Sendo a aplicação $h : B \to B \times U_2$, definida pela regra $h : x \mapsto \big(x, s_0(x)\big)$, contínua pela Proposição 1.6.6, a imagem inversa $h^{-1}(W)$ é aberta em B, $b \in h^{-1}(W) \subset\circ B$. Vamos mostrar que $h^{-1}(W)$ é a vizinhança desejada. Para todo $x \in h^{-1}(W)$, temos $h(x) \in W$ e, portanto, existe $(x', z) \in W'$ tal que $h(x) = g(x', z)$. Mas $h(x) = \big(x, s_0(x)\big)$ e $g(x', z) = \big(x', g_2(x', z)\big)$, implicando $x = x'$ e $s_0(x) = g_2(x, z)$ com $(x, z) \in W'$. De $E\big(x, g_2(x, z)\big) = z$ e $E\big(x, s_0(x)\big) = p_3$, concluímos que $z = p_3$. Agora, $s_0(x) = g_2(x, z)$ significa que $s_0(x) = g_2(x, p_3) = s(x)$. Vemos que $h^{-1}(W)$ serve como a desejada vizinhança de b. Assim, acabamos de mostrar que o conjunto $A := \big\{b \in B \mid s(b) = s_0(b)\big\}$ é aberto, $A \subset\circ B$.

Suponhamos que $s(b) \neq s_0(b)$ para algum $b \in B$. O segmento de reta que liga p_1 e b está inteiramente contido na bola aberta B. Parametrizando este segmento, $c : [0, 1] \to B$, $c(t) := (1 - t)p_1 + tb$, vemos que $p_1 \in c^{-1}(A) \subset\circ [0, 1]$, pois c é contínuo.

Por outro lado, se $r_i \in c^{-1}(A)$, $i \in \mathbb{N}$, é uma sequência convergente em $[0,1]$, então seu limite $r := \lim_{i\to\infty} r_i$ pertence a $c^{-1}(A)$. Realmente, $c(r_i) \in A$ implica $s\big(c(r_i)\big) = s_0\big(c(r_i)\big)$. As aplicações $s \circ c$ e $s_0 \circ c$ são contínuas, implicando $s\big(c(r)\big) = s_0\big(c(r)\big)$ pela Proposição 1.6.5. Concluímos que $r \in c^{-1}(A)$. Pelo Lema 1.4, $c^{-1}(A) \subset_f [0,1]$.

Já que $0 \in c^{-1}(A)$, o subconjunto $c^{-1}(A)$ não é vazio. Pelo Lema 1.12.3, $c^{-1}(A) = [0,1]$. Agora $1 \in c^{-1}(A)$ significa que $s(b) = s_0(b)$, uma contradição ■

2.12.4. Máximos e mínimos condicionados, multiplicadores de Lagrange. Consideremos o seguinte problema. Sejam dadas uma função $\mathbb{R}^m \circ\!\!\supset U \xrightarrow{f} \mathbb{R}$ e uma aplicação $\mathbb{R}^m \circ\!\!\supset U \xrightarrow{g} \mathbb{R}^n$, ambas de classe C^1, e seja $c \in \mathbb{R}^n$. Queremos achar máximos e mínimos de f no *conjunto de nível* S de g definido por $S := \big\{x \in U \mid g(x) = c\big\}$.

Multiplicadores de Lagrange. Seja $p \in S$ um extremo local de f em S. Suponhamos que a matriz jacobiana $D_p g$ tem posto n (isto significa que todas as n linhas da matriz são linearmente independentes). Então existe $\lambda \in \mathbb{R}^n$, chamado *multiplicador de Lagrange,* tal que $D_p f = \lambda(D_p g)$.

Assim, para achar um máximo/mínimo de f sobre um conjunto de nível S *liso* (isto significa que a derivada $D_q g$ de g em cada $q \in S$ tem posto n), em termos de coordenadas, podemos resolver o seguinte sistema de equações em $x_1, \dots, x_m, \lambda_1, \dots, \lambda_n$

$$g_i(x_1,\dots,x_m) = c_i, \quad i = 1,\dots,n, \qquad \frac{\partial f}{\partial x_j}(x_1,\dots,x_m) = \sum_{i=1}^{n} \lambda_i \frac{\partial g_i}{\partial x_j}(x_1,\dots,x_m), \quad j = 1,\dots,m.$$

Cada tal solução produz um ponto $p := (x_1,\dots,x_m) \in S$ que é um candidato a máximo/mínimo de f sobre S. (Sendo S liso, os λ_i's são únicos caso existam.) Se, por exemplo, há finitas soluções, podemos comparar os valores de f nessas e decidir quais dos p's podem constituir um máximo ou mínimo. Caso S seja um compacto, o problema será assim resolvido pelo Lema 1.7.10. Caso contrário, precisamos ainda estudar o compartamento de f no infinito e/ou na fronteira de S.

Demonstração do método de multiplicadores de Lagrange. Da álgebra linear sabemos que $n \le m$ e que existe uma $(n \times n)$-submatriz de $D_p g$ cujo determinante não é nulo. Renomeando as variáveis $x_1,\dots,x_m$ podemos supor que essa submatriz é formada pelas últimas n colunas de $D_p g$. Assim, interpretamos $\mathbb{R}^m$ como $\mathbb{R}^k \times \mathbb{R}^n$, $k := m - n$.

Apliquemos o Teorema 2.12.2 para $p = (p_1, p_2)$, $p_3 := c$ e $E(x,y) := g(x,y)$. O fato que a submatriz M_2 formada pelas últimas n colunas de $D_p g$ tem determinante não-nulo significa que $\det(D''_{(p_1,p_2)} E) \neq 0$. Pela álgebra linear, $D''_{(p_1,p_2)} E$ é um isomorfismo linear. Podemos diminuir U, fazendo-o da forma $U = U_1 \times U_2 \ni (p_1,p_2)$ com $U_1 \subset\!\circ\, \mathbb{R}^k$ e $U_2 \subset\!\circ\, \mathbb{R}^n$. Pelo Teorema 2.12.2, temos uma bola aberta $p_1 \in B \subset\!\circ\, U_1$ e uma aplicação $s : B \to U_2$ de classe C^1 tal que $s(p_1) = p_2$ e $E\big(x, s(x)\big) = p_3$ para todo $x \in B$, ou seja, $g\big(x, s(x)\big) = c$. Portanto, a função $f \circ h : B \to \mathbb{R}$, onde $h(x) := \big(x, s(x)\big)$, tem extremo local para $x = p_1$. Pela primeira parte do Critério 2.11.3, $D_{p_1}(f \circ h) = 0$. Sendo $g \circ h$ uma constante, obtemos $D_{p_1}(g \circ h) = 0$. Pelo Lema 2.2, $(D_p g) \circ (D_{p_1} h) = 0$ e $(D_p f) \circ (D_{p_1} h) = 0$. Na forma matricial em blocos, $D_p f = [v_1\, v_2]$, $D_p g = [M_1\, M_2]$ e $D_{p_1} h = \begin{bmatrix} 1 \\ M \end{bmatrix}$. Logo, $M_1 + M_2 M = 0$ e $v_1 + v_2 M = 0$. Já que a matriz M_2 possui inversa, temos $M = -M_2^{-1} M_1$ e $v_1 = v_2 M_2^{-1} M_1$. Resta definir $\lambda := v_2 M_2^{-1}$, pois $\lambda[M_1\, M_2] = [v_2 M_2^{-1} M_1 \;\, v_2 M_2^{-1} M_2] = [v_1\, v_2]$ ■

Na demonstração do Teorema 1.12.1, precisaremos dos seguintes fatos, curiosos e úteis por si.

2.12.5. Lema (de contração)**.** *Seja V um espaço $\mathbb{R}$-linear munido de uma norma $|\cdot|$ e sejam $p \in V$, $0 < r \in \mathbb{R}$ e $1 > c \in \mathbb{R}$. Denotemos por $\overline{B}(p,r) := \big\{x \in V \mid |x - p| \le r\big\}$ a bola fechada de raio r centrada em p. Suponhamos que uma aplicação $h : \overline{B}(p,r) \to \overline{B}(p,r)$ satisfaça a propriedade $\big|h(x') - h(x)\big| \le c|x' - x|$ para todos $x, x' \in \overline{B}(p,r)$. Então existe um único ponto fixo $q \in \overline{B}(p,r)$ de h, $h(q) = q$.*

Demonstração. Se $|x'-x|<\varepsilon$ para $x,x'\in\overline{B}(p,r)$, então $\big|(h(x')-h(x)\big|<c\varepsilon$. Isto implica que $h\big(B(x,\varepsilon)\cap\overline{B}(p,r)\big)\subset B\big(h(x),c\varepsilon\big)\cap\overline{B}(p,r)$. Assim, h é uma aplicação contínua.

Seja $q_0\in\overline{B}(p,r)$. A sequência definida por indução $q_{i+1}:=h(q_i)$, $i\in\mathbb{N}$, é de Cauchy. Realmente, para todo $k\in\mathbb{N}$, temos

$$|q_0-q_k|\le|q_0-q_1|+|q_1-q_2|+\dots+|q_{k-1}-q_k|\le(1+c+\dots+c^{k-1})|q_0-q_1|\le\tfrac{1}{1-c}|q_0-q_1|.$$

Portanto, para $j\ge i\ge n$, podemos estimar

$$|q_i-q_j|\le c^i|q_0-q_{j-i}|\le\tfrac{1}{1-c}c^i|q_0-q_1|\le\tfrac{c^n}{1-c}|q_0-q_1|.$$

Levando em conta que $\lim\limits_{n\to\infty}\frac{c^n}{1-c}=0$, concluímos que q_i, $i\in\mathbb{N}$, é uma sequência de Cauchy e, portanto, converge. Sendo a bola $\overline{B}(p,r)$ fechada, o limite $q:=\lim\limits_{i\to\infty}q_i$ está em $\overline{B}(p,r)$, $q\in\overline{B}(p,r)$. Sendo h contínua, $h(q)=\lim\limits_{i\to\infty}h(q_i)=\lim\limits_{i\to\infty}q_{i+1}=q$, ou seja, q é um ponto fixo de h.

Se $q'\in\overline{B}(p,r)$ é um outro ponto fixo de h, $h(q')=q'$, então $|q'-q|=\big|h(q')-h(q)\big|\le c|q'-q|$. Uma contradição $\blacksquare$

Vamos interpretar $\mathbb{R}^{n^2}$ como o espaço de todas as $(n\times n)$-matrizes. A função $\det:\mathbb{R}^{n^2}\to\mathbb{R}$ é de classe C^∞, pois se calcula via expressões polinomiais nos coeficientes da matriz. Em particular, det é uma função contínua. Logo, as matrizes *não-degeneradas* (isto é, as matrizes cujo determinante não é nulo) formam um conjunto aberto $\operatorname{GL}_n\mathbb{R}:=\{M\in\mathbb{R}^{n^2}\mid\det M\ne0\}\subset\!\circ\,\mathbb{R}^{n^2}$. Pela álgebra linear, toda matriz $M\in\operatorname{GL}_n\mathbb{R}$ possui inversa $M^{-1}\in\operatorname{GL}_n\mathbb{R}$ e $M^{-1}=\frac{1}{\det M}\operatorname{adj}M$, onde $\operatorname{adj}M$ denota a matriz adjunta a M. No cálculo da matriz adjunta usamos expressões polinomiais nos coeficientes da matriz M. Consequentemente, a aplicação $\operatorname{adj}:\mathbb{R}^{n^2}\to\mathbb{R}^{n^2}$ é de classe C^∞. Chegamos à seguinte

2.12.6. Observação. *A aplicação* $\operatorname{GL}_n\mathbb{R}\xrightarrow{-1}\operatorname{GL}_n\mathbb{R}$ *é de classe* C^∞ $\blacksquare$

2.12.7. Lema. *Sejam* $V\circ\!\supset W\xrightarrow{f}W'\subset\!\circ\,V$ *uma aplicação de classe* C^k, $k\ge1$, *e* $g:W'\to W$ *a aplicação inversa de* f *de classe* C^1, *onde* V,V' *são espaços* $\mathbb{R}$*-lineares de dimensão finita. Entao* g *é de classe* C^k.

Demonstração. Por indução sobre k, podemos supor que g é de classe C^{k-1} com $k\ge2$. Pelo Lema 2.2, $D_{w'}g=(D_{g(w')}f)^{-1}$ para todo $w'\in W'$. O fato que f é de classe C^k, de acordo com a Definição 2.4, significa que a aplicação $Df:W\to\operatorname{Lin}_{\mathbb{R}}(V,V')$, definida pela regra $w\mapsto D_wf$, é de classe C^{k-1}. Sendo g de classe C^{k-1} pela hipótese de indução, concluímos, utilizando a fórmula $D_{w'}g=(D_{g(w')}f)^{-1}$, que a aplicação $w'\mapsto D_{w'}g$ é de classe C^{k-1} pela Observação 2.12.6 $\blacksquare$

Demonstração do Teorema 1.12.1. Compondo f com certas translações, podemos supor sem perda de generalidade que $p=0$ e $f(p)=0$.

Toda aplicação linear coincide com sua derivada em qualquer ponto (em particular, é de classe C^∞), portanto, compondo f com a aplicação linear $(D_pf)^{-1}$, podemos supor sem perda de generalidade que $V=V'$ e que $D_pf=1_V$, pois

$$D_p\big((D_pf)^{-1}\circ f\big)=D_{f(p)}\big((D_pf)^{-1}\big)\circ(D_pf)=(D_pf)^{-1}\circ(D_pf)=1_V$$

pelo Lema 2.2.

A derivada em $p=0$ da aplicação $h_0:U\to V$, definida por $h_0(x):=x-f(x)$, é nula, $D_0h_0=0$. Munimos V de uma norma. Sendo h_0 de classe C^1, os coeficientes da matriz da aplicação linear D_xh_0 são contínuos em $x\in U$. Claro que a norma $|D_xh_0|$ da aplicação linear D_xh_0 é pequena se estes coeficientes são pequenos. Consequentemente, existe $R>0$ tal que $|D_xh_0|<\frac12$ para todo $x\in B(0,2R)\subset U$. Pelo Corolário 2.11.7, $\big|h_0(x')-h_0(x)\big|\le\frac12|x'-x|$ para todos $x,x'\in B(0,2R)$.

Fixemos um $0<r<R$ arbitrário. Mostraremos que, para todo $d\in\overline{B}(0,r)$, existe um único $q\in\overline{B}(0,2r)$ tal que $f(q)=d$. Para isto, apliquemos o Lema 2.12.5 à função $h_d(x):=d+x-f(x)=d+h_0(x)$.

Com efeito, de $d \in \overline{B}(0,r)$ e $x \in \overline{B}(0,2r)$ segue que $\big|h_d(x)\big| \leq |d| + \big|h_0(x)\big| \leq r + \big|h_0(x) - h_0(0)\big| \leq r + \frac{1}{2}|x-0| \leq 2r$, pois $x, 0 \in \overline{B}(0,2r) \subset B(0,2R)$. Assim, $h_d : \overline{B}(0,2r) \to \overline{B}(0,2r)$. Sejam $x, x' \in \overline{B}(0,2r)$. Então $\big|h_d(x') - h_d(x)\big| = \big|h_0(x') - h_0(x)\big| \leq \frac{1}{2}|x'-x|$, pois $\overline{B}(0,2r) \subset B(0,2R)$, verificando deste modo as hipóteses do Lema 2.12.1 para $c := \frac{1}{2}$. Pelo Lema 2.12.5, existe um único $q \in \overline{B}(0,2r)$ tal que $h_d(q) = q$. Resta observar que $h_d(q) = q$ é equivalente a $f(q) = d$.

Acabamos de provar que $f|_{\overline{B}(0,2r)} : \overline{B}(0,2r) \to \overline{B}(0,r)$ é uma bijeção para qualquer $0 < r < R$. É fácil concluir daí que $f|_{B(0,2r)} : B(0,2r) \to B(0,r)$ é uma bijeção para todo $0 < r < R$.

Sendo f de classe C^1, a derivada $D_x f$ é contínua em x. Logo, a função $x \mapsto \det(D_x f)$ é contínua em x. De $D_0 f = 1_V$ segue que $\det(D_0 f) = 1$. Pela continuidade de $\det(D_x f)$, existe $0 < r < R$ tal que $\det(D_x f) \neq 0$ para todo $x \in B(0,2r)$. Isto significa que $D_x f : V \to V$ é um isomorfismo linear para todo $x \in B(0,2r)$.

Denotemos por $g : B(0,r) \to B(0,2r)$ a inversa de $f|_{B(0,2r)} : B(0,2r) \to B(0,r)$. Para quaisquer $x, x' \in B(0,2r)$, temos $|x'-x| = \big|f(x') - f(x) + h_0(x') - h_0(x)\big| \leq \big|f(x') - f(x)\big| + \big|h_0(x') - h_0(x)\big| \leq \big|f(x') - f(x)\big| + \frac{1}{2}|x'-x|$, implicando $|x'-x| \leq 2|y'-y|$, onde $y' := f(x')$ e $y := f(x)$.

Vamos demonstrar que $D_y g = (D_{g(y)} f)^{-1}$ para todo $y \in B(0,r)$. Fixamos $y \in B(0,r)$ e denotamos $x := g(y)$. Precisamos mostrar que

$$\lim_{y \neq y' \to y} \frac{g(y') - g(y) - (D_x f)^{-1}(y'-y)}{|y'-y|} = 0.$$

Fazemos $x' := g(y')$. Sendo f contínua, $x' \to x$ implica $y' \to y$ pela Proposição 1.6.5. Da desigualdade $|x'-x| \leq 2|y'-y|$ demonstrada acima segue que $y' \to y$ implica $x' \to x$. Portanto, podemos trocar no limite y' por $f(x')$. Assim, basta mostrar que $\lim\limits_{x \neq x' \to x} \frac{x'-x-(D_x f)^{-1}\big(f(x')-f(x)\big)}{|x'-x|} \frac{|x'-x|}{|y'-y|} = 0$. Já que $\frac{|x'-x|}{|y'-y|} \leq 2$, reduzimos a tarefa à $\lim\limits_{x \neq x' \to x} (D_x f)^{-1} \frac{(D_x f)(x'-x) - f(x') + f(x)}{|x'-x|} = 0$. Resta lembrar que qualquer aplicação linear é contínua (no nosso caso, a aplicação $(D_x f)^{-1}$) e usar a definição da derivada $D_x f$.

Sendo $D_x f$ contínua em x e sendo $g(y)$ contínua em y (já que g é derivável em todos os pontos $y \in B(0,r)$), a derivada $D_y g = (D_{g(y)} f)^{-1}$ é contínua em y pela Observação 1.12.5. Em outras palavras, g é de classe C^1.

Resta fazer $W := B(0,2r)$ e $W' := B(0,r)$ e usar o Lema 1.12.6 ■

2.12.8. Gradiente e conjuntos de nível. Percebemos que o uso de derivadas foi extremamente útil no estudo de extremos (locais) e, mais geralmente, de pontos críticos de uma função (vide, por exemplo, os Exemplo 2.5, Critério 2.11.3 e itens 2.11.4 e 2.12.4).

Infelizmente, lidando com funções definidas no conjunto de nível $S := \big\{x \in U \mid g(x) = c\big\}$ de uma aplicação $\mathbb{R}^m \circ\!\!\supset U \xrightarrow{g} \mathbb{R}^n$ de classe C^1, não há como fazer derivadas direcionais de uma função $f : S \to \mathbb{R}$, mesmo das bastante suaves, pois, na expressão $v_p f := \lim\limits_{0 \neq t \to 0} \frac{f(p+tv) - f(p)}{t}$, o valor $f(p+tv)$ não faz muito sentido.

Entretanto, se a função f pode ser localmente estendida para uma vizinhança aberta $p \in W \subset\!\!\circ\, U$ do ponto $p \in S$ (ou seja, existe uma função $\hat{f} : W \to \mathbb{R}$ cuja restrição $\hat{f}|_{S \cap W}$ para $S \cap W$ coincide com f, $\hat{f}|_{S \cap W} = f|_{S \cap W}$), este cálculo faz pleno sentido. Mas agora enfrentamos um outro problema: a definição da derivada direcional pode depender de como nós estendemos a função f.

No caso de um conjunto de nível S *liso* no ponto $p \in S$ (isto significa que a derivada $D_p g$ tem posto n), quando discutimos o método de multiplicadores de Lagrange no item 2.12.4, superamos a mencionada dificuldade conseguindo localmente parametrizar S na vizinhaça aberta de $p \in S$ por um subconjunto aberto em $\mathbb{R}^{m-n}$: de fato, nos salvou o teorema da função implícita. Será que poderíamos derivar funções diretamente, sem escolha de uma parametrização?

Intuitivamente, sendo $D_p g$ a melhor aproximação linear de g na proximidade de p, isto é, $g(x) - g(p) = (D_p g)(x-p) +$ um pouco, o espaço tangente a S em p (supostamente a melhor aproximação linear de S perto de p) deveria ser dado pela equação $(D_p g)(x-p) = 0$, pois o conjunto S é dado pela equação

$g(x)-g(p)=0$. Acontece que a descrição heurística do espaço tangente pela equação $(D_pg)(x-p)=0$ é quase adequada (e é adequada em pontos lisos de S).

Por simplicidade, consideremos o caso $n=1$. O fato que D_pg tem posto 1 significa $D_pg\neq 0$, digamos, $\frac{\partial g}{\partial x_m}(p)\neq 0$. Pelo Teorema 2.12.2, obtemos uma função $s:B\to\mathbb{R}$ tal que $g\big(x,s(x)\big)=c$ para todo $x\in B$, onde $p_1\in B\subset\!\circ\,\mathbb{R}^{m-1}$ é uma bola aberta e $p=(p_1,p_2)$. Assim parametrizamos S por B numa vizinhança aberta de p. Em outras palavras, nessa vizinhança, o conjunto de nível S coincide com o gráfico da função s. Intuitivamente, as derivadas direcionais de s parecem ter algo a ver com os vetores tangentes a S em p, isto é, com os vetores tangentes ao gráfico de s em p.

Na presença do produto interno $\langle -,-\rangle$ em $\mathbb{R}^m$, podemos interpretar a derivada D_pg como um vetor $\nabla_pg\in\mathbb{R}^m$, chamado *gradiente* de g em p e dado pela identidade $\langle\nabla_pg,x\rangle=(D_pg)x$ válida para todo $x\in\mathbb{R}^m$. O gradiente ∇_pg indica uma direção na qual a função g cresce mais rápido. Realmente, como foi notado no início da demonstração da Proposição 2.10.3, $v_pg=(D_pg)v$ para todo $v\in\mathbb{R}^m$. Portanto, para qualquer vetor unitário $v\in\mathbb{R}^m$, $|v|=1$, temos $v_pg=(D_pg)v=\langle\nabla_pg,v\rangle=|\nabla_pg|\cos\alpha$, onde α é o ângulo entre v e ∇_pg. Assim, o valor máximo da derivada direcional v_pg com $|v|=1$ se realiza para $v:=\frac{\nabla_pg}{|\nabla_pg|}$ caso $\nabla_pg\neq 0$, e a norma $|\nabla_pg|$ expressa a taxa de crescimento de g na indicada direção. Por outro lado, a equação $(D_pg)(x-p)=0$ do espaço tangente do conjunto de nível S significa nessas circunstâncias que o gradiente ∇_pg é normal (= ortogonal) ao espaco tangente.

2.12.9. Definição. Sejam $S\subset V$ um subconjunto num espaço $\mathbb{R}$-linear V de dimensão finita. Dizemos que uma *função* $S\supset\!\circ\, W\xrightarrow{f}\mathbb{R}$ é *de classe* C^∞ e denotamos $f\in C^\infty(W)$ se localmente f é a restrição de uma função de classe C^∞ definida num subconjunto aberto de V. Isto significa que, para todo $p\in W$, existe uma função $f\in C^\infty(U)$, $U\subset\!\circ\, V$, tal que $p\in S\cap U\subset W$ e $\hat f|_{S\cap U}=f|_{S\cap U}$.

Um *vetor tangente* v a S em $p\in S$ é uma regra que associa um número $vf\in\mathbb{R}$ a cada função f de classe C^∞ definida numa vizinhança aberta de p (ou seja, $S\supset\!\circ\, W\xrightarrow{f}\mathbb{R}$ com $p\in W$) de modo que, para quaisquer funções f_1,f_2 de classe C^∞ definidas em vizinhanças abertas de p, valem as seguintes propriedades:

- $v(r_1f_1+r_2f_2)=r_1(vf_1)+r_2(vf_2)$ para todos $r_1,r_2\in\mathbb{R}$ (vf é $\mathbb{R}$-linear em f);
- $v(f_1f_2)=f_1(p)(vf_2)+f_2(p)(vf_1)$ (vf satisfaz a regra de Leibniz);
- vf depende apenas do comportamento de f numa (pequena) vizinhança de p; mais formalmente, isto significa que $v(f|_{W'})=vf$ se $p\in W'\subset\!\circ\, W$.

Grosso modo, um vetor tangente a S em p é simplesmente uma derivação local de funções definidas na proximidade de p.

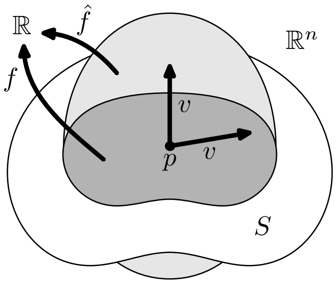
Denotamos por T_pS o conjunto de todos os vetores tangentes a S em p. De fato, T_pS é um espaço $\mathbb{R}$-linear, pois podemos definir $(rv)f:=r(vf)$ e $(v_1+v_2)f:=v_1f+v_2f$ para quaisquer $v,v_1,v_2\in\mathrm{T}_pS$, $r\in\mathbb{R}$ e função f de classe C^∞ definida numa vizinhança aberta de p. (A verificação do fato que $rv,v_1+v_2\in\mathrm{T}_pS$ é imediata.) Chamamos T_pS *espaço tangente* a S em p.

Nas hipóteses do Teorema 2.12.2, o espaço tangente T_pS (dado pela Definição) para o conjunto de nível $S:=\big\{(x,y)\in U_1\times U_2\mid E(x,y)=p_3\big\}$ no ponto $p:=(p_1,p_2)\in S$ é um subespaço $\mathbb{R}$-linear de $\mathrm{T}_p(V_1\times V_2)$ descrito por $\mathrm{T}_pS=\big\{v_p\in\mathrm{T}_p(V_1\times V_2)\mid (D_pE)v=0\big\}$. Portanto, podemos pensar em T_pS como dado pela equação linear $(D_pE)(x-p)=0$ em $V_1\times V_2$ (comprovando assim as considerações heurísticas acima). Com efeito, toda função f definida numa vizinhança aberta suficientemente pequena do ponto p em S tem a forma $f(x,y)=f\big(x,s(x)\big)$, onde $s:B\to U_2$ é a aplicação obtida pelo Teorema 2.12.2. Já que as derivadas em p_1 de funções da forma $f\big(x,s(x)\big)$ definidas em vizinhanças abertas de p_1 são aquelas direcionais pela Proposição 2.10.3, cada vetor tangente a S em p pode ser visto como $(v_1)_{p_1}\in\mathrm{T}_{p_1}B$.

Seja $\hat f(x,y)$ uma função de classe C^∞ definida numa vizinhança aberta de p. Pelo Lema 2.2,

$$(v_1)_{p_1}\hat f\big(x,s(x)\big)=\Big(D_{p_1}\hat f\big(x,s(x)\big)\Big)v_1=(D_p\hat f)\big(v_1,(D_{p_1}s)v_1\big)=(D_p\hat f)(v_1,v_2)=(v_1,v_2)_p\hat f,$$

onde $v_2:=(D_{p_1}s)v_1$. Sendo $E\big(x,s(x)\big)$ uma constante, obtemos $D'_pE+(D''_pE)(D_{p_1}s)=0$ pelo Lema 2.2. Agora vemos que $v_2=(D_{p_1}s)v_1$ é equivalente a $(D_pE)(v_1,v_2)=0$.



No caso geral, um vetor tangente $v_p \in \mathrm{T}_p V$ a V em p é tangente a S, $v_p \in \mathrm{T}_p S$, se e só se a derivada (direcional) $v_p \hat{f}$ de uma extensão $\hat{f}$ de uma função f definida numa vizinhança aberta de p em S não depende da escolha de extensão. Essa propriedade é simplesmente embutida na Definição.

Poderíamos definir vetores tangentes utilizando curvas parametrizadas contidas em S, mas este caminho é menos efetivo.

3. Integração

Para comprar o integral, antes se particiona.
— Elefante na loja de louças

3.1. Integral de Riemann. Chamamos n-*bloco* (*fechado*) qualquer produto cartesiano de n segmentos fechados limitados $B := I_1 \times \cdots \times I_n \subset \mathbb{R}^n$. O *volume* de um n-bloco B é o produto dos comprimentos de seus segmentos, $v(B) := \prod\limits_{i=1}^{n} |I_i|$.

3.1.1. Definição. Seja $I = [a_0, a_k] \subset \mathbb{R}$ um segmento fechado e limitado e sejam $a_0 < a_1 < \cdots < a_k$. Para qualquer $j = 1, \ldots, k$, denotemos $J_j := [a_{j-1}, a_j]$. Então a decomposição $I = \bigcup\limits_{j=1}^{k} J_j$ é dita *partição do segmento* I. Claro que $|I| = \sum\limits_{j=1}^{k} |J_j|$.

Seja $B := I_1 \times \cdots \times I_n \subset \mathbb{R}^n$ um n-bloco e sejam $I_i = \bigcup\limits_{j=1}^{k_i} J_{ij}$, $i = 1, \ldots, n$, partições. Denotemos $B_{j_1 \ldots j_n} := J_{1j_1} \times \cdots \times J_{nj_n}$. Então a decomposição $B = \bigcup\limits_{j_1, \ldots, j_n} B_{j_1 \ldots j_n}$ é uma *partição do n-bloco B*. Também podemos escrever $B = \bigcup\limits_{b \in P} b$, onde a partição P lista todos os $B_{j_1 \ldots j_n}$. De $|I_i| = \sum\limits_{j=1}^{k} |J_{ij}|$, segue $v(B) = \sum\limits_{b \in P} v(b)$.

Dizemos que a partição P' de um n-bloco B é um *refinamento* da partição P de B, e escrevemos $P' \leq P$, se cada n-bloco $b' \in P'$ está contido num n-bloco $b \in P$, $b' \subset b$. Neste caso, cada n-bloco $b \in P$ é particionado nos n-blocos $b' \in P'$ tais que $b' \subset b$.

3.1.2. Observação. *Dadas duas partições P_1 e P_2 de um mesmo n-bloco, existe um refinamento comum P de P_1 e P_2, $P \leq P_1, P_2$* ∎

3.1.3. Definição. Sejam $B \subset \mathbb{R}^n$ um n-bloco, P uma partição de B e $f : B \to \mathbb{R}$ uma função limitada. Para qualquer subconjunto $X \subset B$, denotemos $m_X(f) := \inf\limits_{x \in X} f(x)$ e $M_X(f) := \sup\limits_{x \in X} f(x)$. Nestes termos, façamos $m(f, P) := \sum\limits_{b \in P} m_b(f) \cdot v(b)$ e $M(f, P) := \sum\limits_{b \in P} M_b(f) \cdot v(b)$, as *somas de Riemann, inferior e superior.*

Denotemos por $\underline{\int}_B f := \sup\limits_P m(f, P)$ e $\overline{\int}_B f := \inf\limits_P M(f, P)$ as *integrais inferior* e *superior* de uma função limitada $f : B \to \mathbb{R}$, onde B é um n-bloco e P percorre todas as partições de B.

3.1.4. Lema. *Sejam P' um refinamento da partição P de um n-bloco B e $f : B \to \mathbb{R}$ uma função limitada. Então $m(f, P) \leq m(f, P')$ e $M(f, P') \leq M(f, P)$. Além disso, $m_B(f) \cdot v(B) \leq m(f, P)$ e $M(f, P) \leq M_B(f) \cdot v(B)$.*

Demonstração. A última afirmação segue diretamente da Definição 3.1.3. Por exemplo, $M(f, P) = \sum\limits_{b \in P} M_b(f) \cdot v(b) \leq \sum\limits_{b \in P} M_B(f) \cdot v(b) = M_B(f) \sum\limits_{b \in P} v(b) = M_B(f) \cdot v(B)$.

A primeira afirmação segue da segunda. Realmente, cada n-bloco $b \in P$ é particionado em alguns n-blocos $b' \in P'$. Denotamos tal partição por P'_b, $P'_b := \{b' \in P' \mid b' \subset b\}$. Pela segunda afirmação, $m_b(f) \cdot v(b) \leq m(f, P'_b)$ e $M(f, P'_b) \leq M_b(f) \cdot v(b)$. Logo, $m(f, P) = \sum\limits_{b \in P} m_b(f) \cdot v(b) \leq \sum\limits_{b \in P} m(f, P'_b) = \sum\limits_{b \in P} \sum\limits_{P' \ni b' \subset b} m_{b'}(f) \cdot v(b') = \sum\limits_{b' \in P'} m_{b'}(f) \cdot v(b') = m(f, P')$. De modo semelhante, $M(f, P') \leq M(f, P)$ ∎

3.1.5. Definição. Sejam B um n-bloco e $f : B \to \mathbb{R}$ uma função limitada. Pelo Lema 3.1.4 e pela Observação 3.1.2, $m(f, P_1) \leq M(f, P_2)$ para quaisquer partições P_1 e P_2 de B. Portanto, $\underline{\int}_B f \leq \overline{\int}_B f$.

Dizemos que f é *integrável* se $\int_B f := \underline{\int}_B f = \overline{\int}_B f$; este número é a *integral* de f. Claro que $m(f,P) \leq \int_B f \leq M(f,P)$ se f é integrável e P é uma partição de B.

3.1.6. Proposição. *Sejam B um n-bloco, P uma partição de B, $r \in \mathbb{R}$, $c > 0$ e $f, f_i : B \to \mathbb{R}$, $i \in \mathbb{N}$, funções. Então as seguintes afirmações são válidas.*

1. *Se f_1, f_2 são integráveis e $f_1(x) \leq f_2(x)$ para todo $x \in B$, então $\int_B f_1 \leq \int_B f_2$.*
2. *Se f é integrável, então rf é integrável e $\int_B rf = r\int_B f$.*
3. *Se f é integrável, então $|\int_B f| \leq ||f|| \cdot v(B)$.*
4. *Se f_1, f_2 são integráveis, então $f_1 + f_2$ é integrável e $\int_B (f_1 + f_2) = \int_B f_1 + \int_B f_2$.*
5. *Se f é integrável, então f^+ é integrável, onde $f^+(x) := \max\big(0, f(x)\big)$.*
6. *Se f é integrável, então $|f|$ é integrável e $|\int_B f| \leq \int_B |f|$.*
7. *Se f_1, f_2 são integráveis, então $f_1 f_2$ é integrável.*
8. *Se f é integrável e $\big|f(x)\big| \geq c$ para todo $x \in B$, então $\frac{1}{f}$ é integrável.*
9. *Todas $f|_b$, $b \in P$, são integráveis se e só se f é integrável. Neste caso, $\int_B f = \sum\limits_{b \in P} \int_b f|_b$.*
10. *Se todas f_i's são integráveis e $\lim\limits_{n\to\infty} f_i = f$ é limite uniforme, então f é uma função integrável e $\int_B f = \lim\limits_{i\to\infty} \int_B f_i$. Em outras palavras, limite uniforme comuta com integral.*

Demonstração. 1. Claro que $m_b(f_1) \leq m_b(f_2)$ e $M_b(f_1) \leq M_b(f_2)$ para todo $b \in P$. Logo, $m(f_1,P) \leq m(f_2,P)$ e $M(f_1,P) \leq M(f_2,P)$ para uma partição arbitrária P de B. Daí, $\int_B f_1 \leq \int_B f_2$.
2. Se $r \geq 0$, então $m(rf,P) = r \cdot m(f,P)$ e $M(rf,P) = r \cdot M(f,P)$ implicando que rf é integrável com $\int_B rf = r\int_B f$.

Se $r < 0$, então $m(rf,P) = r \cdot M(f,P)$ e $M(rf,P) = r \cdot m(f,P)$ novamente implicando que rf é integrável com $\int_B rf = r\int_B f$.
3. Tomemos $f_1 := f$ e $f_2 := ||f||$ (constante). Note que $\int_B f_2 = ||f|| \cdot v(B)$. Claro que $f_1(x) \leq f_2(x)$ para todo $x \in B$. Pelo item 1, temos $\int_b f_1 \leq \int_B f_2$, ou seja, $\int_B f \leq ||f|| \cdot v(B)$. Da mesma maneira, tomando $f_1 := -f$, obtemos $-\int_B f \leq ||f|| \cdot v(B)$ pelo item 2. Logo, $|\int_B f| \leq ||f|| \cdot v(B)$.
4. Para todo $b \in B$, temos $m_b(f_1) + m_b(f_2) \leq m_b(f_1 + f_2)$ e $M_b(f_1 + f_2) \leq M_b(f_1) + M_b(f_2)$. Portanto, $m(f_1,P) + m(f_2,P) \leq m(f_1 + f_2, P)$ e $M(f_1 + f_2, P) \leq M(f_1,P) + M(f_2,M)$ para qualquer partição P de B. Sendo f_1, f_2 integráveis, podemos escolher P com as diferenças $M(f_1,P) - m(f_1,P)$ e $M(f_2,P) - m(f_2,P)$ arbitrariamente pequenas. Isto implica o fato desejado.
5. Seja $b \in P$. Se $f(x) \leq 0$ para todo $x \in b$, então $m_b(f^+) = M_b(f^+) = 0$, logo, $M_b(f^+) - m_b(f^+) \leq M_b(f) - m_b(f)$. Se $f(x) \geq 0$ para todo $x \in b$, então $m_b(f^+) = m_b(f)$ e $M_b(f^+) = M_b(f)$, logo, $M_b(f^+) - m_b(f^+) = M_b(f) - m_b(f)$. Se $f(x) < 0 < f(x')$ para alguns $x, x' \in b$, então $m_b(f^+) = 0$, $m_b(f) < 0$ e $M_b(f^+) = M_b(f)$, logo, $M_b(f^+) - m_b(f^+) = M_b(f) \leq M_b(f) - m_b(f)$. Em qualquer caso, $M_b(f^+) - m_b(f^+) \leq M_b(f) - m_b(f)$. Isto implica que $M(f^+,P) - m(f^+,P) \leq M(f,P) - m(f,P)$ para toda partição P de B. Sendo f integrável, escolhendo P, podemos fazer a diferença $M(f,P) - m(f,P)$ arbitrariamente pequena. Portanto, f^+ é integrável.
6. Temos $f = f^+ - f^-$, onde $f^- := (-f)^+$. Pelos itens 2 e 5, f^- é integrável. Pelo item 3, $|f| = f^+ + f^-$ é integrável. Para todo $x \in B$, temos $-f(x) \leq \big|f(x)\big|$ e $f(x) \leq \big|f(x)\big|$. Pelos itens 1 e 2, $-\int_B f \leq \int_B |f|$ e $\int_B f \leq \int_B |f|$, isto é, $|\int_B f| \leq \int_B |f|$.
7. Decompondo $f_1 = f_1^+ - f_1^-$ e $f_2 = f_2^+ - f_2^-$, podemos supor pelos itens anteriores que $f_1(x) \geq 0$ e $f_2(x) \geq 0$ para todo $x \in B$. Sendo f_1, f_2 limitadas, $||f_1||, ||f_2|| \leq M$ para algum $M \in \mathbb{R}$.

Seja $b \in P$. Então $m_b(f_1) \cdot m_b(f_2) \leq m_b(f_1 f_2)$ e $M_b(f_1 f_2) \leq M_b(f_1) \cdot M_b(f_2)$. Portanto,

$$M_b(f_1 f_2) - m_b(f_1 f_2) \leq M_b(f_1) \cdot M_b(f_2) - m_b(f_1) \cdot m_b(f_2) = \big(M_b(f_1) - m_b(f_1)\big) M_b(f_2) +$$

$$+ m_b(f_1)\big(M_b(f_2) - m_b(f_2)\big) \leq M\big(M_b(f_1) - m_b(f_1) + M_b(f_2) - m_b(f_2)\big).$$

Consequentemente, $M(f_1 f_2, P) - m(f_1 f_2, P) \leq M\big(M(f_1,P) - m(f_1,P) + M(f_2,P) - m(f_2,P)\big)$. Assim, o fato que f_1, f_2 são integráveis implica que $f_1 f_2$ é integrável.

8. A função $\frac{1}{f}$ é limitada, pois $\left|\left|\frac{1}{f}\right|\right| \le \frac{1}{c}$. Para todo $b \in P$, temos $m_b\big(\frac{1}{f}\big) = \frac{1}{M_b(f)}$ e $M_b\big(\frac{1}{f}\big) = \frac{1}{m_b(f)}$. Portanto,

$$\Big|M\Big(\frac{1}{f}, P\Big) - m\Big(\frac{1}{f}, P\Big)\Big| \le \sum_{b\in P}\Big|\frac{1}{m_b(f)} - \frac{1}{M_b(f)}\Big| \cdot v(b) = \sum_{b\in P}\frac{M_b(f) - m_b(f)}{\big|m_b(f)\big| \cdot \big|M_b(f)\big|} \cdot v(b) \le$$

$$\le \frac{1}{c^2}\big(M_b(f) - m_b(f)\big) \cdot v(b) = \frac{1}{c^2}\big(M(f,P) - m(f,P)\big).$$

Escolhendo P, fazemos $\frac{1}{c^2}\big(M(f,P) - m(f,P)\big)$ arbitrariamente pequeno. Assim, $\frac{1}{f}$ é integrável.

9. Suponhamos que f é integrável. Seja $b \in P$. Para mostrar que $f|_b$ é integrável, tomemos $\varepsilon > 0$ qualquer. Precisamos encontrar uma partição P_b de b com $M(f|_b, P_b) - m(f_b, P_b) < \varepsilon$. Existe uma partição P' de B tal que $M(f,P') - m(f,P') < \varepsilon$. Podemos achar um refinamento P'' de P', $P'' \le P'$, para o qual b é a união de alguns blocos $b'' \in P''$. Em outras palavras, P'' induz a partição $P_b := \{b'' \in P'' \mid b'' \subset b\}$ de b. Agora,

$$M(f|_b, P_b) - m(f|_b, P_b) = \sum_{b''\in P_b}\big(M_{b''}(f) - m_{b''}(f)\big)v(b'') \le$$

$$\le \sum_{b''\in P''}\big(M_{b''}(f) - m_{b''}(f)\big)v(b'') = M(f,P'') - m(f,P'') \le M(f,P') - m(f,P') < \varepsilon$$

pelo Lema 3.1.4.

Suponhamos que $f|_b$ é integrável para todo $b \in P$. Denotamos por k o número de blocos $b \in P$. Seja $\varepsilon > 0$ qualquer. Precisamos encontrar uma partição P' de B com $M(f,P') - m(f,P') < \varepsilon$. Para cada $b \in P$, existe uma partição P_b de b tal que $M(f|_b, P_b) - m(f|_b, P_b) < \frac{\varepsilon}{k}$. É facil ver que existe uma partição P' de B satisfazendo as seguintes propriedades: todo $b \in P$ é a união de alguns blocos $b' \in P'$ e a partição P'_b de $b \in P$ induzida por P' é mais fina do que P_b, $P'_b \le P_b$. Pelo Lema 3.1.4, $\sum\limits_{b\in P} m(f|_b, P_b) \le \sum\limits_{b\in P} m(f|_b, P'_b) = m(f,P')$ e $M(f,P') = \sum\limits_{b\in P} M(f|_b, P'_b) \le \sum\limits_{b\in P} M(f|_b, P_b)$. Pela Definição 3.1.5, $m(f|_b, P'_b) \le \int_b f|_b \le M(f|_b, P'_b)$ para todo $b \in P$. Logo, $m(f,P') \le \sum\limits_{b\in P}\int_b f|_b \le M(f,P')$ e $M(f,P') - m(f,P') \le \sum\limits_{b\in P}\big(M(f|_b, P_b) - m(f|_b, P_b)\big) < k\frac{\varepsilon}{k} = \varepsilon$. Sendo $\varepsilon > 0$ arbitrário, isto mostra que f é integrável. Mais ainda, já que $m(f,P') \le \int_B f \le M(f,P')$ e $m(f,P') \le \sum\limits_{b\in P}\int_b f|_b \le M(f,P')$ com $M(f,P') - m(f,P') < \varepsilon$, temos $\Big|\int_B f - \sum\limits_{b\in P}\int_b f|_b\Big| < \varepsilon$. Fazendo $\varepsilon > 0$ arbitrariamente pequeno, obtemos $\int_B f = \sum\limits_{b\in P}\int_b f|_b$.

10. De f_i ser limitada e $||f_i - f|| < \infty$ segue que f é limitada.

Antes de nada mais, mostraremos o seguinte fato. Sejam $g_1, g_2 : B \to \mathbb{R}$ funções limitadas e seja $b \in P$. Então $\big|m_b(g_1) - m_b(g_2)\big| \le 4||g_1 - g_2||$ e $\big|M_b(g_1) - M_b(g_2)\big| \le 4||g_1 - g_2||$. Com efeito, basta mostrar a primeira desigualdade, pois a segunda segue da primeira aplicada a $-g_1, -g_2$ no lugar de g_1, g_2. Se $||g_1 - g_2|| = 0$, então $g_1 = g_2$ e a desigualdade é válida. Se $||g_1 - g_2|| > 0$, existem $x_1, x_2 \in b$ tais que

$$g_1(x_1) - m_b(g_1) \le ||g_1 - g_2||, \qquad g_2(x_2) - m_b(g_2) \le ||g_1 - g_2||.$$

Portanto,

$$g_1(x_1) - g_2(x_2) = g_1(x_1) - g_1(x_2) + g_1(x_2) - g_2(x_2) \le g_1(x_1) - m_b(g_1) + ||g_1 - g_2|| \le 2||g_1 - g_2||,$$

$$g_2(x_2) - g_1(x_1) = g_2(x_2) - g_2(x_1) + g_2(x_1) - g_1(x_1) \le g_2(x_2) - m_b(g_2) + ||g_1 - g_2|| \le 2||g_1 - g_2||,$$

pois $m_b(g_1) \le g_1(x_2)$, $g_1(x_2) - g_2(x_2) \le ||g_1 - g_2||$, $m_b(g_2) \le g_2(x_1)$ e $g_2(x_1) - g_1(x_1) \le ||g_1 - g_2||$. Logo, $\big|g_1(x_1) - g_2(x_2)\big| \le 2||g_1 - g_2||$. Consequentemente,

$$\big|m_b(g_1) - m_b(g_2)\big| \le \big|m_b(g_1) - g_1(x_1)\big| + \big|g_1(x_1) - g_2(x_2)\big| + \big|g_2(x_2) - m_b(g_2)\big| \le 4||g_1 - g_2||.$$

Usando o fato que acabamos de mostrar, obtemos

$$\big|m(f_i, P) - m(f, P)\big| \le \sum_{b \in P} \big|m_b(f_i) - m_b(f)\big| \cdot v(b) \le 4||f_i - f|| \cdot v(B),$$

$$\big|M(f, P) - M(f_i, P)\big| \le \sum_{b \in P} \big|M_b(f) - M_b(f_i)\big| \cdot v(b) \le 4||f - f_i|| \cdot v(B),$$

implicando

$$\big|M(f,P) - m(f,P)\big| \le \big|M(f,P) - M(f_i,P)\big| + \big|M(f_i,P) - m(f_i,P)\big| + \big|m(f_i,P) - m(f,P)\big| \le$$

$$\le 8||f_i - f|| \cdot v(B) + \big|M(f_i,P) - m(f_i,P)\big|.$$

Dado $\varepsilon > 0$, pela convergência uniforme $\lim_{n\to\infty} f_i = f$, temos $8||f_i - f|| \cdot v(B) < \varepsilon$ para i suficientemente grande. Para tal i, podemos encontrar uma partição P de B com $\big|M(f_i,P) - m(f_i,P)\big| < \varepsilon$, pois f_i é integrável. Assim deduzimos que f é integrável. Finalmente, $\big|\int_B f_i - \int_B f\big| = \big|\int_B (f_i - f)\big| \le ||f_i - f|| \cdot v(B)$ pelos itens 2, 4 e 3 ∎

3.1.7. Lema. *Qualquer função contínua $f : B \to \mathbb{R}$ é integrável, onde B é um n-bloco.*

Demonstração. Sendo B um compacto, f é uniformemente contínua sobre B pelo Lema 1.7.10. Para qualquer partição P de B, temos $M(f,P) \ge \inf_P M(f,P) = \overline{\int}_B f \ge \underline{\int}_B f := \sup_P m(f,P) \ge m(f,P)$. Assim, para mostrar que $\overline{\int}_B f = \underline{\int}_B f$, basta verificar que, para todo $\varepsilon > 0$, existe uma partição P de B tal que $M(f,P) - m(f,P) \le \varepsilon \cdot v(B)$. Pela continuidade uniforme de f, existe $\delta > 0$ tal que $\big|f(x_1) - f(x_1)\big| < \varepsilon$ se os pontos $x_1, x_2 \in B$ satisfazem $|x_1 - x_2| < \delta$. É fácil escolher uma partição P de B tão fina que $|x_1 - x_2| < \delta$ para todos $x_1, x_2 \in b \in P$. Logo, $M_b(f) - m_b(f) \le \varepsilon$ para todo $b \in P$. Agora, $M(f,P) - m(f,P) = \sum_{b \in P} \big(M_b(f) - m_b(f)\big)v(b) \le \sum_{b \in P} \varepsilon \cdot v(b) = \varepsilon \cdot v(B)$ ∎

3.2. Teorema (Fubini). *Sejam B_1 um n_1-bloco, B_2 um n_2-bloco e $f : B_1 \times B_2 \to \mathbb{R}$ uma função integrável. Para qualquer $x_1 \in B_1$, definamos $\underline{S}(x_1) := \underline{\int} f(x_1, x_2)\,\mathrm{d}x_2$ e $\overline{S}(x_1) := \overline{\int} f(x_1, x_2)\,\mathrm{d}x_2$. Então as funções $\underline{S}, \overline{S} : B_1 \to \mathbb{R}$ são integráveis e $\int_{B_1} \underline{S}(x_1)\,\mathrm{d}x_1 = \int_{B_1} \overline{S}(x_1)\,\mathrm{d}x_1 = \int_{B_1 \times B_2} f$. Em outras palavras,*

$$\int_{B_1 \times B_2} f = \int_{B_1} \Big(\underline{\int}_{B_2} f(x_1, x_2)\,\mathrm{d}x_2\Big)\,\mathrm{d}x_1 = \int_{B_1} \Big(\overline{\int}_{B_2} f(x_1, x_2)\,\mathrm{d}x_2\Big)\,\mathrm{d}x_1 =$$

$$= \int_{B_2} \Big(\underline{\int}_{B_1} f(x_1, x_2)\,\mathrm{d}x_1\Big)\,\mathrm{d}x_2 = \int_{B_2} \Big(\overline{\int}_{B_1} f(x_1, x_2)\,\mathrm{d}x_1\Big)\,\mathrm{d}x_2.$$

Demonstração. Sejam P_1, P_2 partições de B_1, B_2. Então $P_1 \times P_2 := \{b_1 \times b_2 \mid b_1 \in P_1, b_2 \in P_2\}$ é uma partição de $B_1 \times B_2$. (Claro que toda partição de $B_1 \times B_2$ tem essa forma.)

Fixamos temporariamente um $b_1 \in P_1$. Para todos $x_1 \in b_1$ e $b_2 \in P_2$, temos

$$m_{b_1 \times b_2}(f) \le m_{b_2}\big(f(x_1, -)\big), \qquad M_{b_2}\big(f(x_1, -)\big) \le M_{b_1 \times b_2}(f).$$

Portanto,

$$\sum_{b_2 \in P_2} m_{b_1 \times b_2}(f) \cdot v(b_2) \le \sum_{b_2 \in P_2} m_{b_2}\big(f(x_1, -)\big) \cdot v(b_2) = m\big(f(x_1, -), P_2\big) \le \underline{\int} f(x_1, x_2)\,\mathrm{d}x_2 = \underline{S}(x_1),$$

$$\overline{S}(x_1) = \overline{\int} f(x_1, x_2)\,\mathrm{d}x_2 \le M\big(f(x_1, -), P_2\big) \le \sum_{b_2\in P_2} M_{b_2}\big(f(x_1,-)\big)\cdot v(b_2) \le \sum_{b_2\in P_2} M_{b_1\times b_2}(f)\cdot v(b_2),$$

ou seja,

$$\sum_{b_2\in P_2} m_{b_1\times b_2}(f)\cdot v(b_2) \le \underline{S}(x_1), \qquad \overline{S}(x_1) \le \sum_{b_2\in P_2} M_{b_1\times b_2}(f)\cdot v(b_2).$$

Logo,

$$\sum_{b_2\in P_2} m_{b_1\times b_2}(f)\cdot v(b_2) \le m_{b_1}(\underline{S}), \qquad M_{b_1}(\overline{S}) \le \sum_{b_2\in P_2} M_{b_1\times b_2}(f)\cdot v(b_2).$$

Fazendo agora $b_1 \in P_1$ variar, obtemos

$$m(f, P_1\times P_2) = \sum_{b_1\in P_1, b_2\in P_2} m_{b_1\times b_2}(f)\cdot v(b_1\times b_2) = \sum_{b_1\in P_1, b_2\in P_2} m_{b_1\times b_2}(f)\cdot v(b_2)\cdot v(b_1) \le m(\underline{S}, P_1),$$

$$M(\overline{S}, P_1) \le \sum_{b_1\in P_1, b_2\in P_2} M_{b_1\times b_2}(f)\cdot v(b_2)\cdot v(b_1) = \sum_{b_1\in P_1, b_2\in P_2} M_{b_1\times b_2}(f)\cdot v(b_1\times b_2) = M(f, P_1\times P_2).$$

Assim,

$$m(f, P_1\times P_2) \le m(\underline{S}, P_1) \le M(\underline{S}, P_1) \le M(\overline{S}, P_1) \le M(f, P_1\times P_2),$$
$$m(f, P_1\times P_2) \le m(\underline{S}, P_1) \le m(\overline{S}, P_1) \le M(\overline{S}, P_1) \le M(f, P_1\times P_2).$$

Sendo f integrável, com uma escolha da partição $P_1 \times P_2$ de $B_1 \times B_2$, podemos fazer a diferença $M(f, P_1\times P_2) - m(f, P_1\times P_2)$ arbitrariamente pequena.

Daí concluímos que $\underline{S}$ e $\overline{S}$ são integráveis e, mais ainda, que $\int_{B_1} \underline{S}(x_1)\,\mathrm{d}x_1 = \int_{B_1} \overline{S}(x_1)\,\mathrm{d}x_1 = \int_{B_1\times B_2} f$ ■

Pelo Lema 3.1.7, obtemos o

3.2.1. Corolário. *Sejam B_1 um n_1-bloco, B_2 um n_2-bloco e $f : B_1\times B_2 \to \mathbb{R}$ uma função contínua. Então*

$$\int_{B_1\times B_2} f = \int_{B_1}\Big(\int_{B_2} f(x_1,x_2)\,\mathrm{d}x_2\Big)\,\mathrm{d}x_1 = \int_{B_2}\Big(\int_{B_1} f(x_1,x_2)\,\mathrm{d}x_1\Big)\,\mathrm{d}x_2 \ ■$$

3.3. Partição da unidade. Dado um subconjunto $S \subset \mathbb{R}^n$, uma família de subconjuntos abertos $\mathcal{U} := \{U_i \mid i \in I\}$ é dita uma *cobertura aberta* de S se $S \subset \bigcup_{i\in I} U_i$. Escolhendo um subconjunto $I' \subset I$ do conjunto de índices I, obtemos uma *subcobertura* $\mathcal{U}' := \{U_i \mid i \in I'\}$ da cobertura $\mathcal{U}$ se $S \subset \bigcup_{i\in I'} U_i'$. Caso o conjunto de índices seja finito, a *cobertura* se chama *finita.*

Seja $S \subset \mathbb{R}^n$. O *fecho* $\overline{S}$ de S é o menor subconjunto fechado de $\mathbb{R}^n$ que contém S. Sendo fechada a interseção de qualquer família de subconjuntos fechados, é fácil ver que $\overline{S} = \bigcap_{S\subset F\subset_f \mathbb{R}^n} F$.

Seja $\varphi : \mathbb{R}^n \to \mathbb{R}$ uma função. O *suporte* $\operatorname{supp}\varphi$ de φ é o fecho do conjunto $S := \big\{x \in \mathbb{R}^n \mid \varphi(x) \ne 0\big\}$, $\operatorname{supp}\varphi := \overline{S}$.

Precisamos do seguinte fato técnico que deixamos por enquanto sem demonstração.

3.3.1. Lema. *Seja $K \subset \mathbb{R}^n$. Então K é um compacto se e só se de qualquer cobertura aberta de K é possível escolher uma subcobertura finita.*

3.3.2. Proposição (partição da unidade)**.** *Seja $K \subset \mathbb{R}^n$ um compacto e seja $\mathcal{U} := \{U_i \mid i \in I\}$ uma cobertura aberta de K. Então existem funções $\varphi_1, \dots, \varphi_k : \mathbb{R}^n \to [0,1]$ de classe C^∞ tais que $\sum_{j=1}^{k} \varphi_j|_K = 1|_K$ (a função constante 1 sobre K) e, para todo $1 \le j \le k$, existe $i \in I$ tal que $\operatorname{supp}\varphi_j \subset U_i$.*

A coleção de funções descrita na proposição é uma *partição da unidade subordinada* à *cobertura* $\mathcal{U}$ de K.

Demonstração. Por definição, $K \subset \bigcup\limits_{i\in I} U_i$. Logo, para todo $x \in K$, existe $i(x) \in I$ tal que $x \in U_{i(x)}$. Sendo $U_{i(x)}$ aberto em $\mathbb{R}^n$ que contém x, $x \in U_{i(x)} \subset\!\circ\, \mathbb{R}^n$, existe $r_x > 0$ satisfazendo $B(x, 3r_x) \subset U_{i(x)}$. As bolas abertas $B(x, r_x)$, $x \in K$, formam uma cobertura aberta de K, pois cada ponto $x \in K$ está em $\bigcup\limits_{x\in K} B(x, r_x)$ (sendo x o centro da bola $B(x, r_x)$). Pelo Lema 3.3.1, podemos escolher uma subcobertura finita. Assim, $K \subset B(x_1, r_{x_1}) \cup \dots \cup B(x_k, r_{x_k})$ para alguns $x_1, \dots, x_k \in K$.

Pelo Corolário 2.9 aplicado a $r_0 := r_{x_j}$, $r_1 := 2r_{x_j}$ e $p := x_j$, podemos construir uma função $g_j : \mathbb{R}^n \to [0, 1]$ de classe C^∞ tal que $g_j(x) = 1$ se $x \in \overline{B}(x_j, r_{x_j})$ e $g_j(x) = 0$ se $x \notin B(x_j, 2r_{x_j})$. Isto significa que $\operatorname{supp} g_j \subset \overline{B}(x_j, 2r_{x_j}) \subset B(x_j, 3r_{x_j}) \subset U_{i(x_j)}$.

Definamos $\varphi_1 := g_1$ e $\varphi_{j+1} := (1 - g_1) \dots (1 - g_j)g_{j+1}$. Se $g_j(x) = 0$, então $\varphi_j(x) = 0$, implicando que $\operatorname{supp} \varphi_j \subset \operatorname{supp} g_j \subset U_{i(x_j)}$. Sendo as funções g_j's de classe C^∞, as funções φ_j's são de classe C^∞. É imediato que φ_j tem valores em $[0, 1]$. Resta verificar que $\sum\limits_{j=1}^{k} \varphi_j|_K = 1|_K$.

Mostraremos por indução sobre j que $\varphi_1 + \dots + \varphi_j = 1 - (1 - g_1) \dots (1 - g_j)$. Para $j = 1$ isto é óbvio, pois $\varphi_1 = g_1$. Se $\varphi_1 + \dots + \varphi_j = 1 - (1 - g_1) \dots (1 - g_j)$, então

$$\varphi_1 + \dots + \varphi_j + \varphi_{j+1} = 1 - (1 - g_1) \dots (1 - g_j) + \varphi_{j+1} = 1 - (1 - g_1) \dots (1 - g_j) + (1 - g_1) \dots (1 - g_j)g_{j+1} =$$

$$= 1 + (1 - g_1) \dots (1 - g_j)(-1 + g_{j+1}) = 1 - (1 - g_1) \dots (1 - g_j)(1 - g_{j+1}).$$

Consequentemente, $\sum\limits_{j=1}^{k} \varphi_j = 1-(1-g_1) \dots (1-g_k)$. Mas a função $(1-g_1) \dots (1-g_k)$ é nula sobre K, pois cada ponto $x \in K$ pertence a uma bola $B(x_j, r_{x_j})$, $x \in B(x_j, r_{x_j})$, devido à inclusão $K \subset B(x_1, r_{x_1}) \cup \dots \cup B(x_k, r_{x_k})$ e, sendo assim, $g_j(x) = 1$ pela construção de g_j ■

3.4. Mudança de variáveis. Podemos integrar funções sobre regiões que não são necessariamente blocos. Por exemplo, se uma função contínua $f : \mathbb{R}^n \to \mathbb{R}$ tem suporte compacto (por definição, o suporte é fechado; deste modo, pedimos apenas que é limitado), então $\operatorname{supp} f \subset B$ para um certo n-bloco B, e podemos definir $\int_{\mathbb{R}^n} f := \int_B f$. Tal definição independe da escolha do n-bloco que contém o suporte. Realmente, se $\operatorname{supp} f \subset B'$ e $\operatorname{supp} f \subset B''$, onde B', B'' são n-blocos, podemos encontrar um n-bloco B que abrange ambos, $B', B'' \subset B$. Assim, basta verificar que $\int_{B'} f = \int_B f$ caso $B' \subset B$. Para tal verificação, podemos usar, por exemplo, o item 9 da Proposição 3.1.6 (porém, isto é demais para uma afirmação tão óbvia): escolhemos uma partição P de B que induz uma partição P' de B' e observamos que $b \in P \setminus P'$ implica $\int_b f|_b = 0$, pois f é nula em tais blocos devido a $\operatorname{supp} f \subset B'$.

No caso geral, podemos definir a integral de uma função $f : R \to \mathbb{R}$ sobre uma região $R \subset \mathbb{R}^n$ (talvez, de natureza bem complicada) pela fórmula $\int_R f := \int_{\mathbb{R}^n} \chi_R f$, onde

$$\chi_R(x) := \begin{cases} 0 & \text{se} \quad x \notin R \\ 1 & \text{se} \quad x \in R \end{cases}$$

é a *função característica* de R. Claro que, nessa definição, é preciso pedir que a função $\chi_R f : \mathbb{R}^n \to \mathbb{R}$ seja integrável. Além disso, se o suporte $\operatorname{supp}(\chi_R f)$ não for compacto, precisaríamos lidar com integrais impróprias, o que parece impróprio neste curso (apesar de ocorrer com frequência no uso prático de cálculo).

3.4.1. Teorema (mudança de variáveis). *Seja* $\mathbb{R}^n \circ\!\supset U_1 \xrightarrow{t} U_2 \subset\!\circ\, \mathbb{R}^n$ *uma bijeção de classe* C^1 *tal que* $\det(D_x t) \neq 0$ *para todo* $x \in U_1$*. Então*

$$\int_{\mathbb{R}^n} f(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} f\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x$$

para qualquer função contínua $f : \mathbb{R}^n \to \mathbb{R}$ *com suporte compacto contido em* U_2, $\operatorname{supp} f \subset U_2$.

Pelo Teorema 2.12.1 (da função inversa), a aplicação $t^{-1} : U_2 \to U_1$ é de classe C^1 e, em particular, é contínua.

Por enquanto, não é claro se a segunda integral faz sentido.

Um problema, que a função $g(x)$ sob essa integral não é definida sobre todo $\mathbb{R}^n$, não é muito sério. A função $g(x)$ é bem definida e é contínua sobre U_1. Vamos defini-la no subconjunto $U_3 := \mathbb{R}^n \setminus t^{-1}(K)$ como constante nula, onde $K := \operatorname{supp} f$. Note que $K \subset U_2$ implica que $U_1 \cup U_3 = \mathbb{R}^n$ e que $f \circ t$ é nula em $U_1 \cap U_3$, pois $t(x) \notin K$ implica $f\big(t(x)\big) = 0$ para todo $x \in U_1 \cap U_3$. Portanto, a definição de $g(x)$ é correta e $g(x)$ é contínua pela Proposição 1.6.4.

Outro problema é que sabemos integrar apenas funções com suporte compacto. Claro que $\operatorname{supp} g \subset t^{-1}(K)$. Para provar que $t^{-1}(K)$ é um compacto, usamos o Lema 3.3.1. Seja $\mathcal{U} := \{U_i \mid i \in I\}$ uma cobertura aberta de $t^{-1}(K)$, $t^{-1}(K) \subset \bigcup\limits_{i\in I} U_i$. Sem perda de generalidade, podemos supor que $U_i \subset U_1$ para todo $i \in I$. Sendo t^{-1} contínua, concluímos que $t(U_1) \subset\!\!\circ\, U_2$. Aplicando o Lemma 3.3.1 aos compacto K e sua cobertura aberta $K \subset \bigcup\limits_{i\in I} t(U_i)$ podemos escolher uma subcobertura finita, $K \subset t(U_{i_1}) \cup \dots \cup t(U_{i_k})$. Aplicando t^{-1} a essa inclusão, obtemos a inclusão desejada $t^{-1}(K) \subset U_{i_1} \cup \dots \cup U_{i_k}$.

Consideremos alguns casos particulares do teorema.

1. Caso t seja a permutação de duas variáveis ou a aplicação idêntica, o teorema vale devido ao teorema de Fubini (vide o Corolário 3.2.1), pois $\det(D_x t) = \pm 1$ para todo $x \in U_1$.

2. Caso t seja uma translação, $t(x) = x + v$ com $v \in \mathbb{R}^n$, temos $\det(D_x t) = 1$ e $\int_{\mathbb{R}^n} f(x+v)\,\mathrm{d}x = \int_B f(x+v)\,\mathrm{d}x = \int_{B+v} f(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} f(y)\,\mathrm{d}y$, onde $B \subset \mathbb{R}^n$ é um bloco suficientemente grande tal que $\operatorname{supp} f \subset B + v$.

3. Caso $n = 1$. Neste caso, pelo Lema 3.3.1, o compacto $\operatorname{supp} f$ é coberto por um número finito de intervalos abertos contidos em U_2. Deste modo, podemos reduzir o problema ao caso quando $\operatorname{supp} f \subset [a,b] \subset U_2$. Pela hipótese do teorema, $t'(x) \neq 0$ para todo $x \in U_1$. Daí segue que $t : [u,v] \to [a,b]$ é uma função monótona. Caso t seja crescente, temos $t(u) = a$ e $t(v) = b$. Portanto,

$$\int_{\mathbb{R}} f\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x = \int_u^v f\big(t(x)\big) \cdot t'(x)\,\mathrm{d}x = \int_a^b f(y)\,\mathrm{d}y = \int_{\mathbb{R}} f(y)\,\mathrm{d}y$$

pela fórmula da mudança de uma variável. Caso t seja decrescente, temos $t(u) = b$ e $t(v) = a$. Agora,

$$\int_{\mathbb{R}} f\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x = -\int_u^v f\big(t(x)\big) \cdot t'(x)\,\mathrm{d}x = -\int_b^a f(y)\,\mathrm{d}y = \int_{\mathbb{R}} f(y)\,\mathrm{d}y$$

pela fórmula da mudança de uma variável.

4. Caso as componentes de t sejam $t(x_1,\dots,x_n) = (x_1,\dots,x_{m-1},h_m,x_{m+1},\dots,x_n)$, onde a função $h_m : U_1 \to \mathbb{R}$ é de classe C^1, a aplicação t se chama *elementar*. É fácil ver que $\det(D_x t) = \frac{\partial h_m}{\partial x_m}(x)$. Denotando por x' o grupo de todas as variaveis além da variável x_m, pelo teorema de Fubini (vide o Corolário 3.2.1) e pelo caso particular anterior, obtemos

$$\int_{\mathbb{R}^n} f\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} f\big(x', h_m(x', x_m)\big) \cdot \left|\frac{\partial h_m}{\partial x_m}(x', x_m)\right| \mathrm{d}x_m \right) \mathrm{d}x' =$$

$$= \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} f(x', y_m)\,\mathrm{d}y_m \right) \mathrm{d}x' = \int_{\mathbb{R}^n} f(y)\,\mathrm{d}y.$$

3.4.2. Lema. *Seja $\mathbb{R}^n \supset\!\!\circ\, U_1 \xrightarrow{t} \mathbb{R}^n$ uma aplicação de classe C^1 tal que $0 \in U_1$, $t(0) = 0$ e $\det(D_0 t) \neq 0$. Então $t|_{W_1} = p_1 \circ \dots \circ p_n \circ g_n \circ \dots \circ g_1$, onde p_i é a permutação de duas variáveis ou a aplicação idêntica para todo $1 \le i \le n$, $0 \in W_i \subset\!\!\circ\, \mathbb{R}^n$ para todo $1 \le i \le n+1$ e $g_i : W_i \to W_{i+1}$ é uma bijeção elementar de classe C^1 tal que $g_i(0) = 0$ e $\det(D_{w_i} g_i) \neq 0$ para todo $w_i \in W_i$ e qualquer $1 \le i \le n$.*

Demonstração. Denotamos por $\pi_m : \mathbb{R}^n \to \mathbb{R}^m$ a projeção para as primeiras m coordenadas, $0 \le m \le n$. Seja $t_1 := t$. Por indução sobre m, construímos uma aplicação $t_m : U_m \to \mathbb{R}^n$ de classe C^1 tal que $0 \in U_m \subset\!\circ\, \mathbb{R}^n$, $t_m(0) = 0$, $\det(D_0 t_m) \neq 0$ e $\pi_{m-1} \circ t_m = \pi_{m-1}|_{U_m}$. Para $m = 1$, o fato é óbvio, pois $\pi_0 = 0$. Suponhamos agora que o fato é válido para $1 \le m \le n$.

É claro de $\pi_{m-1} \circ t_m = \pi_{m-1}|_{U_m}$ que $t_m(x_1, \dots, x_n) = (x_1, \dots, x_{m-1}, c_m, \dots, c_n)$, onde $c_m, \dots, c_n : U_m \to \mathbb{R}$ são funções de classe C^1. Segue de $\det(D_0 t_m) \neq 0$ que um dos $\frac{\partial c_m}{\partial x_m}(0), \dots, \frac{\partial c_n}{\partial x_m}(0)$ não é nulo. Digamos, $\frac{\partial c_k}{\partial x_m}(0) \neq 0$, $m \le k \le n$. Denotemos por p_m a aplicação que permuta as variáveis x_m e x_k. Então $(p_m \circ t_m)(x_1, \dots, x_n) = (x_1, \dots, x_{m-1}, h_m, \dots)$ e a função $h_m : U_m \to \mathbb{R}$ de classe C^1 satisfaz $h_m(0) = 0$ e $\frac{\partial h_m}{\partial x_m}(0) \neq 0$. A aplicação elementar $g_m : U_m \to \mathbb{R}^n$, dada por suas componentes $x_1, \dots, x_{m-1}, h_m, x_{m+1}, \dots, x_n$, satisfaz $g_m(0) = 0$, $\det(D_0 g_m) = \frac{\partial h_m}{\partial x_m}(0) \neq 0$ e $\pi_m \circ p_m \circ t_m = \pi_m \circ g_m$.

Pelo Teorema 2.12.1 (da função inversa), existem vizinhanças abertas $0 \in V_m \subset\!\circ\, U_m$ e $0 \in U_{m+1} \subset\!\circ\, \mathbb{R}^n$ tais que $g_m|_{V_m} : V_m \to U_{m+1}$ é uma bijeção cuja inversa g_m^{-1} é de classe C^1. Resta definir $t_{m+1} := p_m \circ t_m \circ g_m^{-1} : U_{m+1} \to \mathbb{R}^n$, pois $\pi_m \circ p_m \circ t_m = \pi_m \circ g_m$ implica $\pi_m \circ t_{m+1} = \pi_m \circ p_m \circ t_m \circ g_m^{-1} = \pi_m \circ g_m \circ g_m^{-1} = \pi_m|_{U_{m+1}}$.

Note que $t_{n+1} = 1_{U_{n+1}}$ e que $p_m \circ p_m = 1_{\mathbb{R}^n}$ para todo $1 \le m \le n$. Daí, $p_m \circ t_{m+1} \circ g_m|_{V_m} = t_m|_{V_m}$. Fazendo $W_{n+1} := U_{n+1}$ e $W_m := g_m^{-1}(W_{m+1}) \subset\!\circ\, V_m \subset\!\circ\, U_m$ para cada $1 \le m \le n$, obtemos a decomposição $t|_{W_1} = p_1 \circ \dots \circ p_n \circ g_n \circ \dots \circ g_1$ do lema por indução sobre m ∎

3.4.3. Lema. *Sejam $U_1 \xrightarrow{t_1} U_2 \xrightarrow{t_2} U_3$ bijeções de classe C^1 tais que $\det(D_x t_1) \neq 0$ e $\det(D_y t_2) \neq 0$ para todos $x \in U_1$ e $y \in U_2$, onde $U_1, U_2, U_3 \subset\!\circ\, \mathbb{R}^n$. Suponhamos que*

$$\int_{\mathbb{R}^n} f_1(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} f_1\big(t_1(x)\big) \cdot \big|\det(D_x t_1)\big|\,\mathrm{d}x, \qquad \int_{\mathbb{R}^n} f_2(z)\,\mathrm{d}z = \int_{\mathbb{R}^n} f_2\big(t_2(y)\big) \cdot \big|\det(D_y t_2)\big|\,\mathrm{d}y$$

para quaisquer funções contínuas $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ com suportes compactos contidos respectivamente em U_2 e em U_3, $\operatorname{supp} f_1 \subset U_2$ e $\operatorname{supp} f_2 \subset U_3$. Então

$$\int_{\mathbb{R}^n} f(z)\,\mathrm{d}z = \int_{\mathbb{R}^n} f\big((t_2 \circ t_1)(x)\big) \cdot \Big|\det\big(D_x(t_2 \circ t_1)\big)\Big|\,\mathrm{d}x$$

para qualquer função contínua $f : \mathbb{R}^n \to \mathbb{R}$ com suporte compacto contido em U_3, $\operatorname{supp} f \subset U_3$.

Demonstração. Temos

$$\int_{\mathbb{R}^n} f(z)\,\mathrm{d}z = \int_{\mathbb{R}^n} f\big(t_2(y)\big) \cdot \big|\det(D_y t_2)\big|\,\mathrm{d}y.$$

Pelas observações logo após o Teorema 3.4.1, a função contínua $f_1 : y \mapsto f\big(t_2(y)\big) \cdot \big|\det(D_y t_2)\big|$ tem suporte compacto, $\operatorname{supp} f_1 \subset t_2^{-1}(\operatorname{supp} f) \subset U_2$. Portanto

$$\int_{\mathbb{R}^n} f(z)\,\mathrm{d}z = \int_{\mathbb{R}^n} f_1(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} f_1\big(t_1(x)\big) \cdot \big|\det(D_x t_1)\big|\,\mathrm{d}x =$$

$$= \int_{\mathbb{R}^n} f\Big(t_2\big(t_1(x)\big)\Big) \cdot \big|\det(D_{t_1(x)} t_2)\big| \cdot \big|\det(D_x t_1)\big|\,\mathrm{d}x.$$

Resta usar o Lema 2.2 (regra da cadeia) e propriedades do determinante:

$$\Big|\det\big(D_x(t_2 \circ t_1)\big)\Big| = \Big|\det\big((D_{t_1(x)} t_2) \circ (D_x t_1)\big)\Big| =$$

$$= \big|\det(D_{t_1(x)} t_2) \cdot \det(D_x t_1)\big| = \big|\det(D_{t_1(x)} t_2)\big| \cdot \big|\det(D_x t_1)\big| \;\blacksquare$$

Demonstração do Teorema 3.4.1. Seja $p \in U_1$. Pelo Lema 3.4.2, existe uma vizinhança aberta $p \in W_p \subset\!\circ\, U_1$ tal que $W_p \xrightarrow{t|_{W_p}} t(W_p)$ é a composta de (duas) translações, de permutações de variáveis e de aplicações elementares. Pela consideração de casos particulares e pelo Lema 3.4.3, o teorema vale para $t|_{W_p}$. Temos uma cobertura aberta $U_2 = \bigcup_{p\in U_1} t(W_p)$ de U_2 e, portanto, do compacto $K := \operatorname{supp} f$. Pela Proposição 3.3.2, obtemos uma partição da unidade $\varphi_j : \mathbb{R}^n \to [0,1]$, $1 \le j \le k$, subordinada a essa cobertura. Isto significa que $\sum_{j=1}^{k} \varphi_j|_K = 1|_K$ e que, para todo $1 \le j \le k$, φ_j é de classe C^∞ e existe $p_j \in U_1$ tal que $\operatorname{supp} \varphi_j \subset t(W_{p_j})$.

Denotemos $f_j := \varphi_j \cdot f$. Claro que $\sum_{j=1}^{k} f_j = f$ e que $\operatorname{supp} f_j \subset \operatorname{supp} \varphi_j \cap \operatorname{supp} f$. Logo, f_j tem suporte compacto e $\operatorname{supp} f_j \subset t(W_{p_j})$. Já que o teorema vale para f_j e $t|_{W_{p_j}}$, temos

$$\int_{\mathbb{R}^n} f(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} \sum_{j=1}^{k} f_j(y)\,\mathrm{d}y = \sum_{j=1}^{k} \int_{\mathbb{R}^n} f_j(y)\,\mathrm{d}y = \sum_{j=1}^{k} \int_{\mathbb{R}^n} f_j\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x =$$

$$= \int_{\mathbb{R}^n} \sum_{j=1}^{k} f_j\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x = \int_{\mathbb{R}^n} f\big(t(x)\big) \cdot \big|\det(D_x t)\big|\,\mathrm{d}x$$

pelo item 4 da Proposição 3.1.6 ∎

3.4.4. Coordenadas polares, cilíndricas e esféricas. Para $U_1' := (0,\infty) \times [a, a+2\pi)$ e $U_2 := \mathbb{R}^2 \setminus \{0\}$, as fórmulas $x := r\cos\vartheta$ e $y := r\operatorname{sen}\vartheta$ definem coordenadas *polares* em $\mathbb{R}^2$, ou seja, uma bijeção $t : U_1' \to U_2$. A matriz jacobiana $\begin{bmatrix} \cos\vartheta & -r\operatorname{sen}\vartheta \\ \operatorname{sen}\vartheta & r\cos\vartheta \end{bmatrix}$ de t tem determinante $\det(D_{(r,\vartheta)}t) = r$.

Para $U_1' := (0,\infty) \times [a, a+2\pi) \times \mathbb{R}$ e $U_2 := \mathbb{R}^3 \setminus (0,0,\mathbb{R})$, as fórmulas $x := r\cos\vartheta$, $y := r\operatorname{sen}\vartheta$ e $z := h$ definem coordenadas *cilíndricas* em $\mathbb{R}^3$, ou seja, uma bijeção $t : U_1' \to U_2$. A matriz jacobiana $\begin{bmatrix} \cos\vartheta & -r\operatorname{sen}\vartheta & 0 \\ \operatorname{sen}\vartheta & r\cos\vartheta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ de t tem determinante $\det(D_{(r,\vartheta,h)}t) = r$.

Para $U_1' := (0,\infty) \times [a, a+2\pi) \times [b, b+\pi)$ e $U_2 := \mathbb{R}^3 \setminus \{0\}$, as fórmulas $x := r\operatorname{sen}\alpha\cos\vartheta$, $y := r\operatorname{sen}\alpha\operatorname{sen}\vartheta$ e $z := r\cos\alpha$ definem coordenadas *esféricas* em $\mathbb{R}^3$, ou seja, uma bijeção $t : U_1' \to U_2$. A matriz jacobiana $\begin{bmatrix} \operatorname{sen}\alpha\cos\vartheta & r\cos\alpha\cos\vartheta & -r\operatorname{sen}\alpha\operatorname{sen}\vartheta \\ \operatorname{sen}\alpha\operatorname{sen}\vartheta & r\cos\alpha\operatorname{sen}\vartheta & r\operatorname{sen}\alpha\cos\vartheta \\ \cos\alpha & -r\operatorname{sen}\alpha & 0 \end{bmatrix}$ de t tem determinante $\det(D_{(r,\alpha,\vartheta)}t) = r^2\operatorname{sen}\alpha$.

Alguns exemplos de uso dessas coordenadas. Seja R a região em $\mathbb{R}^2$ dada por $R := \big\{(x,y) \in \mathbb{R}^2 \mid 0 \le x, y,\ a^2 \le x^2+y^2 \le b^2\big\}$, onde $0 \le a < b$. Então

$$\textstyle\int_R \ln(x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_{R'} r\ln(r^2)\,\mathrm{d}\vartheta\,\mathrm{d}r = \int_a^b \big(\int_0^{\frac{\pi}{2}} r\ln(r^2)\,\mathrm{d}\vartheta\big)\,\mathrm{d}r =$$

$$\textstyle= \frac{\pi}{2}\int_a^b r\ln(r^2)\,\mathrm{d}r = \frac{\pi}{4}(a^2 - 2a^2\ln a - b^2 + 2b^2\ln b)$$

pelo teorema de Fubini, onde $R' := \big\{(r,\vartheta) \in \mathbb{R}^2 \mid a \le r \le b,\ 0 \le \vartheta \le \frac{\pi}{2}\big\}$.

O volume da bola $\overline{B}(0,R)$ é igual

$$\textstyle\int_{\overline{B}(0,R)} 1 = \int_0^\pi \Big(\int_0^{2\pi} \big(\int_0^R r^2\operatorname{sen}\alpha\,\mathrm{d}r\big)\,\mathrm{d}\vartheta\Big)\,\mathrm{d}\alpha = \frac{4\pi}{3}R^3.$$

Às vezes precisamos lembrar os métodos de integração de funções de uma variável que aprendemos no cálculo I. Fazendo a substituição padrão $r := 2^{\frac{3}{2}}\frac{t}{1-t^2}$, obtemos

$$\textstyle\int_{\overline{B}(0,1)} \frac{\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}{\sqrt{x^2+y^2+z^2+2}} = \int_0^\pi \Big(\int_0^{2\pi} \big(\int_0^1 \frac{r^2}{\sqrt{r^2+2}}\operatorname{sen}\alpha\,\mathrm{d}r\big)\,\mathrm{d}\vartheta\Big)\,\mathrm{d}\alpha = 4\pi\int_0^1 \frac{r^2}{\sqrt{r^2+2}}\,\mathrm{d}r =$$

$$\textstyle= 4\pi\int_0^1 \frac{r^2}{\sqrt{r^2+2}}\,\mathrm{d}r = 4\pi\int_0^{\sqrt{3}-\sqrt{2}} \big(\frac{2}{(1-t)^3} + \frac{2}{(1+t)^3} - \frac{1}{(1-t)^2} - \frac{1}{(1+t)^2} - \frac{1}{1-t} - \frac{1}{1+t}\big)\,\mathrm{d}t = \ldots$$

4. Variedades e formas

O mal dos mortais são variedades;
em nenhum lugar problemas desta mesma envergadura se encontram.
— *Ésquilo, pai da tragédia*

... superfícies foram inventadas pelo diabo.
— *Wolfgang Pauli*

4.1. Variedades. Grosso modo, uma variedade $S \subset \mathbb{R}^n$ de dimensão m é um subconjunto que pode ser localmente parametrizado por conjuntos abertos de $\mathbb{R}^m$ ou, se falar sobre as variedades com bordo, por abertos no semiespaço em $\mathbb{R}^m$ limitado por um hiperplano em $\mathbb{R}^m$. Entretanto, a existência de uma parametrização local é apenas uma ferramenta: a própria variedade "nem quer saber" que pode ser parametrizada.

4.1.1. Definição. Um subconjunto $S \subset \mathbb{R}^n$ é uma *variedade com bordo* de dimensão m se existe uma cobertura aberta $S \subset \bigcup\limits_{i\in I} W_i$, $W_i \subset\!\circ\, \mathbb{R}^n$, e aplicações $\mathbb{R}^m \circ\!\supset U_i \xrightarrow{\hat\varphi_i} \mathbb{R}^n$ de classe C^∞ tais que a restrição $H^m \cap U_i \xrightarrow{\varphi_i} S \cap W_i$ é uma bijeção para todo $i \in I$ e o posto da derivada $D_{u_i}\hat\varphi_i$ é o máximo possível, $\operatorname{rk}(D_{u_i}\hat\varphi_i) = m$, para todos $i \in I$ e $u_i \in U_i$, onde H^m denota o *semiespaço* em $\mathbb{R}^m$ dado por $H^m := \big\{(x_1,\dots,x_m) \in \mathbb{R}^m \mid x_1 \geq 0\big\}$. Chamamos $\hat\varphi_i$ ou φ_i uma *parametrização* de $S \cap W_i$ ou *parametrização local* de S. Em palavras: uma variedade de dimensão m é um subconjunto em $\mathbb{R}^n$ que possui parametrizações lisas locais por subconjuntos abertos do semiespaço H^m.

Denotemos por $\partial H^m \simeq \mathbb{R}^{m-1}$ o *bordo* de H^m, $\partial H^m := \big\{(x_1,\dots,x_m) \in \mathbb{R}^m \mid x_1 = 0\big\}$, e definamos o *bordo*[6] de S por $\partial S := \bigcup\limits_{i\in I} \varphi_i(\partial H^m \cap U_i)$. Caso $\partial S = \varnothing$, dizemos que S é uma *variedade sem bordo.* Sendo máximo o posto de $D_{u_i}\hat\varphi_i$, $\operatorname{rk}(D_{u_i}\hat\varphi_i) = m$, o posto de $D_{u_i}(\varphi_i|_{\partial H^m\cap U_i})$ também é máximo para todo $u_i \in \partial H^m \cap U_i$, $\operatorname{rk}\big(D_{u_i}(\varphi_i|_{\partial H^m\cap U_i})\big) = m-1$, pois a matriz jacobiana de $\varphi_i|_{\partial H^m\cap U_i}$ em qualquer ponto $u_i \in \partial H^m \cap U_i$ é formada pelas $m-1$ últimas colunas da matriz jacobiana de φ_i em u_i. Daí segue que ∂S é uma variedade sem bordo de dimensão $m-1$, $\partial\partial S = \varnothing$.

Segue imediatamente da definição que qualquer subconjunto aberto $S \cap W$, $W \subset\!\circ\, \mathbb{R}^n$, de uma variedade S de dimensão m é uma variedade de dimensão m.

Seja $S \subset \mathbb{R}^n$ uma variedade de dimensão m. De acordo com a Definição 2.12.9, as funções de classe C^∞ definidas sobre subconjuntos abertos em S são localmente restrições de funções de classe C^∞ definidas sobre conjuntos abertos em $\mathbb{R}^n$. Por outro lado, podemos introduzir funções lisas sobre subconjuntos abertos de uma variedade S usando parametrizações locais. Felizmente, essas classes de funções coincidem:

4.1.2. Lema. *Sejam $S \subset \mathbb{R}^n$ uma variedade de dimensão m, $W \subset\!\circ\, \mathbb{R}^n$ e $f : S\cap W \to \mathbb{R}$ uma função. Escolhemos parametrizações locais $H^m \cap U_i \xrightarrow{\varphi_i} S \cap W_i$ de S, $i \in I$ (como aquelas na Definição 4.1.1). Então $f \in C^\infty(S\cap W)$ se e só se $f|_{S\cap W\cap W_i} \circ \varphi_i \in C^\infty\big(\varphi_i^{-1}(S\cap W\cap W_i)\big)$ para todo $i \in I$. Além disso, a inversa de φ_i é continua para todo $i \in I$.*

Demonstração. Pela Definição 2.12.9, a função $f : S\cap W \to \mathbb{R}$ é de classe C^∞ se e só se, para cada $q \in S \cap W$, existem uma vizinhança aberta $q \in W' \subset\!\circ\, W$ e uma função $\hat f \in C^\infty(W')$ tais que $f|_{S\cap W'} = \hat f|_{S\cap W'}$. Em outras palavras, o fato que f é lisa se verifica localmente. O mesmo se aplica na verificação do fato que $f|_{S\cap W\cap W_i} \circ \varphi_i \in C^\infty\big(\varphi^{-1}(S\cap W\cap W_i)\big)$ para todo $i \in I$.

Assim, considerando a variedade $S \cap W_i$ no lugar de S, podemos supor sem perda de generalidade que a variedade S é toda parametrizada (I tem um único elemento). Em particular, temos uma bijeção $\varphi : H^m \cap U \to S$ com $\varphi(p) = q$ e $p \in H^m \cap U$. Pela álgebra linear, de $\operatorname{rk}(D_p\hat\varphi) = m$ segue que algumas m linhas apropriadas da matriz jacobiana $D_p\hat\varphi$ são linearmente independentes. Isto significa que existe

[6] Tal definição de bordo independe da escolha de parametrizações locais da variedade S, mas não planejamos discutir este fato aqui.

uma projeção linear $\pi : \mathbb{R}^n \to \mathbb{R}^m$ para correspondentes m coordenadas tal que $\det\big(D_p(\pi\circ\hat\varphi)\big) \neq 0$. Diminuindo U, podemos supor pelo Teorema 2.12.1 (da função inversa) que $\pi\circ\hat\varphi : U \to (\pi\circ\hat\varphi)U \subset\!\circ\, \mathbb{R}^m$ é uma bijeção com a inversa $\hat\psi$ de classe C^∞.

Pela Proposição 1.6.4, a verificação que a inversa de φ_i é contínua basta realizar localmente. Seja $H^m\cap U'$ um subconjunto aberto arbitrário em $H^m\cap U$, $U' \subset\!\circ\, U$. Temos que mostrar que $\varphi(H^m\cap U') \subset\!\circ\, S$. Já que $\hat\psi$ é contínua, temos $(\pi\circ\hat\varphi)U' \subset\!\circ\, \mathbb{R}^m$ e, portanto, $\pi^{-1}\big((\pi\circ\hat\varphi)U'\big) \subset\!\circ\, \mathbb{R}^n$ e $S\cap\pi^{-1}\big((\pi\circ\hat\varphi)U'\big) \subset\!\circ\, S$. Logo,

$$S\cap\pi^{-1}\big((\pi\circ\hat\varphi)U'\big) = \Big\{\varphi(x) \mid x\in H^m\cap U,\ \pi\big(\varphi(x)\big)\in(\pi\circ\hat\varphi)U'\Big\} =$$

$$= \big\{\varphi(x) \mid x\in H^m\cap U,\ (\hat\psi\circ\pi\circ\varphi)(x)\in U'\big\} = \varphi(H^m\cap U').$$

Agora, se $f = \hat f|_{S\cap W}$ para alguma função $\hat f\in C^\infty(W)$, então $\hat f\circ\hat\varphi\in C^\infty\big(\hat\varphi^{-1}(W)\big)$, pois $\hat\varphi$ é de classe C^∞. Resta observar que $f\circ\varphi : \varphi^{-1}(S\cap W)\to\mathbb{R}$ é a restrição de $\hat f\circ\hat\varphi$ para $\varphi^{-1}(S\cap W)$.

Reciprocamente, se $f : S\cap W\to\mathbb{R}$ e $f\circ\varphi\in C^\infty\big(\varphi^{-1}(S\cap W)\big)$, então, tomando $S\cap W$ no lugar de S e $\hat\varphi^{-1}(W)$ no lugar de U, podemos supor sem perda de generalidade que $S\subset\hat\varphi(U)\subset W\subset\!\circ\,\mathbb{R}^n$ e $f\circ\varphi\in C^\infty\big(H^m\cap U\big)$. Diminuindo novamente $U\ni p$, podemos encontrar uma função $\hat g\in C^\infty(U)$ tal que $f\circ\varphi = \hat g|_{H^m\cap U}$.

Definamos $\hat f := \hat g\circ\hat\psi\circ\pi\in C^\infty\Big(\pi^{-1}\big((\pi\circ\hat\varphi)U\big)\Big)$, onde $S\subset\pi^{-1}\big((\pi\circ\hat\varphi)U\big)\subset\!\circ\,\mathbb{R}^n$. Resta mostrar que $\hat f|_S = f$. Sendo $\hat\varphi|_{H^m\cap U} = \varphi : H^m\cap U\to S$ uma bijeção, basta observar que $(\hat f\circ\hat\varphi)|_{H^m\cap U} = \hat g\circ\hat\psi\circ\pi\circ\hat\varphi|_{H^m\cap U} = \hat g|_{H^m\cap U} = f\circ\varphi$ ■

4.1.3. Definição. Sejam S_1 e S_2 variedades. Uma aplicação contínua $\varphi : S_1\to S_2$ é *lisa* se, para qualquer função $f\in C^\infty(U_2)$ com $U_2\subset\!\circ\, S_2$, temos $f\circ\varphi\in C^\infty\big(\varphi^{-1}(U_2)\big)$. Em outras palavras, uma aplicação entre variedades é lisa se sua composta com funções lisas no codomínio produz funções lisas no domínio.

É imediato que a composta de aplicações lisas entre variedades é uma aplicação lisa, que a aplicação de inclusão $i : U\hookrightarrow S$ de um subconjunto aberto $U\subset\!\circ\, S$ de uma variedade S é lisa e que as funcões lisas definidas em subconjuntos abertos de uma variedade são aplicações lisas (lembre que um subconjunto aberto de uma variedade é uma variedade). Pela Definição 2.12.9, se $S\subset\mathbb{R}^n$ é uma variedade, então a aplicação de inclusão $S\hookrightarrow\mathbb{R}^n$ é lisa. Pelo Lema 4.1.2, qualquer parametrização local $\varphi_i : H^m\cap U_i\to S\cap W_i$ de uma variedade S de dimensão m (como aquelas na Definição 4.1.1) e sua inversa são aplicações lisas entre variedades.

4.1.4. Observação. *Sejam $\varphi_j : H^m\cap U_j\to S\cap W_j$, $j = 1,2$, parametrizações locais de uma variedade S de dimensão m. Então a aplicação bijetora $\varphi_2^{-1}\circ\varphi_1 : H^m\cap U_1'\to H^m\cap U_2'$ é lisa* (*assim como a sua inversa*), *onde $H^m\cap U_j' := \varphi_j^{-1}(S\cap W_1\cap W_2)$, $j = 1,2$.*

Demonstração. Basta ver que $\varphi_j : H^m\cap U_j'\to S\cap W_1\cap W_2$, $j = 1,2$, são parametrizações locais de S ■

Sejam S uma variedade e $p\in S$. Lembramos (vide a Definição 2.12.9) que os vetores tangentes a S em $p\in S$ são derivações locais de funções de classe C^∞ definidas em vizinhanças abertas de p em S e que, de acordo com as regras $(rv)f := r(vf)$ e $(v_1+v_2)f := v_1f + v_2f$, todos os vetores tangentes a S em p formam um espaço $\mathbb{R}$-linear T_pS chamado espaço tangente a S em p.

4.1.5. Definição. Sejam S_1 e S_2 variedades, $\varphi : S_1\to S_2$ uma aplicação lisa e v um vetor tangente a S_1 em $p\in S_1$. Definamos o vetor φv tangente a S_2 em $\varphi(p)$, chamado *φ-imagem* de v, usando a regra $(\varphi v)f := v(f\circ\varphi)$, onde $f : U_2\to\mathbb{R}$ é uma função lisa arbitrária definida numa vizinhança aberta de $\varphi(p)$ em S_2, $\varphi(p)\in U_2\subset\!\circ\, S_2$.

Essa definição faz pleno sentido. Com efeito, $p\in\varphi^{-1}(U_2)\subset\!\circ\, S_1$ e $f\circ\varphi\in C^\infty\big(\varphi^{-1}(U_2)\big)$ pela Definição 4.1.3. Assim, o número $v(f\circ\varphi)\in\mathbb{R}$ é bem definido. Restringindo f para uma vizinhança menor de $\varphi(p)$, $\varphi(p)\in U_2'\subset\!\circ\, U_2$, vemos que $f|_{U_2'}\circ\varphi = (f\circ\varphi)|_{\varphi^{-1}(U_2')}$. Portanto, $v(f\circ\varphi) = v(f|_{U_2'}\circ\varphi)$.

Para quaisquer funções f_1, f_2 de classe C^∞ definidas em vizinhanças abertas de $\varphi(p)$ em S_2 e todos $r_1, r_2 \in \mathbb{R}$, temos $(r_1 f_1 + r_2 f_2) \circ \varphi = r_1(f_1 \circ \varphi) + r_2(f_2 \circ \varphi)$ e $(f_1 \cdot f_2) \circ \varphi = (f_1 \circ \varphi) \cdot (f_2 \circ \varphi)$. Consequentemente,

$$(\varphi v)(r_1 f_1 + r_2 f_2) = v\big((r_1 f_1 + r_2 f_2) \circ \varphi\big) = v\big(r_1(f_1 \circ \varphi) + r_2(f_2 \circ \varphi)\big) =$$

$$= r_1 v(f_1 \circ \varphi) + r_2 v(f_2 \circ \varphi) = r_1(\varphi v) f_1 + r_2(\varphi v) f_2,$$

$$(\varphi v)(f_1 \cdot f_2) = v\big((f_1 \cdot f_2) \circ \varphi\big) = v\big((f_1 \circ \varphi) \cdot (f_2 \circ \varphi)\big) =$$

$$= (f_1 \circ \varphi)(p) \cdot v(f_2 \circ \varphi) + (f_2 \circ \varphi)(p) \cdot v(f_1 \circ \varphi) = f_1\big(\varphi(p)\big) \cdot (\varphi v) f_2 + f_2\big(\varphi(p)\big) \cdot (\varphi v) f_1,$$

pois v é um vetor tangente. Assim, para φv, temos verificada a Definição 2.12.9.

Pela Definição 2.12.9 de soma e de multiplicação por escalar de vetores tangentes de $\mathrm{T}_p S$, a fórmula $v(f \circ \varphi)$ é linear em v. Assim, obtemos uma aplicação linear $\varphi : \mathrm{T}_p S_1 \to \mathrm{T}_{\varphi(p)} S_2$ *induzida* por φ e denotada pelo mesmo símbolo φ.

4.1.6. Observação. *Sejam* $S_1 \xrightarrow{\varphi_1} S_2 \xrightarrow{\varphi_2} S_3$ *aplicações lisas entre variedades e* $p \in S_1$. *Então a composta* $\varphi_2 \circ \varphi_1$ *induz a aplicação linear* $\varphi_2 \circ \varphi_1 : \mathrm{T}_p S_1 \to T_{(\varphi_2 \circ \varphi_1)p} S_3$ *que coincide com a composta de aplicações lineares* $\varphi_1 : \mathrm{T}_p S_1 \to T_{\varphi_1(p)} S_2$ *e* $\varphi_2 : \mathrm{T}_{\varphi_1(p)} S_2 \to T_{(\varphi_2(\varphi_1(p)))} S_3$.

Demonstração. Pela Definição 4.1.5, $\big(\varphi_2(\varphi_1 v)\big) f = (\varphi_1 v)(f \circ \varphi_2) = v(f \circ \varphi_2 \circ \varphi_1)$ ■

4.1.7. Observação. *Sejam* $i : V \hookrightarrow S$ *a aplicação lisa de inclusão de um subconjunto aberto* $V \subset\!\circ\, S$ *numa variedade* S *e* $p \in V$. *Então* $i : \mathrm{T}_p V \to \mathrm{T}_p S$ *é um isomorfismo.*

Demonstração. A afirmação é uma consequência imediata do fato que os vetores tangentes são derivações `locais` ■

4.1.8. Observação. *Sejam* $i : H^m \hookrightarrow \mathbb{R}^m$ *a aplicação lisa de inclusão do semiespaço* $H^m \subset \mathbb{R}^m$ *e* $p \in H^m$. *Então* $i : \mathrm{T}_p H^m \to \mathrm{T}_p \mathbb{R}^m$ *é um isomorfismo.*

Demonstração. Seja $v \in \mathrm{T}_p H^m$. Se $iv = 0$, então para qualquer função lisa $f : U \to \mathbb{R}$ definida numa vizinhança aberta U de p em $\mathbb{R}^m$, $p \in U \subset\!\circ\, \mathbb{R}^m$, temos $vf|_{H^m \cap U} = 0$. Mas toda função lisa definida numa vizinhança aberta de p em H^m numa certa vizinhança aberta (menor) tem a forma $f|_{H^m \cap U}$. Isto implica $v = 0$.

Por outro lado, temos derivadas direcionais $v_p \in \mathrm{T}_p H^m$ bem definidas para os vetores $v \in \mathbb{R}^m$ cuja primeira coordenada é não-negativa. É fácil ver que $iv_p = v_p$. Em outras palavras, $i\mathrm{T}_p H^m$ contém todos os vetores de $\mathrm{T}_p \mathbb{R}^m$ a menos de sinal. Mas $-\mathrm{T}_p H^m = \mathrm{T}_p H^m$, pois $\mathrm{T}_p H^m$ é um espaço $\mathbb{R}$-linear ■

4.1.9. Corolário. *Seja* S *uma variedade de dimensão* m. *Então* $\dim_{\mathbb{R}} \mathrm{T}_p S = m$ *para todo* $p \in S$.

Demonstração. Pelo Observação 4.1.7, podemos supor que a variedade S é parametrizada, ou seja, que temos uma bijeção lisa $\varphi : H^m \cap U \to S$ cuja inversa ψ também é lisa. Pela Observação 4.1.6, ambas as compostas das aplicações lineares $\psi : \mathrm{T}_p S \to T_{\psi(p)}(H^m \cap U)$ e $\varphi : T_{\psi(p)}(H^m \cap U) \to \mathrm{T}_p S$ são idênticas. Resta observar que $\dim_{\mathbb{R}} T_{\psi(p)}(H^m \cap U) = m$ pelas Observações 4.1.7 e 4.1.8 ■

4.1.10. Lema. *Seja* $\varphi : H^m \cap U \to S \cap W$ *uma parametrização local de uma variedade* $S \subset \mathbb{R}^n$, *onde* $U \subset\!\circ\, \mathbb{R}^m$ *e* $W \subset\!\circ\, \mathbb{R}^n$, *e seja* $q \in H^m \cap U$. *Então, em termos das coordenadas* $(x_1, \dots, x_m)$ *em* $\mathbb{R}^m$ *e* $(y_1, \dots, y_n)$ *em* $\mathbb{R}^n$, *denotando por* $\varphi_1, \dots, \varphi_n$ *as componentes de* φ, *a aplicacação linear* φ : $\mathrm{T}_q(H^m \cap U) \to \mathrm{T}_{\varphi(q)} S$ *é dada por* $\varphi \frac{\partial}{\partial x_j}\big|_q = \sum\limits_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(q) \frac{\partial}{\partial y_i}\big|_{\varphi(q)}$.

Demonstração. Seja $f \in C^\infty(V)$ uma função definida numa vizinhança aberta $\varphi(q) \in V \subset\!\circ\, S \cap W$. Então $\Big(\varphi \frac{\partial}{\partial x_j}\big|_q\Big) f = \frac{\partial}{\partial x_j}\big|_q (f \circ \varphi) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial y_i}\big(\varphi(q)\big) \frac{\partial \varphi_i}{\partial x_j}(q) = \Big(\sum\limits_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(q) \frac{\partial}{\partial y_i}\big|_{\varphi(q)}\Big) f$ pelo Lema 2.2 (regra da cadeia) ■

4.1.11. Lema. *Sejam S uma variedade, $p \in V \subset\!\circ\, S$ e $f \in C^\infty(V)$. Então existe uma função $\tilde f \in C^\infty(S)$ com suporte compacto contido em V, $\operatorname{supp} \tilde f \subset V$, que coincide com f numa vizinhança aberta $p \in V' \subset\!\circ\, V$, $\tilde f|_{V'} = f|_{V'}$.*

Demonstração. Temos uma parametrização local $\varphi : H^m \cap U \to S \cap W$ com $p \in S \cap W$, onde $S \subset \mathbb{R}^n$ e $W \subset\!\circ\, \mathbb{R}^n$. Tomando $V \cap W$ no lugar de V, podemos supor sem perda de generalidade que $V \subset S \cap W$.

Basta construir uma função $f' \in C^\infty(V)$ que coincide com f numa vizinhança aberta $p \in V' \subset\!\circ\, V$, $f'|_{V'} = f|_{V'}$, e cujo suporte é compacto e contido em V, $\operatorname{supp} f' \subset V$, pois podemos definir $\tilde f(y) := \begin{cases} f'(y) & \text{se } y \in V \\ 0 & \text{se } y \in V'' \end{cases}$, onde $V'' := S \setminus \operatorname{supp} f \subset\!\circ\, S$ (por definição, $\operatorname{supp} f' \subset_f S$) e $S = V \cup V''$. Realmente, $f'|_{V \cap V''} = 0$ implica que a definição de $\tilde f$ é correta e o fato que $\tilde f \in C^\infty(S)$ se verifica localmente.

Sendo φ e sua inversa contínuas, um subconjunto $K \subset H^m \cap U$ é um compacto se e só se $\varphi(K) \subset S \cap W$ é um compacto pelo Lema 3.3.1. Assim, podemos supor que $V = S = H^m \cap U$.

Para alguns $0 < r_0 < r_1$, temos $B(p, 2r_1) \subset U$ e, devido ao Corolário 2.9, existe uma função lisa $g : \mathbb{R}^m \to [0, 1]$ tal que $\operatorname{supp} g \subset \overline{B}(p, r_1)$ e $g|_{\overline{B}(p,r_0)} = 1$. Façamos $f' := fg|_V$. O subconjunto $H^m \cap \overline{B}(p, r_1) \subset H^m \cap B(p, 2r_1) \subset H^m \cap U = V$ é um compacto em $\mathbb{R}^m$ e $\operatorname{supp} f' \subset H^m \cap \overline{B}(p, r_1)$, pois $\operatorname{supp} g \subset \overline{B}(p, r_1)$. Logo, $\operatorname{supp} f'$ é um compacto. Definindo $V' := H^m \cap B(p, r_0)$, vemos que $g|_{\overline{B}(p,r_0)} = 1$ implica $f'|_{V'} = f|_{V'}$ ■

4.1.12. Definição. Todo espaço $\mathbb{R}$-linear V de dimensão finita admite duas orientações. Cada escolha de uma base linear $b_1, \dots, b_m \in V$ detemina uma orientação. Se $L : V \to V$ é uma transformação linear, a orientação relacionada com a base $Lb_1, \dots, Lb_m$ é a mesma se $\det L > 0$, e é a oposta se $\det L < 0$. Uma *orientação* de uma variedade S é uma escolha de orientações de todos os espaços tangentes $\mathrm{T}_p S$, $p \in S$, que depende de p de maneira contínua, ou seja, que é localmente a mesma. Assim, para orientar uma variedade S de dimensão m, podemos localmente parametrizar S por subconjuntos abertos do semiespaço $\mathbb{H}^m$ munido de uma orientação fixa de modo que quaisquer duas parametrizações locais que envolvem um mesmo ponto $p \in S$ providenciam em $\mathrm{T}_p S$ a mesma orientação. Dadas parametrizações locais de $S \subset \mathbb{R}^n$, isto é, bijeções lisas $\varphi_i : H^m \cap U_i \to S \cap W_i$, $i \in I$, com inversas lisas, onde $U_i \subset\!\circ\, \mathbb{R}^m$ e $W_i \subset\!\circ\, \mathbb{R}^n$ para todo $i \in I$ e $S \subset \bigcup\limits_{i \in I} W_i$, definamos bijeções $\varphi_j^{-1} \circ \varphi_i : H^m \cap U_{ij} \to H^m \cap U_{ji}$, $i, j \in I$, fazendo $H^m \cap U_{ij} := \varphi_i^{-1}(S \cap W_i \cap W_j)$, como na Observação 4.1.4. Pedindo que $\det\big(D_q(\varphi_j^{-1} \circ \varphi_i)\big) > 0$ para todos $q \in U_{ij}$ e $i, j \in I$, obtemos uma orientação induzida em S. (Note que $D_q(\varphi_j^{-1} \circ \varphi_i)$ é de fato a aplicação linear $\varphi_j^{-1} \circ \varphi_i : \mathrm{T}_q(H^m \cap U_{ij}) \to \mathrm{T}_{\varphi_j^{-1}(\varphi_i(q))}(H^m \cap U_{ji})$.)

4.1.13. Definição. Seja S uma variedade. Um *campo* F sobre S é uma aplicação $F : C^\infty(S) \to C^\infty(S)$ tal que $F(r_1 f_1 + r_2 f_2) = r_1(F f_1) + r_2(F f_2)$ e $F(f_1 f_2) = (F f_1) f_2 + f_1(F f_2)$ para todos $f_1, f_2 \in C^\infty(S)$ e $r_1, r_2 \in \mathbb{R}$. Sejam F, F_1, F_2 campos sobre S e seja $g \in C^\infty(S)$. Como antes (vide a Definição 2.10.4), as regras $(gF)f := g(Ff)$, $(F_1 + F_2)f := F_1 f + F_2 f$ e $[F_1, F_2]f := F_1(F_2 f) - F_2(F_1 f)$, onde $f \in C^\infty(S)$, definem campos gF, $F_1 + F_2$ e $[F_1, F_2]$ sobre S.

4.1.14. Observação. *Sejam F_1, F_2 campos sobre uma varidade S e $f \in C^\infty(S)$. Então $[fF_1, F_2] = f[F_1, F_2] - (F_2 f)F_1$.*

Demonstração. Para qualquer $g \in C^\infty(S)$, temos

$$[fF_1, F_2]g = (fF_1)(F_2 g) - F_2\big((fF_1)g\big) = f\big(F_1(F_2 g)\big) - F_2\big((f(F_1 g)\big) =$$

$$= f\big(F_1(F_2 g)\big) - (F_2 f)(F_1 g) - f\big(F_2(F_1 g)\big) = f\big([F_1, F_2]g\big) - \big((F_2 f)F_1\big)g = \big(f[F_1, F_2] - (F_2 f)F_1\big)g \; ■$$

4.1.15. Lema. *Todo campo F sobre uma variedade S pode ser visto como uma família $F_p \in \mathrm{T}_p S$, $p \in S$, de vetores tangentes, isto é, $(Ff)(p) = F_p f$ para todos $f \in C^\infty(S)$ e $p \in S$. Podemos restringir F para qualquer subconjunto aberto $V \subset\!\circ\, S$, obtendo um campo $F|_V$ sobre a variedade V.*

Demonstração. De fato, repetimos aqui a demonstração da Proposição 2.10.6.

Seja $p \in S$. Observemos que $(Ff_1)(p) = (Ff_2)(p)$ se as funções $f_1, f_2 \in C^\infty(S)$ coincidem numa vizinhança aberta $p \in V \subset\circ S$, $f_1|_V = f_2|_V$. Com efeito, pelo Lema 4.1.11, podemos encontrar uma vizinhança aberta $p \in V' \subset\circ V$ e uma função $\tilde{f} \in C^\infty(S)$ tais que $\operatorname{supp} \tilde{f} \subset V$ e $\tilde{f}|_{V'} = 1$. Portanto, $(f_1 - f_2)\tilde{f} = 0$ e $\tilde{f}(Ff_1 - Ff_2) + (f_1 - f_2)F\tilde{f} = 0$. Calculando os valores em p, obtemos $(Ff_1)(p) = (Ff_2)(p)$.

Sejam $p \in V \subset\circ S$ e $f \in C^\infty(V)$. Definamos $F_p f := (F\tilde{f})(p)$, onde $\tilde{f} \in C^\infty(S)$ é uma função obtida pelo Lema 4.1.11. Pela observação acima, $F_p f$ independe da escolha de $\tilde{f}$. Pelo mesmo motivo, $F_p f$ depende apenas do comportamento de f numa (pequena) vizinhança aberta de p. Do modo semelhante, a linearidade de F e a regra de Leibniz para F implicam facilmente as correspondentes propriedades de F_p. Além disso, $(Ff)(p) = F_p f$ para qualquer $f \in C^\infty(S)$.

Seja $V \subset\circ S$. Para $f \in C^\infty(V)$ e qualquer $p \in V$, definamos $(F|_V f)(p) := F_p f$. A verificação que $F|_V f \in C^\infty(V)$ se realiza localmente: pelo Lema 4.1.11, existem uma vizinhança aberta $p \in V' \subset\circ V$ e uma função $\tilde{f} \in C^\infty(S)$ tais que $\tilde{f}|_{V'} = f|_{V'}$. De $F\tilde{f} \in C^\infty(S)$ segue, pela observação acima, que $F|_V f$ é lisa sobre V'. A linearidade e a regra de Leibniz para $F|_V$ se verificam diretamente ■

4.1.16. Lema. *Seja $v \in \mathrm{T}_p S$ um vetor tangente a uma variedade S num ponto $p \in S$. Então existe um campo F sobre S tal que $F_p = v$.*

Demonstração. Se S admite uma parametrização $\varphi : H^m \cap U \to S$, o fato segue das Observações 4.1.7 e 4.1.8 e das Proposições 2.10.3 e 2.10.6. Assim, no caso geral, podemos supor que temos um campo F' sobre V com $F'_p = v$, onde $p \in V \subset\circ S$. Pelo lema 4.1.11, encontramos uma vizinhança $p \in V' \subset\circ V$ e uma função $\tilde{f} \in C^\infty(S)$ com suporte compacto tais que $\operatorname{supp} \tilde{f} \subset V$ e $\tilde{f}|_{V'} = 1$. O campo $F'' := (\tilde{f}|_V)F'$ sobre V se estende a um campo F sobre S, fazendo $F_y := F''_y$ se $y \in V$ e $F_y := 0$ se $y \in S \setminus V$. Realmente, a linearidade e a regra de Leibniz para F definido pela regra $(Fg)(y) := F_y g$ se verificam diretamente. Basta então entender que $Fg \in C^\infty(S)$ para todo $g \in C^\infty(S)$. Essa tarefa se realiza localmente: $(Fg)|_V = F''(g|_V) \in C^\infty(V)$ e $(Fg)|_{V''} = 0$, onde $V'' := S \setminus \operatorname{supp} \tilde{f} \subset\circ S$ e $S = V \cup V''$. Resta notar que $F_p = v$ ■

4.1.17. Lema. *Sejam S uma variedade, $p \in V \subset\circ S$ e F' um campo sobre V. Então existem uma vizinhança aberta $p \in V' \subset\circ V$ e um campo F sobre S tais que $F|_{V'} = F'|_{V'}$.*

Demonstração. Como na demonstração do lema anterior, encontramos pelo Lema 4.1.11 uma vizinhança $p \in V' \subset\circ V$ e uma função $\tilde{f} \in C^\infty(S)$ com suporte compacto tais que $\operatorname{supp} \tilde{f} \subset V$ e $\tilde{f}|_{V'} = 1$ e estendemos o campo $(\tilde{f}|_V)F'$ sobre V a um campo F sobre S ■

4.1.18. Lema. *Sejam F um campo sobre uma variedade S de dimensão m e $p \in S$ tais que $F_p = 0$. Então existem campos $F_1, \dots, F_{m+1}$ sobre S e funções $f_1, \dots, f_{m+1} \in C^\infty(S)$ tais que $F = \sum\limits_{i=1}^{m+1} f_i F_i$ e $f_i(p) = 0$ para todo $1 \leq i \leq m+1$.*

Demonstração. Se tomar no lugar de S uma vizinhança aberta $p \in V \subset\circ S$ que admite uma parametrização, pelas Proposição 2.10.6 e Observações 4.1.7 e 4.1.8, existem campos $F'_1, \dots, F'_m$ sobre V e funções $f'_1, \dots, f'_m \in C^\infty(V)$ tais que $f'_i(p) = 0$ para todo $1 \leq i \leq m$ e $F|_V = \sum\limits_{i=1}^{m} f'_i F'_i$. Diminuindo a vizinhança aberta V, pelos Lemas 4.1.11 e 4.1.16, podemos estender as funções e os campos. Assim, obtemos campos $F_1, \dots, F_m$ sobre S e funções $f_1, \dots, f_m \in C^\infty(S)$ tais que $F_{m+1}|_{V'} = 0$ e $f_i(p) = 0$ para todo $1 \leq i \leq m$, onde $F_{m+1} := F - \sum\limits_{i=1}^{m} f'_i F'_i$ e $p \in V' \subset\circ V$. Resta fazer $f_{m+1} = 1 - \tilde{f}$, onde $\tilde{f} \in C^\infty(S)$ com $\operatorname{supp} \tilde{f} \subset V'$ e $f(p) = 1$, pois $f_{m+1}F_{m+1} = F_{m+1}$ ■

4.2. Formas. Seja S uma variedade. Uma 0-forma sobre S é simplesmente uma funcão lisa sobre S. O conceito de 1-forma sobre S é dual ao de campo sobre S (por exemplo, o gradiente de uma função lisa é um campo de vetores cotangentes, ou seja, uma 1-forma). Em seguida, precisaremos de um conceito mais geral.

4.2.1. Definição. Sejam S uma variedade e $k \in \mathbb{N}$. Por definição, uma *k-forma* ω sobre S é algo que produz uma função $\omega(F_1, \dots, F_k) \in C^\infty(S)$ para quaisquer k campos $F_1, \dots, F_k$ sobre S satisfazendo as seguintes identidades[7]

$$\omega(F_1, \dots, F_{i-1}, F_i + F_i', F_{i+1}, \dots, F_k) =$$

$$\omega(F_1, \dots, F_{i-1}, F_i, F_{i+1}, \dots, F_k) + \omega(F_1, \dots, F_{i-1}, F_i', F_{i+1}, \dots, F_k),$$

$$\omega(F_1, \dots, F_{i-1}, fF_i, F_{i+1}, \dots, F_k) = f\omega(F_1, \dots, F_{i-1}, F_i, F_{i+1}, \dots, F_k)$$

para quaisquer $1 \le i \le k$, $f \in C^\infty(W)$ e campos $F_1, \dots, F_{i-1}, F_i, F_i', F_{i+1}, \dots, F_k$ sobre S.

Sejam ω, ω' k-formas sobre S, $F_1, \dots, F_k$ campos sobre S e $f \in C^\infty(S)$. Então as fórmulas

$$(f\omega)(F_1, \dots, F_k) := f\omega(F_1, \dots, F_k), \qquad (\omega + \omega')(F_1, \dots, F_k) := \omega(F_1, \dots, F_k) + \omega'(F_1, \dots, F_k)$$

definem as k-formas $f\omega$ e $\omega + \omega'$ sobre S.

4.2.2. Lema. *Seja ω uma k-forma sobre uma variedade S. Então, para quaisquer campos $F^1, \dots, F^k$ sobre S e todo $p \in S$, o número $\omega(F^1, \dots, F^k)(p) \in \mathbb{R}$ depende apenas de p e dos vetores tangentes $F^1_p, \dots, F^k_p \in \mathrm{T}_pS$, $\omega(F^1, \dots, F^k)(p) = \omega(F^1_p, \dots, F^k_p)$. Além disso, podemos restringir ω para qualquer subconjunto aberto $V \subset\circ S$, obtendo uma k-forma $\omega|_V$ sobre a variedade V.*

Demonstração. Para mostrar a primeira afirmação, podemos, mantendo os vetores tangentes $F^1_p, \dots, F^k_p$, trocar um por um os campos $F^1, \dots, F^k$. Pela primeira identidade na Definição 4.2.1, basta entender que $\omega(F^1, \dots, F^k)(p) = 0$ se F^i_p para algum $1 \le i \le k$. Isto segue do Lema 4.1.18 e da Definição 4.2.1.

Note que, pelo Lema 4.1.16, o número $\omega(v_1, \dots, v_k) \in \mathbb{R}$ é definido para todos $v_1, \dots, v_k \in \mathrm{T}_pS$.

A segunda afirmação segue da primeira: para quaisquer campos $F^1, \dots, F^k$ sobre V e todo $p \in V$, definimos $\omega(F^1, \dots, F^k)(p) := \omega(F^1_p, \dots, F^k_p)$ e verificamos localmente que $\omega(F^1, \dots, F^k) \in C^\infty(V)$ usando o Lema 4.1.17 e a primeira afirmação ■

4.2.3. Definição. Seja $\varphi : S_1 \to S_2$ uma aplicação lisa entre variedades e ω uma k-forma sobre S_2. Para todos $p \in S_1$ e $v_1, \dots, v_k \in \mathrm{T}_pS_1$, definamos $\varphi^*\omega(v_1, \dots, v_k) := \omega(\varphi v_1, \dots, \varphi v_k)$. Pelos Definição 4.2.1 e Lema 4.2.2, essa fórmula é correta e $\varphi^*\omega(v_1, \dots, v_k)$ é linear em cada v_i (para p fixo). Sejam $F^1, \dots, F^k$ campos sobre S_1. Para todo $p \in S_1$, façamos $\varphi^*(F^1, \dots, F^k) = \varphi^*\omega(F^1_p, \dots, F^k_p)$

[7]Rigorosamente falando, uma k-forma deve ainda ser *alternante,* o que será exigido mais tarde.

Índice

Se alguém não está conseguindo lembrar estes assuntos,
talvez ainda não nasceu ou não tem
sua (seu) namorada (namorado) na platéia.

0-forma 36-4b
k-forma (não alternada) 37-1
n-bloco fechado 23-4
aplicação composta 2-15b
aplicação contínua 2-16
aplicação de classe C^0 8-9b
aplicação de classe C^1 8-6b
aplicação de classe C^∞ 8-4b
aplicação de classe C^k 8-5b
aplicação derivável num ponto 7-10
aplicação derivável 8-9b
aplicação elementar 29-10b
aplicação induzida de espaços tangentes 34-10
aplicação lisa entre variedades 33-20
aplicação pequena em comparação com $|h|$ 7-12
axioma de Hausdorff 1-19
bloco fechado 4-19b
bola aberta 1-9
bordo de uma variedade 32-20
bordo do semiespaço 32-19
caminho de classe C^0 5-4b
caminho de classe C^1 6-1
caminho de classe C^∞ 6-2
caminho de classe C^k 6-1
caminho diferenciável 5-2b
caminho 5-3b
campo sobre uma variedade 35-13b
campo vetorial suave 11-3b
cobertura aberta 27-15b
cobertura finita 27-13b
compacto 4-21
componentes de uma aplicação 2-17b
comutador de campos 12-4, 35-10b
conjunto de nível 18-9
convergência uniforme de funções 4-10b
coordenadas cilíndricas 31-15
coordenadas esféricas 31-13b
coordenadas polares 31-12
curva integral de um campo 14-2
curva parametrizada 5-3b
derivada de uma aplicação num ponto 7-18
derivada direcional 11-16
desigualdade triangular 1-6
equação linear do espaço tangente 20-1b, 21-9b
espaço tangente 11-12
fecho de um subconjunto 27-12b
função característica de um subconjunto 28-8b
função de classe C^∞ sobre subconjunto 21-19
função de projeção 3-17b
função integrável 5-1, 24-1
função que preserva limites 3-1
função uniformemente contínua 4-18
gráfico de uma função 21-7
gradiente de uma função 21-10
imagem de um vetor tangente 33-6b
imagem inversa 2-17
integral de Riemann 23-4, 24-1, 28-18b, 28-11b
integral imprópria 28-7b
integral inferior 23-13b
integral superior 23-13b
limite de uma sequência 2-8
limite uniforme 4-9b
mínimo local de uma função 9-1
mínimos condicionados 18-7
máximo local de uma função 9-2
máximos condicionados 18-7
matriz hessiana de uma função 9-13b
matriz jacobiana 8-15, 8-12b
matriz não-degenerada 19-15
matriz simétrica positiva 9-19b
mudança de variáveis 28-19, 28-4b
multiplicadores de Lagrange 18-13
norma de uma aplicação linear 16-18b
norma de uma função 4-16b
norma em $\mathbb{R}^n$ 6-18
normas equivalentes 6-19b
operador diferencial de ordem 1 11-3b
orientação de um espac co linear 35-19
orientacão de uma variedade 35-21
parametrização de uma variedade 32-16
parametrização lisa local 32-18
parametrização local de uma variedade 32-17
partição da unidade subordinada 27-2b
partição da unidade 27-16b, 27-5b, 27-2b
partição de um n-bloco 23-11
partição de um segmento 23-8
permutação de variáveis 29-16
polinômio de Taylor de uma função 14-16b, 14-2b

ponto crítico de uma função 15-20b
ponto de sela no gráfico de uma função 15-11b
ponto liso de um conjunto de nível 20-8b
posto de uma matriz 18-12
propriedade local de uma aplicação 2-2b
refinamento de uma partição 23-14
regra da cadeia 7-13b
regra de Leibniz 11-4, 11-1b
restrição de uma aplicação 2-21b
semiespaço 32-15
sequência convergente 2-7
soma de k-formas 37-9
soma de Riemann inferior 23-15b
soma de Riemann superior 23-15b
soma de campos 12-3, 35-10b
subcobertura 27-14b
subconjunto aberto 1-8
subconjunto fechado 1-10
subconjunto limitado 4-20
suporte de uma função 27-10b
supremo 4-15b, 4-2b
teorema da função implícita 17-3, 17-11, 17-14
teorema da função inversa 16-7b, 16-2b
teorema de Fubini 26-12b
topologia induzida 1-6b
topologia padrão em $\mathbb{R}^n$ 1-20
translação 29-18
variedade com bordo 32-12
variedade sem bordo 32-20
variedade 32-12
vetor aceleração 6-7
vetor tangente a U em $p \in U$ 10-1b
vetor tangente a um subconjunto 21-22
vetor velocidade 6-6
vizinhança aberta 1-18
volume de um n-bloco 23-5